



(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2019-0003939
(43) 공개일자 2019년01월10일

- (51) 국제특허분류(Int. Cl.)
C12P 7/44 (2006.01) C12N 15/52 (2006.01)
C12N 9/88 (2006.01)
- (52) CPC특허분류
C12P 7/44 (2013.01)
C12N 15/52 (2013.01)
- (21) 출원번호 10-2018-7028248
- (22) 출원일자(국제) 2017년03월01일
심사청구일자 없음
- (85) 번역문제출일자 2018년10월01일
- (86) 국제출원번호 PCT/US2017/020263
- (87) 국제공개번호 WO 2017/151811
국제공개일자 2017년09월08일
- (30) 우선권주장
62/302,558 2016년03월02일 미국(US)

- (71) 출원인
피티티지씨 이노베이션 아메리카 코포레이션
미국, 메사추세츠 01801 우번, 42 커밍스 파크
- (72) 발명자
실러스, 라이언
미국, 메사추세츠 01890, 포레스트 서클 윈체스터 17
스펜서, 미셸
미국, 메사추세츠 01463, 오크 힐 스트리트 페페렐 82
(뒷면에 계속)
- (74) 대리인
허용록

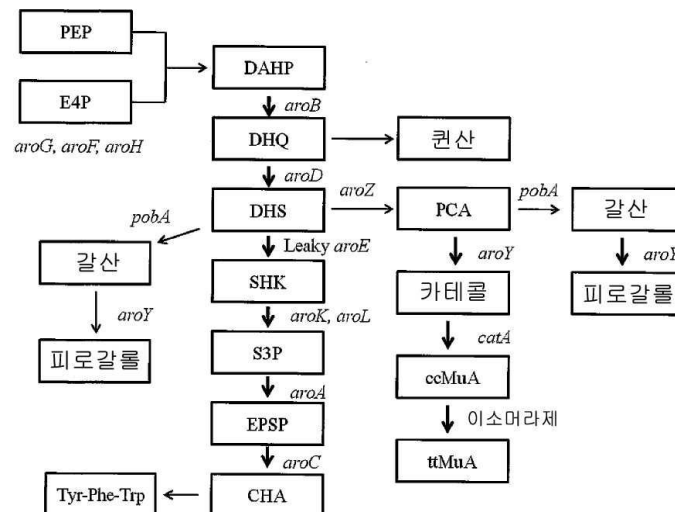
전체 청구항 수 : 총 30 항

(54) 발명의 명칭 유전자 조작된 미생물로부터 개선된 뮤콘산 생산

(57) 요약

본 발명의 주제는 발효에 의한 뮤콘산의 생물학적 생산의 수율 및 역가의 개선이다. 뮤콘산 경로에 관여하는 하나 이상의 효소의 증가된 활성은 뮤콘산의 생산 증가로 이어진다.

대표도 - 도2



(52) CPC특허분류

C12N 9/88 (2013.01)
C12Y 101/01025 (2013.01)
C12Y 205/01054 (2013.01)
C12Y 207/01002 (2013.01)
C12Y 401/01031 (2013.01)
C12Y 401/01063 (2013.01)
C12Y 604/01001 (2013.01)

(72) 발명자

요쿰, 알., 로저스

미국, 메사추세츠 02420, 오차드 스트리트 렉싱턴
4

헤르만, 테론

미국, 메사추세츠 02474, 워렌 스트리트 알링턴
151

러셀, 유다니

미국, 메사추세츠 02476, 뉴포트 스트리트 알링턴
72

명세서

청구범위

청구항 1

3,4-디히드록시벤조산 탈탄산효소를 인코딩하는 적어도 하나의 외인성 유전자 및 상기 외인성 3,4-디히드록시벤조산 탈탄산효소의 활성을 증가시키는 단백질을 코딩하는 적어도 하나의 다른 유전자를 포함하는 비-방향족 탄소원으로부터 류콘산을 생산하는 유전자 조작된 미생물.

청구항 2

제 1항에 있어서, 상기 유전자 조작된 미생물은 72시간에 적어도 60g/L의 류콘산을 생산하는 것을 특징으로 하는, 유전자 조작된 미생물.

청구항 3

제 1항에 있어서, 상기 유전자 조작된 미생물은 박테리아인 것을 특징으로 하는, 유전자 조작된 미생물.

청구항 4

제 1항에 있어서, 상기 유전자 조작된 미생물은 대장균(*Escherichia coli*)인 것을 특징으로 하는, 유전자 조작된 미생물.

청구항 5

제 1항에 있어서, 상기 3,4-디히드록시벤조산 탈탄산효소의 활성의 증가는 *kpdB* 유전자 또는 이의 동족체에 의해 인코딩된 KpdB 단백질의 발현에 의해 달성되고, 상기 *kpdB* 유전자의 상기 동족체는 KpdB 단백질과 적어도 25% 아미노산 동일성을 갖는 단백질을 인코딩하는 것을 특징으로 하는, 유전자 조작된 미생물.

청구항 6

제 1항에 있어서, 상기 3,4-디히드록시벤조산 탈탄산효소의 활성의 증가는 *UbiX* 유전자 또는 이의 동족체에 의해 인코딩된 UbiX 단백질의 발현에 의해 달성되고, 상기 *UbiX* 유전자의 상기 동족체는 UbiX 단백질과 적어도 25% 아미노산 동일성을 갖는 단백질을 인코딩하는 것을 특징으로 하는, 유전자 조작된 미생물.

청구항 7

제 1항에 있어서, 누출성 쉬키메이트 탈수소효소를 인코딩하는 유전자를 더 포함하고, 상기 누출성 쉬키메이트 탈수소효소는 방향족 아미노산 및 비타민에 대해 프로토포피(protothrophy)를 부여하나 상기 방향족 아미노산 또는 비타민의 현저한 분비를 야기하지 않는 것을 특징으로 하는, 유전자 조작된 미생물.

청구항 8

제 1항에 있어서, *aroZ*, *qa-4*, *asbF*, *aroY*, *quiC* 및 *catAX*로 이루어진 군으로부터 선택된 하나 이상의 외인성 유전자를 더 포함하고, 상기 외인성 유전자는 류콘산 경로에서 기능하는 단백질을 인코딩하는 것을 특징으로 하는, 유전자 조작된 미생물.

청구항 9

제 1항에 있어서, *aroB*, *aroD*, *aroF*, *aroG*, *aroH*, *tktA*, *talB*, *rpe*, 및 *rpi*로 이루어진 군으로부터 선택된 하나 이상의 외인성 유전자를 더 포함하고, 상기 외인성 유전자는 쉬킴산 경로에서 기능하는 단백질을 인코딩하는 것을 특징으로 하는, 유전자 조작된 미생물.

청구항 10

제 1항에 있어서, *tyrR* 유전자 또는 이의 동족체에 의해 인코딩된 방향족 아미노산 생합성의 음성 조절자 단백질의 활성은 상당히 감소되거나 또는 제거되는 것을 특징으로 하는, 유전자 조작된 미생물.

청구항 11

제 1항에 있어서, 페닐알라닌에 의한 저해에 대해 상당히 내성이 있는 DAHP 합성효소를 인코딩하는 *aroG** 유전자를 더 포함하는, 유전자 조작된 미생물.

청구항 12

제 11항에 있어서, DAHP 합성효소를 인코딩하는 *aroG** 유전자는 *aroG*20-893*, *aroG*20-897*, *aroG*20-899*, *aroG*20-901*, *aroG*111*, *aroG*211*, *aroG*212*, *aroG*311*, *aroG*312*, *aroG*411*, *aroG*412*, 및 *aroG*511*로 이루어진 군으로부터 선택되는 것을 특징으로 하는, 유전자 조작된 미생물.

청구항 13

제 1항에 있어서, 포스포에놀 피루베이트 의존성 포스포트랜스퍼라제 시스템 및 글루코오스 이입을 위한 GalP 단백질 기반 시스템의 활성은 감소되고, 유전자 조작된 미생물은 기능성 글루코오스-촉진 확산 단백질을 인코딩하는 외인성 유전자를 더 포함하는 것을 특징으로 하는, 유전자 조작된 미생물.

청구항 14

제 13항에 있어서, 기능성 글루코오스-촉진 확산 단백질은 *glf* 유전자 또는 이의 동족체에 의해 코딩된 Glf 단백질인 것을 특징으로 하는, 유전자 조작된 미생물.

청구항 15

제 13항에 있어서, 글루코키나제를 인코딩하는 외인성 유전자를 더 포함하는, 유전자 조작된 미생물.

청구항 16

피루베이트 카복실라제 효소를 인코딩하는 적어도 하나의 외인성 유전자를 포함하는 비-방향족 탄소원으로부터 뮈콘산을 생산하는 유전자 조작된 미생물.

청구항 17

제 16항에 있어서, 상기 유전자 조작된 미생물은 72시간에 적어도 60g/L의 뮈콘산을 생산하는 것을 특징으로 하는, 유전자 조작된 미생물.

청구항 18

제 16항에 있어서, 상기 유전자 조작된 미생물은 박테리아인 것을 특징으로 하는, 유전자 조작된 미생물.

청구항 19

제 16항에 있어서, 상기 유전자 조작된 미생물은 대장균(*Escherichia coli*)인 것을 특징으로 하는, 유전자 조작된 미생물.

청구항 20

제 16항에 있어서, 상기 피루베이트 카복실라제 효소를 인코딩하는 외인성 유전자는 *사카로마이세스 세레비지애* (*Saccharomyces cerevisiae*)로부터의 *pyc* 유전자 또는 이의 동족체에 의해 코딩되며, 상기 *pyc* 유전자의 상기 동족체는 *사카로마이세스 세레비지애*의 *pyc* 유전자에 의해 코딩된 카복실라제 효소와 적어도 25% 아미노산 동일성을 갖는 단백질을 인코딩하는 것을 특징으로 하는, 유전자 조작된 미생물.

청구항 21

제 16항에 있어서, 포스포에놀 피루베이트 카복실라제 활성의 감소 또는 제거를 위한 돌연변이를 더 포함하는, 유전자 조작된 미생물.

청구항 22

제 16항에 있어서, *사카로마이세스 세레비지애*로부터의 외인성 *pyc* 유전자는 포스포에놀 피루베이트 카복실라제

를 인코딩하는 *ppc* 유전자의 유전자좌에 삽입되는 것을 특징으로 하는, 유전자 조작된 미생물.

청구항 23

제 16항에 있어서, *aroZ*, *qa-4*, *asbF*, *aroY*, *quiC* 및 *catAX*로 이루어진 군으로부터 선택된 하나 이상의 외인성 유전자 또는 이의 동족체를 더 포함하고, 상기 외인성 유전자는 류콘산 경로에서 기능하는 단백질을 인코딩하는 것을 특징으로 하는, 유전자 조작된 미생물.

청구항 24

제 16항에 있어서, *aroB*, *aroD*, *aroF*, *aroG*, *aroH*, *tktA*, *talB*, *rpe*, 및 *rpi*로 이루어진 군으로부터 선택된 하나 이상의 외인성 유전자 또는 이의 동족체를 더 포함하고, 상기 외인성 유전자는 쉬킴산 경로에서 기능하는 단백질을 인코딩하는 것을 특징으로 하는, 유전자 조작된 미생물.

청구항 25

제 16항에 있어서, 3,4-디히드록시벤조산 탈탄산효소를 인코딩하는 적어도 하나의 외인성 유전자 또는 이의 동족체, 및 상기 외인성 3,4-디히드록시벤조산 탈탄산효소의 활성을 증가시키는 단백질을 코딩하는 적어도 하나의 다른 유전자를 더 포함하는, 유전자 조작된 미생물.

청구항 26

제 25항에 있어서, 상기 3,4-디히드록시벤조산 탈탄산효소의 활성의 증가는 *kpdB* 유전자 또는 이의 동족체에 의해 인코딩된 KpdB 단백질의 발현에 의해 달성되고, 상기 *kpdB* 유전자의 동족체는 KpdB 단백질과 적어도 25% 아미노산 동일성을 갖는 단백질을 인코딩하는 것을 특징으로 하는, 유전자 조작된 미생물.

청구항 27

제 25항에 있어서, 상기 3,4-디히드록시벤조산 탈탄산효소의 활성의 증가는 *UbiX* 유전자 또는 이의 동족체에 의해 인코딩된 UbiX 단백질의 발현에 의해 달성되고, 상기 *UbiX* 유전자의 동족체는 UbiX 단백질과 적어도 25% 아미노산 동일성을 갖는 단백질을 인코딩하는 것을 특징으로 하는, 유전자 조작된 미생물.

청구항 28

제 16항에 있어서, 포스포에놀 피루베이트 의존성 포스포트랜스퍼라제 시스템 및 글루코오스 이입을 위한 GalP 단백질 기반 시스템의 활성은 감소되고, 유전자 조작된 미생물은 기능성 글루코오스-축진 확산 단백질을 인코딩하는 외인성 유전자를 더 포함하는 것을 특징으로 하는, 유전자 조작된 미생물.

청구항 29

제 28항에 있어서, 기능성 글루코오스-축진 확산 단백질은 *glf* 유전자에 의해 코딩된 Glf 단백질인 것을 특징으로 하는, 유전자 조작된 미생물.

청구항 30

제 29항에 있어서, 글루코키나제를 인코딩하는 외인성 유전자를 더 포함하는, 유전자 조작된 미생물.

발명의 설명

기술 분야

[0001] 관련 출원에 대한 상호-참조

[0002] 본 출원은 2016년 3월 2일에 출원된 미국 가출원 번호 제 62/302,558호의 우선권을 주장한다.

[0003] 기술분야

[0004] 본 발명은 중심 방향족 생합성 경로에서 유전자 조작된 생촉매를 이용한 재생 가능한 화학 원료를 생산하는 분야에 있다. 보다 구체적으로는, 본 발명은 유전자 변형된 생촉매를 이용한 재생 가능한 탄소 자원으로부터 류콘산의 생산을 개선하는 방법을 제공한다.

배경 기술

- [0005] 아디프산은 주요 범용 화합물이며, 나일론 6,6 및 폴리우레탄의 생산에 사용된다. 아디프산은 현재 석유화학 원료에서 유래된다. 아디프산의 현재 합성은 환경적으로 유해한 아질산(nitrous acid)을 방출한다 (Xie et al., 2014). 대안적으로, 아디프산은 화학적 수소화에 의해 뮈콘산의 3개의 이성질체 (시스, 시스; 시트, 트랜스; 트랜스, 트랜스 이성질체) 중 어느 것으로부터 제조될 수 있다. 미생물과 함께 발효에 의해 재생 가능한 자원으로부터 뮈콘산을 생산한 다음, 수소화 공정을 통해 아디프산을 생산하는 것이 바람직한데, 그 이유는 아디프산으로의 이러한 경로가 종래의 석유화학적 경로보다 더 환경 친화적이기 때문이다 (Niu, Draths and Frost, 2002; Frost and Draths, 1997). 많은 기타 화합물들은 1,6-헥산 디올, 3-헥센디카복실산, 1,6-헥산디아민, 및 테레프탈산을 포함하나 이에 한정되지 않는 하나 이상의 뮈콘산 이성질체의 화학적 전환에 의해 제조될 수 있다.
- [0006] 국제특허출원 공개번호 WO 2011/017560는 뮈코네이트 경로를 갖는 생촉매 및 이들 생촉매를 이용한 뮈콘산의 생산 방법을 주장한다. 요컨대, 이 공개된 특허 출원은 뮈콘산의 생산을 위한 4가지 상이한 경로를 개시한다. 뮈콘산 생산을 위한 제 1 경로는 석시닐-CoA 및 아세틸-CoA로 시작한다. 뮈콘산 생산을 위한 제 2 경로는 피루베이트 및 말로네이트 세미알데히드로 시작한다. 뮈콘산 생산을 위한 제 3 경로는 피루베이트 및 석시닉 세미알데히드로 시작한다. 뮈콘산 생산을 위한 제 4 경로는 리신으로 시작한다. 이 특허 출원에서 제안된 뮈콘산 생산을 위한 모든 이러한 경로는 컴퓨터 모델링에 기초하며, 이러한 생촉매가 뮈콘산에 대해 상업적으로 허용가능한 생산성 및 수율로 제조될 수 있는지 여부는 아직 밝혀지지 않았다.
- [0007] 유전자 조작된 *E. coli* 시스템을 이용한 시스, 시스-뮈콘산을 생산하는 발효 경로는 과학 문헌 (Niu et al., 2002; Frost and Draths, 1997) 및 특허 문헌 (US 5,487,987; US 5,616,496; WO 2011/085311 A1)에 기재되어 있다. 그러나, 뮈콘산 생산을 위한 선행기술 공정은 고가의 중간 성분 (방향족 아미노산 및 비타민) 및 화학 유도제와 같은 중요한 단점, 및 산업 생산에 요구되는 것보다 낮은 수율로 인해 어려움을 겪었다. 최근의 미국특허출원 공개번호 US2015/0044755는 관련 유전자의 항시성 발현, 개선된 이중 유전자 및 신규한 "누출성(leaky)" AroE 효소를 포함하는 뮈콘산 생산을 위한 개선된 생촉매를 제공하며, 전체 참조로 본 명세서에 포함된다. 본 발명은 미국특허출원 공개번호 US2015/0044755에 기재된 뮈콘산 생촉매의 추가 개선점을 제공한다.
- [0008] 본 명세서에 포함된 모든 특허, 특허출원, 간행물, 서열 및 기타 공개된 자료는 참고 문헌으로 포함된다.

선행기술문헌

특허문헌

- [0009] (특허문헌 0001) 미국특허 번호 4,480,034
 (특허문헌 0002) 미국특허 번호 4,535,059
 (특허문헌 0003) 미국특허 번호 4,588,688
 (특허문헌 0004) 미국특허 번호 4,608,338
 (특허문헌 0005) 미국특허 번호 4,681,852
 (특허문헌 0006) 미국특허 번호 4,753,883
 (특허문헌 0007) 미국특허 번호 4,833,078
 (특허문헌 0008) 미국특허 번호 4,968,612
 (특허문헌 0009) 미국특허 번호 5,168,056
 (특허문헌 0010) 미국특허 번호 5,272,073
 (특허문헌 0011) 미국특허 번호 5,487,987
 (특허문헌 0012) 미국특허 번호 5,616,496
 (특허문헌 0013) 미국특허 번호 6,600,077
 (특허문헌 0014) 미국특허 번호 6,180,373

(특허문헌 0015) 미국특허 번호 6,210,937
 (특허문헌 0016) 미국특허 번호 6,472,169
 (특허문헌 0017) 미국특허 번호 6,613,552
 (특허문헌 0018) 미국특허 번호 6,962,794
 (특허문헌 0019) 미국특허 번호 7,244,593
 (특허문헌 0020) 미국특허 번호 7,638,312
 (특허문헌 0021) 미국특허 번호 7,790,431
 (특허문헌 0022) 미국특허 번호 8,871,489
 (특허문헌 0023) 미국특허 번호 9,017,976
 (특허문헌 0024) 미국특허출원 공개번호 US 2009/0191610 A1
 (특허문헌 0025) 미국특허출원 공개번호 US 2010/0314243 A1
 (특허문헌 0026) 미국특허출원 공개번호 US 2013/0337519 A1
 (특허문헌 0027) 미국특허출원 공개번호 US. 2014/0234923A1
 (특허문헌 0028) 미국특허출원 공개번호 US 2015/0044755 A1
 (특허문헌 0029) 미국특허출원 공개번호 US2016/0017381 A1
 (특허문헌 0030) 유럽특허출원 번호 86300748.0
 (특허문헌 0031) 국제특허출원 공개번호 WO 2011/017560
 (특허문헌 0032) 국제특허출원 공개번호 WO 2011/085311
 (특허문헌 0033) 국제특허출원 공개번호 WO 2011/123154
 (특허문헌 0034) 국제특허출원 공개번호 WO2013/116244

비특허문헌

[0010]

(비특허문헌 0001) Altschul, S. F., Gish, W., Miller, W., Myers, E. W., and Lipman, D. J. (1990) Basic local alignment search tool, J Mol Biol 215, 403-410.
 (비특허문헌 0002) Altschul, S. F., Madden, T. L., Schaffer, A. A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D. J. (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs, Nucleic Acids Res 25, 3389-3402.
 (비특허문헌 0003) Aussel, Laurent, Fabien Pierrel, Laurent Loiseau, Murielle Lombard, Marc Fontecave, and Frederic Barras. 2014. "Biosynthesis and Physiology of Coenzyme Q in Bacteria." Biochimica et Biophysica Acta - Bioenergetics 1837 (7): 1004-11.
 (비특허문헌 0004) Baba, Tomoya, Takeshi Ara, Miki Hasegawa, Yuki Takai, Yoshiko Okumura, Miki Baba, Kirill A Datsenko, Masaru Tomita, Barry L Wanner, and Hirotada Mori. 2006. "Construction of Escherichia Coli K-12 in-Frame, Single-Gene Knockout Mutants: The Keio Collection." Molecular Systems Biology 2: 2006.0008. doi:10.1038/msb4100050.
 (비특허문헌 0005) Barbe, V., Vallenet, D., Fonknechten, N., Kreimeyer, A., Oztas, S., Labarre, L., Cruveiller, S., Robert, C., Duprat, S., Wincker, P., Ornston, L. N., Weissenbach, J., Marliere, P., Cohen, G. N., and Medigue, C. (2004) Unique features revealed by the genome sequence of Acinetobacter sp. ADP1, a versatile and naturally transformation competent bacterium, Nucleic Acids Res 32, 5766-5779.

- (비특허문헌 0006) Bird, J. A. and Cain, R. B. (1968) cis-cis-muconate, the product inducer of catechol 1,2-oxygenase in *Pseudomonas aeruginosa*. *Biochem. J.* 109, 479-481.
- (비특허문헌 0007) Bongaerts, J., Kramer, M., Muller, U., Raven, L. and Wubbolts, M. (2001) Metabolic engineering for microbial production of aromatic acids and derived compounds. *Met. Eng.* 3, 289-300.
- (비특허문헌 0008) Chandran, S. S., Yi, J., Draths, K. M., von Daeniken, R., Weber, W. and Frost, J. W. (2003) Phosphoenolpyruvate availability and the biosynthesis of shikimic acid. *Biotechnol. Prog.* 19, 808-814.
- (비특허문헌 0009) Chen, R., Hatzimanikatis, V., Yap, W. M. G. J., Potma, P. W. and Bailey, J. E. (1997) Metabolic consequences of phosphotransferase (PTS) mutation in a phenylalanine-producing recombinant *Escherichia coli*. *Biotechnol. Prog.* 13, 768-775.
- (비특허문헌 0010) Chen, K., Dou, J., Tang, S., Yang, Y., Wang, H., Fang, H. and Zhou, C. (2012) Deletion of the *aroK* gene is essential for high shikimic acid accumulation through the shikimate in *E. coli*. *Bioresour. Technol.* 119, 141-147.
- (비특허문헌 0011) Choi, W. J., Lee, E. Y., Cho, M. H., and Choi, C. Y. (1997) Enhanced production of cis, cis-muconate in a cell-recycle bioreactor. *J. Fermentation and Bioengineering.* 84, 70-76.
- (비특허문헌 0012) Curran, Kathleen A., John M. Leavitt, Ashty S. Karim, and Hal S. Alper. 2013. "Metabolic Engineering of Muconic Acid Production in *Saccharomyces Cerevisiae*." *Metabolic Engineering* 15 (1): 55-66.
- (비특허문헌 0013) de Berardinis, V., Vallenet, D., Castelli, V., Besnard, M., Pinet, A., Cruaud, C., Samair, S., Lechaplais, C., Gyapay, G., Richez, C., Durot, M., Kreimeyer, A., Le Fevre, F., Schachter, V., Pezo, V., Doring, V., Scarpelli, C., Medigue, C., Cohen, G. N., Marliere, P., Salanoubat, M., and Weissenbach, J. (2008) A complete collection of single-gene deletion mutants of *Acinetobacter baylyi* ADP1, *Mol Syst Biol* 4, 174.
- (비특허문헌 0014) Draths, K. M., Pompliano, D. L., Conley, D. L., Frost, J. W., Berry, A., Disbrow, G. L., Staversky, R. J., and Lievens, J. C. (1992) Biocatalytic Synthesis of Aromatics from D-Glucose - the Role of Transketolase, *Journal of the American Chemical Society* 114, 3956-3962.
- (비특허문헌 0015) Draths, K. M., and Frost, J. W. (1995) Environmentally Compatible Synthesis of Catechol from D-Glucose, *Journal of the American Chemical Society* 117, 2395-2400.
- (비특허문헌 0016) Elsemore, D. A., and Ornston, L. N. (1995) Unusual ancestry of dehydratases associated with quinate catabolism in *Acinetobacter calcoaceticus*, *J Bacteriol* 177, 5971-5978.
- (비특허문헌 0017) Escalante, A., Calderon, R., Valdivia, A., de Anda, R., Hernandez, G., Ramirez, O. T., Gosset, G. and Bolivar, F. (2010) Metabolic engineering for the production of shikimic acid in an evolved *Escherichia coli* strain lacking the phosphoenolpyruvate: carbohydrate phosphotransferase system. *Microbial Cell Factories* 9, 21-33.
- (비특허문헌 0018) Escalante, Adelfo, Rocio Calderon, Araceli Valdivia, Ramon de Anda, Georgina Hernandez, Octavio T Ramirez, Guillermo Gosset, and Francisco Bolivar. 2010. "Metabolic Engineering for the Production of Shikimic Acid in an Evolved *Escherichia Coli* Strain Lacking the Phosphoenolpyruvate: Carbohydrate Phosphotransferase System." *Microbial Cell Factories* 9 (Ccm): 21. doi:10.1186/1475-2859-9-21.
- (비특허문헌 0019) Flores, N., Xiao, J., Berry, A., Bolivar, F. and Valle, F. (1996) Pathway engineering for the production of aromatic compounds in *Escherichia coli*. *Nature Biotechnol.* 14, 620 - 623.
- (비특허문헌 0020) Fox, D. T., Hotta, K., Kim, C. Y., and Koppisch, A. T. (2008) The missing link in petrobactin biosynthesis: *asbF* encodes a (-)-3-dehydroshikimate dehydratase, *Biochemistry* 47, 12251-

12253.

(비특허문헌 0021) Ger, Y., Chen, S., Chiang, H., and Shiuan, D. (1994) A Single Ser-180 Mutation Desensitizes Feedback Inhibition of the Phenylalanine-Sensitive 3-Deoxy-D-Arabinose-7-Phosphate (DAHP) Synthetase in *Escherichia coli*, *J Biochem* 116, 986-990.

(비특허문헌 0022) Grant, D. J., and Patel, J. C. (1969) The non-oxidative decarboxylation of p-hydroxybenzoic acid, gentisic acid, protocatechuic acid and gallic acid by *Klebsiella aerogenes* (*Aerobacter aerogenes*), *Antonie Van Leeuwenhoek* 35, 325-343.

(비특허문헌 0023) Hansen, E. H., Moller, B. L., Kock, G. R., Bunner, C. M., Kristensen, C., Jensen, O. R., Okkels, F. T., Olsen, C. E., Motawia, M. S., and Hansen, J. (2009) De novo biosynthesis of vanillin in fission yeast (*Schizosaccharomyces pombe*) and baker's yeast (*Saccharomyces cerevisiae*), *Appl Environ Microbiol* 75, 2765-2774.

(비특허문헌 0024) Horwitz, Andrew A., Jessica M. Walter, Max G. Schubert, Stephanie H. Kung, Kristy Hawkins, Darren M. Platt, Aaron D. Hernday, et al. 2015. "Efficient Multiplexed Integration of Synergistic Alleles and Metabolic Pathways in Yeasts via CRISPR-Cas." *Cell Systems*. 1(1): 88-96.

(비특허문헌 0025) Hu, Changyun, Peihong Jiang, Jianfeng Xu, Yongqing Wu, and Weida Huang. 2003. "Mutation Analysis of the Feedback Inhibition Site of Phenylalanine-Sensitive 3-Deoxy-D-Arabinose-Heptulosonate 7-Phosphate Synthase of *Escherichia coli*." *Journal of Basic Microbiology* 43 (5): 399-406.

(비특허문헌 0026) Hu, C., Jiang, P., Xu, J., Wu, Y., and Huang, W. (2003) Mutation analysis of the feedback inhibition site of phenylalanine-sensitive 3-deoxy-D-arabinose-heptulosonate 7-phosphate synthase of *Escherichia coli*, *J Basic Microbiol* 43, 399-406.

(비특허문헌 0027) Iwagami, S. G., Yang, K., and Davies, J. (2000) Characterization of the protocatechuic acid catabolic gene cluster from *Streptomyces* sp. strain 2065, *Appl Environ Microbiol* 66, 1499-1508.

(비특허문헌 0028) Jantama, K., Haupt, M. J., Svoronos, S. A., Zhang, X., Moore, J. C., Shanmugam, K. T., and Ingram, L. O. (2008a) Combining metabolic engineering and metabolic evolution to develop nonrecombinant strains of *Escherichia coli* C that produce succinate and malate, *Biotechnol Bioeng* 99, 1140-1153.

(비특허문헌 0029) Jantama, K., Zhang, X., Moore, J. C., Shanmugam, K. T., Svoronos, S. A., and Ingram, L. O. (2008b) Eliminating side products and increasing succinate yields in engineered strains of *Escherichia coli* C, *Biotechnol Bioeng* 101, 881-893.

(비특허문헌 0030) Jimenez, Natalia, Jose Antonio Curiel, Ines Reveron, Blanca de las Rivas, and Rosario Munoz. 2013. "Uncovering the *Lactobacillus Plantarum* WCFS1 Gallate Decarboxylase Involved in Tannin Degradation." *Applied and Environmental Microbiology* 79 (14): 4253-63.

(비특허문헌 0031) Johnson, C. W., Salvachua, D., Khanna, P., Peterson, D. J. and Beckham, G. (2016) Enhancing muconic acid production from glucose and lignin-derived aromatic compound via increased protocatechuic acid decarboxylase activity. *Metabolic Engineering Communication*. 3: 111-119.

(비특허문헌 0032) Kaneko, A., Ishii, Y., and Kirimura, K. (2011) High-yield production of cis, cis-muconic acid from catechol in aqueous solution by biocatalyst. *Chem. Lett.* 40, 381-383.

(비특허문헌 0033) Kikuchi, Y., Tsujimoto, K., and Kurahashi, O. (1997) Mutational analysis of the feedback sites of phenylalanine-sensitive 3-deoxy-D-arabinose-heptulosonate-7-phosphate synthase of *Escherichia coli*, *Appl Environ Microbiol* 63, 761-762.

(비특허문헌 0034) Kojima, Y., Fujisawa, H., Nakazawa, A., Nakazawa, T., Kanetsuna, F., Taniuchi, H., Nozaki, M., and Hayaishi, O. (1967) Studies on pyrocatechase. I. Purification and spectral properties, *J Biol Chem* 242, 3270-3278.

- (비특허문헌 0035) Kramer, M., Bongaerts, J., Bovenberg, R., Kremer, S., Muller, U., Orf, S., Wubbolts, M. and Raeven, L. (2003) Metabolic engineering for microbial production of shikimic acid. *Metabol. Eng.* 5, 277-283.
- (비특허문헌 0036) Lerner, C. G., and Inouye, M. (1990) Low copy number plasmids for regulated low-level expression of cloned genes in *Escherichia coli* with blue/white insert screening capability, *Nucleic Acids Res* 18, 4631.
- (비특허문헌 0037) Li, K. and Frost, J.W. (1999) Microbial synthesis of 3-dehydroshikimic acid: A comparative analysis of D-xylose, L-arabinose, and D-glucose carbon sources. *Biotechnol. Prog.* 15, 876-883.
- (비특허문헌 0038) Lin, H., Ravishankar V Vadali, R. V., George N Bennett, G. N. and San, K. Y. (2004) Increasing the Acetyl-CoA Pool in the Presence of Overexpressed Phosphoenolpyruvate Carboxylase or Pyruvate Carboxylase Enhances Succinate Production in *Escherichia Coli*. *Biotechnology Progress* 20 (5): 1599-1604.
- (비특허문헌 0039) Lin, Fengming, Kyle L. Ferguson, David R. Boyer, Xiaoxia Nina Lin, and E. Neil G. Marsh. 2015. "Isofunctional Enzymes PAD1 and UbiX Catalyze Formation of a Novel Cofactor Required by Ferulic Acid Decarboxylase and 4-Hydroxy-3-Polyprenylbenzoic Acid Decarboxylase." *ACS Chemical Biology* 11(4): 1137-1144.
- (비특허문헌 0040) Lu, J. L., and Liao, J. C. (1997) Metabolic engineering and control analysis for production of aromatics: Role of transaldolase, *Biotechnol Bioeng* 53, 132-138.
- (비특허문헌 0041) Lupa, Boguslaw, Delina Lyon, Moreland D. Gibbs, Rosalind a. Reeves, and Juergen Wiegel. 2005. "Distribution of Genes Encoding the Microbial Non-Oxidative Reversible Hydroxyarylic Acid Decarboxylases/phenol Carboxylases." *Genomics* 86 (3): 342-51.
- (비특허문헌 0042) Lutke-Eversloh, T., and Stephanopoulos, G. (2007) L-tyrosine production by deregulated strains of *Escherichia coli*, *Appl Microbiol Biotechnol* 75, 103-110.
- (비특허문헌 0043) Mizuno, S., Yoshikawa, N., Seki, M., Mikawa, T., and Imada, Y. (1988) Microbial production of cis, cis- muconic acid from benzoic acid. *Appl Microbiol Biotechnol.* 28, 20-25.
- (비특허문헌 0044) Nakazawa, A., Kojima, Y., and Taniuchi, H. (1967) Purification and properties of pyrocatechase from *Pseudomonas fluorescens*, *Biochim Biophys Acta* 147, 189-199.
- (비특허문헌 0045) Neidhardt, F. C., and Curtiss, R. (1996) *Escherichia coli* and *Salmonella* : cellular and molecular biology, Vol. 22nd ed., ASM Press, Washington, D.C.
- (비특허문헌 0046) Neidle, E. L., and Ornston, L. N. (1986) Cloning and expression of *Acinetobacter calcoaceticus* catechol 1,2-dioxygenase structural gene *catA* in *Escherichia coli*, *J Bacteriol* 168, 815-820.
- (비특허문헌 0047) Niu, W., Draths, K. M., and Frost, J. W. (2002) Benzene-free synthesis of adipic acid, *Biotechnol Prog* 18, 201-211.
- (비특허문헌 0048) Parker, C., Barnell, W. O., Snoep, J. L., Ingram, L. O., and Conway, T. (1995) Characterization of the *Zymomonas mobilis* glucose facilitator gene product (*glf*) in recombinant *Escherichia coli*: examination of transport mechanism, kinetics and the role of glucokinase in glucose transport, *Mol Microbiol* 15, 795-802.
- (비특허문헌 0049) Parsek, M. R., Shinabarger, D. .L., Rithmel, R. K. and Chakrabarty, A. M. (1992) Roles of CatR and cis, cis-Muconate in activation of the *catBC* operon, which is involved in benzoate degradation in *Pseudomonas putida*. *J. Bacteriol.* 174, 7798-7806.
- (비특허문헌 0050) Patnaik, R. and Liao, J. C. (1994) Engineering of *Escherichia coli* central metabolism for aromatic metabolite with near theoretical yield. *App. Env. Microbiol.* 60, 3903-3908.

- (비특허문헌 0051) Payne, Karl a. P., Mark D. White, Karl Fisher, Basile Khara, Samuel S. Bailey, David Parker, Nicholas J. W. Rattray, et al. 2015. "New Cofactor Supports α,β -Unsaturated Acid Decarboxylation via 1,3-Dipolar Cycloaddition." *Nature* 522 (7557): 497-501.
- (비특허문헌 0052) Pflieger, B. F., Kim, Y., Nusca, T. D., Maltseva, N., Lee, J. Y., Rath, C. M., Scaglione, J. B., Janes, B. K., Anderson, E. C., Bergman, N. H., Hanna, P. C., Joachimiak, A., and Sherman, D. H. (2008) Structural and functional analysis of AsbF: origin of the stealth 3,4-dihydroxybenzoic acid subunit for petrobactin biosynthesis, *Proc Natl Acad Sci U S A* 105, 17133-17138.
- (비특허문헌 0053) Perez-Pantoja, D., De la Iglesia, R., Pieper, D. H., and Gonzalez, B. (2008) Metabolic reconstruction of aromatic compounds degradation from the genome of the amazing pollutant-degrading bacterium *Cupriavidus necator* JMP134, *FEMS Microbiol Rev* 32, 736-794.
- (비특허문헌 0054) Perez-Pantoja, D., Donoso, R., Agullo, L., Cordova, M., Seeger, M., Pieper, D. H., and Gonzalez, B. (2011) Genomic analysis of the potential for aromatic compounds biodegradation in Burkholderiales, *Environ Microbiol.* 14.5 (2012): 1091-1117.
- (비특허문헌 0055) Pittard, J. and Wallace, B. J. (1966) Distribution and function of genes concerned with aromatic biosynthesis in *Escherichia coli*. *J. Bacteriol.* 91, 1494-1508.
- (비특허문헌 0056) Polen, T., Spelberg, M. and Bott, M. (2013) Toward bitechnological produciton of adipic acid and precursors from biorenewables, *J. Biotechnol.* 167(2): 75-84.
- (비특허문헌 0057) Rutledge, B. J. (1984) Molecular characterization of the qa-4 gene of *Neurospora crassa*, *Gene* 32, 275-287.
- (비특허문헌 0058) Schirmer, F., and Hillen, W. (1998) The *Acinetobacter calcoaceticus* NCIB8250 mop operon mRNA is differentially degraded, resulting in a higher level of the 3' CatA-encoding segment than of the 5' phenolhydroxylase-encoding portion, *Mol Gen Genet* 257, 330-337.
- (비특허문헌 0059) Shumilin, I. A., Kretsinger, R. H., and Bauerle, R. H. (1999) Crystal structure of phenylalanine-regulated 3-deoxy-D-arabino-heptulosonate-7-phosphate synthase from *Escherichia coli*, *Structure* 7, 865-875.
- (비특허문헌 0060) Shumilin, I. A., Zhao, C., Bauerle, R., and Kretsinger, R. H. (2002) Allosteric inhibition of 3-deoxy-D-arabino-heptulosonate-7-phosphate synthase alters the coordination of both substrates, *J Mol Biol* 320, 1147-1156.
- (비특허문헌 0061) Shumilin, I. A., Bauerle, R., Wu, J., Woodard, R. W., and Kretsinger, R. H. (2004) Crystal structure of the reaction complex of 3-deoxy-D-arabino-heptulosonate-7-phosphate synthase from *Thermotoga maritima* refines the catalytic mechanism and indicates a new mechanism of allosteric regulation, *J Mol Biol* 341, 455-466.
- (비특허문헌 0062) Shumkova, E. S., Solyanikova, I. P., Plotnikova, E. G. and Golovleva, L. A. (2009) Phenol degrdation by *Rhodococcus opacus* Strain 1G. *App. Biocehm. Microbiol.* 45, 43-49.
- (비특허문헌 0063) Sietmann, R., Uebe, R., Boer, E., Bode, R., Kunze, G., and Schauer, F. (2010) Novel metabolic routes during the oxidation of hydroxylated aromatic acids by the yeast *Arxula adenivorans*, *J Appl Microbiol* 108, 789-799.
- (비특허문헌 0064) Smith, M. R. and Ratledge, C. (1989) Quantitative biotransformation of catechol to cis, cis-muconate. *Biotech. Lett.* 11, 105-110.
- (비특허문헌 0065) Snoep, J. L., Arfman, N., Yomano, L. P., Fliege, R. K., Conway, T., and Ingram, L. O. (1994) Reconstruction of glucose uptake and phosphorylation in a glucose-negative mutant of *Escherichia coli* by using *Zymomonas mobilis* genes encoding the glucose facilitator protein and glucokinase, *J Bacteriol* 176, 2133-2135.
- (비특허문헌 0066) Sonoki, Tomonori, Miyuki Morooka, Kimitoshi Sakamoto, Yuichiro Otsuka, Masaya

- Nakamura, Jody Jellison, and Barry Goodell. 2014. Enhancement of Protocatechuate Decarboxylase Activity for the Effective Production of Muconate from Lignin-Related Aromatic Compounds. *Journal of Biotechnology* 192 (Part A): 71-77.
- (비특허문헌 0067) Sprenger, G. A. (1995) Genetics of pentose-phosphate pathway enzymes of *Escherichia coli* K-12, *Arch Microbiol* 164, 324-330.
- (비특허문헌 0068) Sprenger, G. A., Schorken, U., Sprenger, G., and Sahm, H. (1995a) Transketolase A of *Escherichia coli* K12. Purification and properties of the enzyme from recombinant strains, *Eur J Biochem* 230, 525-532.
- (비특허문헌 0069) Sprenger, G. A., Schorken, U., Sprenger, G., and Sahm, H. (1995b) Transaldolase B of *Escherichia coli* K-12: cloning of its gene, *talB*, and characterization of the enzyme from recombinant strains, *J Bacteriol* 177, 5930-5936.
- (비특허문헌 0070) Stroman, P., Reinert, W. R., and Giles, N. H. (1978) Purification and characterization of 3-dehydroshikimate dehydratase, an enzyme in the inducible quinic acid catabolic pathway of *Neurospora crassa*, *J Biol Chem* 253, 4593-4598.
- (비특허문헌 0071) Tang, J., Zhu, X., Lu, J. and Liu, P. (2012) Recruiting alternative glucose utilization pathways for improving succinate production. *App Microbiol Biotechnol* DOI 10.1007/s00253-012-434.1.
- (비특허문헌 0072) Tateoka, T., and Yasuda, I. (1995) 3-Dehydroshikimate dehydratase in mung bean cultured cells, *Plant Cell Reports* 15, 212-217.
- (비특허문헌 0073) Vemuri, G. N., M. A. Eiteman, and E. Altman. 2002. "Effects of Growth Mode and Pyruvate Carboxylase on Succinic Acid Production by Metabolically Engineered Strains of *Escherichia Coli*." *Applied and Environmental Microbiology* 68 (4): 1715-27.
- (비특허문헌 0074) Weaver, L. M., and Herrmann, K. M. (1990) Cloning of an *aroF* allele encoding a tyrosine-insensitive 3-deoxy-D-arabino-heptulosonate 7-phosphate synthase, *J Bacteriol* 172, 6581-6584.
- (비특허문헌 0075) Weber, C., Bruckner, C., Weinreb, S., Lehr, C., Essl, C. and Bole, E. (2012) Biosynthesis of *cis, cis*-muconic acid and its aromatic precursors catechol and protocatechuic acid, from renewable feedstocks by *Saccharomyces cerevisiae*, *App Environ Microbiol*. 78, 8421-8430.
- (비특허문헌 0076) Wheeler, K. A., Lamb, H. K., and Hawkins, A. R. (1996) Control of metabolic flux through the quinate pathway in *Aspergillus nidulans*, *Biochem J* 315 (Pt 1), 195-205.
- (비특허문헌 0077) White, Mark D., Karl A. P. Payne, Karl Fisher, Stephen A. Marshall, David Parker, Nicholas J. W. Rattray, Drupad K. Trivedi, et al. 2015. "UbiX Is a Flavin Prenyltransferase Required for Bacterial Ubiquinone Biosynthesis." *Nature* 522 (7557): 502-6.
- (비특허문헌 0078) Wu, C-M., Wu, C-C., Su, C-C., Lee, S-N., Lee, Y-A. and Wu, J-Y. (2006) Microbial synthesis of *cis, cis*-muconic acid from benzoate by *Sphingobacterium* sp. Mutants. *Biochem. Eng. J.* 29, 35-40.
- (비특허문헌 0079) Xie, N., Tang, H., Feng, J., Tao, F., Ma, C. and Xu, P. (2009) Characterization of benzoate degradation by newly isolated bacterium *Pseudomonas* sp. XP-M2. *Biochem. Eng. J.* 46, 79-82.
- (비특허문헌 0080) Xie, N., Hong Liang, Ri-Bo Huang, and Ping Xu. 2014. "Biotechnological Production of Muconic Acid: Current Status and Future Prospects." *Biotechnology Advances* 32 (3): 615-22.
- (비특허문헌 0081) Yi, J., Draths, K. M., Li, K. and Frost, J. W. (2003) Altered Glucose Transport and Shikimate Pathway Product Yields in *E. coli*. *Biotechnol. Prog.* 2003, 19, 1450-1459.
- (비특허문헌 0082) Yoshikawa, N., Mizuno, S., Ohta, K., and Suzuki, M. (1990) Microbial production of *cis, cis*-muconic acid. *J. Biotechnol.* 14, 203-210.

발명의 내용

해결하려는 과제

- [0011] 본 발명은 비-방향족 탄소원, 예를 들어 글루코오스, 수크로오스, 글리세롤 및 셀룰로오스 가수분해물 (cellulosic hydrolysates)을 포함하나 이에 한정되지 않는 당 및 탄수화물로부터 시작하여 시스, 시스-뮤콘산을 생산하는 유전자 조작된 미생물을 제공한다.
- [0012] 본 발명의 일 실시예에서, 방향족 아미노산 생합성의 음성 조절자의 활성은 유전적으로 조작된다. 본 발명의 일 측면에서, 음성 조절자 TyrR의 활성은 TyrR 단백질을 코딩하는 *tyrR* 유전자의 발현을 조절함으로써 상당히 감소된다. 본 발명의 다른 측면에서, 음성 조절자 TyrR의 활성은 미생물의 염색체 DNA에서 *tyrR* 유전자를 결실시키거나 또는 불활성화시킴으로써 완전히 제거될 수 있다.
- [0013] 본 발명의 또 다른 실시예에서, 방향족 아미노산 경로에서 특정 대사산물에 의한 특정 효소의 피드백 저해는 유전자 조작을 통해 극복된다. 대부분의 야생형 *E. coli* 균주에서, 방향족 아미노산 경로의 초반에 기능하는 데옥시아라비노-헵툴로손네이트 7-포스페이트 합성효소 (deoxyarabino-heptulosonate 7-phosphate synthase; "DAHPSynthase")는 3개의 상이한 유전자, 즉 *aroG*, *aroF* 및 *aroH*에 의해 인코딩되는 것으로 알려진 3개의 상이한 동질효소(isozymes)로 존재한다. 이러한 3개의 유전자 각각에 의해 인코딩된 단백질은 방향족 아미노산 경로의 하나 이상의 대사산물에 의해 피드백 저해를 받는다. 본 발명의 일 측면에서, 야생형 *aroG* 유전자는 미생물 세포 (microbial cell) 내에서 방향족 아미노산 경로의 하나 이상의 대사산물에 의한 피드백 저해에 대해 내성이 있는 AroG 단백질을 코딩하는 변형된 *aroG* 유전자로 대체된다. 이러한 AroG 단백질의 피드백 내성 형태는 "AroG^{FBR}"라고 한다. 본 발명의 다른 측면에서, 야생형 *aroF* 유전자는 미생물 세포 내에서 방향족 아미노산 경로의 하나 이상의 대사산물에 의한 피드백 저해에 대해 내성이 있는 AroF 단백질을 코딩하는 *aroF* 유전자로 대체된다 (AroF^{FBR}). 본 발명의 또 다른 측면에서, 야생형 *aroH* 유전자는 미생물 세포 내에서 방향족 아미노산 경로의 하나 이상의 대사산물에 의한 피드백 저해에 대해 내성이 있는 AroH 단백질을 코딩하는 *aroH* 유전자로 대체된다 (AroH^{FBR}). 본 발명의 또 다른 측면에서, 시스, 시스-뮤콘산의 상업적 생산을 위해 선택된 생축매는 DAHP 합성효소에 대해 하나 이상의 피드백 내성 동질효소를 가질 수 있다.
- [0014] 본 발명의 또 다른 실시예에서, 미생물 세포 내에서 중심 방향족 생합성 경로에 관여하는 하나 이상의 효소의 활성이 향상된다. 본 발명의 일 측면에서, 방향족 경로 및/또는 뮤콘산 경로의 작동에 관여하는 하나 이상의 효소의 활성 향상은 유전자 조작을 통해 달성된다. 본 발명의 바람직한 측면에서, 단백질 AroF, AroG, AroH, AroB, TktA, TalB, AroZ, QutC, Qa-4, AsbF, QuiC, AroY, Rpe, Rpi, Pps, CatA 및 CatX 또는 이의 동족체 (homologs) 또는 유사체를 코딩하는 하나 이상의 유전자의 발현이 향상되어, 상기 단백질의 증가된 활성을 야기한다. Rpe는 리불로오스-5-포스페이트 에피머라제(ribulose-5-phosphate epimerase)이고, Rpi는 리불로오스-5-포스페이트 이소머라제(isomerase)이며, Pps는 포스포에놀 피루베이트 합성효소이다 (Neidhardt and Curtiss, 1996). 만일 숙주 균주가 효모, 예를 들어, *사카로마이세스 세레비지애*(*Saccharomyces cerevisiae*), 또는 사상균(filamentous fungus), 예를 들어, *뉴로스포라 크라싸*(*Neurospora crassa*) 이면, *S. 세레비지애*의 경우 쉬키메이트 경로(shikimate pathway)에서 반응을 촉진시키는 몇몇 효소들은 ARO1 유전자에 의해 인코딩된 Aro1p라고 불리는 하나의 큰 단백질 또는 폴리펩티드 내로 혼합될 수 있다. Aro1p는 AroB, AroD, AroE, AroK (또는 AroL), 및 AroA의 기능을 겸비한다. 이와 같이, 본 발명의 목적을 위해, Aro1p 및 ARO1, 또는 이의 일부는 치환체로서 사용될 수 있거나, 또는 AroB, AroD, AroE, AroK, 및/또는 AroA에 더하여 사용될 수 있다.
- [0015] 본 발명의 또 다른 실시예에서, 박테리아 세포 내에서 에리트로오스-4-포스페이트를 통한 분출(flux)은 펜토오스 포스페이트 경로의 작동에서 효소를 과발현시킴으로써 향상된다. 본 발명의 일 측면에서, 트랜스알돌라제 효소, 예를 들어 *talB* 또는 *talA* 유전자에 의해 코딩된 트랜스알돌라제 효소의 발현은 유전자 변형에 의해 향상된다. 본 발명의 다른 측면에서, 트랜스케톨라제 효소를 인코딩하는 유전자, 예를 들어, *tktA* 유전자의 발현은 유전자 조작에 의해 향상된다. 본 발명의 또 다른 측면에서, 리불로오스-5-포스페이트 에피머라제 및 리불로오스-5-포스페이트 이소머라제 중 하나 또는 둘 모두를 인코딩하는 유전자의 발현은 유전자 조작에 의해 향상된다.
- [0016] 본 발명의 다른 실시예에서, 방향족 아미노산 경로의 기능에 필요한 포스포에놀 피루베이트(PEP)는 유전자 조작을 통해 증가된다. 본 발명의 일 측면에서, PEP 풀(pool)의 사용에 대한 경쟁은 글루코오스 흡수를 위한 PEP 독

립 시스템과 글루코오스의 흡수를 위한 PEP-의존성 포스포트랜스퍼라제 시스템 (PTS)의 제거 및/또는 보완을 통해 감소된다. 본 발명의 또 다른 측면에서, GalP 기반 당 흡수 시스템은 미생물 세포 내에서 ATP를 보존할 목적으로 불활성화된다. 본 발명의 또 다른 측면에서, PTS 시스템 및 GalP 기반 당 흡수 시스템 (Δ PTS/ Δ galP) 둘 다의 기능에 결함이 있는 미생물 세포에서, 당 흡수는 Glf (글루코오스 확산을 촉진하는 단백질)를 인코딩하는 외인성 유전자 또는 Glf 및 Glk (글루코키나제) 단백질을 모두 인코딩하는 외인성 유전자를 도입함으로써 달성된다. 본 발명의 또 다른 실시예에서, PEP의 이용률은 *pps*와 같은 PEP 합성효소를 인코딩하는 유전자의 발현을 증가시킴으로써 증가된다.

[0017] 본 발명의 또 다른 실시예에서, 3,4-디히드록시벤조산 탈탄산효소 (3,4-dihydroxybenzoic acid decarboxylase; AroY)의 활성이 향상된다. 이 실시예의 일 측면에서, AroY의 발현은 유전자 조작에 의해 향상된다. 본 발명의 다른 측면에서, UbiX, KpdB, Elw, Kox, Lpl 및 이들의 동족체를 포함하는 군으로부터 선택된 AroY에 대한 부속 단백질로서 작용하는 단백질의 발현은 유전자 조작에 의해 증가되어, 3,4-디히드록시벤조산 탈탄산효소 활성의 증가를 야기한다.

[0018] 본 발명의 다른 실시예에서, PEP의 이용률은 포스포에놀피루베이트 카복실라제 (Ppc) 활성의 감소 또는 제거에 의해 증가된다. 본 발명의 일 측면에서, 특히 Ppc 활성이 제거될 때, 피루베이트 카복실라제 (Pyc) 활성이 증가되고 및/또는 Ppc 활성으로 대체된다.

과제의 해결 수단

[0019] 본 명세서에 포함된 모든 특허, 특허출원, 간행물, 서열 및 기타 공개된 자료는 참조로 포함된다.

[0020] 본 특허출원에서 사용된 바와 같이, 문구 "예를 들어" 또는 "~와 같은"는 당면의 주제에 대해 하나 이상의 방법, 접근법, 해결책 또는 물질의 구성이 있음을 나타내기 위한 것이며, 주어진 예는 그 예에 한정되는 것을 의미하지 않는다.

[0021] 용어 "이종(heterologous)"은 유기체에서 자연적으로 또는 천연적으로 발견되지 않으나, 유전공학에 의해, 예를 들어 형질전환, 교배, 또는 형질도입 (transduction)에 의해 유기체에 도입될 수 있는 유전자 또는 단백질을 나타낸다. 이종 유전자는 염색체 내로 통합(삽입)되거나 또는 플라스미드 상에 함유될 수 있다. 용어 "외인성(exogenous)"은 유전공학에 의해, 예를 들어 형질전환, 교배, 형질도입, 또는 돌연변이유발에 의해 활성을 증가, 감소 또는 제거하기 위한 목적으로 유기체 내로 도입되거나 또는 유기체에서 변경되는 유전자 또는 단백질을 나타낸다. 외인성 유전자 또는 단백질은 이종일 수 있거나, 또는 숙주 유기체에 고유하나, 하나 이상의 방법, 예를 들어, 돌연변이, 결실, 프로모터의 변화, 종결자의 변화, 중복, 또는 염색체 또는 플라스미드에서 하나 이상의 추가 복제의 삽입에 의해 변경된 유전자 또는 단백질일 수 있다. 따라서, 예를 들어, 만일 *aroB* 유전자의 제 2 복제가 천연 부위와 완전히 다른 염색체의 부위에서 삽입되면, 제 2 복제는 외인성일 것이다.

[0022] 본 발명에서 사용된 용어 "미생물"은 발효 공정을 통해 시스, 시스-뮤콘산의 상업적 생산에 사용될 수 있는 박테리아, 고세균(archaea), 효모, 조류 및 사상균을 포함한다. 용어 "유전자 조작된 미생물"은 자연에 존재하지 않으나 본 특허 출원에 기재된 바와 같이 하나 또는 다른 유전자 변형을 이용하여 생성된 미생물을 나타낸다.

[0023] 명명법의 경우, 유전자 또는 코딩 영역은 일반적으로 이탤릭체로 된 소문자, 예를 들어 "*aroZ*"로 명명되며, 유전자에 의해 인코딩된 효소 또는 단백질은 동일한 문자이나, 첫 글자는 대문자이고 이탤릭체는 사용하지 않는 것으로 명명될 수 있다, 예를 들어 "AroZ". 또한, 효소 또는 단백질은 더 기술적인 이름으로 언급될 수 있다, 예를 들어, AroZ는 3-데히드로쉬키메이트 탈수효소(3-dehydroshikimate dehydratase)라고도 한다. 특정 촉매 활성을 갖는 효소의 일 예를 인코딩하는 유전자 또는 코딩 영역은, 역사적으로 다른 기원 때문에 또는 유전자가 상이한 종으로부터 유래되기 때문에, 다양한 이름을 가질 수 있다. 예를 들어, *바실러스 안트라시스(Bacillus anthracis)*로부터의 3-데히드로쉬키메이트 탈수효소를 인코딩하는 유전자는 *aroZ* 대신에 *asbF*로 명명될 수 있으며, *아스페르길루스 니둘란스(Aspergillus nidulans)*로부터의 동일한 유전자는 *qutC*로 명명될 수 있고, *뉴로스 포라 크라싸(Neurospora crassa)*로부터의 동일한 유전자는 *qa-4*로 명명될 수 있으며, *아시네토박터 바일리(Acinetobacter baylyi)*로부터의 동일한 유전자는 *quiC*로 명명될 수 있다.

[0024] "플라스미드"는 미생물의 염색체 또는 염색체들과는 별개로, 염색체보다 상당히 더 작고, 염색체 또는 염색체들로부터 별도로 복제하는 원형 또는 선형 DNA 분자를 의미한다. "플라스미드"는 세포당 약 하나의 복제 또는 세포당 하나 이상의 복제에 존재할 수 있다. 미생물 세포 내에서 플라스미드의 유지는 일반적으로 항생제 선택을 필요로 하지만, 영양요구성의 보완이 사용될 수도 있다.

- [0025] 박테리아 세포와 관련하여 본 발명에서 사용된, 용어 "염색체" 또는 "염색체 DNA"는 플라스미드보다 상당히 더 크고 임의의 항생제 선택을 필요로하지 않는 원형 DNA 분자이다.
- [0026] 용어 "발현 카세트" 또는 "카세트"는 적어도 프로모터, 및 효소 또는 다른 단백질을 코딩하는 유전자 또는 영역을 함유하는 염색체 또는 플라스미드의 일부일 수 있는 DNA 서열을 의미하며, 따라서 코딩 영역은 프로모터에 의해 발현되고, 효소 또는 단백질은 DNA 서열을 함유하는 숙주 세포에 의해 생성된다. "발현 카세트"는 코딩 영역이 자연적으로 코딩 영역과 관련되지 않은 프로모터로부터 발현되도록, 적어도 부분적으로 합성되거나, 또는 유전공학 방법에 의해 구축될 수 있다. 임의로, "발현 카세트"는 자연적으로 코딩 영역과 관련된 종결자일 수도 있거나 또는 종결자가 아닐 수도 있는 전사 종결자를 함유할 수 있다. "발현 카세트"는 하나 이상의 단백질에 대한 코딩 영역을 가질 수 있으며, 이 경우, 이것은 오페론, 또는 합성 오페론이라고 불릴 수 있다.
- [0027] 유전자 또는 코딩 영역의 "과발현"은 동일하거나 또는 유사한 성장 조건 하에서 숙주 미생물의 야생형 버전에서 확인된 수준보다 높은 수준으로 숙주 미생물에서 그 유전자 또는 코딩 영역에 의해 인코딩된 효소 또는 단백질을 생산하는 것을 의미한다. 이는 하기 방법들 중 하나 이상에 의해 달성될 수 있다: 1) 더 강한 프로모터의 도입, 2) 번역 시작 코돈의 약 4 내지 10 염기 상류에 위치한 5'-AGGAGG의 DNA 서열과 같은 더 강한 리보솜 결합 부위의 도입, 3) 종결자 또는 더 강한 종결자의 도입, 4) 코딩 영역의 하나 이상의 부위에서 코돈의 선택의 개선, 5) mRNA 안정성의 개선, 및 6) 염색체에 다수의 복제를 도입하거나 또는 다중복제 플라스미드 상에 카세트를 놓음으로써, 유전자의 복제수의 증가. 과발현되는 유전자로부터 생산된 효소 또는 단백질은 "과잉생산된"이라고 한다. "과발현된" 유전자 또는 "과잉생산된" 단백질은 숙주 미생물에 대해 고유한 것일 수 있으며, 또는 상이한 유기체로부터 유전공학 방법에 의해 숙주 미생물에 이식된 것일 수 있다. 이 경우, 효소 또는 단백질 및 효소 또는 단백질을 인코딩하는 유전자 또는 코딩 영역은 "외래" 또는 "이종"이라고 한다. 외래 또는 이종 유전자 및 단백질은 유전자 조작되지 않은 숙주 유기체에 존재하지 않기 때문에, 이들은 의미상 과발현되고 과잉생산된다.
- [0028] 제 1 유전자, DNA 서열 또는 단백질의 "동족체"는 상기 제 1 유전자, DNA 서열 또는 단백질의 생물학적 기능과 유사한 생물학적 기능을 수행하고, 기본 매개 변수를 이용하여 서열 비교를 위한 BLAST 컴퓨터 프로그램 (Altschul et al, 1990; Altschul et al, 1997)에 의해 측정된 바와 같이 (단백질 서열을 비교하거나 또는 유전자 서열로부터 유래된 단백질 서열을 비교할 때) 상기 제 1 유전자 또는 단백질과 적어도 25% 서열 동일성을 가지는, 제 2 유전자, DNA 서열 또는 단백질이며, 결실 및 삽입을 허용한다. *E. coli aroG* 유전자의 동족체의 예는 *살모넬라 타이피뮤리움(Salmonella typhimurium)*으로부터의 *aroG*일 것이다.
- [0029] 상동성에 의해 매우 먼 거리에 있는 2개의 효소 또는 단백질은 동일한 생화학적 기능을 수행할 수 있으나, 단지 서로 상대적으로 약하게 상동일 수 있다. 예를 들어, *E. coli* K-12 (GenBank NP_416129)로부터의 FumA 푸마라제는 이들의 중첩 영역 상의 *클로스트리디움 보툴리눔(Clostridium botulinum)* (GenBank GAE03909.1)으로부터의 푸마라제와 약 26.9% 상동이며, 이들의 중첩 영역 상의 *피로코쿠스(Pyrococcus)* sp. ST04 (GenBank AKF23146.1)로부터의 푸마라제 베타 아단위와 약 25.1% 상동이다. 다른 예로서, DHS를 PCA로 전환시키는 작용을 하는 *클렙시엘라(Klebsiella)* AroZ 및 *뉴로스포라 크라싸(Neurospora crassa)* Qa-4 효소는 29.3% 동일하다. 따라서, 대사 경로의 유전공학의 경우, 이종 효소 또는 단백질의 중요한 특징은 효소 또는 단백질에 의해 수행되는 기능 또는 반응이지, 근원 유기체(source organism) 또는 정확한 아미노산 서열이 아니기 때문에, 우리는 "상동" 효소 또는 단백질 또는 "동족체"를 이들의 아미노산 서열의 동일성에서 25% 이상인 임의의 쌍의 효소 또는 단백질을 포함하는 것으로 정의한다, 예를 들어 리프만-피어슨 방법(Lipman-Pearson method) (Ktuple = 2, 갭 페널티(Gap penalty) = 4, 및 갭 길이 페널티 = 12)으로 LaserGene 12 (DNASTar, Madison, WI) MegAlign 프로그램을 이용하여 정렬을 위한 기본 매개변수를 사용하여 나타낸 것처럼 간격을 허용한다.
- [0030] 제 1 유전자, DNA 서열 또는 단백질의 "유사체"는 상기 제 1 유전자, DNA 서열 또는 단백질의 생물학적 기능과 유사한 생물학적 기능을 수행하는 제 2 유전자, DNA 서열 또는 단백질이나, 서열 비교를 위한 BLAST 컴퓨터 프로그램 (Altschul et al, 1990; Altschul et al, 1997)에 의해 측정된 바와 같이 (단백질 서열을 비교하거나 또는 유전자 서열로부터 유래된 단백질 서열을 비교할 때) 상기 제 1 유전자, DNA 서열 또는 단백질과 25% 미만의 서열 동일성이 있는 경우, 결실 및 삽입을 허용한다. *클렙시엘라 뉴모니아(Klebsiella pneumoniae)* AroZ 단백질의 유사체의 예는 *아스페르길루스 니둘란스(Aspergillus nidulans)*로부터의 QutC 단백질일 것인데, 그 이유는 두 단백질 모두 3-데히드로쉬키메이트 탈수효소 반응을 촉진시키는 효소이나, 두 효소 또는 각각의 유전자 간에 유의한 서열 상동성이 없기 때문이다. 당해 기술분야의 통상의 기술자는 특정 생물학적 기능을 갖는 많은 효소 및 단백질, 예를 들어 DAHP 합성 효소 또는 3-데히드로쉬키메이트 탈수효소가 많은 상이한 유기체에서 동족체 또는 유사체로서 확인될 수 있다는 것을 알 것이고, 이러한 효소 또는 단백질 군의 구성원은 동일한 기능

을 공유하기 때문에, 이들이 구조적으로 약간 또는 상당히 상이할 수 있을지라도, 동일한 군의 상이한 구성원은 많은 경우에 현재의 유전공학 방법을 이용하여 동일한 생물학적 기능을 수행하는데 사용될 수 있다. 따라서, 예를 들어, AorZ 효소 및 QutC 효소는 동일한 반응인 DHS 탈수효소를 촉진하므로, 어느 하나가 적당한 상황에서 시스, 시스-뮤콘산의 생산을 야기할 것이며, 궁극적으로 어느 것을 사용할 것인지에 대한 선택은 유사한 발효 조건 하에서 시스, 시스-뮤콘산의 더 높은 역가를 야기하는 것을 선택함으로써 이루어질 수 있다.

[0031] 용어 "비-방향족 탄소원" 또는 "비-방향족 화합물"은 탄소 및/또는 에너지의 공급원으로서 본 발명의 미생물을 공급하는데 사용될 수 있는 탄소-함유 화합물을 의미하며, 여기서 화합물은 벤젠과 관련된 6-원 고리를 함유하지 않는다. 비-방향족 탄소원의 예는 글루코오스, 자일로오스, 락토오스, 글리세롤, 아세테이트, 아라비노오스, 갈락토오스, 만노오스, 말토오스, 또는 수크로오스를 포함한다. "방향족 화합물"은 벤젠과 관련된 하나 이상의 6-원 고리를 함유하는 화합물이다. 방향족 화합물의 예는 카테콜, 또는 1,2-디히드록시 벤젠이다. 비-방향족 탄소원의 공급원으로서 글루코오스를 이용하여 뮤콘산을 생산하기 위해 선택된 미생물은, 특허 문헌 [미국특허 번호 8,871,489 및 미국특허출원 공개번호 US2013/0337519A1 및 US2014/0234923A]에 제공된 유전공학 기술을 이용하여 글리세롤, 수크로오스 및 자일로오스와 같은 다른 유형의 비-방향족 탄소원을 사용하도록 추가로 조작될 수 있다.

[0032] 용어 "강한 항시성 프로모터(strong constitutive promoter)"는 DNA 서열 또는 RNA 폴리머라제에 의해 전사되는 유전자의 일반적으로 상류 (통상적인 5'에서 3' 방향으로 나타낸 경우 유전자의 5' 측면)에 놓이고, 상기 DNA 서열 또는 유전자가 임의의 적당한 분석 절차에 의해 직접 또는 간접적으로 용이하게 검출되는 수준에서 RNA 폴리머라제에 의한 전사에 의해 발현되도록 하는 DNA 서열을 의미한다. 적당한 분석 절차의 예는 1) 정량적 역전사효소 + PCR, 2) 인코딩된 효소의 효소 분석, 3) 쿠마시 블루-염색된 단백질 겔, 또는 4) 상기 전사의 결과로서 간접적으로 생산되는 대사산물의 측정가능한 생산, 및 전사, 대사산물, 또는 유도 화합물질의 수준을 특이적으로 조절하는 단백질의 존재 또는 부재에 상관없이 발생하는 이러한 측정가능한 전사를 포함한다. "강한 항시성 프로모터"가 아닌 프로모터의 예는 *E. coli*의 P_{lac} 프로모터인데, 그 이유는 락토오스 또는 유도제 IPTG의 부재 하에 억제제에 의해 억제되기 때문이다. 당해 기술분야에서 잘 알려진 방법을 이용함으로써, "강한 항시성 프로모터"는 천연 프로모터 (그렇지 않으면 DNA 서열 또는 유전자의 상류에 자연적으로 존재하는 프로모터)를 대체하기 위해 사용될 수 있어, 플라스미드 또는 염색체 중 어느 하나에 놓일 수 있고 천연 프로모터로부터의 수준보다 높은 수준에서 원하는 DNA 서열 또는 유전자의 발현 수준을 제공하는 발현 카세트를 야기한다. 강한 항시성 프로모터는 종 또는 속에 대해 특이적일 수 있으나, 종종 박테리아의 강한 항시성 프로모터는 멀리 떨어져 있는 박테리아에서 잘 기능할 수 있다. 예를 들어, *바실러스 서브틸리스*(*Bacillus subtilis*)로부터의 프로모터 또는 *바실러스 서브틸리스*(*B. subtilis*) 상에서 정상적으로 성장하는 파지는 *E. coli*에서 잘 기능할 수 있다. "강한 항시성 프로모터"는 시스, 시스-뮤콘산의 선행기술 생산에서 사용된 P_{tac} 와 같은 유도성 프로모터와 상당히 상이하며, 일반적으로 원하는 수준의 기능을 위해서 고가의 화합물질 또는 기타 환경 변화를 필요로 한다 (Niu et al., 2002). 강한 항시성 프로모터의 예는 *바실러스 서브틸리스* 파지 SP01 및 *coli* 파지 람다 P_R 로부터의 P_{15} , P_{26} 이다.

[0033] "돌연변이"는 DNA 서열의 임의의 변화로, DNA 서열을 관련 야생형 서열과 상이하게 만드는 것이다. "돌연변이"는 단일 염기 변화, 결실, 삽입, 치환, 틀이동 (frameshift), 역위, 복제 또는 DNA 서열의 다른 유형의 변화를 포함할 수 있다. 일반적으로 "돌연변이"는 기능에 부정적인 영향을 미치거나 또는 유전자 또는 유전자 산물의 활성을 감소시키는 변화를 나타낸다, 그러나, 본 명세서에서, 용어 "돌연변이"는 유전자 또는 유전자 산물의 활성을 증가시키는 변화를 나타낼 수도 있다. 예를 들어, *aroG* 유전자의 피드백 내성 돌연변이는 페닐알라닌과 같은 저해제의 존재 하에 AroG의 활성을 증가시킨다. 프로모터를 상이하고 더 강한 프로모터로 대체하면 유전자 또는 유전자 산물의 활성을 증가시킬 수 있는 돌연변이가 야기된다. "삭제돌연변이(null mutation)"는 대부분의 또는 모든 유전자의 결실과 같은 돌연변이이며, 유전자의 기능을 효과적으로 제거하는 돌연변이이다. "돌연변이체"는 하나 이상의 돌연변이를 포함하는 균주 또는 분리 균주이다.

[0034] 시스, 시스-뮤콘산 (본 명세서에서 간단히 "뮤콘산"이라고 함)의 생물학적 생산은 방향족 아미노산 경로로부터의 탄소의 방향 전환(redirection)에 기초한다. 방향족 아미노산 및 비타민의 본래 생산은 대사산물인 에리트르오스-4-포스페이트 (erythrose-4-phosphate; E4P) 및 포스포에놀피루베이트 (PEP)를 필요로 한다. 방향족 아미노산 합성의 제 1 개입 단계는 효소 3-데옥시-아라비노-헵툴로네이트 7-포스페이트 (DAHP) 합성효소에 의해 촉진된다. *E. coli*에서, 이 단계는 3개의 상이한 동질효소인 AroG, AroF, 또는 AroH에 의해 수행될 수 있다. 각각의 이들 효소는 억제제 단백질 TyrR에 의한 전사 수준 및 경로의 말단 생성물인 페닐알라닌, 티로신 및 트립토

관으로부터의 저해에 의한 단백질 수준에서 각각 조절된다. 류콘산의 생산은 방향족 아미노산 경로인 테히드로쉬킴산(dehydroshikimic acid; DHS)의 중간체로부터 진행되며, 3개의 이중 효소인 테히드로쉬키메이트 탈수효소(AroZ), 3,4-디히드로시벤조에이트 탈탄산효소(AroY) 및 카테콜 1,2-디옥시게나제(CatA)의 발현을 필요로 한다. 이 경로는 도 1 및 도 2에 나타난다.

[0035] "류코닉 경로" 또는 "류콘산 경로"는 DHS로부터 PCA로, 카테콜로, 시스, 시스-류콘산으로의 생화학적 경로를 나타내고, "류코닉 경로 유전자"는 류코닉 경로의 단계를 촉진하는 효소를 인코딩하거나, 또는 상기 효소들, 예를 들어 *aroZ*, *aroY*, *catA*, *catX*, 및 *qutC* 중 하나의 활성을 향상시키는 역할을 하는 보조 기능을 인코딩하는 유전자이다(도 3). DHS는 3-테히드로쉬키메이트의 약어이고, PCA는 프로토크아테큐산(protocatechuic acid)의 약어이다. "류코닉 플라스미드"는 하나 이상의 류코닉 경로 유전자를 함유하는 플라스미드이다.

[0036] 본 발명에서 사용된 유전자 조작은 도 1에 나타난 바와 같이 많은 미생물 세포에 존재하는 방향족 아미노산 및 방향족 비타민(또는 비타민-유사) 생합성의 공통 경로를 중심으로 한다. 도 1에 나타난 방향족 아미노산 생합성의 공통 경로는 "쉬킴산" 또는 "쉬키메이트" 경로, "코리스민산(chorismic acid)" 또는 "코리스메이트(chorismate)" 경로, 또는 "중심 방향족" 또는 "중심 방향족 생합성" 경로라고 할 수 있다.

[0037] 방향족 아미노산, 페닐알라닌, 티로신 및 트립토판의 생산을 위한 미생물의 유전공학에 관한 상당한 양의 발표된 연구물이 있다(US 4,681,852, US 4,753,883, US 6,180,373, 유럽특허출원 86300748.0). 방향족 아미노산의 생산 방법은 피드백 내성 효소(AroF, AroG, PheA, TyrA), 전사의 억제제 탈조절(*tyrR*), 프로모터 강도의 증가(P_{tac} , P_{lac}) 및 하나 이상의 유전자의 복제수의 증가(*tktA*)의 다양한 조합을 이용하는 것을 포함한다. 상기한 유전자 변형의 많은 특정 조합에 따라 류콘산 생산에 적합한 생축매를 얻을 수 있다.

[0038] 본 명세서에 전체 참조로 포함된, 국제특허출원 공개번호 W02013/116244에 개시된 바에 따르면, 유전자 조작된 미생물은 원하는 표현형을 얻기 위해 필요한 특정 외인성 또는 이중 유전자를 가지고 있지만, 류콘산을 생산하기 위해 임의의 외인성 플라스미드를 함유할 필요는 없다. 본 발명의 바람직한 실시예에서, 미생물에 도입된 외인성 유전자는 염색체 DNA에 안정적으로 통합된다. 이러한 외인성 유전자의 염색체 DNA 통합의 결과로, 외인성 DNA를 가지는 플라스미드를 유지하기 위한 항생제 또는 다른 선택 방법의 사용에 대한 필요성이 완전히 제거되었다. 또한, 화학 유도제를 필요로 하지 않는 강한 프로모터는 글루코오스와 같은 탄소원에서 시스, 시스-류콘산으로의 경로의 작동에 필요한 유전자를 발현하는데 사용된다.

[0039] 외인성 코딩 서열이 염색체 DNA에 통합될 때, 이것은 유전자좌에 통합되며, 그 결실은 어떠한 악영향을 야기하지 않는 것으로 보고되었다. 예를 들어, *ydeM* 으로도 알려진 *E. coli* 박테리아의 물리적 위치 0039에서의 코딩 영역은 지질단백질(lipoprotein)로 주석을 달고, 결실 시에 악영향을 미치지 않는 것으로 입증되었다. 유사하게, *nlpA* 로도 알려진 *E. coli* 박테리아의 물리적 위치 2160에서의 코딩 영역은 라디칼 SAM 도메인 단백질로 주석을 달고, 결실 시에 악영향을 미치지 않는 것으로 입증되었다. 본 발명에서, $P_R-catAX$ 의 복제는 *E. coli* 박테리아의 물리적 위치 0039에서 삽입되었고, P_R-aroG^{FBR} 의 복제는 *E. coli* 박테리아의 물리적 위치 2160에서 삽입되었다. 본 명세서에 기재된 다른 예에서, 삽입은 녹아웃 또는 결실에 유리한 유전자, 예컨대 원치 않는 기능을 암호화하는 유전자, 예를 들어 *ptsI* 유전자 또는 *tyrR* 유전자에서 이루어진다.

[0040] 방향족 아미노산 생합성 경로는 많은 미생물, 특히 *E. coli*에 대해 잘 알려져 있다(Neidhart and Curtiss 1996). 야생형 세포에서, 경로는 피드백 저해 및 전사의 억제에 의해 엄격하게 조절된다. 제 1 개입 단계는 테옥시-아라비노-헵톨로소네이트 7-포스페이트(DAHP) 합성효소에 의해 촉진되며, 거기에는 *aroF*, *aroG* 및 *aroH*에 의해 인코딩되는 3개의 동질효소가 있다. 3개의 동질효소인 AroF, AroG 및 AroH는 방향족 아미노 생합성 경로의 생성물, 즉 티로신, 페닐알라닌 및 트립토판에 의해 피드백 저해된다. AroF, AroG 및 AroH의 피드백 내성 돌연변이체는 잘 알려져 있다(Hu et al. 2003; Lutke-Eversloh and Stephanopoulos 2007). 본 발명의 일 측면은 방향족 아미노산 생합성 경로의 산물에 의한 피드백 저해에 대해 내성이 있는 AroF, AroG 및 AroH 효소 단백질을 발현하기 위하여 *aroF*, *aroG* 및 *aroH* 유전자의 피드백 내성 대립유전자(alleles)의 사용을 포함한다. 피드백 저해에 대해 내성이 있는 AroF, AroG 및 AroH 효소 단백질은 $AroF^{FBR}$, $AroG^{FBR}$ and $AroH^{FBR}$ 라고 한다.

[0041] 방향족 경로에 관련된 오페론 중 몇몇의 전사는 *tyrR* 유전자에 의해 인코딩된 억제자 또는 *trpR* 유전자에 의해 인코딩된 억제자 중 어느 하나, 또는 둘 모두에 의해 조절된다(Neidhardt and Curtiss 1996). 특히 중요한 것은, TyrR 단백질이 하나 이상의 방향족 아미노산과 결합할 때, TyrR 단백질에 의한 *aroG* 및 *aroF*의 전사의 음성 조절이다. 본 발명의 일 측면은 숙주 박테리아 균주의 염색체로부터 이들 유전자를 제거함으로써 *tyrR* 또는

trpR 유전자에 의한 음성 조절의 제거를 포함한다.

- [0042] 본 발명은 류코산 생산에 적합한 생촉매의 유전적 요소의 특정 조합, 예를 들어, 과잉생산된 피드백 내성 *AroG*, 과잉생산된 피드백 내성 *AroF*, 과발현된 *tktA*, 과발현된 *talA*, 강한 항시성 프로모터로부터 *aroZ*, *aroY* 및 *catAX* (또는 이의 유사체 또는 동족체)를 발현하기 위해 염색체로 통합된 카세트, 및 방향족 아미노산 및 비타민에 대해 프로토포피(prototrophy)를 부여하나 원치 않는 방향족 화합물의 현저한 분비를 야기하지 않는 *AroE* 효소를 인코딩하는 유전자로서 정의한 누출성 *aroE* 대립유전자의 다양한 조합을 교시하나, 이에 한정되지 않는다.
- [0043] 본 명세서에 개시된 균주 구조의 모든 구체적인 예는 야생형 대장균 (*Escherichia coli*) C 균주 (ATCC 8739), 또는 대장균 W 균주(ATCC 9637)에 기초한다. 그러나, 이 시점에서, 본 명세서에 개시된 유전적 요소의 발현 카세트 또는 적당한 유사체 및 동족체가 임의의 다른 적당한 미생물, 예컨대 임의의 다른 적당한 *E. coli* 균주 및 발효 공정을 통한 류코산의 상업적 생산에 사용될 수 있는 박테리아, 고세균, 효모, 조류 및 사상균의 다른 종에 결합될 수 있다는 것을 인식해야 한다.
- [0044] *E. coli*에서, 글루코오스로부터의 방향족 아미노산 생합성 경로는 펜토오스 포스페이트 경로 (pentose phosphate pathway; PPP)의 비-산화적 분지(non-oxidative branch)로 시작한다. 비-산화적 펜토오스 포스페이트 경로에서 4개의 주요 효소는 트랜스케톨라제, 트랜스알돌라제, 리불로오스-5-포스페이트 에피머라제 및 리불로오스-5-포스페이트 이소머라제이다. 이들 효소는 핵소오스 또는 펜토오스 당으로부터 에리트로오스 4-포스페이트 (E4P)의 형성을 야기하는 반응을 촉진한다. *E. coli*에서 E4P의 이용률을 증가시키기 위하여, 트랜스케톨라제를 인코딩하는 *tktA* 유전자가 과발현될 수 있다 (Niu et al., 2002). 유사하게, 트랜스알돌라제 유전자의 과발현도 어떤 상황에서 E4P의 이용률을 증가시키는 것으로 예상된다 (Bongaerts et al., 2001). 본 발명의 또 다른 측면에서, 트랜스케톨라제 및 트랜스알돌라제 유전자의 발현은 트랜스케톨라제 및 트랜스알돌라제 효소의 활성의 증가를 야기하는 유전자 조작을 통해 향상된다. 본 발명의 또 다른 측면에서, PPP의 비-산화적 분지를 통한 분출은 리불로오스-5-포스페이트 에피머라제 및 리불로오스-5-포스페이트 이소머라제를 과잉생산함으로써 증가된다.
- [0045] 공통 방향족 아미노산 경로에서 제 1 개입 단계 및 가장 엄격하게 조절된 반응은 (*aroG*, *aroF*, 및 *aroH*에 의해 인코딩된) DAHP 합성효소에 의해 테옥시아라비노-헵톨로소네이트 7-포스페이트 (DAHP)를 생산하기 위한 포스포에놀피루베이트 (PEP)와 E4P의 축합이다. *E. coli*에 의해 소비된 D-글루코오스는 부분적으로 PPP를 통해 그리고 부분적으로 해당 작용(glycolysis)을 통해 방향족 생합성으로 유도된다. 트랜스케탈로오스 (*tktA*) 및 DAHP 합성효소의 동질효소 (*aroG*)가 이들의 복제수를 증가시킴으로써 이들의 발현을 증가시키는 플라스미드로 형질전환을 통해 증폭되면, 방향족 경로로의 글루코오스의 흐름은 크게 증가된다 (Niu et al., 2002). 본 발명의 바람직한 측면에서, 외인성 *aroG* 및 *tktA* 유전자는 트랜스케톨라제 및 DAHP 합성효소의 표소 활성의 증폭을 위해 염색체 DNA 내로 통합된다.
- [0046] 본 발명의 다른 실시예에서, 미생물 세포 내의 PEP를 통한 분출은 PEP의 분출을 다른 경로로 감소시킴으로써 DAHP의 합성에 이용가능한 PEP를 증가시켜 개선된다. 많은 속의 박테리아 세포는 포스포트랜스퍼라제 시스템 (PTS)을 이용하여 세포막을 가로질러 글루코오스의 수송에서 PEP를 소비하며, 여기서 하나의 PEP 분자는 박테리아 외부막을 가로질러 수송된 글루코오스의 모든 분자에 대해 소비된다. PEP-의존성 PTS를 비-PEP 의존성 (PEP 비의존성) 글루코오스 흡수 기전으로 대체하거나 또는 보완함으로써, 미생물 세포 내에서 방향족 아미노산 생합성 경로에 이용가능한 PEP의 풀 크기를 증가시키는 것이 가능하다. 예를 들어, 당 흡수용 PTS 시스템은 GalP-기반 당 흡수 시스템 또는 Glf/Glk 단백질에 기반한 당 수송체 시스템에 의해 대체되거나 또는 보완될 수 있다 (Chandran et al., 2003; Yi et al., 2003). 본 발명의 바람직한 측면에서, 미생물 세포 내에서 PEP 풀을 보존하기 위해 당 흡수용 PTS 시스템을 삭제하는 것 외에도, GalP 기반 당 흡수 시스템은 미생물 세포 내에서 ATP를 보존하기 위해 불활성화된다. PTS 시스템 및 GalP 기반 당 흡수 시스템 둘 모두의 기능에 결합이 있는 미생물 세포 (Δ PTS/ Δ galP)에서, 당 흡수는 Glf를 코딩하는 외인성 유전자 (글루코오스 촉진 확산 단백질) 또는 Glf 및 Glk (글루코키나제) 단백질을 모두를 인코딩하는 외인성 유전자를 도입함으로써 달성될 수 있다. 본 발명에서 사용된 바와 같이, 용어 기능성 글루코오스-촉진 확산 단백질은 임의의 Glf 단백질 및 Glf와 기능적으로 동등하고 촉진 확산에 의해 당을 미생물 세포 내로 수송시키는 기능을 하는 임의의 다른 단백질을 나타낸다. 본 발명의 일 측면에서, 글루코오스 촉진자 단백질 Glf를 코딩하는 유전자는 Δ PTS/ Δ galP인 미생물 세포 내로 도입되고, 미생물 세포 내로 수송된 글루코오스는 내인성 글루코오스 키나제에 의해 인산화된다. 본 발명의 다른 측면에서, Glf 및 Glk 단백질을 모두를 코딩하는 유전자는 Δ PTS/ Δ galP인 미생물 세포 내로 도입된다. 본 발명의 바

람직한 측면에서, 미생물 세포 내로 도입된 외인성 *glf* 및 *glk* 유전자는 숙주 염색체 DNA 내로 통합된다.

[0047] 본 발명의 다른 실시예에서, 성장 및 에너지를 위한 탄소원이 글루코오스신생합성(gluconeogenesis)을 필요로 하는 경우 (예를 들어, 탄소원이 아세테이트 또는 석시네이트인 경우), PEP 풀은 세포 내에 이미 존재하는 카복실화 효소, 예를 들어 *E. coli*에서 *pck*에 의해 인코딩되는 PEP 카복시키나제의 활성을 증가시킴으로써, 또는 외인성 카복실화 효소를 도입함으로써 증가될 수 있다. 바람직한 실시예에서, 카복실화 효소를 코딩하는 도입된 외인성 유전자는 숙주 염색체 내로 안정되게 통합된다. 카복실화 효소를 코딩하는 유전자는 다양한 미생물 종으로부터 유래될 수 있다. 카복실화 효소를 코딩하는 유전자는 추가로 유전자 조작을 받을 수 있어서, 시스, 시스-뮤콘산 생산을 위한 생촉매 내에서 카복실화 효소의 발현이 현저히 향상된다.

[0048] PEP는 방향족 경로에 대한 2개의 주요 대사산물 중 하나이며, Ppc 활성의 감소 또는 제거는 방향족 경로 및 뮤콘산 생산을 위한 PEP를 보존한다. Ppc는 TCA 회로의 중간체인 옥살로아세테이트를 형성하는 보충대사 반응(anaplerotic reaction)을 촉진시킨다. Ppc 활성은 최소 배지에서 야생형 *E. coli* 및 일부 다른 유기체에는 필수적이거나, 다른 것에는 결여되어 있다. Ppc가 없는 효모 같은 일부 유기체는 Pyc를 이용하여 옥살로아세테이트를 보충한다. Ppc 활성의 감소 또는 결여는 PEP 대신에 피루베이트로부터 옥살로아세테이트로의 대체 경로를 제공함으로써 Pyc 활성으로 보완될 수 있는데, 이는 뮤콘산 생산 및 중간체로 PEP를 필요로 하는 방향족 화합물과 같은 다른 화합물의 생산을 위한 PEP의 이용률을 증가시킨다. 본 명세서에 개시된 적어도 하나의 예에서, Ppc를 Pyc로 대체하면 PEP로부터 OAA로의 분출을 감소시킬 수 있으며, 이는 차례로 중심 방향족 경로에 대해 PEP를 보존한다.

[0049] 본 발명의 또 다른 실시예에서, 미생물 세포 내의 PEP 풀은 PEP를 기질로 사용하는 PykA 및 PykF와 같은 피루베이트 키나제 효소의 활성을 감소시키거나 또는 제거함으로써 증가된다.

[0050] DAHP로부터, 방향족 아미노산 경로는 여러 중간체를 거쳐 3개의 방향족 아미노산, 즉 L-티로신 (L-Tyr), L-페닐알라닌 (L-Phe), 및 L-트립토판 (L-Trp)의 생합성을 위한 분지점(branch point)인 코리스메이트 (CHA)로 진행한다.

[0051] 공통 방향족 아미노산 경로의 초기 단계에서, 3-데히드로퀴네이트 (3-dehydroquinate; DHQ) 합성효소 (AroB)는 DAHP로부터 포스페이트 기를 제거하여, DHQ의 형성을 야기한다. 효소 DHQ 탈수효소 (AroD)는 DHQ로부터 물 분자를 제거하여, 3-데히드로쉬키메이트 (DHS)의 형성을 야기하고, 이어서 쉬키메이트 탈수소효소 (AroE)에 의해 쉬키메이트 (SHK)로 환원된다. 쉬키메이트 키나제 I/II (AroK, AroL)는 쉬키메이트를 쉬키메이트 3-포스페이트 (S3P)로 인산화한다. S3P를 PEP와 축합시켜, 5-에놀피루보일쉬키메이트 3-포스페이트 (EPSP)의 형성을 야기한다. EPSP의 형성은 EPSP 합성효소 (AroA)에 의해 매개된다. EPSP로부터의 포스페이트 기는 코리스메이트 합성효소 (AroC)에 의해 제거되어, 코리스메이트 (CHA)의 형성을 야기한다.

[0052] 도 2에 나타난 바와 같이, 방향족 아미노산 경로는 *aroE* 유전자의 돌연변이로 인해 3-데히드로쉬키메이트 (DHS)에서 쉬키메이트 (SHK)로의 전환 수준에서 차단될 수 있어서, DHS의 축적을 야기한다 (Niu et al., 2002). 외인성 *aroZ* 유전자의 도입은 DHS를 프로토크테퀴에이트 (protocatechuate; PCA)로 전환시키는 역할을 한다. 이어서 PCA는 AroY 효소에 의해 매개된 탈카복실화 반응을 통해 카테콜로 전환된다. 카테콜은 궁극적으로 *catA* 유전자 산물의 작용을 통해 시스-시스 뮤콘산 (ccMuA)으로 전환된다. ccMuA는 말레일 아세토아세테이트 이소머라제에 의해 작용하여 트랜스-트랜스 뮤콘산 (ttMuA)을 얻을 수 있다. DHS로부터 ccMuA 및/또는 ttMuA로의 생합성 경로는 뮤콘산 경로라고 한다. DHS의 ccMuA로의 전환을 담당하는 3개의 상이한 유전자는 다양한 미생물 종으로부터 얻어질 수 있고, 대장균과 같은 뮤콘산 생산을 위해 선택된 미생물 내로 도입될 수 있다. 본 발명의 바람직한 실시예에서, 뮤콘산 경로에 관여하는 단백질을 코딩하는 외인성 유전자는 숙주 염색체 DNA 내로 통합된다.

[0053] 방향족 아미노산 경로를 시스, 시스-뮤콘산의 생산으로의 방향 전환에서, *aroE* 유전자의 돌연변이가 중요하다. *aroE* 유전자는 완전히 불활성화되어, 뮤콘산 생산을 위해 기재된 *E. coli*의 WN1/pWN2.248 균주로 수행된 것처럼, 방향족 아미노산의 생합성에서 완전한 차단을 야기할 수 있다 (Niu et al., 2002). WN1/pWN2.248 *E. coli* 균주 및 관련 균주의 중요한 단점은, *aroE* 유전자의 완전한 불활성화로 인해 이 균주가 페닐알라닌, 티로신 및 트립토판과 같은 방향족 아미노산, 및 상기 언급한 방향족 비타민 또는 비타민-유사 화합물에 대해 영양 요구성이 되었다는 것이다. 그 결과, 시스, 시스-뮤콘산의 생산을 위한 균주의 성장 동안 이 균주는 이러한 화합물 (또는 쉬키메이트와 같은 공통 중간체)의 외인성 첨가를 필요로 하며, 이에 의하여 이러한 균주를 이용하는 시스, 시스-뮤콘산의 상업적 생산 비용을 상당히 증가시킨다. 방향족 아미노산의 외인성 공급원에 대한 이러한 의존성을 극복하기 위한 신규한 방법은, *aroE*에서 누출성 돌연변이가 있는 균주를 사용하는 것이다. 누출성 *aroE* 돌연변이체는 상당량의 DHS를 축적하는 동안 쉬킴산에 제한된 탄소 흐름을 허용할 것이며, 이는 AroZ 효소

의 작용에 의해 PCA로의 전환에 이용 가능하다. 따라서, *aroE*의 누출성 돌연변이체 형태의 사용은 여전히 탄소의 흐름을 시스, 시스-뮤콘산으로 전환시키면서, 외인성 방향족 아미노산에 대한 의존성을 제거할 것이다.

[0054] DHS의 시스, 시스-뮤콘산으로의 전환에 필수적인 AroZ, AroY 및 CatA 단백질의 합성을 코딩하는 유전자는 많은 미생물 중 중 어느 하나로부터 유래될 수 있다. 일 실시예에서, 이러한 외인성 유전자는 개발중인 생축매의 숙주 염색체 내로 통합된다. 바람직한 실시예에서, 생축매 내에서 이러한 외인성 유전자의 발현은 어떠한 유도제도 필요 없는 항시성 프로모터에 의해 유도된다.

[0055] 효소 3-데히드로쉬키메이트 탈수효소 (AroZ; EC 4.2.1.118)는 중간체 프로토키테큐에이트의 생합성에 필요하다. 본 명세서에서, "AroZ"는 3-데히드로쉬키메이트 탈수효소 반응을 촉진시키는 임의의 효소를 나타낸다. 선행기술에서, 이 효소는 *클렙시엘라 뉴모니아*(*Klebsiella pneumoniae*) 균주 A170-40 (ATCC25597)의 *aroZ* 유전자로부터 발현된다 (Niu et al., 2002; Draths and Frost, 1995). 그러나, AroZ의 특이적 활성은 유기체에 따라 0.1 내지 261 micromoles/min/mg으로 다양하다 (Wheeler et al, 1996; Fox et al, 2008; Pflieger et al, 2008), 따라서 *K. 뉴모니아* 보다 더 높은 특이적 활성을 갖는 유기체, 예를 들어 *아시네토박터 바일리* (*Acinetobacter baylyi*), *아스페르길루스 니둘란스*(*Aspergillus nidulans*) (Wheeler et al, 1996) [현재 *에머리셀라 니둘란스* (*Emericella nidulans*)로도 알려짐], 또는 *뉴로스포라 크라싸*(*Neurospora crassa*) (Rutledge, 1984; Stroman et al, 1978), 또는 *포도스포라 안세리나*(*Podospora anserina*) [*포도스포라 파우시세타*(*Podospora pauciseta*)로도 알려짐 (Hansen et al, 2009)]로부터의, *asbF* (Fox et al, 2008; Pflieger et al, 2008), *qutC* (Wheeler et al, 1996), *qa-4* (Rutledge, 1984), 및 *quiC*로도 알려진 *aroZ* 유전자를 발현시킴으로써 현저한 개선이 이루어질 수 있다.

[0056] 일 특정 예로서, 3-데히드로쉬키메이트 탈수효소를 인코딩하는 *N. 크라싸*로부터의 *qa-4* 유전자에 대한 코딩 서열은 임의의 몇몇 잘 알려진 방법, 예를 들어 전체 유전자 DNA 합성, cDNA 클로닝, 또는 게놈 DNA 클로닝 및 PCR 또는 합성 DNA 링커 합성의 조합에 의해 얻어질 수 있다. *qa-4* 유전자에 인트론이 없으므로, 코딩 영역은 게놈 DNA로부터 PCR에 의해 얻어질 수 있다 (Rutledge, 1984). *qa-4* 효소의 단백질 서열 (서열번호 4) 및 천연 유전자의 DNA 서열 (서열번호 5)은 알려져 있다.

[0057] 대안적으로, 발현 카세트는 *A. 니둘란스*로부터의 3-데히드로쉬키메이트 탈수효소에 대해 구축될 수 있다. *A. 니둘란스*로부터의 *QutC* 효소에 대한 코딩 서열은 임의의 몇몇 잘 알려진 방법, 예를 들어 전체 유전자 DNA 합성, cDNA 클로닝, 또는 게놈 DNA 클로닝 및 PCR 또는 합성 DNA 링커 합성의 조합에 의해 얻어질 수 있다. *QutC*의 단백질 서열 (서열번호 6) 및 인트론을 함유하지 않은 천연 유전자의 DNA 서열 (서열번호 7; GenBank accession number M77665.1)은 알려져 있다. 발현 카세트는 DNA 합성에 의해, 또는 게놈 클로닝 및 PCR의 조합에 의해 얻어질 수 있어서, *QutC* 효소는 *E. coli*에서 정확하게 생산될 수 있다. *E. coli*에서 강한 항시성 프로모터로부터 *QutC*에 대한 코딩 서열을 발현함으로써, 염색체에 통합된 유전자의 1 또는 2 복제로부터 충분한 발현이 얻어질 수 있어서, 선행기술 (Niu et al., 2002)에 개시되어 있고 불안정성을 야기할 수 있는 다중복제 플라스미드 상의 발현 카세트의 2 이상의 복제의 유지를 필요없게 한다. 상기 기재된 방법을 사용하여 일반적으로 원하는 효소를 코딩하는 DNA 서열을 얻을 수 있고, 그 다음 코딩 서열을 사용하여 *E. coli* 또는 다른 적당한 미생물 숙주 유기체에서 기능하도록 설계된 발현 카세트를 구축할 수 있다.

[0058] AroZ의 특이적 활성은, 실시예 4에 기재된 바와 같이, 보다 강한 프로모터 및/또는 리보솜 결합 부위(ribosome binding site; RBS)가 코딩 영역 앞에 도입되어 있는, 개선된 발현 카세트를 구축함으로써 선행기술 (Niu et al., 2002)의 단백질 서열을 사용함으로써 개선될 수도 있다.

[0059] *클렙시엘라 뉴모니아* 균주 A170-40로부터의 AroZ (3-데히드로쉬키메이트 탈수효소)를 인코딩하는 *aroZ* 유전자는 선행기술에 기재된 바와 같이 얻어질 수 있다. 유전자의 DNA 서열 및 주변 DNA(surrounding DNA)는 당해 기술분야에 잘 알려진 방법에 의해 결정될 수 있다. *aroZ*와 같은 본 발명의 이종 유전자는 천연 DNA 서열을 이용하여 발현 카세트 내에 포함될 수 있거나, 또는 의도된 숙주 유기체에 대한 코돈 최적화 서열로 합성될 수 있다. *aroZ* 유전자는 활성 *aroZ* 유전자를 함유하는 임의의 다른 미생물, 예를 들어, *K. 뉴모니아* 균주 342, *아시네토박터* *Sp. ADP1* (*아시네토박터 바일리* ADP1), *바실러스 튜링겐시스*(*Bacillus thuringiensis*), *에머리셀라 니둘란스*(*Emericella nidulans*), *어위니아 아밀로보라*(*Erwinia amylovora*), *슈도모나스 푸티다* W619(*Pseudomonas putida* W619), *뉴로스포라 크라싸*(*Neurospora crassa*), *아스페르길루스 니둘란스*(*Aspergillus nidulans*) 및 기타 다수에서 기재된 바와 같이 클로닝될 수 있다 (Draths and Frost, 1995).

[0060] 효소 프로토키테큐에이트 탈탄산효소 (AroY; EC 4.1.1.63)는 중간체 카테콜의 생합성을 위해 필요하다. 본 명세서에서, "AroY"는 프로토키테큐에이트 탈탄산효소 반응을 촉진하는 임의의 효소를 나타낼 것이다.

선행기술에서, 이 효소는 다중복제 플라스미드 상에서 *클렙시엘라 뉴모니아* 균주 A170-40 (ATCC25597)의 *aroY* 유전자로부터 발현된다 (Niu et al., 2002). 그러나, 다시 한번, 공정에서의 개선은 숙주 유기체의 염색체에 통합된 발현 카세트의 1 또는 2 복제로부터 효소를 충분히 생산함으로써 얻어질 수 있다. 이는 선행기술의 *K. 뉴모니아* AroY 효소의 특이적 활성보다 더 높은 특이적 활성을 갖는 AroY 효소를 자연적으로 생산하는 유기체로부터 *aroY* 유전자를 얻음으로써, 또는 실시예 4에 기재된 바와 같이, 강한 항시성 프로모터 및/또는 강한 RBS를 사용하는 발현 카세트를 구축하여 *K. 뉴모니아* AroY의 발현 수준을 증가시킴으로써 달성될 수 있다. *K. 뉴모니아* 균주 A170-40으로부터의 AroY에 대한 단백질 서열은 서열번호 8로 제시된다. 해당 유전자인 *aroY*는 상기 기재된 대로 클로닝될 수 있거나 (Draths and Frost, 1995), 또는 단백질 서열에 기초하여, 의도된 숙주 유기체에 대한 최적화 코돈으로 합성될 수 있다.

[0061] *aroY* 유전자는 동족체 또는 유사체를 함유하는 임의의 다른 미생물, 예를 들어, *K. 뉴모니아* 균주 NCTC418 (ATCC15380), *클렙시엘라 뉴모니아* 342, 및 *아르술라 아데니니보란스*(*Arxula adenivorans*)로부터 얻어질 수 있다 (Sietmann et al, 2010). *클렙시엘라 뉴모니아* 342로부터의 *aroY* 유전자의 DNA 서열 및 주변 DNA는 서열번호 9로 제시된다.

[0062] 효소 카테콜 1,2-디옥시게나제 (CatA; EC 1.13.11.1)는 시스, 시스-뮤콘산 생합성의 마지막 단계에서 필요하다. 본 명세서에서, "CatA"는 카테콜 1,2-디옥시게나제 반응을 촉진하는 임의의 효소를 나타낼 것이다. 선행기술에서, 이 효소는 다중복제 플라스미드 상에서 *아시네토박터 칼코아세트익스*(*Acinetobacter calcoaceticus*) 균주 ADP1의 *catA* 유전자로부터 발현된다 (Niu et al., 2002). 공급원 균주인 *아시네토박터 칼코아세트익스* 균주 ADP1은, 명백하게 *아시네토박터* Sp. ADP1 및 *아시네토박터 바일*이 ADP1으로 재명명되었다 (Neidle and Ornston, 1986; Barbe et al, 2004; de Berardinis et al, 2008). 이 선행기술의 예에서, *catA* 유전자는 락토오스 또는 유도제로서 IPTG (이소프로필티오갈락토시드) 중 어느 하나를 필요로 하는 P_{tac} 프로모터로부터 발현되었다. 이들 화합물은 상업적 발효에서 사용하기에는 너무 비싸다, 그래서 다시, 고가의 유도제에 대한 필요성을 제거하고 발현 카세트를 염색체에 통합시킴으로써 더 안정한 균주를 생성하기 위해 공정의 상당한 개선이 필요하다. 이것은 다른 실시예에서 상기 기재된 바와 같은 강한 항시성 프로모터, 강한 RBS, 및/또는 더 안정한 mRNA를 사용하는 *catA* 유전자에 대한 발현 카세트를 구축함으로써 달성될 수 있다.

[0063] *아시네토박터 바일*이 ADP1으로부터의 *catA* 유전자의 DNA 서열 및 주변 서열은 서열번호 10으로 제시된다. 동일한 균주로부터의 CatA에 대한 단백질 서열은 서열번호 11로 제시된다. 바람직한 실시예에서, *catA*에 대한 발현 카세트는, *catA* 유전자의 발현 수준을 증가시키기 위하여, *catA*로부터 하류에 자연적으로 존재하는 하나 또는 2개의 추가 개방형 판독 프레임(open reading frames)을 함유한다 (Schirmer and Hillen, 1998). 많은 다른 유기체, 예를 들어 *슈도모나스 아빌라* (*Pseudomonas arvilla*), *슈도모나스 플루오레센스*(*Pseudomonas fluorescens*) (Nakazawa et al, 1967; Kojima et al, 1967), *스트렙토마이세스* Sp. 균주 2065 (Iwagami et al, 2000), *쿠프리아비두스 네카토르* 335T(*Cupriavidus necator* 335T) 및 기타 다수 (Perez-Pantoja et al, 2008)는 *catA* 유전자에 대한 공급원일 수 있다.

[0064] 시스, 시스-뮤콘산으로의 탄소의 흐름을 개선하기 위하여, 누출성 *aroE* 돌연변이체를 이용하여 DHS로부터 쉬키메이트 (SHK)로의 탄소의 흐름을 감소시키는 것 이외에, 방향족 아미노산 경로에서 분지하는 특정 다른 경로를 차단할 필요가 있다. *아시네토박터* 및 *슈도모나스* 속의 일부 박테리아는 DHS를 갈산으로 전환하는 효소인 p-히드록시벤조에이트 히드록실라제를 인코딩하는 *pobA*라는 유전자를 함유한다. PobA 동족체 또는 유사체가 *E. coli*에서 발견되지는 않았지만, DHS를 생산하도록 조작된 *E. coli*의 균주는 측정가능한 양의 갈산을 분비하므로 (Li and Frost, 1999), 이러한 효소는 *E. coli*에 존재할 가능성이 있다. 또한, DHS로부터 유래된 PCA는 *pobA* 유전자에 의해 코딩된 p-히드록시벤조에이트 히드록실라제 (PobA) 효소의 작용에 의해 갈산으로 전환될 수 있다. 이렇게 생산된 갈산은 이후에 피로갈롤(pyrogallol)로 전환될 수 있다. 개선된 시스, 시스-뮤콘산을 위해 선택된 생축매의 갈산 및 피로갈롤로의 탄소 흐름을 차단하기 위한 하나의 방법은, 유전자 조작을 통해 p-히드록시벤조에이트 히드록실라제 (PobA) 단백질의 활성을 차단하거나 또는 감소시키는 것이다. 유사하게, DHS의 전구체인 DHQ는 *aroE*에 의해 코딩된 쉬키메이트 탈수소효소에 의해 작용하여, 퀴산의 생산을 야기할 수도 있다. 본 발명의 실시예에서, 누출성 AroE 돌연변이체 효소는 DHQ를 퀴산으로 전환시키는 것에 대한 무능력 또는 감소된 능력을 추가로 선택하거나 또는 스크리닝한다.

[0065] 시스, 시스-뮤콘산 대신 트랜스, 트랜스-뮤콘산의 생산에는 몇가지 장점이 있다. 트랜스, 트랜스-뮤콘산은 테레프탈산의 생산을 위해 에틸렌과의 딜스-알더 반응(Diels Alder reaction)에서 시스, 시스 뮤콘산보다 바람직하다. 유전자 조작된 방향족 경로와 함께 생축매는 화학 전환 공정을 이용하여 세포 외부의 트랜스, 트랜스-뮤콘

산으로 전환될 수 있는 시스, 시스-뮤콘산을 생산한다. 반면에, 말레일아세토아세테이트 이소머라제 또는 유사한 이소머라제 효소를 생촉매에 도입함으로써, 박테리아 생촉매 내에서 시스, 시스-뮤콘산을 트랜스, 트랜스-뮤콘산으로 전환하는 것이 가능하다.

[0066] 일 실시예에서, 본 발명은 3,4-디히드록시벤조산 (프로토카테큐산 또는 PCA)의 카테콜로의 전환에 관여하는 AroY 단백질의 활성을 향상시키는 유전적 방법을 제공한다. AroY 단백질의 활성은 글루코오스의 뮤콘산으로의 생물학적 전환에서 주요 제한 및 병목으로 확인되었다 (Horwitz et al 2015; Weber et al 2012; Curran et al 2013; Sonoki et al 2014). AroY는 많은 박테리아 사이에 널리 퍼져있는 비-산화적 탈탄산효소의 종류에 속하며 매우 다양한 기질을 이용한다 (Lupa et al 2005). 비-산화적 탈탄산효소를 인코딩하는 많은 이들 유전자는 B, C, D 형의 유전자를 인코딩하는 3개의 유전자 오페론으로 구성된다. "C" 형 유전자는 AroY와 같은 탈탄산효소를 인코딩하는 반면, B 및 D 형 유전자의 특정 기능은 이들이 C 형 탈탄산효소의 완전한 활성을 실현하는데 필요한 것으로 종종 나타났지만 알려지지 않았다 (Lupa et al 2005; Jimenez et al 2013; Lin et al 2015; Sonoki et al 2014). 웨버(Weber) 등 (2012)은 *클렙시엘라 뉴모니아* 또는 *세디멘티박터 히드록시벤조이쿠스* (*Sedimentibacter hydroxybenzoicus*)의 B, C 및 D 유전자를 높은-복제수 효모 벡터 pRS4K-HKT7에 클로닝하고, 외부 첨가된 PCA가 있는 배지에서 발효를 수행하여 이들 유전자 군으로부터 PCA 탈탄산효소 활성의 존재를 결정하였다; 그러나, C 유전자에 의해 인코딩된 PCA 탈탄산효소가 이들 유전자 군의 B 또는 D 유전자에 의존적이었는지 여부를 결정하기 위한 어떠한 노력도 없었으며, B, C 및 D 유전자를 포함하는 이들 유전자 군의 발현이 글루코오스와 같은 비-방향족 탄소원을 사용하여 뮤콘산 생산을 향상시킬 수 있었는지 여부를 결정하려는 어떠한 시도도 없었다.

[0067] 본 명세서에 개시된 특정 예에서 사용되는 AroY를 인코딩하는 유전자는 *클렙시엘라 뉴모니아*로부터 유래되나, 이 *aroY* 유전자의 국소 유전자 구조는 B 또는 D 형 효소를 인코딩하는 전사적으로 결합된 유전자를 나타내지 않는다. 이전 연구 결과에 따르면, *클렙시엘라 뉴모니아* (4-히드록시벤조산 탈탄산효소의 오페론 부분)의 다른 B 형 유전자인 *kpdB*의 포함은 리그닌-관련 방향족 화합물로부터 뮤콘산을 생산할 때 AroY 활성을 증가시킬 수 있다 (Sonoki et al 2014). 그러나, 리그닌은 화학 물질의 복잡한 혼합물이며, 뮤콘산과 같은 순수한 생성물을 생산하기 위해서는 고가의 하류 정화 공정을 필요로 한다. 존슨(Johnson) 등 (2016)은 *엔테로박터 클로아카* (*Enterobacter cloacae*)로부터의 AroY (PCA 탈탄산효소)를 발현하는 *슈도모나스 푸티다* (*Pseudomonas putida*) 균주에서 KpdB와 89.3%의 서열 동일성을 갖는 *E. 클로아카*로부터의 EcdB 단백질의 공-발현이 탄소원으로 글루코오스를 사용하여 54시간 발효 종료시에 1.44 g/L 내지 4.92 g/L의 뮤콘산 생산을 증가시켰음을 나타내었다; 그러나, 이 *슈도모나스 푸티다* 기반 시스템에서 뮤콘산 수율은 여전히 극단적으로 (0.077 mol/mol)로 나타났으며, 따라서 이 *슈도모나스 푸티다* 기반 시스템은 상업적 규모의 적용에 적합하지 않다. 따라서, 탄수화물 및 기타 비-방향족 화합물 (상기 참조)과 같은 저렴하고 순수한 비-방향족 탄소원으로부터 개선된 뮤콘산 생산 공정에 대한 필요성이 여전히 존재한다. 최근 연구에 따르면, KpdB의 동족체인 UbiX가 프레닐화 플라빈 모노뉴클레오티드 보조인자(prenylated flavin mononucleotide cofactor)를 생산하며, 이 보조인자가 유비퀴논 형성에서 UbiD의 탈카복실화 활성을 지지한다는 것을 나타내었다 (White et al 2015; Payne et al 2015). UbiX 및 이의 동족체의 이전 특성이 4-히드록시-3-폴리프레닐벤조에이트 탈탄산효소, 히드록시벤조에이트 탈탄산효소, 아단위 B 단백질, 페놀산 탈탄산효소, 또는 페닐아크릴산 탈탄산효소와 같은 단백질을 나타낼지라도, 이들 단백질은 이제 더 정확하게 플라빈 프레닐트랜스퍼라제로 주석을 달았다.

[0068] 뮤콘산의 생산에서 추가 제한은 전구체 대사산물인 PEP와 E4P의 제한에 기인할 수 있다. 이들 대사산물의 소비를 제거하거나 또는 감소시키는 변화가 본 발명에서 생성물 형성을 개선시키는 것으로 나타났다. 예를 들어, PEP 이용률은 천연 글루코오스 수송인 포스포트랜스퍼라제 시스템 (PTS)을 대체 시스템 (*지모모나스 모빌리스* (*Zymomonas mobilis*)의 Glf-Glk)으로 대체하여 향상시켰다. *E. coli*의 PTS의 본래 사용은 글루코오스의 수송 및 인산화를 위해 PEP의 이용을 필요로 한다. Glf-Glk 시스템은 글루코오스의 수송 및 인산화를 위해 촉진 확산기 및 글루코키나제 + ATP를 각각 이용한다. PTS 대신 Glf-Glk의 사용은 여러 가지 방향족 산물에 대한 수율과 역가를 향상시키는 것으로 나타났다. PEP는 다양한 생물학적 반응을 위한 기질이다. PTS에 의한 글루코오스 수송 이외에, PEP는 피루베이트 키나제의 기질이며, ATP 및 피루베이트의 생산을 야기하여 해당 작용에서 생성된 ATP의 절반을 제공한다. 피루베이트 키나제 (*pykA* 또는 *pykF*)를 인코딩하는 유전자의 불활성화는 방향족 경로에서 생성물의 수율을 증가시키는 것으로 나타났다 (Escalante et al. 2010). 또한, PEP는 포스포에놀피루베이트 카복실라제 (Ppc)에 의한 옥살로아세테이트의 생산을 위한 기질이다. *E. coli*에서, 이것은 필수 유전자이고 (Baba et al., 2006), *ppc*에 대해 결실된 균주는 최소 배지에서 성장할 수 없다. 포스포에놀피루베이트 카복실라제를 가지고 있지 않으나 대신 피루베이트로부터 피루베이트 카복실라제 (Pyc)로 옥살로아세테이트를 보충하는 다른 유기체, 예를 들어, *S. 세레비지애*가 있다. *E. coli*에서 피루베이트 카복실라제의 첨가가 연구되었으나, 임의의

방향족 산물에 대해서가 아닌 석시네이트 생산의 개선에 대해서만 연구되었다 (Lin et al. 2004; Vemuri, Eiteman, and Altman 2002).

도면의 간단한 설명

도 1. *E. coli*에서 방향족 아미노산 생합성의 경로.

도 2. *E. coli*에서 류콘산 생합성의 경로.

도 3. 3-데옥시-아라비노-헵툴로네이트 7-포스페이트 합성효소(DHAP synthase)의 류콘산으로의 전환 반응 단계.

도 4. 총 류콘산 및 생화학적 중간체의 HPLC 분석에 사용된 기준을 나타내는 크로마토그래프.

도 5. 류콘산 이성질체의 HPLC 분석에 사용된 기준을 나타내는 크로마토그래프.

도 6. 플라스미드 pCP32AMP, pCP14 및 pCP54로 형질전환된 *E. coli* 균주 MYR34에서 DHS의 생산에 대한 역가. *E. coli*의 MYR34 균주는 *aroE* 유전자의 결실을 가진다. 플라스미드 pCP32AMP는 DAHP 합성효소를 코딩하는 *aroG* 유전자를 발현한다. 플라스미드 pCP14는 DHQ 합성효소를 코딩하는 *aroB* 유전자를 발현한다. 플라스미드 pCP54는 *aroB* 및 *aroG* 유전자 모듈을 발현한다.

도 7. 플라스미드 pCP32AMP 및 pCP54로 형질전환된 *E. coli* 균주 MYR34 및 MYR170에서 DHS의 생산에 대한 역가. *E. coli*의 MYR34 균주는 *aroE* 유전자의 결실을 가진다. MYR170 균주는 숙주 염색체 DNA의 *ack* 유전자좌 (locus)에서 통합된 P₁₅ 프로모터의 조절 하에 *aroE* 유전자의 결실 및 *aroB* 유전자의 제 2 복제를 가진다. 플라스미드 pCP32AMP는 DAHP 합성효소를 코딩하는 *aroG* 유전자를 발현한다. 플라스미드 pCP54는 *aroB* 및 *aroG* 유전자를 모두 발현한다.

도 8. 플라스미드 pMG37 단독으로 형질전환되거나 또는 pMG37 및 pCP32AMP 플라스미드와 함께 형질전환된 *E. coli* 균주 MYR34 및 MYR170에서 시스, 시스-류콘산의 생산에 대한 역가. *E. coli*의 MYR34 균주는 *aroE* 유전자의 결실을 가진다. MYR170 균주는 숙주 염색체 DNA의 *ack* 유전자좌에서 통합된 P₁₅ 프로모터의 조절 하에 *aroE* 유전자의 결실 및 *aroB* 유전자의 제 2 복제를 가진다. 플라스미드 pCP32AMP는 DAHP 합성효소를 코딩하는 *aroG* 유전자를 발현한다. 플라스미드 pMG37는 류콘산 경로에서 기능하는 단백질을 코딩하는 *aroZ*, *aroY*, 및 *catAX* 유전자들을 발현한다.

도 9. *aroG* 유전자 단독을 발현하는 플라스미드 (pCP32AMP) 또는 *aroG* 및 *tktA* 유전자를 동시에 발현하는 플라스미드 (pCP50)로 형질전환된 *E. coli*의 MYR170 균주에서 DHS의 생산에 대한 역가. MYR170 균주는 숙주 염색체 DNA의 *ack* 유전자좌에서 통합된 P₁₅ 프로모터의 조절 하에 *aroE* 유전자의 결실 및 *aroB* 유전자의 제 2 복제를 가진다.

도 10. 플라스미드 pCP32AMP 및 pCP50으로 형질전환된 *E. coli*의 MYR34 및 MYR170 균주로부터의 DHS 수율. DHS 수율은 소비된 글루코오스의 그래프 당 생산된 DHS의 그래프로 계산된다. 플라스미드 pCP32AMP는 *aroG* 유전자를 발현하고, pCP50은 *aroG* 및 *tktA*를 발현한다. 박테리아 균주 MYR34는 *aroE* 유전자의 결실을 가진다. *E. coli*의 MYR170 균주는 MYR34로부터 유래되고, 염색체 DNA 상의 *ack* 유전자좌에서 통합된 추가의 *aroB* 유전자를 가진다.

도 11. 플라스미드 pCP32AMP 및 pCP50으로 형질전환된 *E. coli*의 MYR170 및 MYR261 균주로부터의 DHS 역가. 플라스미드 pCP32AMP는 *aroG* 유전자를 발현하고, pCP50은 *aroB* 및 *tktA* 유전자를 발현한다. MYR170 균주는 숙주 염색체 DNA의 *ack* 유전자좌에서 통합된 P₁₅ 프로모터의 조절 하에 *aroE* 유전자의 결실 및 *aroB* 유전자의 제 2 복제를 가진다. *E. coli*의 MYR261 균주는 *E. coli*의 MYR170 균주로부터 유래된다. *E. coli*의 MYR261 균주는 염색체 DNA의 *poxB* 유전자좌에서 통합된 이의 천연(native) 프로모터를 갖는 *tktA* 유전자의 제 2 복제를 가진다.

도 12. 쉬킵산 생합성 경로에서 DAHP 합성효소를 코딩하는 *aroG*를 발현하는 플라스미드 pCP32AMP 및 류콘산 경로에서 기능하는 단백질을 코딩하는 *aroZ*, *aroY* 및 *catAX* 유전자를 발현하는 플라스미드 pMG37로 형질전환된 *E. coli* 균주 MYR170, MYR261 및 MYR305에서 류콘산 및 아세트산 생산. MYR170 균주는 숙주 염색체 DNA의 *ack* 유전자좌에서 삽입된 P₁₅ 프로모터의 조절 하에 *aroE* 유전자의 결실 및 *aroB* 유전자의 추가 복제를 가진다. MYR261 및 MYR305는 MYR170 균주의 유도체이다. MYR261은 숙주 염색체 DNA 상의 *poxB* 유전자좌에서 통합된 *tktA* 유전자의 추가 복제를 가지며, MYR305는 숙주 염색체 DNA 상의 *poxB* 유전자좌에서 결실을 가진다.

도 13. *E. coli* 균주 MYR34에 의해 생산된 내인성 DHS의 류콘산으로의 전환. *E. coli*의 균주 MYR34는 쉬키메

이트 탈수소효소(shikimate dehydrogenase)를 코딩하는 *aroE* 유전자의 결실을 가진다. 결과적으로, DHS의 축적이 있다. 균주 MYR34가 류콘산 경로에서 기능하는 단백질을 코딩하는 *aroZ*, *aroY* 및 *catAX* 유전자를 발현하는 플라스미드로 형질전환된 경우, DHS가 류콘산으로 전환된다. 그러나, MYR34 균주가 임의의 외인성 유전자가 없는 공 플라스미드 벡터 (pCL1921)로 형질전환된 경우, DHS의 류콘산으로의 전환은 발생하지 않는다.

도 14. DHS가 류콘산 경로로 전환하는 능력에 대한 *aroZ* 유사체의 비교. 3개의 상이한 *aroZ* 유사체, 즉, *아시네토박터* *sp.* ADP1으로부터의 *quiC*, *바실러스* *튜링겐시스*(*Bacillus thuringiensis*)로부터의 *asbF*, 및 *뉴로스포라* *그라싸* (*Neurospora crassa*)로부터의 *qa-4*는 P₁₅ 및 람다 P_R 프로모터로부터 각각 *catAX* 및 *aroY* 유전자를 발현하는 저-복제 플라스미드(low-copy plasmid)에서 P₂₆ 프로모터 하에서 클로닝되었다. 이러한 3개의 상이한 플라스미드 구조체는 형질전환을 통해 MYR34에서 발현되었고, 생산된 류콘산의 양이 측정되었다.

도 15. *catAX*, *aroY* 및 *quiC*의 단일 복제는 *E. coli*의 MYR170 균주 내에 염색체로 통합되어(Δ *aroE*, Δ *ack::P₁₅-aroB*), MYR352를 야기하였다(서열번호: 41). 또한, MYR170은 류콘산 경로의 작동에 필요한 모든 유전자를 가진 저-복제 플라스미드 pMG37로 형질전환되어, MYR219 균주를 야기하였다. MYR352 및 MYR219 모두는 YE_p24 (중-복제 공백터(medium-copy empty vector)) 또는 pCP32AMP (중-복제 *aroG* 발현 플라스미드) 또는 pCP50 (중-복제 *aroG* 및 *tkfA* 발현 플라스미드)로 형질전환되었고, 생산된 PCA, 카테콜 및 류콘산의 양은 HPLC 방법을 사용하여 정량되었다.

도 16. *catAX*의 발현을 증가시킴으로써 MYR352에서 카테콜 축적의 제거. MYR352는 *aroY* 단백질을 발현하는 플라스미드 (pMG27), 또는 *quiC* 단백질을 발현하는 플라스미드 (pMG39), 또는 3개의 류콘산 경로 유전자, 즉, *catAX*, *aroY* 및 *quiC*를 모두 발현하는 플라스미드 (pMG37), 또는 류콘산 경로에서 단지 2개의 유전자, 즉, *catAX* 및 *aroY*를 발현하는 플라스미드 (pMG33)로 형질전환되었다. *catAX* 단독의 과발현은 카테콜의 축적을 방지하기에 충분하였다.

도 17. 글루코오스의 이입(importing)을 위해 상이한 시스템을 이용한 균주의 성장. *ptsHI* 및 *galP*의 결실 (MYR31)은 최소 글루코오스 배지에서 성장의 결핍을 야기하는 반면, *glf* 및 *glk* 유전자의 도입 (MYR217)은 다시 성장을 가져온다. 대조 균주 MYR34는 Δ *aroE*이지만, 그렇지 않으면 야생형이다. 3개의 방향족 아미노산 및 3개의 방향족 비타민을 배지에 첨가하여 영양요구성(auxotrophic) 균주의 성장을 허용하였다.

도 18. *E. coli*의 MYR34 및 MYR217 균주에서 DHS 생산. DHS의 생산을 야기하는 플라스미드로 형질전환된 경우, 글루코오스 이입을 위해 *glf-glK*를 이용하는 MYR217은, 포스포트랜스퍼라제 시스템 (PTS)을 이용하는 MYR34의 형질전환체보다 더 높은 역가의 DHS를 생산하였다.

도 19. 7 리터 발효기에서 *E. coli*의 MYR428 균주에 의한 류콘산의 생산. Δ *aroE* Δ *ackA::P₁₅-aroB* Δ *poxB::tkfA*의 유전자형을 갖는 *E. coli*의 MYR261를 플라스미드 pCP32AMP 및 pMG37로 형질전환시켜, *E. coli*의 MYR428 균주를 생성하였다.

도 20. 류콘산을 생산하도록 유전자 조작된 *E. coli* 균주 MYR993 및 2개의 MYR993 유도체인 MYR993 Δ *ubiX* 및 MYR993 Δ *ubiD*에서 류콘산 및 PCA 생산. MYR993 Δ *ubiX* *E. coli* 균주는 *ubiX* 유전자의 코딩 영역을 카나마이신 내성을 코딩하는 카세트로 대체함으로써 MYR993 균주로부터 유래되었다. MYR993 Δ *ubiD* *E. coli* 균주는 *ubiD* 유전자의 코딩 영역을 카나마이신 내성을 코딩하는 카세트로 대체함으로써 MYR993 균주로부터 유사하게 유래되었다.

도 21. UbiX의 다양한 동족체의 상대적 활성 측정. UbiX 동족체의 활성 측정은 290nm에서의 흡광도 (A₂₉₀)의 감소로 측정된 PCA의 탈카복실화로부터 수행되었다. 5개의 상이한 UbiX 동족체, 즉 KpdB, UbiX, Elw, Kok 및 Lpl은 본 연구에서 사용되었다.

도 22. 낮은 또는 높은 수준의 *kpdB* 유전자 발현을 갖는 *E. coli* 균주에서 류콘산 및 PCA 생산. 외인성 *kpdB* 유전자를 갖지 않는 *E. coli* 균주 MYR1305를 부모 균주로 사용하였고, P₂₆ 프로모터 또는 *E. coli* *pgi* 프로모터로부터 발현된 *kpdB* 유전자를 갖는 저 복제 플라스미드로 형질전환시켰다. P₂₆ 프로모터로부터의 유전자 발현은 비교적 낮은 수준으로 예상되는 반면, *pgi* 프로모터로부터의 유전자 발현은 비교적 높은 수준으로 예상된다.

도 23. *E. coli* 균주 MYR1674 및 이의 유도체 MYR1772에서 류콘산 생산. MYR1772는 *ppc*의 코딩 영역과 프로모터를 P_R-*pyc* 유전자로 대체함으로써 MYR1674로부터 유래되었다. P_R은 콜리파지 람다로부터의 강한 우측 프로모터

의 약자이다.

발명을 실시하기 위한 구체적인 내용

- [0070] 이 특허 출원의 명세서에는 뮤콘산의 효율적인 생산을 위한 미생물 균주의 구축과 관련된 발명의 여러 가지 측면을 제공한다. 당업자는 본 발명의 여러 가지 측면을 편집하여 뮤콘산의 생산에 매우 높은 효율성을 갖는 생축매를 구축할 수 있다.
- [0071] **실험 부분**
- [0072] **총설**
- [0073] **균주 및 접종물(inoculum) 제조:** 본 발명에서 사용된 박테리아 균주의 목록은 표 1 및 표 2에 제공된다. 본 발명에서 사용된 플라스미드의 목록은 표 3에 제공된다. 본 명세서에 개시된 균주 구조의 모든 구체적인 예는 야생형 *E. coli* C 균주 (ATCC 8739), 또는 *E. coli* K-12 균주(YMC9 또는 MM294)로부터 유래되나, 본 명세서에 개시된 유전적 요소는 임의의 다른 적당한 *E. coli* 균주에 결합될 수 있으며, 본 명세서에 개시된 유전적 요소의 발현 카세트 또는 적당한 유사체 및 동족체는 임의의 다른 적당한 미생물, 예를 들어 발효 공정을 통한 시스, 시스-뮤콘산의 상업적 생산에 사용될 수 있는 박테리아, 고세균, 효모, 조류 및 사상균의 다른 종에 결합될 수 있다.
- [0074] *E. coli* C는 AM1 무기 배지(mineral media)에서 10% 글루코오스를 발효시킬 수 있다. AM1 배지는 2.63 g/L $(\text{NH}_4)_2\text{HPO}_4$, 0.87 g/L $\text{NH}_4\text{H}_2\text{PO}_4$, 1.5 mM MgSO_4 , 1.0 mM 베타인, 및 1.5 ml/L 미량 원소를 함유한다. 미량 원소는 1000X 스톡(stock)으로서 제조되고 하기 성분들을 함유한다: 1.6 g/L FeCl_3 , 0.2 g/L $\text{CoCl}_2 \cdot 6\text{H}_2\text{O}$, 0.1 g/L CuCl_2 , 0.2 g/L $\text{ZnCl}_2 \cdot 4\text{H}_2\text{O}$, 0.2 g/L NaMoO_4 , 0.05 g/L H_3BO_3 , 및 0.33 g/L $\text{MnCl}_2 \cdot 4\text{H}_2\text{O}$. 발효액의 pH는 1.0 - 10.0 M KOH 또는 1.0 - 9.0 M 수산화암모늄을 사용하여 7.0으로 유지된다.
- [0075] **발효:** 유전자 조작되고 -80°C 냉동고에 저장된 *E. coli* 균주의 40% 글리세롤 스톡으로부터 신선한 NBS-2% 글루코오스 (Jantama et al., 2008a) 플레이트 상에 도말하여 발효를 시작하였다. (만약 있다면) 플라스미드는 한천 플레이트 및 액체 배지에 적당한 항생제(들)을 포함하는 것으로 유지된다. 암피실린 (나트륨 염)을 150 mg/L로, 스펙티노마이신 HCL을 100 mg/L로, 테트라사이클린 HCL을 15 mg/L로, 및 카나마이신 설페이트를 50 mg/L로 사용하였다. 24 내지 48시간 후 (37°C), 단일 콜로니를 진탕 플라스크(shake flask) 내 25 ml의 동일한 배지에서 채취한다. 세포가 약 1.0의 OD_{600} 으로 성장할 때까지 37°C 에서 200 rpm으로 진탕한 후, 배양물을 얼음으로 냉각시키고 동일한 부피의 무균 80% 글리세롤을 첨가하였다. 그 다음, 발효를 위한 접종물로서 사용하기 위하여, 2ml 분취액을 -80°C 에서 동결시켰다. 용어 "역가"는 발효액의 단위 부피당 생산된 발효 산물의 양을 의미하며, 용어 "수율"은 생산된 발효 산물 대 소비된 탄소원의 비율 (g/g 또는 mol/mol)을 의미한다.
- [0076] **세포 성장:** Thermo Electronic Spectronic 20 분광 광도계를 이용하여 550 nm (OD_{550}) 또는 600 nm (OD_{600})에서 광학 밀도를 측정하여 세포 질량을 평가하였다.
- [0077] **쉬김산 경로 및 뮤콘산 경로에서 중간체의 분석:** 시스, 시스-뮤콘산 및 시스, 트랜스-뮤콘산, 및 기타 생화학적 중간체를 포함하는 발효액에서 생산된 총 뮤콘산을, Waters Alliance 기기를 구비한 HPLC로 분석하고, 시그마-알드리치로부터 구입한 표준을 이용하여 210 nm에서 흡광도 또는 45° 에서 굴절률을 모니터링하였다. 컬럼은 BioRad Aminex HPX-87H로, 50°C 에서 이동상으로 8 mM 황산을 0.6 ml/min의 유속으로 40분 동안 수행하였다. 구입한 표준 (시그마-알드리치)의 크로마토그램은 도 4에 나타내었다. HPLC를 준비하기 위해, 발효 시료를 0.05 M 인산 칼륨 완충액, pH 7.0에서 10배 또는 100배 희석하여, 시스, 시스-형태의 뮤콘산을 시스, 트랜스-형태로 이성질체화하는 것로부터 보호한다.
- [0078] 뮤콘산의 이성질체를 분리하기 위하여, 상기와 같이 제조된 시료를 제 2 HPLC 시스템에서 행하였다. 기기는 Agilent 1200 HPLC이었고, 컬럼은 인산으로 pH 3.0으로 조정된 30% 메탄올 내 50mM KH_2PO_4 를 이동상으로 30°C 에서 작동하는 Agilent Eclipse XDB-C18, 4.6 x 150 mm이었다. 유속은 278 nm에서 흡광도에 의한 검출로, 4분 동안 1 ml/min이었다. 시스, 트랜스-뮤콘산 표준은 시스, 시스-뮤콘산을 물에 용해시켜, HPLC 피크가 완전히 새로운 위치로 이동될 때까지, 실온에서 약 2시간 동안 자발적 산 촉매 이성질체화를 거쳐 만들었다. 다른 표준은 시그마-알드리치에서 구입하였다. 표준들이 나타내는 크로마토그램은 도 5에 나타내었다.
- [0079] **발효 공정을 위한 뮤콘산 생산 배지의 조성:** 발효 배지의 각 리터에는 50 ml/L의 1M KH_2PO_4 , 10 ml의 200g/L 시

트르산 + 25 g/L 구연산철(Ferric citrate), 1.2 ml의 98% 황산, 및 한 방울의 소포제(Antifoam) 204를 함유한다. 이들 성분은 충분한 물과 혼합하여 하기의 다른 성분을 추가할 수 있는 공간을 허용한다. 고압 멸균 처리 후, 하기 성분들을 첨가하였다: 10, 20, 30 또는 40 ml의 50% 글루코오스 (5, 10, 15, 또는 20 g/l 최종을 얻음), 2 ml의 1M MgSO₄, 1 ml의 0.1M CaCl₂, 10 ml의 1000X 미량 원소 (Jantama et al. 2008a), 및 필요한 경우, 1, 2, 4, 또는 8 ml의 50 g/L 페닐알라닌 + 50 g/L 티로신 + 50 g/L 트립토판 (0.5, 0.1, 0.2, 또는 0.4 g/l 최종을 얻음), 10 ml의 1 g/L p-히드록시벤조산 + 1 g/L p-아미노벤조산 + 1 g/L 2,3-디히드록시벤조산 (마지막 3개의 화합물은 방향족 "비타민" 또는 "비타민-유사 화합물"이라고 함), 및, 필요에 따라, 1 ml의 150 mg/ml 암피실린(나트륨 염) 및/또는 1 ml의 100 mg/ml 스펙티노마이신 HCl. 방향족 아미노산과 비타민은 필요하지 않았고, 기능성 AroE 단백질을 발현하는 균주에는 사용되지 않았다.

[0080] 진탕 플라스크의 경우, NBS 염 (Jantama et al. 2008a) + 0.2 M MOPS 완충액, pH 7.4는 상기 기재된 예비-고압 멸균 혼합물로 대체하였지만, 글루코오스 및 다른 첨가제는 동일하였다. 유가식(fed batch) 발효의 경우, 공급 병은 600 g/L의 무수 글루코오스, 및 필요한 경우, 32 ml/L의 50 g/L 페닐알라닌 + 50 g/L 티로신 + 50 g/L 트립토판을 함유하였다. 9M NH₄OH를 염기로 사용하여 발효 배지의 pH를 유지하였다. 방향족 아미노산과 비타민은 필요하지 않았고, 기능성 AroE 단백질을 발현하는 균주에는 사용되지 않았다.

[0081] DCU 조절기 또는 Biocommand Software에 의해 조절된 pH, DO, 온도, 글루코오스 및 공급 속도를 가진 7L New Brunswick 과학적 발효기에서 유가식 발효를 수행하였다. 배지는 50 mM K₂HPO₄, 20 mM K₂SO₄, 3 mM MgSO₄ 및 미량 원소를 함유하였다. 미량 원소는 50X 스톡으로서 제조되고 하기 성분들을 함유한다: 1.6 g/L FeCl₃, 0.1 g/L CuCl₂, 0.2 g/L ZnCl₂ · 4H₂O, 0.2 g/L NaMoO₄, 0.05 g/L H₃BO₃, 및 0.55 g/L MnCl₂ · 4H₂O. 온도를 37°C로 유지하였고, 9N 암모늄수로 pH를 7.0으로 유지하였다. 폭기 (Aeration)는 0.5 vvm이었고, 임펠러의 속도를 750 rpm에서 1200 rpm으로 자동 증가시켜 용존 산소 (DO)를 30%로 유지하였다. 배지 내 초기 글루코오스 농도는 약 20 내지 25 g/L이었다. 농도가 5 g/L 이하로 떨어지면, 공급 글루코오스 용액을 발효기에 첨가하였다. 초기 글루코오스 공급 속도는 4 g/L/hr이었고, 7 g/L/hr의 공급 속도로 48시간까지 증가시킨 후, 7 g/L/hr로 유지시켰다.

[0082] **뮤콘산 경로 유전자를 발현하는 플라스미드의 구축:** DHS를 뮤콘산으로 전환하는데 필요한 3개의 이중 유전자를 단독으로 또는 저-복제 플라스미드인 pCL1921와 조합하여 클로닝하였다 (Lerner and Inouye, 1990). pCL1921의 DNA 서열은 표 7에서 서열번호 20으로 제시된다. 간략하게, *catAX*, *aroY* 및 *aroZ* 유사체 또는 동족체의 코딩서열은 *E. coli*에서 발현을 위해 코돈-최적화되었고, 상업적으로 합성되었다 (GeneArt, Invitrogen). 그 다음, 독특한 리보솜-결합 부위를 갖는 정방향 프라이머 및 각 유전자에 대해 독특한 종결자 서열을 갖는 역방향 프라이머를 이용하여 이들 서열을 PCR 증폭시켰다. 결과의 PCR 단편을 제한 효소로 절단하고, 표준 분자 클로닝 과정에 의해 독특한 항시성 프로모터 서열의 하류를 클로닝하였다. 프로모터 서열은 이전에 기재된 공급원 DNA 서열로부터 PCR 증폭 (미국특허출원 20090191610; 미국특허 7244593)한 다음, 제한효소 절단 및 표준 분자 클로닝으로 클로닝하였다. 프로모터-RBS-코딩서열-종결자 서열은 함께 발현 카세트를 구성하였다. 그 다음, 개별 발현 카세트를 조합하여 하나, 둘 또는 3개 모두 뮤콘산 경로 유전자를 발현하는 플라스미드를 생성하였다.

[0083] 본 발명은 하기의 예를 이용하여 더 예시된다: 그러나, 단독으로 또는 이들의 임의의 조합으로 본 명세서에 제공된 예는 본 발명의 범위 또는 실시예를 제한하는 것으로 해석되어서는 안된다. 마지막에 제공되는 청구범위는 발명의 범위를 정의한다. 당업자는 청구범위에서 정의된 본 발명의 범위를 명확하게 이해할 수 있다. 당업자는 본 발명의 정신 및 범위를 벗어나지 않고 본 발명에 의해 제공되는 기술적 해결 수단을 변형 또는 변화시킬 수 있다.

[0084] 실시예 1

[0085] *aroG* 및 *aroF* 유전자의 발현의 증가

[0086] *E. coli*의 *tyrR* 유전자는 다수의 잘 알려진 방법들, 예를 들어 화학적 또는 방사선 돌연변이유발 및 스크리닝 (예를 들어, PCR 및 DNA 시퀀싱에 의함) 또는 유사체 내성에 대한 선택 (예를 들어, 4-플루오로티로신에 대한 내성), 트랜스포존 돌연변이유발, 박테리오파지 Mu 돌연변이유발, 또는 형질전환 중 어느 하나에 의해 돌연변이될 수 있다. 바람직한 실시예에서, *tyrR* 유전자의 돌연변이는 삭제돌연변이 (검출가능한 활성을 남기지 않는 돌연변이)이고, 더 바람직한 실시예에서는, *tyrR* 유전자의 적어도 일부는 결실된다. 이것은 선행 DNA 분자를 이용한 2-단계의 형질전환 방법을 이용하여 달성될 수 있다 (Jantama et al, 2008a; Jantama et al, 2008b). 제 1 단계에서, *cam*^R, *sacB* 카세트는 *tyrR* 유전자좌에 통합되어 클로람페니콜 내성을 선택하는 이중 재조합에 의해

tyrR 개방형 판독 프레임의 대부분 또는 모두를 대체한다. 제 2 단계에서, *tyrR* 유전자의 결실된 버전을 포함하는 선형 DNA는 LB와 같은 농축 배지에서 5% 수크로오스에 대한 내성을 선택하는 이중 재조합에 의해 통합된다. 정확한 결실은 진단 중합효소연쇄반응 (PCR)에 의해 동정되고 확인된다. *tyrR*의 결실의 목적은 *aroG* 및 *aroF*의 발현을 증가시키는 것이다. 유사한 결과를 얻는 대안적인 방법은 *aroG* 및/또는 *aroF* 앞에 있는 천연 프로모터를 강한 항시성 프로모터로 대체하고, 필요하다면, 전사 종결자를 추가하는 것이다. 이것이 어떻게 수행되는지에 대한 더 상세한 내용은 일반적으로 하기 실시예 4에 제시된다.

[0087] *tyrR* 결실이 *aroLM*과 같은 유전자의 원치 않는 과발현을 야기할 수 있기 때문에, TyrR 단백질에 의한 AroG 및 AroF 활성의 억제를 극복하기 위한 상기 기재된 2개의 방법 중 후자가 바람직하다 (Neidhardt and Curtiss, 1996). 이것이 어떻게 수행되는지에 대한 더 상세한 내용은 일반적으로 하기 실시예 4에 제시된다.

[0088] 실시예 2

[0089] 피드백 내성 AroG 및 AroF

[0090] 피드백 내성 AroG 효소 (3-데옥시-D-아라비노헵톨로소네이트-7-포스페이트 합성효소 또는 DAHPS)를 야기하는 *aroG* 유전자의 돌연변이는 당해 기술분야에 잘 알려져 있다 (Shumilin et al, 1999; Kikuchi et al, 1997; Shumilin et al, 2002). 또한, 이러한 돌연변이를 생성, 동정 및 특징화하는 방법도 잘 알려져 있다 (Ger et al., 1994, Hu et al., 2003). 바람직한 돌연변이는 페닐알라닌에 의한 저해에 대해 완전한 내성을 야기하는 돌연변이이다. 임의의 알려진 공개된 피드백 내성 돌연변이는 다수의 잘 알려진 방법 중 임의의 방법에 의해 염색체에 함유된 *aroG* 유전자 내로 또는 플라스미드 상에 도입될 수 있으며, 이의 일 예는 원하는 돌연변이가 PCR 프라이밍(priming) 올리고뉴클레오타이드의 일부로서 합성되는 돌연변이유발 PCR이다 (Hu et al., 2003). 돌연변이의 정확한 도입은 DNA 시퀀싱으로 확인된다. *E. coli* C로부터의 야생형 *aroG* 유전자의 서열은 서열번호 18로 제시된다. 바람직한 돌연변이는 코돈 150을 CCA에서 CTA로 변화시킴으로써, AroG의 아미노산 150을 프롤린에서 류신으로 변화시키는 점 돌연변이이다 (Hu et al, 2003). 더 바람직한 실시예에서, *E. coli*에서 바람직한 코돈인 코돈 150은 CCA에서 CTG로 변한다. 인코딩된 DAHP 합성효소가 페닐알라닌에 의한 저해에 대해 최대 3 mM 까지 완전히 내성을 가지며 야생형 효소와 유사한 특이적 활성을 가지기 때문에, *aroG*의 이러한 특정 대립유전자가 바람직하다 (Hu et al., 2003).

[0091] 추가의 피드백 내성 *aroG* 대립유전자는 돌연변이유발 및 하나 이상의 페닐알라닌 유사체, 예를 들어 베타-2-티에닐알라닌, p-플루오로페닐알라닌, p-클로로페닐알라닌, o-플루오로페닐알라닌, 및 o-클로로페닐알라닌에 대한 내성에 대해 선택한 다음, 내성을 야기하는 돌연변이가 *aroG* 유전자에 결합되는 것을 입증함으로써 얻어질 수 있다 (Ger et al., 1994; 미국특허 4,681,852). *aroG*와의 결합은 페닐알라닌의 존재 및 부재 하에 DNA 시퀀싱 또는 효소 분석에 의해 직접적으로 (Ger et al., 1994), 또는 파지 매개 형질 도입 및 *aroG* 유전자좌에서 또는 근처에서 선택되거나 또는 반대로 선택될 수 있는 유전자 마커에 대한 선택에 의해 간접적으로 입증될 수 있다 (미국특허 4,681,852). 이러한 유전자 마커는 *aroG* 유전자 자체의 결실 또는 점 돌연변이일 수 있거나, 또는 *E. coli*의 경우 *nadA*와 같은 임의의 적당히 가깝게 결합된 유전자의 돌연변이일 수 있다. *E. coli*의 예를 들면, 돌연변이유발 및 페닐알라닌 유사체 내성에 대해 선택한 후, 개별 돌연변이체 또는 돌연변이체의 풀은 3개의 모든 DAHP 합성효소 유전자인 *aroG*, *aroF*, 및 *aroH*에 대해 결실되고, 적당한 최소 배지 상에서 성장을 위해 선택된 천연 수령자(recipient)로의 P1 매개 형질도입을 위한 공여자(donors)로서 사용될 수 있다. 그러면, 형질도입체는 원하는 유전자(들)의 돌연변이에 대해 풍부해질 것이다. 대안적으로, 돌연변이유발 및 유사체 내성에 대해 선택한 후, 개별 돌연변이체 또는 돌연변이체의 풀은 니코틴아미드가 결합된 적당한 최소 배지 상에서의 성장을 위해 다시 선택되는 *nadA* 유전자의 삭제돌연변이를 함유하는 천연 수령자로서의 P1 매개 형질도입을 위한 공여자로서 사용될 수 있다. 또 다른 방법은 *aroG* 유전자 근처, 예를 들어 *nadA* 유전자에서 삽입된 트랜스포존, 예를 들어 Tn10을 함유하는 배경 균주에서 내성 돌연변이체를 선택하는 것이다. 유사체 내성 돌연변이체로부터, 테트라사이클린 또는 다른 적당한 항생제 내성에 대해 선택하는 상기 트랜스포존을 함유하지 않는 배경 균주로의 P1 형질도입은 원하는 *aroG* 돌연변이를 풍부하게 할 것이다. 모든 이러한 방법에서, 피드백 내성은 효소 분석 및 유전자의 DNA 시퀀싱에 의해 궁극적으로 확인된다. 우리는 피드백 저해에 대해 내성이 있는 *aroG*의 대립유전자를 *aroG**로 나타낼 것이다.

[0092] 균주 WM191 ($\Delta tyrR$, $\Delta aroF$)은 YMC9 (ATCC 33927)로부터 유래되었다. 2 단계의 유전자 대체 방법 (Jantama et al., 2008a)을 사용하여 *tyrR* 및 *aroF*에서 깨끗한 결실을 도입하였으며, 균주 WM191을 얻었다. 다음으로, *nadA::Tn10* 대립유전자를 CAG12147 (CGSC 7351, Coli Genetic Stock Center, Yale University)로부터 형질도입시켜, 균주 WM189 ($\Delta tyrR$, $\Delta aroF$, *nadA::Tn10*)를 얻었다. 선택은 LB + 테트라사이클린 HCl (15 mg/l) 상에서

이루어졌다. 균주 RY890 ($\Delta tyrR::kan$, *aroF363*)은 P1 형질도입에 의한 3 단계로 MM294 (ATCC 33625)로부터 유래되었다. 공여자 균주는 순서대로 JW1316-1 (CGSC 9179, Coli Genetic Stock Center, Yale University), NK6024 (CGSC 6178, Coli Genetic Stock Center, Yale University), 및 AB3257 (CGSC 3257, Coli Genetic Stock Center, Yale University)이었고, 3개의 선택은 순서대로 LB + 카나마이신 설페이트 (50 mg/l), LB + 테트라사이클린 히드로클로라이드 (15 mg/l), 및 티아민 HCl (5 mg/l)과 함께 NBS 최소 글루코오스 (Jantama et al., 2008a)이었다.

[0093] WM189를 약 20% 생존하도록 자외선으로 돌연변이화시키고, o-플루오로페닐알라닌 (1 mM), 티아민 (5 mg/l) 및 니코틴아미드 (1 mM)를 함유하는 NBS 최소 글루코오스 배지 (Jantama et al., 2008a) 상에 플레이트팅하였다. 여러 플레이트 각각으로부터의 콜로니를 별도의 폴에 모으고, P1vir 용해물(lysates)을 각각의 폴 상에서 제조하였다. 이 용해물을 사용하여 WM191을 LB 배지에서 테트라사이클린 내성 (15 mg/l)으로 형질도입시키고, 얻어진 콜로니를 o-플루오로페닐알라닌 (1 mM), 티아민 (5 mg/l) 및 니코틴아미드 (1 mM)를 함유하는 NBS 최소 글루코오스 배지에 복제 플레이트팅하였다. 테트라사이클린 및 유사체 모두에서 생존한 콜로니 복제는 *aroG* 내에 피드백 내성 돌연변이를 함유하는 것으로 추정되었다. DNA 시퀀싱을 위하여, 5개의 독립 풀로부터 8개의 개별 콜로니를 선택하였다. *aroG* 코딩 영역은 중합효소연쇄반응에 의해 증폭되고 서열화되었다. 표 4에 나타난 결과는, 8개의 균주 각각이 그들의 *aroG* 유전자에 점 돌연변이를 함유하고 있음을 나타내었다. 대립유전자들 중 일부는 공개된 대립유전자와 동일하였지만, 일부는 신규하였다.

[0094] 상기 기재된 폴 중 하나로부터의 P1vir 용해물을 사용하여 상기 기재된 복제 플레이트팅에 의해 RY890 (*aroG* 야생형 대립유전자를 가짐)을 테트라사이클린 내성 및 o-플루오로페닐알라닌 (0.3 mM)에 대한 내성으로 형질도입시켰다. RY893, RY897, RY899, 및 RY901로 명명된 4개의 콜로니를 DNA 시퀀싱을 위해 선택하였고(표 4), 다시, 대립유전자 중 2개는 공개된 대립유전자와 동일하였으나, 2개는 신규하였다. 4개의 균주에 대해 동질유전자(isogenic)이나 야생형 *aroG* 유전자를 함유하는 균주 RY902는, CAG12147로부터의 형질도입에 의해 대조균으로서 구축되었다. 이러한 5개의 균주는 진탕 플라스크 내의 25 ml NBS 최소 글루코오스 (15 g/l) + 티아민 HCl (5 mg/l) 및 니코틴아미드 (1 mM)에서 밤새도록 성장시켰다. 결과의 세포를 원심분리에 의해 수확하고, 재현탁시켜 10 ml의 물로 행구고, 재-원심분리하고, 0.5 ml의 50 mM 인산칼륨, pH 7.0에서 재현탁시켰다. 현탁된 세포를 3 방울의 클로로포름으로 볼텍싱(vortexing)하여 용해시키고, 하기 변형과 함께 문헌 (Hu et al., 2003)에 기재된 방법과 유사한 방법을 이용하여 DAHP 합성효소 활성화에 대하여 조 용해물을 분석하였다. 인산염 완충액은 50 mM (최종 농도), pH 7.0이었고, 최종 에리트로오스-4-포스페이트 농도는 2 mM이었으며, 최종 포스포에놀 피루베이트 농도는 5 mM이었고, 배양 온도는 30°C이었으며, 반응은 10분에서 중지하였다. 우리는 1 mU를 밀리그램 단백질당 분당 1 nMole의 DAHP를 생산하는 활성으로 정의한다. 피드백 내성을 시험하기 위해, 각각의 조 용해물을 18 mM의 최종 농도의 페닐알라닌과 함께 또는 없이 분석하였다. 분석 결과는 표 5에 나타내었다. 효소는 페닐알라닌에 대해 다양한 특이적 활성 및 내성을 나타내었으나, 시험된 모든 선택된 돌연변이체 버전은 야생형 대조균보다 현저히 더 내성이었다.

[0095] 상기 기재된 RY893, RY899, RY901, 및 RY902로부터의 *aroG* 대립유전자는, 하기와 같이 류콘산 생산 배경 균주로 도입되었다. *aroG** 및 *aroGwt* 공여자 균주로부터의 P1vir 용해물을 사용하여, MYR219 (*E. coli* C, $\Delta aroE$, $\Delta ack::P_{15}-aroB$, pMG37)를 테트라사이클린 HCl 내성(15 mg/l)에 형질도입하여, 각각 새로운 균주 RY903, RY909, RY911, 및 RY912를 얻었다. 그 다음, 각각의 이들 균주를 JW1316-1의 P1vir 용해물을 이용하여 카나마이신 설페이트 내성 (50 mg/l)에 형질도입하여, $\Delta tyrR::kan$ 대립유전자를 도입하고, 각각 균주 RY913, RY919, RY921, 및 RY922를 얻었다. 스펙티노마이신 선택은 류코닉 플라스미드를 유지하기 위하여 전반적으로 유지되었다. 결과의 4개의 균주는 진탕 플라스크 내의 20 g/l 글루코오스, 0.2 M MOPS 완충액, pH 7.4, 니코틴아미드 (1 mM), 페닐알라닌 (100 mg/l), 티로신 (100 mg/l), 트립토판 (100 mg/l), p-히드록시벤조산 (1 mg/l), p-아미노벤조산 (1 mg/l), 2,3-디히드록시벤조산 (1 mg/l), 페놀 레드 (10 mg/l), 및 황산암모늄 (1 g/l)의 보충제를 함유하는 25 ml NBS 최소 배지 (Jantama et al., 2008a)에서 37°C에서 48시간 동안 성장시켰다. pH는 진탕 플라스크에 요구되는 1.0 M KOH의 1 ml 분취액을 수동으로 첨가하여, pH 7.0 표준에 대해 페놀 레드의 색으로부터 눈으로 추정되는 7에 가깝게 유지시켰다. 생산된 류콘산은 상기 기재된 바와 같이 HPLC로 분석하였고, 결과는 표 6에 나타내었다. 피드백 내성 *aroG** 대립유전자를 함유하는 3개의 균주 모두는 야생형 *aroG* 대립유전자를 함유하는 동질유전자 균주보다 더 많은 류콘산을 생산하였다. 본 명세서에 개시된 별도의 실험에서, 다중복제 플라스미드 pCP32AMP 상에 *aroG*를 함유하는 균주 MYR205는 진탕 플라스크에서 1.5 g/l의 류콘산을 생산하였다. 따라서, 본 발명자들은 $\Delta tyrR$ 및 단일 복제 염색체 *aroG**의 조합이 동질유전자 *aroG* 플라스미드 함유 균주에 비해 더 잘 수행하여 진탕 플라스크에서 류콘산을 생산할 수 있음을 나타내었다. 플라스미드 대립유전자에 비해 염색체 대

립유전자의 내재하는 우수한 유전적 안정성에 더하여, 플라스미드 내에 보유하기 위한 선택 배지에 대한 필요성의 완화는, 본 명세서에 기재된 신규한 균주를 대규모의 상업적 발효에 더 적합하게 만든다. 또한, 류콘산 경로 유전자의 발현에는 화학 유도제가 필요하지 않았다. 따라서, 상기 기재된 본 발명의 균주는 선행기술의 균주에 비해 개선되며 (Niu et al., 2002), 이들 모두는 바람직하지 않은 다중복제 플라스미드 상의 DAHP 합성효소의 과발현을 위한 유전자를 함유한다.

[0096] AroG에 대해 상기 기재된 것과 유사한 방식으로, 티로신에 의한 피드백 저해에 대해 내성이 있는 AroF 또는 AroH 동질효소를 야기하는 돌연변이는 플라스미드 상에 또는 염색체 내에 도입될 수 있다. 바람직한 돌연변이는 코돈 148을 CCG에서 CTG로 변화시켜 (Weaver et al., 1990), *aroF**라는 유전자를 제공함으로써, AroF의 아미노산 148을 프롤린에서 류신으로 변화시키는 점 돌연변이이다. *aroF**의 다른 대립유전자는 *aroG** 대립유전자에 대해 상기 기재된 것과 유사한 방식으로, 티로신 유사체 (예를 들어, o-플루오로티로신, m-플루오로티로신, p-플루오로페닐알라닌 등)에 대한 내성으로 분리될 수 있다. *aroF** 대립유전자는, 예를 들어, JW2584 (CGSC 10051, Coli Genetic Stock Center, Yale University)와 같은 균주에서와 같이 가까이 결합된 $\Delta yfiR::kan$ 에서 트랜스포존 또는 카나마이신 내성 삽입과의 결합에 의해 선택되고, 풍부해지며, 형질도입될 수 있다.

[0097] 실시예 3

[0098] 염색체 DNA로부터 *aroE*의 결실 및 류콘산 생산

[0099] 이 실시예에서, 류콘산 경로에서 기능하는 단백질을 코딩하는 유전자의 발현 뿐만 아니라 다중복제 플라스미드 상에서 *aroB* 및 *aroG*의 과발현의 효과를 연구하였다. 이 연구에서, 부모 균주로서 쉬키메이트 탈수소효소를 코딩하는 *aroE* 유전자의 결실을 함유하는 균주 MYR34를 사용하였다. *aroE*의 염색체 복제의 결실은 상기 실시예 1에 기재된 것과 유사한 방식으로 수행되었다. MYR34가 쉬김산 경로에서 기능하는 DAHP 합성효소 단백질을 코딩하는 *aroG* 유전자를 과발현하는 플라스미드 pCP32AMP로 형질전환되었을 때, DHS의 축적이 유의하게 증가하였다. MYR34가 항시성 프로모터로부터 *aroB*를 발현하는 플라스미드로 형질전환되었을 때, DHS의 축적이 유의하게 증가하지 않았다. 그러나, *E. coli* 균주 MYR34가 *aroB* 및 *aroG* 유전자 모두를 발현하는 플라스미드로 형질전환된 경우, *aroG* 단독으로 형질전환된 MYR34에서 관찰된 것보다 DHS의 축적이 증가하여, DHS 생산에서 *aroB*가 2차 병목으로 시사되었다 (도 6).

[0100] 도 7에서 제시된 실험에서, 숙주 염색체 DNA에 통합된 *aroB* 유전자의 추가 복제의 효과를 시험하였다. MYR34로부터 유래된 *E. coli* 균주 MYR170에서, P₁₅ 프로모터의 조절 하에 *aroB* 유전자의 추가 복제는 *ack* 유전자좌에서 숙주 염색체 내로 통합되었다. MYR170 균주가 pCP32AMP 플라스미드로 형질전환된 경우, 동일한 플라스미드로 형질전환된 MYR34 균주에서 검출된 DHS 축적과 비교할 때 DHS 축적이 약간 증가하였다. MYR170에서 DHS의 축적의 약간의 증가는 숙주 염색체 DNA 내에 통합된 *aroB* 유전자의 추가 복제에 기인할 수 있다. MYR170이 *aroB* 및 *aroG* 유전자 모두 발현하는 pCP54로 형질전환된 경우, DHS 축적의 추가 증가로 인해, DHS 생산에서 *aroB*가 2차 병목으로 시사되었다.

[0101] 도 8은 *E. coli* 균주 MYR34 및 MYR170에 의한 류콘산 생산의 결과를 제공한다. *aroB* 및 *aroG* 유전자의 과발현이 있는 *aroE* 결실 균주 MYR34 및 MYR170에서 DHS의 축적이 있음을 입증하였고, 류콘산 생산 경로에서 기능하는 단백질을 코딩하는 "류코닉 경로" 유전자의 발현이 DHS를 시스, 시스-류콘산으로 전환시킬 수 있는지 알아보기 위한 노력이 이루어졌다. 이러한 실험에서, *E. coli* 균주 MYR34 및 MYR170은 플라스미드 pM37 단독으로 형질전환되거나, 또는 플라스미드 pMG37 및 pCP32AMP 모두로 형질전환되었다. 플라스미드 pMG37은 류콘산 경로에서 기능하는 단백질을 코딩하는 *aroZ*, *aroY* 및 *catAX* 유전자를 발현한다. MYR34 및 MYR170 모두에서의 류콘산 생산은, pMG37 플라스미드만으로 형질전환된 이들 두 균주에서의 류콘산 생산에 비해, 이들 박테리아 균주가 플라스미드 pCP32AMP 및 pMG37 모두로 형질전환될 때 증가하였으며, 이는 이들 균주에서 *aroB* 발현이 시스, 시스-류콘산 생산에 대해 병목이라는 것을 시사한다.

[0102] 실시예 4

[0103] *tktA*의 과발현

[0104] *tktA*에 의해 인코딩된 트랜스케톨라제는 펜토오스 포스페이트 경로에서 주요 효소이고, 류콘산의 생산에서 주요 중간체 중 하나인 에리트로오스-4-포스페이트의 생산을 제한하는 것으로 생각된다. 다중복제 플라스미드 상에 천연 프로모터를 갖는 유전자를 도입함으로써 (Sprenger et al., 1995, 1995a), 트랜스케톨라제를 인코딩하는 *tktA*의 과발현은, 방향족 경로로의 분출을 개선시키는 것으로 알려져 있다 (Draths et al., 1992). 그러나, 이러한 플라스미드는 불안정하고, 종종 유지를 위해 항생제 선택을 필요로 한다. 선행기술에서 또 다른 방법은

tktA 유전자의 하나의 추가 복제를 숙주 균주의 염색체에 첨가하는 것이었다 (Niu et al., 2002). 그러나, 천연 프로모터를 갖는 *tktA*의 하나의 추가 복제는 에리트로오스-4-포스페이트로 방향족 경로를 포화시키기에 충분하지 못하며, 이는 이의 천연 프로모터가 이상에 매우 가깝지 않기 때문이다. 따라서, 공정은 상당한 개선이 필요하다.

[0105] *tktA*의 개선된 과발현은, 염색체에서 천연 *tktA* 프로모터를 강한 항시성 프로모터, 예를 들어 *바실러스 서브틸리스* 파지 SP01로부터의 P₁₅ 또는 P₂₆ 프로모터 (각각 서열번호 1 및 서열번호 2), 또는 박테리오파지 람다로부터의 P_R 프로모터 (서열번호 3)로 대체함으로써 얻어질 수 있다. 이것은, 제 1 단계에서 *cam^R*, *sacB* 카세트를 사용하여 천연 염색체 *tktA* 프로모터를 대체하는 것을 제외하고는, 실시예 1에 기재된 바와 같이 2 단계로 수행된다. 제 2 단계에서, 강한 항시성 프로모터는 상기 강한 항시성 프로모터의 하류측에서 *tktA* 코딩 영역의 5' 말단의 적어도 50 염기들 및 상류측에서 천연 *tktA* 프로모터의 바로 상류에 적어도 50 염기쌍의 동족체가 측면에 위치해 있는 강한 항시성 프로모터를 포함하는 선형 DNA로 형질전환시키고, 수크로오스 내성을 선택함으로써 도입된다. 이러한 발현 카세트로부터의 개선된 발현은 발현 카세트로부터 전사되는 mRNA의 안정성을 증가시킴으로써 달성된다. mRNA 안정성의 개선은 mRNA의 5' 말단, mRNA의 3' 말단, 또는 둘 모두에 줄기 고리 구조(stem loop structure)를 추가함으로써 달성된다. 줄기-고리 구조는 종종 로-독립적 전사 종결자(rho-independent transcription terminator)에 의해 자연적으로 종결되는 mRNA의 말단에 존재하지만, 그렇지 않은 경우, 로-독립적 전사 종결자는 잘 알려진 유전공학 방법 (결찰, PCR 등)에 의해 DNA 서열에 첨가될 수 있다. 이러한 종결자는 3 이상의 염기의 "고리(loop)"에 의해 분리된 다음 T가 풍부한 하나 이상의 염기의 영역 (테옥시티미딘)에 의해 분리된 각 반복에서 4 내지 20 염기의 역위 반복(inverted repeat)으로 구성될 수 있다. 역위 반복은 G 및 C (테옥시구아니딘 및 테옥시시티딘)가 풍부하다. 유사하게, 줄기-고리는 전사의 시작점으로부터 바로 하류이나, 리보솜 결합 부위 앞에 상기 기재된 줄기-고리를 함유하나 T-풍부 영역은 포함하지 않는, DNA 서열을 삽입함으로써 mRNA의 5' 말단에 구축될 수 있다. 이러한 예는 P₁₅ 프로모터 (서열번호 1)와 관련하여 제시된다.

[0106] *tktA* 유전자의 과발현이 쉬김산 경로를 통한 탄소의 흐름에 미치는 영향의 분석에서, *E. coli* 균주 MYR170을 부모 균주로 사용하였다. MYR170은 쉬키메이트 탈수소효소를 코딩하는 *aroE* 유전자의 결실 및 *ack* 유전자좌에서 *aroB* 유전자의 추가 복제를 가진다.

[0107] 도 9, 10 및 11에 기재된 실험에서, 2개의 상이한 플라스미드, 즉 pCP32AMP 및 pCP50을 사용하였다. 플라스미드 pCP32AMP는 이의 천연 프로모터로부터 DAHP 합성효소 *aroG* 유전자만을 발현하고, 플라스미드 pCP50은 *aroG* 유전자와 함께 이의 천연 프로모터로부터 트랜스케톨라제 유전자 *tktA*를 발현한다. *aroE* 결실 및 염색체 DNA의 *ack* 유전자좌에서 통합된 P₁₅ 프로모터의 조절 하에 *aroB* 유전자의 추가 복제를 갖는, MYR170은 pCP32AMP 및 pCP50 플라스미드로 개별적으로 형질전환시켰다. 도 8에 나타난 바와 같이, DHS 축적은, *aroG* 유전자만을 발현하는 *E. coli* 세포와 비교하였을 때, *tktA* 유전자와 함께 *aroG* 유전자의 발현으로 더 증가하였다.

[0108] 도 10은 2개의 상이한 균주, 즉 플라스미드 pCP32AMP 또는 pCP50으로 형질전환된 MYR34, MYR170에서 DHS 수율에 대한 자료를 제공한다. *aroE* 유전자 결실을 갖는 MYR34 균주는 소비된 글루코오스의 그램당 0.1g의 DHS를 생산하였다. 이 균주가 *aroG* 유전자 과발현을 갖는 pCP32AMP 플라스미드로 형질전환된 경우, MYR34에서 DHS 수율은 소비된 글루코오스의 그램당 0.15g의 DHS로 증가하였다. MYR170은 *ack* 유전자좌에서 삽입된 *aroB* 유전자의 추가 복제를 가진다. 이러한 *aroB* 유전자의 추가 복제의 존재의 결과로서, pCP32AMP로 형질전환된 MYR170 균주에서 DHS 생산에 대한 수율은 pCP32AMP로 형질전환된 MYR34 균주에서 언급된 DHS 수율보다 약간 더 높았다. 따라서, MYR170에서 *aroB*의 추가 복제의 존재는 쉬김산 경로를 통한 탄소 흐름의 증가를 야기하였다. MYR170 균주가 *aroG* 및 *tktA* 유전자 모두를 발현하는 플라스미드 pCP50으로 형질전환되었을 때, DHS 수율의 추가 증가가 관찰되었다. 따라서, *tktA*의 추가 복제의 존재는 쉬김산 경로를 통한 탄소 흐름의 증가를 설명하였다. 보다 구체적으로는, 추가 *aroB* 및 *tktA* 유전자의 존재의 효과는 DHS 수율에 대해 부가적인 효과를 야기하였다.

[0109] 도 11에 기재된 실험에서 사용된 MYR261을 유전자 조작하여, *poxB* 유전자좌에서 MYR170의 염색체 DNA 내로 *tktA* 유전자의 추가 복제를 통합하였다. MYR261 균주에서 원하는 유전자 대체 (*poxB::tktA*)를 PCR을 통해 확인하였다. MYR261은 pCP32AMP (*aroG* 과발현) 플라스미드 또는 pCP50 (*aroG* 및 *tktA* 과발현) 플라스미드 중 어느 하나로 형질전환되었다. 대조균주로서, MYR170은 pCP32AMP 플라스미드로 형질전환되었다. 도 11에 나타난 결과에서 알 수 있듯이, MYR261의 염색체 DNA에서 *tktA* 유전자의 추가 복제의 존재는, pCP32AMP 플라스미드로 형질전환된 MYR170 균주에서 관찰된 DHS 생산에 대한 역가와 비교할 때, pCP32AMP 플라스미드를 이용한 DHS 생산에 대한 역가를 증가시켰다. 트랜스케톨라제를 과발현하는 플라스미드 pCP50으로 형질전환될 때, MYR261 균주

에서 트랜스케톨라제 수준의 추가 증가는 DHS 생산에 대한 역가의 추가 증가를 야기하였다. *poxB*에 의해 인코딩된 효소, PoxB, 또는 피루베이트 옥시다제는 반응 생성물로서 아세테이트를 생산한다. 이와 같이, 본 명세서에 기재된 바와 같이 *tktA*의 삽입으로 인한 *poxB*의 결실은 아세테이트 생산에 대한 잠재적인 활성 경로를 제거한다. 유사하게, 하기 실시예 12에 기재된 바와 같이, 동시에 *P₁₅aroB*의 삽입 및 AckA를 인코딩하는 *ackA*의 결실, 또는 아세테이트 키나제는 아세테이트로의 또 다른 잠재적인 활성 경로를 제거한다. 아세테이트의 생산은 일반적으로 발효에서 바람직하지 않다 (Jantama et al., 2008b). 따라서, 이러한 결실은 아세테이트 생산을 감소시키는데 유용할 수 있다.

[0110] 도 12는 플라스미드 pCP32AMP 및 pMG37로 형질전환한 후에 *E. coli*의 MYR170, MYR261 및 MYR305 균주에서 뮌콘산 및 아세트산 생산에 대한 역가를 제공한다. MYR305는 염색체 DNA로부터 *poxB* 유전자를 결실시킴으로써 MYR170으로부터 유래되는 반면, MYR261은 *poxB* 유전자가 *tktA* 유전자의 추가 복제를 삽입함으로써 불활성화된 MYR170 유도체이다. 상기한 바와 같이, 플라스미드 pCP32AMP는 *E. coli* 균주 MYR170, MYR261 및 MYR305에서 *aroE* 유전자의 결실로 인해 DHS의 축적을 야기하는 쉬깁산 생합성 경로에서 기능하는 DAHP 합성효소 단백질을 코딩하는 *aroG* 유전자를 발현한다. 도 2에 예시된 바와 같이, 뮌콘 경로 유전자, 즉 플라스미드 pMG37 상에서 *aroZ*, *aroY* 및 *catAX*의 발현으로, DHS는 시스, 시스-뮌콘산으로 전환된다. MYR261 균주에서 *aroB* 유전자 및 *tktA* 유전자의 추가 복제가 존재하면, 아세트산의 축적의 감소와 함께 뮌콘산의 생산이 약간 증가하였다.

[0111] 실시예 5

[0112] *talA* 또는 *talB*의 과발현

[0113] *talB* 유전자는 *E. coli*에서 우세한 트랜스알돌라제를 인코딩하나, *talA* 유전자는 소수의 트랜스알돌라제를 인코딩한다. 트랜스알돌라제의 과잉생산은 방향족 경로로의 분출을 개선하는 것으로 알려져 있다 (Lu and Liao, 1997; Sprenger, 1995; Sprenger et al., 1995b). 선행기술에서, 이것은 이의 천연 프로모터로부터 다중복제 플라스미드 상의 *tal* 유전자 (현재 *talB* 유전자로 알려짐)의 과발현에 의해 달성되었다 (Lu et al., 1997; Sprenger et al., 1995b). 그러나, 이러한 플라스미드는 불안정하고, 유지를 위해 항생제 선택을 필요로 한다. 따라서, 개선된 공정이 필요하다. *talB*의 개선된 과발현은, 염색체에서 천연 *talB* 프로모터를 강한 항시성 프로모터, 예를 들어 *바실러스 서브틸리스* 파지 SP01로부터의 *P₁₅* 또는 *P₂₆* 프로모터 (각각 서열번호 1 및 서열번호 2), 또는 박테리오파지 람다로부터의 *P_R* 프로모터 (서열번호 3)로 대체함으로써 얻어질 수 있다. 이것은, 제 1 단계에서 *cam^R*, *sacB* 카세트를 사용하여 천연 염색체 *talB* 프로모터를 대체하는 것을 제외하고는, 실시예 1에 기재된 바와 같이 2 단계로 수행된다. 제 2 단계에서, 강한 항시성 프로모터는 상기 강한 항시성 프로모터의 하류측에서 *talB* 코딩 영역의 5' 말단의 적어도 50 염기들 및 상류측에서 천연 *talB* 프로모터의 바로 상류에 적어도 50 염기쌍의 동족체가 측면에 위치해 있는 강한 항시성 프로모터를 포함하는 선형 DNA로 형질전환시키고, 수크로오스 내성을 선택함으로써 도입된다. *talA* 유전자는 유사한 방법으로 과발현될 수도 있으나, *talB* 유전자가 우세한 활성을 인코딩하므로 *talB* 유전자를 과발현시키는 것이 바람직하다 (Sprenger, 1995; Sprenger et al., 1995b). 과발현을 위해 설계된 발현 카세트의 구축에 대한 상세한 내용은 실시예 4를 참조하라.

[0114] 실시예 6

[0115] *aroZ*, *aroY* 및 *catAX* 유전자의 발현

[0116] *E. coli*에 의해 생산된 내인성 DHS의 뮌콘산으로의 전환을 입증하기 위하여, 이중 유전자인 *아시네토박터 sp. ADP1*으로부터의 *catAX*, *클렙시엘라 뉴모니아*로부터의 *aroY*, 및 *아시네토박터 sp. ADP1*으로부터의 *quiC*를 저-복제 플라스미드인 pCL1921 내의 강한 항시성 프로모터 (각각, *P₁₅*, *P_R* 및 *P₂₆*) 하에서 클로닝하여 (Lerner and Inouye, 1990), '뮌콘 플라스미드' pMG37을 생성하였다. 공백터 (pCL1921) 또는 pMG37을 가지는 MYR34 균주 유도체는 2% 글루코오스를 함유하는 진탕 플라스크 배지 (방향족 아미노산 및 비타민이 보충된 NBS 최소 배지)에서 17시간 동안 37°C에서 성장시켰다. 상청액을 모으고 HPLC로 분석하였다. DHS의 축적을 나타내는 MYR34/pCL1921과는 대조적으로, MYR34/pMG37은 뮌콘산의 생산을 나타낸다 (도 13). 상당량의 DHS, 또는 PCA 및 카테콜과 같은 중간체 산물은 후자의 균주에서 검출되지 않았으며, 이는 pMG37로부터 발현된 이중 유전자가 기능적이고 충분하다는 것을 시사한다.

[0117] 실시예 7

[0118] *aroZ* 동족체의 비교

[0119] 3개의 상이한 *aroZ* 동족체 및 유사체는 DHS를 류콘산 생산 경로로 전환시키는 능력을 비교하였다 (도 14). *아시네토박터* *sp. ADP1*으로부터의 *quiC*, *바실러스* *튜링겐시스*로부터의 *asbF*, 및 *뉴로스포라 크라썸*로부터의 *qa-4*는 *AroZ*-유사 활성을 갖는 단백질을 인코딩하는 것으로 보고되었다 (Elsemore and Ornstons, 1995; Fox et al, 1995; Rutledge, 1984). 각각의 이들 유전자는 *E. coli*에서 발현하기 위해 코돈-최적화되었고, GeneArt (Invitrogen)에 의해 합성되었으며, P₁₅ 및 P_R 프로모터 각각으로부터 *catAX* 및 *aroY* 유전자를 발현하는 저-복제 '뮤코닉 플라스미드' 내의 강한 항시성 P₂₆ 프로모터 하에서 클로닝되었다. MYR34/pCL1921, MYR34/pMG37 (*aroZ*로서 *quiC*를 갖는 뮤코닉 플라스미드), MYR34/pMG47 (*aroZ*로서 *asbF*를 갖는 뮤코닉 플라스미드), 및 MYR34/pMG70 (*aroZ*로서 *qa-4*를 갖는 뮤코닉 플라스미드)은 2% 글루코오스, 방향족 아미노산 및 방향족 비타민을 함유하는 최소 배지가 있는 진탕 플라스크에서 48시간 동안 37℃에서 성장시켰다. 상청액을 모으고 HPLC로 분석하였다. 예상한 바와 같이, 공백터로 형질전환된 MYR34는 DHS를 축적하였고 류콘산을 생산하지 않았다. 2개의 *aroZ* 동족체 및 하나의 유사체는 DHS를 류콘산 생산으로 전환시키는 기능을 하였으나, 다양한 정도로 기능을 하였다. *quiC* 유전자를 발현하는 MYR34 유도체가 가장 강력하였으며, 미량의 DHS 잔류가 있는 류콘산으로 거의 100%의 DHS 전환을 나타내었다. 균류 *aroZ* 동족체인 *qa-4*를 발현하는 MYR34 유도체는, DHS의 류콘산으로의 약 80%에 육박하는 전환 및 20% DHS 잔류를 나타내었다. 마지막으로, *asbF* 유전자를 발현하는 MYR34 유도체는 DHS의 류콘산으로의 단지 50% 전환 및 50% DHS 잔류를 나타내었다. 종합하면, 우리의 진탕 플라스크 분석 조건 하에서, *quiC* 유전자의 발현 및/또는 활성은 다른 *aroZ* 동족체의 발현 및/또는 활성에 비해 가장 높은 것으로 나타났다.

[0120] 실시예 8

[0121] *catAX*, *aroY* 및 *quiC*의 염색체 통합

[0122] 류콘산은 *adhE* 유전자좌에서 항시성 프로모터로부터 발현된 *catA-X*, *aroY* 및 *quiC*의 염색체로 통합된 단일 복제만을 함유하는 균주에 의해 생산될 수 있다.

[0123] 높은 DHS 생산자인 MYR170 ($\Delta aroE$, $\Delta ack::P_{15}-aroB$)는 염색체에서 *adhE* 유전자좌에서 류콘산 경로 유전자를 통합하는데 사용된 숙주 균주였다 (서열번호 41). 결과의 균주 MYR352를 플라스미드 YEp24 (중-복제, 공백터), pCP32AMP (중-복제, 천연 프로모터로부터 발현된 *aroG*), 또는 pCP50 (중-복제, 각각의 천연 프로모터로부터 발현된 *aroG* 및 *tktA*)로 형질전환시켜, 유도체 균주를 생성하였다. 후자의 2개의 플라스미드를 사용하여 DHS 생산을 증가시켰다. 균주를 상기 기재된 바와 같이 2% 글루코오스를 함유하는 진탕 플라스크 배지에서 72시간 동안 37℃에서 성장시켰다. 상청액을 72시간 동안 모으고 HPLC로 분석하였다. 예상한 바와 같이, *aroG* 및 *aroG/tktA*로 형질전환된 MYR352 유도체는 공백터 대조군에 비해 총 생성물 형성에서 전반적인 증가를 나타내었다 (도 15). 모든 MYR352 형질전환체는 류콘산의 측정가능한 역가를 생산하였으며, 이는 류콘산이 통합된 "뮤코닉 경로" 유전자만을 함유하고 유전자 발현의 공급된 화학 유도체가 없는 균주에 의해 생산될 수 있음을 처음으로 입증한 것이다.

[0124] 이들 MYR352 유도체 균주에서 생산된 모든 DHS가 최종 산물인 류콘산으로 전환된 것은 아니다. 대신, 상당량의 카테콜 축적이 있었으며 (도 15), 이는 *catAX*의 발현 또는 활성이 염색체 상의 단일 복제로부터 발현될 때 제한적이라는 것을 암시한다. 주요 축적 중간체가 카테콜이기 때문에, *quiC* 및 *aroY* 유전자 발현 및/또는 활성은 류콘산 합성을 위한 MYR352 배경 균주에서 충분할 것으로 보인다.

[0125] MYR352 균주 유도체를 유사한 MYR219 균주 유도체와 병행하여 비교하였다. MYR219 균주는 MYR170 균주와 동일하나, 류콘산 경로 유전자를 발현하는 저-복제 플라스미드 pMG37을 함유한다. 따라서, MYR352 및 MYR219 균주 간의 주된 차이점은 류콘산 경로 유전자의 용량과 관련이 있다 (각각 1 복제 vs 약 5 복제). MYR352 유도체 균주와는 대조적으로, MYR219 유도체 균주는 카테콜 또는 다른 중간체의 축적이 거의 없었으며, 최종 산물인 류콘산을 성공적으로 생산하였다. 종합하면, 이러한 결과는 MYR352와 같은 균주에서 *catAX* 활성을 증가시킬 필요가 있음을 나타낸다.

[0126] 실시예 9

[0127] *catAX*의 발현

[0128] MYR352 균주에서 카테콜의 축적 및 류콘산의 비효율적인 생산은 *catAX* 유전자 산물(들)의 용량 및/또는 활성의 제한에 기인한다. 상기 기재된 바와 같이, MYR352는 강한 항시성 프로모터 하에서 $\Delta aroE$, $\Delta ack::P_{15}-aroB$, 및 *catAX*, *aroY* 및 *quiC* 유전자의 염색체로 통합된 단일 복제를 함유한다. 이 균주를 중-복제 공백터 대조군

(YE_p24) 또는 *aroG/tktA* 발현 플라스미드 (pCP50)로 형질전환시켜, 방향족 아미노산 합성 경로로의 탄소 흐름을 증가시키고 다량의 DHS를 생산하였다. 상기 기재된 바와 같이, 2% 글루코오스가 보충된 진탕 플라스크 배지에서 72시간 동안 37℃에서 형질전환된 균주의 성장은 카테콜 중간체의 축적을 야기하였다. 이 결과는 *catAX* 활성이 MYR352에서 불충분할 수 있음을 시사하였다. 이 가설을 확인하기 위하여, MYR352/pCP50에서 카테콜 축적을 완화시키기 위해 저-복제 플라스미드로부터 발현된 하나 이상의 류콘산 경로 유전자의 능력을 시험하였다 (도 16). 구체적으로는, MYR352/pCP50을 저-복제 공백터 대조군 (pCL1921) 또는 류콘산 생산 경로의 3개의 유전자, 2개의 유전자, 또는 하나의 유전자 모두를 발현하는 플라스미드로 더 형질전환시켰다. 유도체 균주를 상기 기재된 바와 같이 진탕 플라스크 실험에서 분석하였다. (pMG27로부터의) *aroY* 단독 또는 (pMG39로부터의) *quiC* 단독의 용량 증가는 카테콜 축적을 완화시키지 못했지만, (pMG37로부터의) 류콘산 경로 유전자 또는 (pMG33으로부터의) *catAX*와 *aroI*를 모두 함께 발현시키면 카테콜을 류콘산으로 성공적으로 전환시켰다. 또한, (pMG31로부터의) *catAX* 단독의 발현은 류콘산의 생산 및 카테콜의 축적을 방지하기에 충분하였다.

[0129] 실시예 10

[0130] 누출성 *aroE* 돌연변이의 구축

[0131] 시스, 시스-류콘산을 생산하기 위한 선행기술 공정에서, 숙주 균주는 삭제돌연변이인, *aroE353*으로 명명된 *aroE* 유전자의 돌연변이를 함유한다. 그 결과, 균주는 쉬키메이트 경로로부터 제조된 방향족 아미노산 (페닐알라닌, 티로신, 및 트립토판) 및 방향족 비타민 (p-히드록시벤조산, p-아미노벤조산, 및 2,3-디히드록시벤조산)의 공급을 필요로 한다. 방향족 아미노산은 상업적으로 매력적인 공정에 공급하기에는 너무 비싸다. 따라서, 선행기술 공정은 상당한 개선이 필요하다. 이것은 *aroE* 유전자의 누출성 버전을 도입하여 수행될 수 있으며, 우리는 *aroE**라고 부를 것이다. 누출성 돌연변이는 삭제 표현형(null phenotype)을 야기하는 *aroE* 코딩 서열에서 하나의 아미노산을 변화시키는 미스센스(missense) 돌연변이를 먼저 생성함으로써 얻어진다. 이것은 상기 열거된 6개의 방향족 화합물의 동시 영양요구성에 대해 임의의 형태의 돌연변이유발 및 스크리닝에 의해 달성될 수 있다. 바람직한 방법은, *Taq* DNA 폴리머라제를 사용하고, 야생형 *E. coli* C 게놈 DNA를 주형으로 사용하고, *aroE* 코딩 영역의 상류에 약 1000 염기쌍 및 하류에 1000 염기쌍을 혼성화하는 PCR 올리고뉴클레오타이드 프라이머를 사용하는 오류-유발(error-prone) PCR 돌연변이유발에 의해 돌연변이체 *aroE* 유전자의 풀을 생성하는 것이다. 결과의 선형 DNA 분자의 풀을 사용하여 시스, 시스-류콘산을 생산하고, *aroE* 코딩 영역을 대체하는 통합된 *cam^R*, *sacB* 카세트를 함유하고 (관련 예에 대해 실시예 4 참조) 수크로오스 내성을 선택하는 *E. coli* C 유도체를 형질전환시킨다. 그 다음, 형질전환체는 클로람페니콜 내성을 잃어 버리고 상기 열거된 6개의 방향족 화합물을 필요로 하는 영양요구성 균주(auxotrophs)에 대해 스크리닝 한다. 몇몇 독립적인 영양요구성 균주를 선택하고, 6개의 방향족 화합물이 없는 최소 글루코오스 플레이트 상에 (최소 글루코오스 배지에서 평균) 약 10^7 , 10^8 , 또는 10^9 세포를 플레이팅하여 복귀능력(revertability)에 대해 시험한다. 플레이트 상에서 콜로니를 생기게 하는 복귀돌연변이체(Revertants)는 시스, 시스-류콘산의 생산을 위해 선택되고 시험되나, 상당한 수준의 방향족 아미노산을 생산하지 않는다. 이러한 복귀돌연변이체 중에는, AroE 효소가 성장을 위해 충분한 방향족 아미노산 및 비타민을 제공하지만 이러한 방향족 화합물의 잉여는 제공하지 않도록, *aroE* 유전자에 하나 이상의 돌연변이를 가지는 균주가 있을 것이다. 누출성 *aroE* 돌연변이체를 얻는 또 다른 방법은, *aroE353* 및 *aroE24* (둘 다 Coli Genetic Stock Center at Yale University, New Haven, CT, USA로부터 얻을 수 있음)와 같은 고전적인 복귀성(revertable) *aroE* 돌연변이체 중 하나를 시스, 시스-류콘산 생산 균주에 도입하고, 상기 기재된 복귀돌연변이체를 선택하는 것이다.

[0132] 실시예 11

[0133] 촉진 확산에 의한 글루코오스의 이입

[0134] 방향족 경로의 제 1 개입 단계에서 기질 중 하나는 포스포에놀피루베이트 (PEP)이다. PEP는 포스페이트의 공급원이고, 박테리아 포스포트랜스퍼라제 시스템 (PTS)에 의해 글루코오스 및 일부 다른 당을 이입하기 위한 에너지이기도 하다. 따라서, 박테리아가 PTS-의존성 당에서 성장하는 경우, PTS 및 PEP의 방향족 경로 사이에는 경쟁이 있다. 이와 같이, 방향족 경로로의 분출 증가에서 상당한 개선은 PTS를 제거하고 당 흡수를 위한 대안적인 경로를 제공함으로써 달성될 수 있다. 이 문제에 대한 한 가지 해결책은 PTS를 글루코오스 흡수에 상당히 잘 작동하는 양성자 공동수송체(proton symporter)인 *E. coli* GalP 투과효소(permease)로 대체하는 것이다 (미국특허 6,692,794). 그러나, 양성자 공동수송체는 여전히 투과효소를 구동시키는데 필요한 양성자 구배를 유지하기 위해 에너지를 사용한다. 이와 같이, 공정의 추가 개선을 위한 필요성이 존재한다.

- [0135] 자일로오스와 같은 일부 당은 ATP (아데노신 트리포스페이트)의 가수분해로부터 에너지를 끌어내는 수송체 단백질(transporter protein)에 의해 이입될 수 있다. 다시 한번, 만일 에너지-의존성 수송체가 적은 에너지를 필요로 하는 수송체로 대체될 수 있다면, ATP에 내재하는 에너지가 다른 유익한 용도로 보존될 수 있기 때문에 개선이 이루어질 수 있다.
- [0136] 당의 이입을 위해 에너지를 소비하지 않는 촉진 확산 수송체를 사용함으로써 상당한 개선이 이루어질 수 있다 (Parker et al., 1995; Snoep et al., 1994). 예를 들어, *glf* 유전자에 의해 인코딩된 *지모모나스 모빌리스* (*Zymomonas mobilis*)로부터의 글루코오스 촉진자는 3-데히드로쉬키메이트(DHS) 생산 균주에서 PTS 대신에 또는 이에 더하여 사용될 수 있다 (Yi et al., 2003). 그러나, 이들 균주는 글루코오스 이입을 위해 여전히 적어도 부분적으로 GalP에 의존한다. GalP는 글루코오스의 이입을 위해 양성자 구배의 형태로 에너지를 필요로 하기 때문에, 미코닉 생산 균주에 대한 글루코오스 이입의 효율을 개선시킬 필요가 있다.
- [0137] 또한, *Z. 모빌리스*로부터의 *glf* + 글루코키나제 유전자인 *glk*의 발현을 위한 카세트는 강한 항시성 프로모터, 예를 들어 P_{26} 과 함께 결합될 수 있다. 이 카세트는 원하는 화합물, 이 경우 시스, 시스-뮤콘산의 생산을 방해하지 않을 위치에서 숙주 균주의 게놈에 통합될 수 있다. *E. coli* 염색체에서 이러한 위치의 예는 트레오닌 분해 오페론인 *tdcABCDEFGF*이다. 만일, 성장 배지가 트레오닌을 함유하지 않으면, 이 오페론은 필요하지 않거나 또는 발현되지 않는다, 그래서 그 오페론에서 발현 카세트의 삽입은 대사를 방해하지 않는다.
- [0138] 상기 기재된 개선을 달성하기 위하여, 실시예 1에 개시된 방법과 유사한 방법을 이용하여 PTS 기능을 인코딩하는 유전자들 중 하나 이상을 결실시킨다. 예를 들어, *ptsH*, *ptsI*, *crr*, 또는 *ptsG* 중 하나 이상을 결실시킬 수 있다. 다음으로, 미국특허 번호 8,871,489에 기재된 공정을 이용하여 *galP*를 결실시킨다. 그 다음, P_{26} -*glf*, *glk* 카세트는 실시예 1에 기재된 것과 유사한 2 단계로 도입될 수 있다. 제 1 단계에서, cam^R , *sacB* 카세트는 pAC21로부터 유래된 선형 DNA (서열번호: 15)를 사용하고 클로람페니콜 (30 mg/1) 내성을 선택하여, *tdc* 오페론에서 통합된다. 제 2 단계에서, P_{26} -*glf*, *glk* 카세트는 pAC19로부터 유래된 선형 DNA (서열번호 15)를 사용하고 수크로오스 내성을 선택하고 클로람페니콜 민감도를 스크리닝하여, *tdc* 오페론에서 통합되고, 이 경우 최소 글루코오스 배지에서의 성장이 개선되었다.
- [0139] 글루코오스의 촉진 확산이 *E. coli*에서 종래의 글루코오스 이입 시스템을 대체할 수 있는지를 시험하기 위하여, *ptsHI* 유전자 및 *galP* 유전자를 MYR34 ($\Delta aroE$)로부터 결실시킨 다음, pAC19로부터 유래된 선형 DNA (서열번호 15)를 이용하여 P_{26} -*glf*, *glk* 카세트를 *tdc* 오페론에서 통합시켜, 균주 MYR217을 얻었다. MYR217은 필요한 3개의 방향족 아미노산 및 3개의 방향족 비타민-유사 화합물로 보충된 최소 글루코오스 배지에서 상당히 잘 성장한다 (도 17). 그러나, *ptsHI* 및 *galP*의 결실을 함유하나 *glf*, *glk* 카세트를 함유하지 않는 균주 MYR31은, 임의의 측정가능한 성장을 나타내지 않았다 (도 17). 따라서, 촉진 확산은 우리의 배경 균주에서 2개의 종래의 글루코오스 이입 시스템을 대체하기에 충분하다.
- [0140] 촉진 확산이 방향족 경로로부터 유래된 화합물을 생산하는데 유용한지를 시험하기 위하여, MYR34 및 MYR217을 pCP54 (*aroG*, *aroB*) 및 pCP55 (*aroG*, *aroB*, *tktA*)로 형질전환시켰다. 진탕 플라스크에서 방향족 중간체 3-데히드로쉬키메이트 (DHS)의 생산은 이러한 2개의 균주에 대해 비교되었다 (도 18). pCP54 또는 pCP55의 경우, 촉진 확산을 이용한 균주는 종래의 글루코오스 이입 시스템을 이용한 균주보다 훨씬 더 많은 DHS를 생산하였다. DHS의 생산은 조작된 *E. coli* 균주에서 뮤콘산 생산의 좋은 대용품이므로, 우리는 글루코오스의 촉진 확산이 뮤콘산 생산에 유용한 개선이라고 결론지을 수 있다.
- [0141] **실시예 12**
- [0142] *aroB* 유전자의 과발현
- [0143] *aroB* 유전자의 발현은 시스, 시스-뮤콘산 생산의 속도를 제한하는 것으로 보고되었다 (Niu et al., 2002). 선행 기술에서, 이것은 이른바 *aroB* 유전자의 제 2 복제와 이의 천연 프로모터를 통합시킴으로써 해결되었다. 그러나, *aroB* 유전자의 천연 프로모터 및 리보솜 결합 부위가 전혀 이상적이지 않기 때문에, 이것은 *aroB* 제한을 완화시키기에 불충분하다. 따라서, 공정은 상당한 개선이 필요하다.
- [0144] *aroB*의 개선된 과발현은, 염색체에서 천연 *aroB* 프로모터를 강한 항시성 프로모터, 예를 들어 *바실러스 서브틸리스* 파지 SP01로부터의 P_{15} 또는 P_{26} 프로모터 (각각 서열번호 1 및 서열번호 2), 또는 박테리오파지 람다로부터의 P_R 프로모터 (서열번호 3)로 대체함으로써 얻어질 수 있다. 이것은, 제 1 단계에서 cam^R , *sacB* 카세트를 사

용하여 천연 염색체 *aroB* 프로모터 및/또는 리보솜 결합 부위를 대체하는 것을 제외하고는, 실시예 4에 기재된 바와 같이 2 단계로 수행된다. 제 2 단계에서, 강한 항시성 프로모터는 강한 항시성 프로모터, 뒤이어 리보솜 결합 부위, 및 상기 강한 항시성 프로모터의 하류측에서 ATG 시작 코돈을 포함하는 *aroB* 코딩 서열의 5' 말단의 적어도 50 염기쌍 및 상류측에서 천연 *aroB* 프로모터의 바로 상류에 적어도 50 염기쌍의 동족체를 포함하는 선형 DNA로 형질전환시키고, 수크로오스 내성을 선택함으로써 도입된다. 유사한 방법을 이용하여 강한 프로모터를 도입하는 것 외에 또는 대신에, 강한 리보솜 결합 부위, 예를 들어 AGGAGG는 *aroB*의 ATG 번역 시작 코돈의 약 4 내지 10 염기쌍 상류에 도입될 수 있다. 이러한 합성 카세트, 예를 들어 P_{15} -*aroB* 카세트의 복제는 천연 *aroB* 유전자좌와는 완전히 다른 유전자좌, 예를 들어 *ack* 유전자좌에서 염색체에 통합될 수 있다. 실시예 4에서와 같이, *ack* 유전자의 동시 결실 및 *poxB* 유전자의 결실은, 발효 중 원치 않는 아세테이트의 형성을 감소시키는데 도움을 줄 수 있다.

[0145] **실시예 13**

[0146] 펜토오스 포스페이트 경로의 산화적 분지를 통한 분출 감소

[0147] 방향족 경로에서 제 1 개입 단계에 필요한 에리트로오스-4-포스페이트는 펜토오스 포스페이트 경로 (PPP)의 비-산화적 부분으로부터 유래된다. 탄소가 PPP에 들어갈 수 있는 2개의 상이한 경로가 있다. 제 1 경로는, (*zwf* 유전자에 의해 인코딩되는) 효소 글루코오스-6-포스페이트 탈수소효소, (*pgl* 유전자로부터 인코딩되는) 6-포스포글루코노락토나제 및 (*gnd* 유전자에 의해 인코딩되는) 6-포스포글루코네이트 탈수소효소에 의해 글루코오스-6-포스페이트로부터 리불로오스-5-포스페이트를 얻는다. 이들 3 단계 중 마지막에서, 하나의 탄소는 CO₂로 소실된다. PPP로의 이 경로를 PPP의 산화적 분지라고 한다. 그 다음, 리불로오스-5-포스페이트는 이소머라제, 에피머라제, 트랜스케톨라제 및 트랜스알돌라제의 작용에 의해 다양한 다른 당 포스페이트로 전환된다. 리불로오스-5-포스페이트로 시작하는 이 가역반응의 군은 PPP의 비-산화적 분지라고 한다. 탄소가 PPP에 들어갈 수 있는 제 2 경로는 프룩토오스-6-포스페이트 및 글리세르알데히드-3-포스페이트 (둘 다 엠덴-마이아호프 경로(Embden-Meyerhof pathway)로부터 유래되고, 해당작용으로도 알려짐)를 통해 이루어지며, 이들은 트랜스알돌라제 및 트랜스케톨라제에 의해 혼합되고 재배열되어 다양한 다른 당 포스페이트를 얻고, 그 중 하나는 에리트로오스-4-포스페이트이다. 만일 탄소가 이 제 2 경로를 통해 PPP에 들어가면, CO₂는 소실되지 않는다. 글루코오스로부터 시스, 시스-뮤콘산의 수율을 개선하기 위하여, PPP의 산화적 분지를 차단하여 CO₂의 소실을 방지할 수 있으며, PPP에 들어가는 모든 탄소는 프룩토오스-6-포스페이트 및 글리세르알데히드-3-포스페이트로부터 비-산화적 경로를 거쳐야 한다. PPP의 산화적 분지의 차단은, *tyrR* 유전자를 결실시키기 위해 실시예 1에 개시된 방법과 유사한 2-단계의 방법을 이용하여 *zwf* 유전자를 결실시킴으로써 달성된다.

[0148] **실시예 14**

[0149] 방향족 경로로의 PEP를 통한 분출 증가

[0150] PEP는 시스, 시스-뮤콘산에 대한 경로 상에서 중간체의 속도를 제한하지 않도록 하는 것이 바람직하다. 이는 다른 실시예들에서 상기 기재된 *pps* 유전자의 과발현 카세트를 통합함으로써 달성되는, 효소 PEP 합성효소에 의해 피루베이트의 PEP로의 재생을 증가시킴으로써 달성된다. 또 다른 방법은 *E. coli*에서 *pykA* 및 *pykF* 유전자에 의해 인코딩되는 피루베이트 키나제에 의해 PEP의 소비를 제한하는 것이다. 이 경우, 방법은 효소(들)의 활성을 감소시키는 것이다. 이는 피루베이트 키나제 (미국특허 번호 9,017,976 및 *tyrR*의 경우 실시예 1에 기재됨)를 인코딩하는 하나 이상의 유전자들을 결실시키거나, 또는 이들 유전자 중 하나 이상의 발현 강도를 감소시킴으로써, 예를 들어, 프로모터, 리보솜 결합 부위 또는 코딩 서열을 돌연변이화함으로써 달성되어, 피루베이트 키나제 활성의 수준이 감소된다. 예를 들어, *E. coli pykA* 유전자 앞의 RBS는 5'CGGAGTATTACATG이다. ATG 번역 시작 코돈은 밑줄그었다. 이 서열은 CaGAGTATTACATG, CaaAGTATTACATG, CaatGTATTACATG, CaataTATTACATG 등으로 돌연변이화될 수 있으므로, RBS 서열은 한번에 하나의 염기 변화에 의해 AGGAGG의 컨센서스 RBS (consensus RBS)를 덜 유사하게 만든다. 각각의 돌연변이화된 버전은 야생형을 대체하는 *pykA* 유전자좌에서의 염색체 내로 도입된 다음, 시스, 시스-뮤콘산 생산 수준의 개선을 위해 측정한다.

[0151] **실시예 15**

[0152] 수크로오스 상에서 성장 부여

[0153] *E. coli* C로부터 유래된 균주는 유일한 탄소원으로 수크로오스 상에서 성장하지 않는다. 그러나, 이들은 국제특허출원 공개번호 WO2012/082720 및 미국특허출원 공개번호 US2013/0337519에 개시된 바와 같이 유전자 조작될

수 있으며, 전체 참조로 본 명세서에 포함된다. 이와 같이, 시스, 시스-뮤콘산 생산 균주는 상기한 출원에 개시된 바와 같이 수크로오스 상에서 성장하도록 조작될 수 있다.

[0154] 실시예 16

[0155] 시스, 시스-뮤콘산의 개선된 생산자

[0156] 실시예 1 내지 15에 기재된 모든 특징은 차례로 특징을 도입함으로써 *E. coli*의 하나의 균주에 결합될 수 있다. 결과의 균주는 개선된 시스, 시스-뮤콘산 생산자를 포함한다. 그 다음, 결과의 균주는 제 1 복제의 위치로부터 떨어진 위치에서 한번에 하나씩 상기 기재된 각각의 과발현 카세트의 제 2 복제를 통합함으로써 더욱 개선될 수 있다. 편리하고 안전한 위치의 예는 리보솜 RNA를 인코딩하는 *rrfF*의 종결자로부터 바로 하류에 있는 *BsrB1* 제한 부위이다. 원하는 카세트를 무딘(blunt) 선형 DNA로서 플라스미드 pMH17F (서열번호 17)의 독특한 *BsrB1* 부위에 결합시킨다. 예로는 *catAX* 발현 카세트를 결합하여 *pcatAX*로 명명된 플라스미드를 얻는다. 동시에, *cam^R*, *sacB* 카세트는 무딘 단편으로서 pMH17F에 결합시켜 pMH28F (서열번호 19)를 얻는다. PCR 또는 제한 효소 절단에 의해 pMH28로부터 유래된 선형 DNA를 사용하여 *rrfF* 부위에 *cam^R*, *sacB* 카세트를 놓는다. 다음으로, PCR 또는 제한효소 절단에 의해 *pcatAX*로부터 유래된 선형 DNA를 사용하여 수크로오스 상에서의 선택을 이용하여 *rrfF* 유전자좌에서 *catAX* 카세트의 제 2 복제를 도입한다. 그 다음, 결과의 균주를 시스, 시스-뮤콘산 생산에 대한 이의 조부모 균주와 비교하여, *catAX*가 제한 단계임을 확인한다. 유사한 방법으로, 실시예 2 내지 15의 각각의 카세트를 속도 제한 단계에 대해 시험하였다. 만일 단계가 속도 제한인 것으로 확인되면, 관련 카세트의 하나 이상의 추가 복제는 염색체에서 또 다른 적당한 위치에서 통합되어, 플라스미드 또는 유도성 프로모터를 필요로 하지 않고 시스, 시스-뮤콘산 생산에서 더욱 개선을 야기한다.

[0157] 실시예 17

[0158] 발효에 의한 시스, 시스-뮤콘산의 생산

[0159] 시스, 시스-뮤콘산은 상기 실시예 1 내지 15에 개시된 유전자 조작된 미생물에 의해 생산될 수 있다. 성장 배지는 광범위하게 다양할 수 있으며, 미생물의 적당한 성장을 지지하는 임의의 배지일 수 있다. 바람직한 배지는 무기염 및 비-방향족 탄소원, 예를 들어 글루코오스, 자일로오스, 락토오스, 글리세롤, 아세테이트, 아라비노오스, 갈락토오스, 만노오스, 말토오스 또는 수크로오스를 함유하는 최소 배지이다 (바람직한 최소 성장 배지의 예는 상기 참조). 조작된 미생물 및 성장 배지의 각각의 조합에 대해, 시스, 시스-뮤콘산을 생산하기 위한 적당한 조건은, 온도, pH, 통기율, 및 화합물 또는 pH를 유지하는데 사용되는 화합물과 같은 발효 매개변수가 체계적으로 변화되는 일상적인 실험에 의해 결정된다. 시스, 시스-뮤콘산이 생산됨에 따라, pH가 너무 낮아지는 것을 방지하기 위하여 하나 이상의 화합물을 발효기에 공급해야 한다. 산을 중화하기 위한 바람직한 화합물은 암모늄, 나트륨, 칼륨, 칼슘, 마그네슘의 산화물, 수산화물, 탄산염 및 중탄산염과 같은 알칼리성 염, 또는 이러한 알칼리성 염 중 2개 이상의 조합을 포함한다.

[0160] 7 리터 발효기에서 *E. coli*의 MYR428 균주에 의한 뮤콘산 생산은 도 19에 나타내었다. $\Delta aroE \Delta ackA::P_{15}-aroB \Delta poxB::tktA$ 의 유전자형을 갖는 *E. coli*의 MYR261 균주를 플라스미드 pCP32AMP 및 pMG37로 형질전환시켜, MYR428을 생성하였다. MYR428은 48시간 동안 글루코오스를 공급하면서 상기 기재된 7리터 발효기에서 성장시켰다. 최종 뮤콘산 역가는 16 g/l 이었다 (도 19 참조).

[0161] 발효가 완료된 후, 세포를 응집, 원심분리, 및/또는 여과로 제거한 다음, 시스, 시스-뮤콘산을 하나 이상의 차후 단계, 예를 들어 침전, 결정화, 전기투석, 크로마토그래피 (이온 교환, 소수성 친화도, 및/또는 크기 기반), 미세여과, 나노여과, 역삼투, 및 증발의 조합에 의해 정화된 액으로부터 정제시킨다.

[0162] 실시예 18

[0163] 3,4-디히드록시벤조산 (PCA) 탈탄산효소 (AroY) 활성의 개선

[0164] 표 2에 제공된 유전자형을 갖는 *E. coli* 균주 MYR993을 부모 균주로 사용하여, *ubiX* 유전자 또는 *ubiD* 유전자 중 하나의 결실을 갖는 균주를 생성하였다. *ubiX*의 결실을 갖는 *E. coli* 균주를 구축할 때, *ubiX* 유전자의 각 말단에 45 bp의 상동성을 갖는 프라이머 MS608 및 MS609를 이용하여 카나마이신 내성 카세트를 증폭시켰다. *ubiD*의 결실을 갖는 *E. coli* 균주를 구축할 때, *ubiD* 유전자의 각 말단에 45 bp의 상동성을 갖는 프라이머 MS604 및 MS605를 이용하여 카나마이신 내성 카세트를 증폭시켰다. PCR 산물을 컬럼 정제 (QIAquick PCR Purification Kit, Qiagen)하고 사용하여, 이전에 개발된 방법 (Datsenko KA, Wanner BL (2000) One-step

inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. Proc Natl Acad Sci USA 97: 6640-6645)을 이용한 *E. coli* 균주 MYR993을 형질전환시켜 (표 2), *ubiX* 유전자 또는 *ubiD* 유전자 중 하나의 결실을 갖는 *E. coli* 균주를 생산하였다 (표 2 - MYR993 Δ ubiX 및 MYR993 Δ ubiD). 결실 균주는 호흡에서 손상될 것으로 예상되므로, 글루코오스를 LB 선택 플레이트에 첨가하여 발효성 성장을 제공하였다.

[0165] *E. coli* 균주 MYR993, MYR993 Δ ubiX 및 MYR993 Δ ubiD를, 탄소원으로 글루코오스를 이용하여, 5g K₂HPO₄, 3.5g KH₂HPO₄, 3.5g (NH₄)₂HPO₄, 1 mM MgSO₄, 0.1 mM CaCl₂, 미량 원소 (1.6mg FeCl₃ · 6H₂O, 0.2mg CoCl₂ · 6H₂O, 0.1mg CuCl₂ · 2H₂O, 0.2mg ZnCl₂, 0.2mg Na₂MoO₄ · 2H₂O, 0.05mg H₃BO₃, 0.55mg MnCl₂ · 4H₂O), 및 0.2 M MOPS 완충액 (모두 리터당)으로 구성된 배지에서 48시간 동안 37°C에서 250 rpm으로 250 ml 진탕 플라스크 내에서 25 ml 배양물로 성장시켰다. 48시간의 성장이 끝날 때, 배양 상청액은 뮤콘산 및 함량을 분석하였다. 도 20에 나타난 결과에서 알 수 있듯이, 부모 균주 MYR993은 PCA가 거의 없는 배양 배지에서 주로 뮤콘산을 축적하는 반면, *E. coli* 균주 MYR993 Δ ubiX는 PCA만 축적하고 뮤콘산은 검출되지 않았다. 반면에, *E. coli* 균주 MYR993 Δ ubiD는 뮤콘산 및 PCA의 축적의 감소를 나타내었다. 결론은 UbiX 단백질이 AroY (PCA 탈탄산효소) 활성화에 필요하다는 것이다.

[0166] 실시예 19

[0167] in vitro 분석에서 UbiX 동족체의 활성의 비교

[0168] in vitro 분석을 통해 UbiX 및 4개의 이의 동족체의 활성을 비교하였다. 이 in vitro 분석에서, AroY 단백질을 발현하는 *E. coli* 균주로부터의 용해물을 UbiX 단백질 또는 이의 동족체를 발현하는 다른 *E. coli* 균주로부터의 용해물과 결합시키고, 결합된 용해물을 기질로서 PCA를 소비하는 능력에 대해 분석하였다. 뮤콘산 생산 *E. coli* 균주에서, PCA는 AroY 단백질에 의해 탈카복실화되어 카테콜을 생성하고, 차례로 CatA 단백질에 의해 뮤콘산으로 전환된다. AroY 단백질의 탈카복실화 활성은 UbiX 또는 이의 동족체 중 하나의 존재에 의해 향상될 것으로 예상되며, UbiX 또는 이의 동족체의 효율에 따라, 분석 용액 내의 PCA는 상이한 속도로 소비될 것이다.

[0169] 이 in vitro 분석에서, UbiX 및 4개의 이의 동족체, 즉 *클렙시엘라 뉴모니아*의 *kpdB* 유전자 (*kpdB*)에 의해 코딩된 KpDB (서열번호 42), *E. coli* #의 *elw* 유전자에 의해 코딩된 Elw (서열번호 46), *클렙시엘라 옥시토카*의 *kox* 유전자 (*kox*)에 의해 코딩된 Kox (서열번호 48) 및 *락토바실러스 플란타룸*(*Lactobacillus plantarum*)의 *lpl* 유전자 (*lpl*)에 의해 코딩된 Lpl (서열번호 50)을 포함한다. 마지막 3개의 동족체의 명명은 본 발명에 제공된 간단히 임시 명명이다. AroY, UbiX, KpDB, Elw, Kox 및 Lpl은 저 복제 플라스미드 (SC101 복제 기원) 상의 강한 항시성 람다 파지 프로모터 P_R (서열번호 3)로부터 발현되었다. 유전자 클로닝을 위해, 플라스미드 pCAT350 (서열번호 55) 및 pCP165 (서열번호 56)를 사용하였다. AroY 플라스미드의 구축에서, 프라이머 RP712 및 RP714를 사용하여 *aroY* 유전자를 증폭시켰으며, 프라이머 MS461 및 MS346을 사용하여 pCAT350 플라스미드 백본을 증폭시켰다. 결과의 PCR 산물을 겔찰하여 AroY 단백질을 과발현하는 플라스미드를 얻었다. KpDB 플라스미드의 구축에서, 프라이머 RP731 및 RP732를 사용하여 *kpdB* 유전자를 증폭시켰으며, 프라이머 MS461 및 MS346을 사용하여 pCAT350 플라스미드 백본을 증폭시켰다. 결과의 PCR 산물을 겔찰하여 KpDB 단백질을 과발현하는 플라스미드를 얻었다. UbiX 플라스미드의 구축에서, 프라이머 MS669 및 MS666을 사용하여 *ubiX* 유전자를 증폭시켰으며, 프라이머 MS461 및 RP607을 사용하여 pCAT350 플라스미드 백본을 증폭시켰다. 결과의 PCR 산물을 겔찰하여 UbiX 단백질을 과발현하는 플라스미드를 얻었다. Elw 플라스미드의 구축에서, 프라이머 MS676 및 MS680을 사용하여 *elw* 유전자를 증폭시켰으며, 프라이머 MS461 및 MS621을 사용하여 pCP165 플라스미드 백본을 증폭시켰다. 결과의 PCR 산물을 겔찰하여 Elw 단백질을 과발현하는 플라스미드를 얻었다. Kox 플라스미드의 구축에서, 프라이머 MS686 및 MS684를 사용하여 *kox* 유전자를 증폭시켰으며, 프라이머 MS461 및 MS621을 사용하여 pCP165 플라스미드 백본을 증폭시켰다. 결과의 PCR 산물을 겔찰하여 Kox 단백질을 과발현하는 플라스미드를 얻었다. Lpl 플라스미드의 구축에서, 프라이머 MS692 및 MS691을 사용하여 *lpl* 유전자를 증폭시켰으며, 프라이머 MS461 및 MS621을 사용하여 pCP165 플라스미드 백본을 증폭시켰다. 결과의 PCR 산물을 겔찰하여 Lpl 단백질을 과발현하는 플라스미드를 얻었다. 모든 단편은 NEBuilder HiFi DNA Assembly Cloning Kit를 이용하여 클로닝할 수 있는 20 bp의 상동성을 함유하였고, NEB5 α *E. coli* 세포 (New England Biolabs)에 클로닝되었다.

[0170] 플라스미드 클로닝 후, AroY 활성을 입증하기 위하여 in vitro 효소 분석법이 개발되었다. 1 mL의 밤새도록 LB 성장시킨 배양물을 하향 회전(spun down)시키고, 200 μ L의 박테리아 단백질 추출 시약 (B-PER) (Thermo Fisher Scientific)에 재현탁시켰다. 회전 혼합기에서 5분 배양한 후, 탁상용 원심분리기에서 13,000 rpm으로 시료를 원심분리하여, 세포 잔해물을 제거하였다. 정화된 조 용해물 상청액을 새로운 튜브로 옮기고 얼음에 저장하였다. 20 μ L의 AroY 과발현 용해물을, 100 mM 인산염 나트륨 완충액 pH 6.4, 25 mM MgCl₂ 및 1 mM 프로토

카테큐산 (Sigma-Aldrich)을 함유하는 150 μ L 반응물 (최종 부피)에서, 20 μ L의 UbiX 또는 동족체 용해물과 혼합시켰다. 290 nm에서의 흡광도를 60분 동안 매분 판독하였다. AroY 활성은 A290에서 PCA의 소실을 모니터링 함으로써 측정되었다. UbiX와 이의 동족체의 상대적 활성은 도 21에 나타내었다. 시험된 모든 UbiX 동족체는 AroY 활성을 개선시켰으나, 시험된 특정 동족체에 따라 다양한 효소 활성의 변화가 나타났다. 가장 높은 AroY 활성은 KpdB를 사용하여 달성되었지만, 가장 낮은 활성은 *락토바실러스* 동족체에서 관찰되었다. 가장 높은 활성이 항상 최적일 수 없기 때문에, 나타난 다양한 활성을 사용하여 뮤코닉 경로 성능을 개선시켰다.

[0171] 실시예 20

[0172] kpdB 발현 수준이 AroY의 활성에 미치는 영향

[0173] *in vitro* 분석에서 KpdB 단백질의 발현이 AroY 단백질의 활성을 향상시킴을 입증하였고, 뮤콘산 생산 생촉매 내 KpdB 단백질의 발현 수준이 뮤콘산 생산 수준에 영향을 미치는지를 결정하기 위한 노력이 이루어졌다. 이 실험에서, 뮤콘산 생산 균주 MYR1305를 상이한 수준으로 KpdB를 발현하는 다양한 플라스미드로 형질 전환시켰다. MYR1305의 형질전환은 3개의 상이한 플라스미드로 수행되었다. 실험 대조군에서, MYR1305는 KpdB 단백질을 코딩하는 임의의 유전자가 없는 대조 플라스미드 pCP140으로 형질전환시켰다. pCP140의 DNA 서열은 서열번호 57로 제시된다. 간략하게, P₁₅ 프로모터 하에서 *E. coli* 코돈-최적화된 *catAX*, 천연 *E. coli* 프로모터 하에서 *E. coli tktA*, P₁₅ 프로모터 하에서 *E. coli aroB*, 및 P₂₆ 프로모터 (서열번호 2) 하에서 *E. coli aroD*를 발현하기 위해, pCP140을 구축하였다. MYR1305를 형질전환시키는데 사용된 제 2 플라스미드 pCP169는 P₂₆ 프로모터 하에서 kpdB 유전자를 추가로 갖는 pCP140의 유도체이다. MYR1305를 형질전환시키는데 사용된 제 3 플라스미드 pCP170은 *E. coli pgi* 프로모터 (서열번호 52) 하에서 kpdB 유전자를 추가로 갖는 pCP140의 유도체이다. 낮은 수준 발현은 P₂₆ 프로모터를 사용하여 달성되는 반면, 높은 발현은 *E. coli pgi* 프로모터를 사용하여 달성되었다. pCP169 및 pCP170은 먼저 2 세트의 프라이머를 이용하여 2개의 단편으로 pCP140 플라스미드를 증폭 시킴으로써 구축되었다 (제 1 단편에 대한 PCR 프라이머 RP607 및 RP677, 및 제 2 단편에 대한 PCR 프라이머 RP671 및 RP664). 2개의 작은 PCR 산물은 하나의 큰 PCR 산물보다 플라스미드 구축을 더 용이하게 한다. 플라스미드 pCP169의 구축에서, 프라이머 RP702 및 RP783을 이용하여 P₂₆ 프로모터를 증폭시켰으며, 프라이머 RP781 및 RP780을 이용하여 kpdB를 증폭시켰다. 플라스미드 pCP170의 구축에서, 프라이머 RP700 및 RP784를 이용하여 *pgi* 프로모터를 증폭시켰으며, 프라이머 RP779 및 RP780을 이용하여 kpdB를 증폭시켰다. 모든 PCR 산물은 NEBuilder HiFi DNA Assembly Cloning Kit를 이용하여 클로닝을 가능하게 하기 위해 20 bp 상동성 중첩을 가졌다. 도 22에 나타낸 바와 같이, kpdB를 발현하는 균주는 외인성 kpdB 유전자 발현이 없는 대조 균주보다 더 높은 수준의 뮤콘산을 생산하였다. 또한, PCA 축적은 외인성 kpdB 유전자를 발현하는 균주에서 제거되었다. 생산된 뮤콘산의 수준은 kpdB 유전자의 발현이 증가함에 따라 증가하지 않았으며, 이는 외인성 kpdB 유전자가 낮은 발현 수준으로 발현되더라도 포화 수준의 활성이 달성되었음을 시사한다.

[0174] 실시예 21

[0175] *ppc* 돌연변이체의 보완 및 뮤콘산 형성에 대한 효과

[0176] 포스포에놀피루베이트 카복실라제 (*ppc*) 유전자가 결실된 박테리아 균주를 이용하여 뮤콘산 형성을 위한 PEP의 증가된 이용률을 연구하였다. *E. coli* 균주 MYR1674를 유전자 조작하여 뮤콘산 생산을 위한 생촉매로 사용하였다. MYR1674는 탄소와 에너지 원으로서 글루코오스를 함유한 최소 배지에서 성장할 수 있으며 뮤콘산을 생산할 수 있다. 그러나, *ppc* 유전자가 MYR1674에서 결실된 경우, 결과의 균주 MYR1674 Δppc 는 글루코오스를 함유한 최소 배지에서 성장할 수 없으며 LB (Luria Broth)와 같은 농축 배지에서만 생존할 수 있다. MYR1674 Δppc 의 최소 배지에서의 성장 능력의 소실은 *E. coli*의 MYR1674 Δppc 균주에서 원래의 *ppc* 유전자좌에서 피루베이트 카복실라제를 코딩하는 *pyc* 유전자를 삽입함으로써 회복될 수 있다.

[0177] *사카로마이세스 세레비지에*로부터의 피루베이트 카복실라제 (*pyc*) 유전자 (서열번호 53)는, *E. coli ppc* 프로모터 및 종결자의 측면 상동성을 함유하는 프라이머 MS1383 및 MS1384를 이용하여 클로닝하였다. 최소 배지에서 강한 성장을 용이하게 하기 위하여, 강한 항시성 우측 램다 파지 프로모터 P_R이 요구되었다. 프라이머 MS1429 및 MS1430을 이용하여 P_R 프로모터를 증폭시켰고, 결과의 PCR 산물을 사용하여 내인성 *ppc* 프로모터를 대체하였다. $\Delta ppc::P_R-pyc$ 유전자좌의 최종 뉴클레오티드 서열은 서열번호 58에 나타내었다. *S. 세레비지에 pyc* 유전자는 임의의 *E. coli* 유전자와 밀접한 관련이 없기 때문에 선택되었으며, 상이한 코돈 사용으로 인한 예상된 낮은 발현은 뮤콘산 생산을 위한 PEP 보존에 유익할 수 있다. 피루베이트 카복실라제를 함유하는 많은 유기체가

있으며, 피루베이트 카복실라제 활성을 갖는 임의의 동족체 또는 유사체가 사용될 수 있다. 원래 *ppc* 유전자좌에서 $\Delta ppc::P_R-*pyc*$ 를 통합하여 MYR1674로부터 유래된 새로운 균주 MYR1772는 클로닝된 *S. 세레비지에* *pyc* 유전자의 기능성을 확인하는 최소 배지에서 생존할 수 있었다. MYR1772 및 이의 부모 MYR1674 균주는 진탕 플라스크 실험에서 이들의 류콘산 생산능력을 비교하였다. 도 23에 나타난 결과에 알 수 있듯이, MYR1772 균주는 부모 MYR1674 보다 류콘산 생산의 높은 역가를 생산하였으며, 이는 내인성 포스포에놀피루베이트 카복실라제 효소를 외인성 피루베이트 카복실라제 효소로 대체하는 이점이 있음을 입증하였다.

[0178] 실시예 22

[0179] 류콘산 생산의 측정

[0180] 박테리아 균주 MYR814, MYR993, MYR1536, MYR1557, MYR1570, MYR1595, MYR1630, MYR1674 및 MYR1772를 진탕 플라스크 배양물에서 밤새도록 성장시켰으며, 류콘산 생산에 대한 역가 및 수율을 결정하였다. 또한, 성장물은 600 nm에서 흡광도를 측정하여 결정하였으며, 다양한 박테리아 균주의 상대적 성장을 표 8에 나타내었다. "++"로 표시되는 박테리아 성장은 야생형 *E. coli* 균주에서 보이는 성장과 유사한 성장을 나타낸다. "+"로 표시되는 박테리아의 성장은 약한 성장을 나타낸다. 중간 성장은 "++"로 표시된다. 특정 균주가 약한 성장을 나타낼 때, 그 균주는 성장을 개선하기 위해 5일 동안 밤새도록 전이를 거치게 되며, 각 전이는 약 10 세대 또는 두 배가 된다.

[0181] 박테리아 균주 MYR814, MYR1570, MYR1630 및 MYR1674를 유가식 방식으로 7 리터 발효기에서 성장시켰으며, 류콘산 생산에 대한 역가 및 수율을 결정하였다 (표 9). 류콘산의 역가 및 수율에 대해 시험한 박테리아 균주는 부산물을 거의 생산하지 않았다. 예를 들어, 유가식 발효의 72시간 후의 박테리아 균주는 부산물로서 0.08g/L의 PCA 및 0.07g/L의 푸마레이트만을 나타내었다.

표 1

표 1. 본 발명에서 사용된 박테리아 균주	
박테리아 균주	유전자형 / 설명
ATCC8739	<i>Escherichia coli</i> "C" 야생형
MYR34	ATCC8739 $\Delta aroE$
MYR170	ATCC8739 $\Delta aroE$, $\Delta ackA::P_{15}aroB$
MYR261	ATCC8739 $\Delta aroE$, $\Delta ackA::P_{15}aroB$, $\Delta poxB::tklA$
MYR305	ATCC8739 $\Delta aroE$, $\Delta ackA::P_{15}aroB$, $\Delta poxB$
MYR31	ATCC8739 $\Delta ptsHI$, $\Delta galP$
MYR217	ATCC8739 $\Delta ptsHI$, $\Delta galP$, $\Delta tdc::glf-glk$, $\Delta aroE$
MYR352	ATCC8739 $\Delta aroE$, $\Delta ackA::P_{15}aroB$, $\Delta adhE::P_{15-catAX}$, P_R-aroY , $P_{26-quiC}$
RY903	$\Delta aroE$, $\Delta ackA::P_{15}aroB$, pMG37, <i>aroG</i> *20-893
RY909	$\Delta aroE$, $\Delta ackA::P_{15}aroB$, pMG37, <i>aroG</i> *20-899
RY911	$\Delta aroE$, $\Delta ackA::P_{15}aroB$, pMG37, <i>aroG</i> *20-901
RY912	$\Delta aroE$, $\Delta ackA::P_{15}aroB$, pMG37, <i>aroGwt</i>
RY913	$\Delta aroE$, $\Delta ackA::P_{15}aroB$, pMG37, <i>aroG</i> *20-893, $\Delta tyrR::kan$
RY919	$\Delta aroE$, $\Delta ackA::P_{15}aroB$, pMG37, <i>aroG</i> *20-899, $\Delta tyrR::kan$
RY921	$\Delta aroE$, $\Delta ackA::P_{15}aroB$, pMG37, <i>aroG</i> *20-901, $\Delta tyrR::kan$
RY922	$\Delta aroE$, $\Delta ackA::P_{15}aroB$, pMG37, <i>aroGwt</i> , $\Delta tyrR::kan$

[0182]

표 2

표 2. 본 발명에서 사용된 박테리아 균주		
균주명	유전자형/설명	부모
MYR802	Δ ackA::P15-aroB Δ adhE::P15-catAX + Pr-aroY + P26-quiC Δ pflB::P15-catAX Δ tyrR aroG ^{FBR} Δ poxB::tkfA Δ ptsHI Δ tdc::P26-glf-glk Δ galP aroE G105M	
MYR814	Δ ackA::P15-aroB Δ adhE::P15-catAX + Pr-aroY + P26-quiC Δ pflB::P15-catAX Δ tyrR aroG ^{FBR} Δ poxB::tkfA Δ ptsHI Δ tdc::P26-glf-glk Δ galP aroE G105M [P15-catA-catX + tkfA + P15-aroB + P26-aroD in pCL1921 backbone]	MYR802
MYR993	Δ ackA::P15-aroB Δ adhE::Pr-aroY + P26-quiC Δ pflB::Pr-catAX Δ tyrR aroG ^{FBR} Δ poxB::tkfA Δ ptsHI Δ tdc::P26-glf-glk Δ galP aroE G105M	MYR802
MYR993 Δ ubiX	Δ ackA::P15-aroB Δ adhE::Pr-aroY + P26-quiC Δ pflB::Pr-catAX Δ tyrR aroG ^{FBR} Δ poxB::tkfA Δ ptsHI Δ tdc::P26-glf-glk Δ galP aroE G105M Δ ubiX::Kan ^R	MYR993
MYR993 Δ ubiD	Δ ackA::P15-aroB Δ adhE::Pr-aroY + P26-quiC Δ pflB::Pr-catAX Δ tyrR aroG ^{FBR} Δ poxB::tkfA Δ ptsHI Δ tdc::P26-glf-glk Δ galP aroE G105M Δ ubiD::Kan ^R	MYR993
MYR1305	Δ ackA::P15-aroB Δ adhE::P15-catAX + Pr-aroY + P26-quiC Δ pflB::P15-catAX Δ tyrR aroG ^{FBR} Δ poxB::tkfA Δ ptsHI Δ tdc::P26-glf-glk Δ galP Δ aroE Δ zwf	MYR802
MYR1305 pCP140	Δ ackA::P15-aroB Δ adhE::P15-catAX + Pr-aroY + P26-quiC Δ pflB::P15-catAX Δ tyrR aroG ^{FBR} Δ poxB::tkfA Δ ptsHI Δ tdc::P26-glf-glk Δ galP Δ aroE Δ zwf [Pr-catA-catX + tkfA + P15-aroB + P26-aroD in pCL1921 backbone]	MYR1305
MYR1305 pCP169	Δ ackA::P15-aroB Δ adhE::P15-catAX + Pr-aroY + P26-quiC Δ pflB::P15-catAX Δ tyrR aroG ^{FBR} Δ poxB::tkfA Δ ptsHI Δ tdc::P26-glf-glk Δ galP Δ aroE Δ zwf [Pr-catA-catX + tkfA + P15-aroB + P26-aroD + P26-kpdB in pCL1921 backbone]	MYR1305
MYR1305 pCP170	Δ ackA::P15-aroB Δ adhE::P15-catAX + Pr-aroY + P26-quiC Δ pflB::P15-catAX Δ tyrR aroG ^{FBR} Δ poxB::tkfA Δ ptsHI Δ tdc::P26-glf-glk Δ galP Δ aroE Δ zwf [Pr-catA-catX + tkfA + P15-aroB + P26-aroD + Pr-cat-kpdB in pCL1921 backbone]	MYR1305
MYR1536	Δ ackA::P _{acpp} -aroB Δ adhE::P15-qa4 Δ tyrR aroG ^{FBR} Δ poxB::Pr-tkfA Δ tdc::P26-glf-glk Δ galP::P15-ubiX Δ pflB::Pr-CatAX P _{acpp} -aroD Δ mgsA::P _{piu} -qa4 Δ ptsHI::Pr-aroY Δ 0039::Pr-catAX	MYR802
MYR1557	Δ ackA::P _{acpp} -aroB Δ adhE::P15-qa4 Δ tyrR aroG ^{FBR} Δ poxB::Pr-tkfA Δ tdc::P26-glf-glk Δ galP::P15-ubiX Δ pflB::Pr-CatAX P _{acpp} -aroD Δ mgsA::P _{piu} -qa4 Δ ptsHI::Pr-aroY Δ 0039::Pr-catAX (Evolved version of MYR1536 for faster growth)	MYR1536
MYR1570	Δ ackA::P _{acpp} -aroB Δ adhE::P15-qa4 Δ tyrR aroG ^{FBR} Δ poxB::Pr-tkfA Δ tdc::P26-glf-glk Δ galP::P15-ubiX Δ pflB::Pr-CatAX P _{acpp} -aroD Δ mgsA::P _{piu} -qa4 Δ ptsHI::Pr-aroY Δ 0039::Pr-catAX Δ zwf	MYR1536
MYR1595	Δ ackA::P _{acpp} -aroB Δ adhE::P15-qa4 Δ tyrR aroG ^{FBR} Δ poxB::Pr-tkfA Δ tdc::P26-glf-glk Δ galP::P15-ubiX Δ pflB::Pr-CatAX P _{acpp} -aroD Δ mgsA::P _{piu} -qa4 Δ ptsHI::Pr-AroY Δ 0039::Pr-catAX Δ zwf Δ pykF	MYR1557
MYR1630	Δ ackA::P _{acpp} -aroB Δ adhE::P15-qa4 Δ tyrR aroG ^{FBR} Δ poxB::Pr-tkfA Δ tdc::P26-glf-glk Δ galP::P15-ubiX Δ pflB::Pr-CatAX P _{acpp} -aroD Δ mgsA::P _{piu} -qa4 Δ ptsHI::Pr-aroY Δ 0039::Pr-catAX Δ zwf Δ pykF Δ 2160::Pr-aroG ^{FBR}	MYR1595
MYR1674	Δ ackA::P _{acpp} -aroB Δ adhE::P15-qa4 Δ tyrR aroG ^{FBR} Δ poxB::Pr-tkfA Δ tdc::P26-glf-glk Δ galP::P15-ubiX Δ pflB::Pr-CatAX P _{acpp} -aroD Δ mgsA::P _{piu} -qa4 Δ ptsHI::Pr-aroY Δ 0039::Pr-catAX Δ zwf Δ pykF Δ 2160::Pr-aroG ^{FBR} (Evolved version of MYR1630 for faster growth)	MYR1630
MYR1772	Δ ackA::P _{acpp} -aroB Δ adhE::P15-qa4 Δ tyrR aroG ^{FBR} Δ poxB::Pr-tkfA Δ tdc::P26-glf-glk Δ galP::P15-ubiX Δ pflB::Pr-CatAX P _{acpp} -aroD Δ mgsA::P _{piu} -qa4 Δ ptsHI::Pr-aroY Δ 0039::Pr-catAX Δ zwf Δ pykF Δ 2160::Pr-aroG ^{FBR} Δ ppe::Pr-pyc	MYR1674

[0183]

표 3

표 3. 본 발명에서 사용된 플라스미드	
박테리아 플라스미드	유전자형 / 설명
YEp24	2 μ 효모 기원, <i>URA3</i> , <i>Tc^R</i> , pMB1 레플리콘, <i>Ap^R</i>
pCP32AMP	2 μ 효모 기원, <i>URA3</i> , <i>Tc^R</i> , pMB1 레플리콘, <i>Ap^R</i> , <i>aroG</i>
pCP14	2 μ 효모 기원, <i>URA3</i> , <i>Tc^R</i> , pMB1 레플리콘, <i>Ap^R</i> , <i>P₁₅aroB</i>
pCP54	2 μ 효모 기원, <i>URA3</i> , <i>Tc^R</i> , pMB1 레플리콘, <i>Ap^R</i> , <i>P₁₅aroB</i> , <i>aroG</i>
pCP50	2 μ 효모 기원, <i>URA3</i> , <i>Tc^R</i> , pMB1 레플리콘, <i>Ap^R</i> , <i>aroG</i> , <i>tklA</i>
pCP55	2 μ 효모 기원, <i>URA3</i> , <i>Tc^R</i> , pMB1 레플리콘, <i>Ap^R</i> , <i>aroG</i> , <i>aroB</i> , <i>tklA</i>
pCL1921	pSC101 레플리콘, <i>Spc^R</i>
pMG27	pSC101 레플리콘, <i>Spc^R</i> , <i>P_R-aroY</i>
pMG31	pSC101 레플리콘, <i>Spc^R</i> , <i>P₁₅-catAX</i>
pMG33	pSC101 레플리콘, <i>Spc^R</i> , <i>P₁₅-catAX</i> , <i>P_R-aroY</i>
pMG37	pSC101 레플리콘, <i>Spc^R</i> , <i>P₁₅-catA-CatX</i> , <i>P_R-aroY</i> , <i>P₂₆-quiC</i>
pMG39	pSC101 레플리콘, <i>Spc^R</i> , <i>P₂₆-quiC</i>
pMG47	pSC101 레플리콘, <i>Spc^R</i> , <i>P₁₅-catAX</i> , <i>P_R-aroY</i> , <i>P₂₆-asbF</i>
pMG70	pSC101 레플리콘, <i>Spc^R</i> , <i>P₁₅-catAX</i> , <i>P_R-aroY</i> , <i>P₂₆-qa-4</i>

[0184]

표 4

표 4. 페닐알라닌 피드백 저해에 대해 내성을 야기하는 <i>aroG</i> * 돌연변이체 대립유전자			
균주	대립유전자 번호	뉴클레오타이드 돌연변이	아미노산 돌연변이
RY893	<i>aroG</i> *20-893	C449T	Pro150Leu
RY897	<i>aroG</i> *20-897	C449T	Pro150Leu
RY899	<i>aroG</i> *20-899	T538C	Ser180Pro
RY901	<i>aroG</i> *20-901	C438T	Pro150Ser
MYR450	<i>aroG</i> *111	C55T	Pro19Ser
MYR451	<i>aroG</i> *211	G533A	Gly178Glu
MYR452	<i>aroG</i> *212	C540T	Ser180Phe
MYR453	<i>aroG</i> *311	36 내지 44bp 염기쌍의 결실	Glu-Ile-Lys 의 결실
MYR454	<i>aroG</i> *312	C632T	Ser211Phe
MYR455	<i>aroG</i> *411	T29A	Ile10Asn
MYR456	<i>aroG</i> *412	G533A	Gly178Glu
MYR457	<i>aroG</i> *511	C448T	Pro150Ser

[0185]

표 5

표 5. 다양한 재조합 <i>E. coli</i> 균주로부터의 조 추출물의 AroG 활성 측정			
균주	<i>aroG</i> * 대립유전자	특이적 활성 mU (1 mU = 분당 밀리그램 단백질당 1 nM 생성물)	페닐알라닌에 대한 내성 활성의 %
RY893	<i>aroG</i> *20-893	62	34
RY897	<i>aroG</i> *20-897	55	77
RY899	<i>aroG</i> *20-899	92	113
RY901	<i>aroG</i> *20-901	78	76
RY902	<i>aroG</i> wild type	38	6
RY890	<i>aroG</i> wild type	54	7

[0186]

표 6

표 6. 피드백 내성 <i>aroG</i> * 대립유전자를 함유하는 균주에 의한 진탕 플라스크에서 류콘산 생산	
균주	류콘산 역가 g/l
RY913	3.04
RY919	3.11
RY921	2.99
RY922	1.45

[0187]

표 7

표 7. 서열 정보		
번호	명명	설명
1	서열번호 1	전사 시작 부위로부터 바로 하류에 추가된 줄기 및 고리(stem and loop)를 갖는, <i>바실러스 서브틸리스</i> 파지 SP01 로부터의 P ₁₅ 프로모터
2	서열번호 2	<i>바실러스 서브틸리스</i> 파지 SP01 로부터의 P ₂₆ 프로모터
3	서열번호 3	<i>Escherichia coli</i> 파지로부터의 P _R 프로모터
4	서열번호 4	<i>qa-4</i> 유전자에 의해 인코딩된 <i>뉴로스포라 크라썸</i> 로부터의 3-데히드로쉬키메이트 탈수효소의 단백질 서열
5	서열번호 5	<i>뉴로스포라 크라썸</i> 로부터의 <i>qa-4</i> 유전자의 제논 DNA 서열 + 주변 서열
6	서열번호 6	<i>gutC</i> 유전자에 의해 인코딩된 <i>아스페르길루스 니들란스</i> 로부터의 3-데히드로쉬키메이트 탈수효소의 단백질 서열
7	서열번호 7	<i>아스페르길루스 니들란스</i> 로부터의 <i>gutC</i> 유전자의 제논 DNA 서열 + 주변 서열
8	서열번호 8	<i>클렙시엘라 뉴모니아</i> ATCC25597 로부터의 프로터카테큐에이트 탈탄산효소(AroY)의 단백질 서열
9	서열번호 9	<i>클렙시엘라 뉴모니아</i> 342의 <i>aroY</i> 유전자의 DNA 서열 + 2 킬로염기의 주변 DNA 서열
10	서열번호 10	410 염기의 상류 서열 및 2 개의 개방형 판독 프레임 하류를 포함하는, <i>아시네토박터 바일리</i> ADP1 로부터의 <i>catA</i> 유전자의 DNA 서열
11	서열번호 11	<i>아시네토박터 바일리</i> ADP1 로부터의 CatA (카테콜 1,2-디옥시게나제)의 단백질 서열
12	서열번호 12	<i>아시네토박터 sp. ADP1</i> 로부터의 <i>quiC</i> (3-데히드로쉬키메이트 탈수효소) 유전자의 DNA 서열
13	서열번호 13	<i>아시네토박터 sp. ADP1</i> 로부터의 <i>quiC</i> (3-데히드로쉬키메이트 탈수효소) 유전자의 코돈-최적화 DNA 서열
14	서열번호 14	<i>아시네토박터 sp. ADP1</i> 로부터의 QuiC (3-데히드로쉬키메이트 탈수소효소)의 단백질 서열
15	서열번호 15	플라스미드 pAC21 의 DNA 서열
16	서열번호 16	플라스미드 pAC19 의 DNA 서열
17	서열번호 17	플라스미드 pMH17F 의 DNA 서열
18	서열번호 18	야생형 <i>aroG</i> 유전자의 코딩 영역의 DNA 서열
19	서열번호 19	플라스미드 pMH28F 의 DNA 서열

[0188]

표 7. 서열 정보 - 계속		
번호	명명	설명
20	서열번호. 20	플라스미드 pCL1921 의 DNA 서열
21	서열번호. 21	플라스미드 pMG27 의 DNA 서열
22	서열번호. 22	플라스미드 pMG31 의 DNA 서열
23	서열번호. 23	플라스미드 pMG33 의 DNA 서열
24	서열번호. 24	플라스미드 pMG37 의 DNA 서열
25	서열번호. 25	플라스미드 pMG39 의 DNA 서열
26	서열번호. 26	플라스미드 pMG47 의 DNA 서열
27	서열번호. 27	플라스미드 pMG70 의 DNA 서열
28	서열번호. 28	플라스미드 pCP32AMP 의 DNA 서열
29	서열번호. 29	플라스미드 pCP14 의 DNA 서열
30	서열번호. 30	플라스미드 pCP50 의 DNA 서열
31	서열번호. 31	플라스미드 pCP54 의 DNA 서열
32	서열번호. 32	플라스미드 pCP55 의 DNA 서열
33	서열번호. 33	플라스미드 YEP24 의 DNA 서열
34	서열번호. 34	결실된 <i>aroE</i> 영역의 DNA 서열
35	서열번호. 35	통합된 카세트 $\Delta ack::P_{13}aroB$ 의 DNA 서열
36	서열번호. 36	$\Delta poxB$ 영역의 DNA 서열
37	서열번호. 37	통합된 카세트 $\Delta poxB::tklA$ 의 DNA 서열
38	서열번호. 38	$\Delta ptsHI$ 영역의 DNA 서열
39	서열번호. 39	통합된 카세트 $\Delta tdc::glf-glk$ 의 DNA 서열
40	서열번호. 40	$\Delta galP$ 영역의 DNA 서열
41	서열번호. 41	MYR352 $\Delta adhE::P_{15-catAX}$, P_K-aroY , $P_{26-quiC}$
42	서열번호. 42	클렙시엘라 뉴모니아가 <i>kpdB</i> 유전자의 뉴클레오티드 서열
43	서열번호. 43	클렙시엘라 뉴모니아가 <i>KpdB</i> 단백질의 아미노산 서열

[0189]

표 7. 서열 정보 - 계속

번호	명명	설명
44	서열번호. 44	<i>Escherichia coli</i> ubiX 유전자의 뉴클레오티드 서열
45	서열번호. 45	<i>Escherichia coli</i> 의 UbiX 단백질의 아미노산 서열
46	서열번호. 46	<i>Escherichia coli</i> W 균주 elw 유전자의 뉴클레오티드 서열
47	서열번호. 47	<i>Escherichia coli</i> W 균주의 Elw 단백질의 아미노산 서열
48	서열번호. 48	클렙시엘라 옥시토카 kox 유전자의 뉴클레오티드 서열
49	서열번호. 49	클렙시엘라 옥시토카의 Kox 단백질의 아미노산 서열
50	서열번호. 50	락토바실러스 플란타룸 ipl 유전자의 뉴클레오티드 서열
51	서열번호. 51	락토바실러스 플란타룸의 Lpl 단백질의 아미노산 서열
52	서열번호. 52	Pgi 프로모터의 뉴클레오티드 서열
53	서열번호. 53	사카로마이세스 세레비자에 pyc 유전자의 뉴클레오티드 서열
54	서열번호. 54	사카로마이세스 세레비자에의 Pyc 단백질의 아미노산 서열
55	서열번호. 55	플라스미드 pCAT350 의 DNA 서열
56	서열번호. 56	플라스미드 pCP165 의 DNA 서열
57	서열번호. 57	플라스미드 pCP140 의 DNA 서열
58	서열번호. 58	$\Delta ppc::Pa-pyc$ 의 뉴클레오티드 서열
59	서열번호. 59	acpP 프로모터의 뉴클레오티드 서열
60	서열번호. 60	η IU 프로모터의 뉴클레오티드 서열

[0190]

서열 번호	프라이머 명	서열
61	MS604	AACGCCGTATAATGGGCGCAGATTAAGAGGCTACAGTGGGCTTACATGGCGA TAGCTAGA
62	MS605	TGTCGGATCGATAAATAGGGCAAAACAAACGCGCATCCCGGAAAACGATTCC GAAGCCCA
63	MS608	AAAGTCTGCCTGCAAGTCTGACAGGGCAACTATTTGTGGGCTTACATGGCGAT AGCTAGA
64	MS609	TTGCAAAATTGCCCTGAAACAGGGCAACAGCGGAGTCCCGGAAAACGATTCC GAAGCCCA
65	MS461	GGCTATATTCCTTATCTAGATTAGT
66	MS346	GTCTGACAGGTGCCGGATTTTCATAT
67	RP712	TCTAGATAAAGGAATATAGCCATGACCGCACCGATTTCAGGATCTGC
68	RP714	AAATCCGGCACCTGTCAGACTTATTTTGCGCTACCCTGGTTTTT
69	RP731	CATGTACTAATCTAGATAAGGAATATAGCCATGAAACTGATTATTGGGATGAC GGGGGCC
70	RP732	GCCGGATATGAAATCCGGCACCTGTCAGACTTATTCGATCTCCTGTGCAAATT GTTCTGC
71	MS669	TCTAGATAAAGGAATATAGCCATGAAACGACTCATTGTAGGCATCA
72	MS666	ACCGAACAGGCTTATGTCCAGATAGCAGGTATAGCGGTTGAATCG
73	RP607	TGGACATAAGCCTGTTCCGGTTTCGT
74	MS621	TTAGATTTGACTGAAATCGTACAGT
75	MS676	TCTAGATAAAGGAATATAGCCATGAAACTGATCGTCGGGATGACAG
76	MS680	ACGATTTTCAGTCAAATCTAATTATTCATTCTCCTGAGAAAAATTC
77	MS686	TCTAGATAAAGGAATATAGCCATGACGGCACGCATCATCATTGGTA
78	MS684	ACGATTTTCAGTCAAATCTAATTAATTAACGTAAGCTCGCCTTCA
79	MS692	TCTAGATAAAGGAATATAGCCATGAAACGAATIGTTGTGGGAATCA
80	MS691	ACGATTTTCAGTCAAATCTAATTAATCCCCCTCCCAACGGCGATCA
81	RP677	CGACGTTGTAAAACGACGGCCAGTG
82	RP671	TTAATCGCCTTGCAGCACATCCCCC
83	RP664	ACGAACCGAACAGGCTTATGTCCA
84	RP702	GCCGTCGTTTTACAACGTCGGATCCGCCTACCTAGCTTCCAAGAA
85	RP783	CCTACAATGAGTCGTTTCATTAGGTTTTCTCAACCCGGGAGCGT
86	RP781	CCCGGGTTGAGGAAAACCTAATGAAACGACTCATTGTAGGCATCA
87	RP780	ATGTGCTGCAAGGCGATTAAAGATAGCAGGTATAGCGGTTGAATCG
88	RP700	GCCGTCGTTTTACAACGTCGAGCGGGGCGGTTGTCAACGATGGGG
89	RP784	CCTACAATGAGTCGTTTCATGGCTATATTCCTCCTCTGCATGAGA
90	RP779	TGCAGAGGAGGAATATAGCCATGAAACGACTCATTGTAGGCATCA
91	MS1383	GATGGGGTGTCTGGGGTAATATGTCGCAAGAAAAATTCGCCGGCT
92	MS1384	GGGTTTGCAGAAGAGGAAGATCATGCCTTAGTTTCAACAGGAACT
93	MS1429	ATTCTGCTATTATTCGTTTCGTTAAATCTATACCAGCAAGGGAT
94	MS1430	GCGAATTTTCTTTGCGACATGGCTATATTCCTTATCTAGATTAGT

[0191]

표 8

표 8. 진탕 플라스크 배양물에서 다양한 <i>E. coli</i> 균주의 성장, 및 류콘산 생산에 대한 역가 및 수율			
균주명	역가 (류콘산 그램 /리터)	수율 (류콘산 그램 /글루코오스 그램)	성장
MYR814	2	0.16	+++
MYR993	1.9	0.15	+++
MYR1536	2.7	0.22	+
MYR1557	3	0.25	+++
MYR1570	4.5	0.38	++
MYR1595	4.5	0.38	++
MYR1630	4.8	0.4	+
MYR1674	3.8	0.32	+++
MYR1772	5	0.42	+++

[0192]

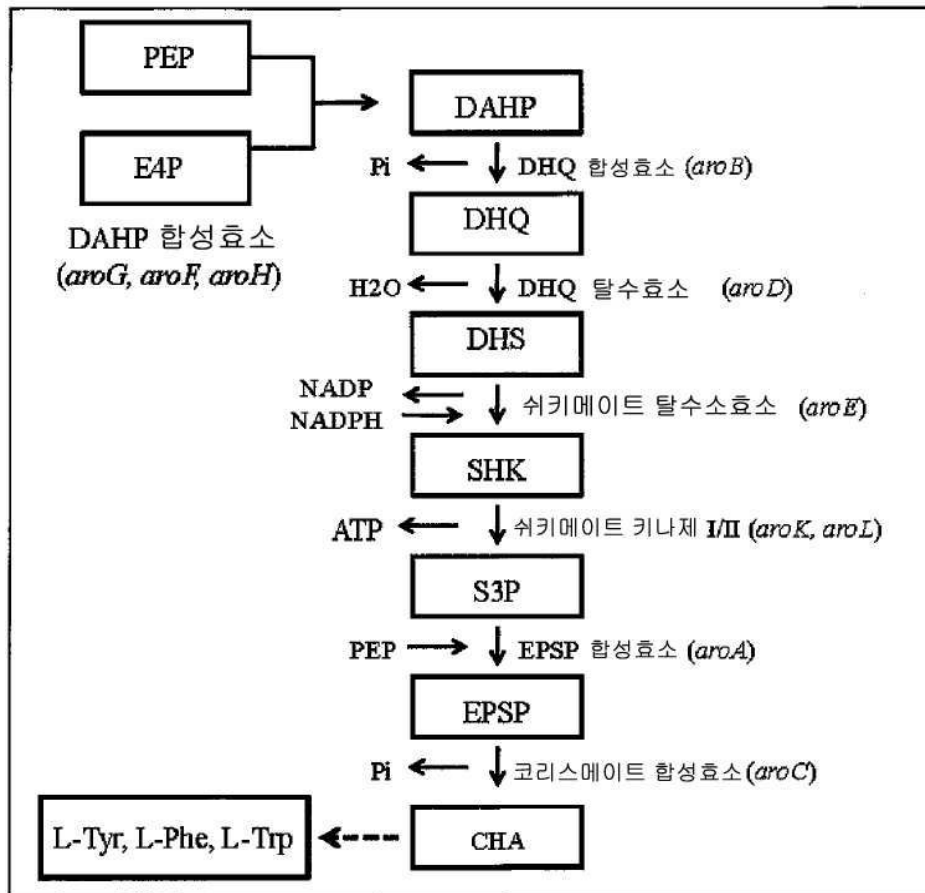
표 9

표 9. 유가식 배양으로 다양한 <i>E. coli</i> 균주의 류콘산 생산에 대한 역가 및 수율			
균주명	역가 (류콘산 그램 /리터)	수율 (류콘산 그램 /글루코오스 그램)	시간 (hours)
MYR814	30.9	0.20	48
MYR1570	49.0	0.36	48
MYR1630	58.3	0.47	48
MYR1630	69.5	0.42	72
MYR1674	81.5	0.43	72

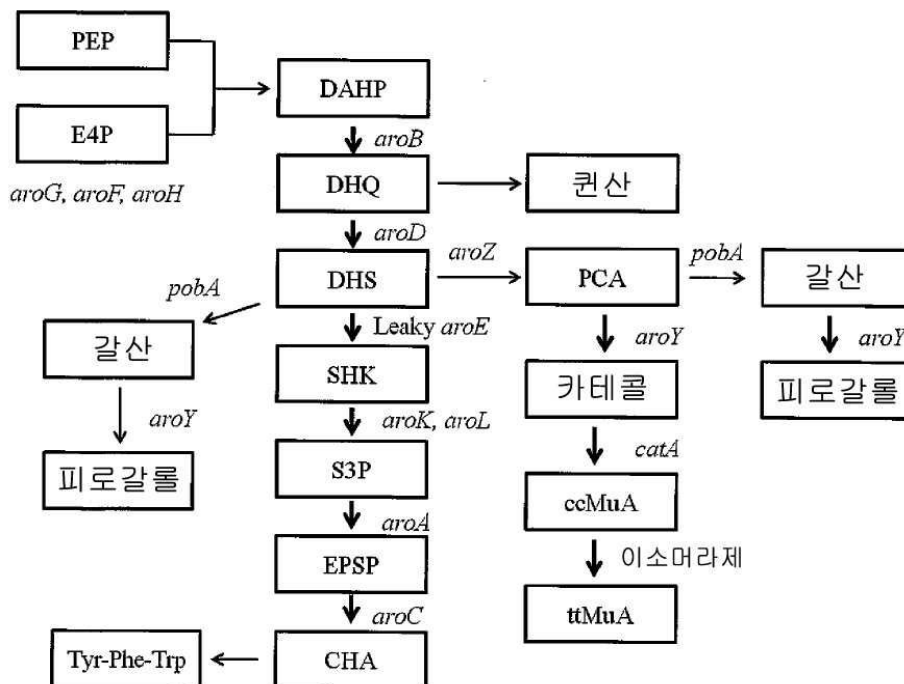
[0193]

도면

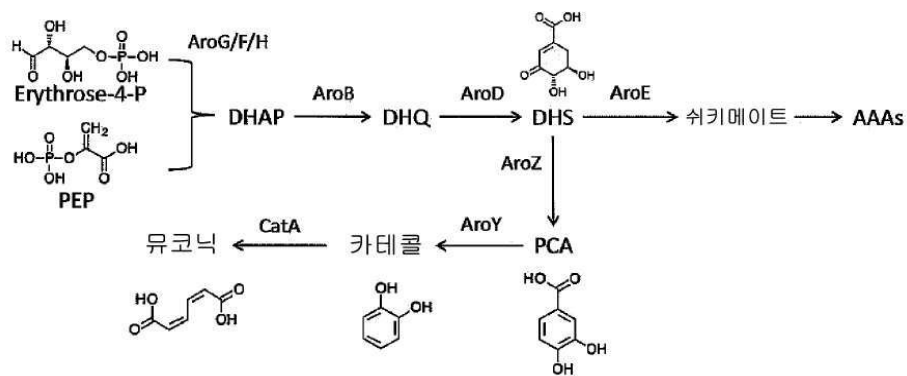
도면1



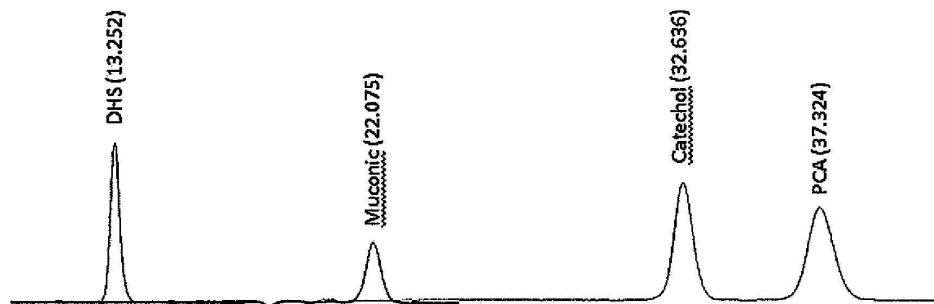
도면2



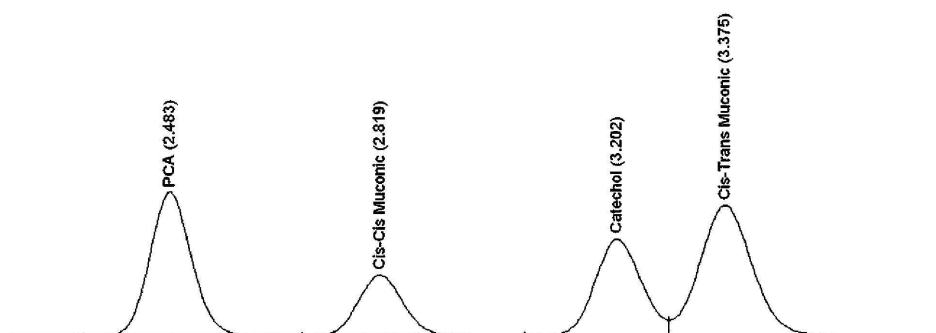
도면3



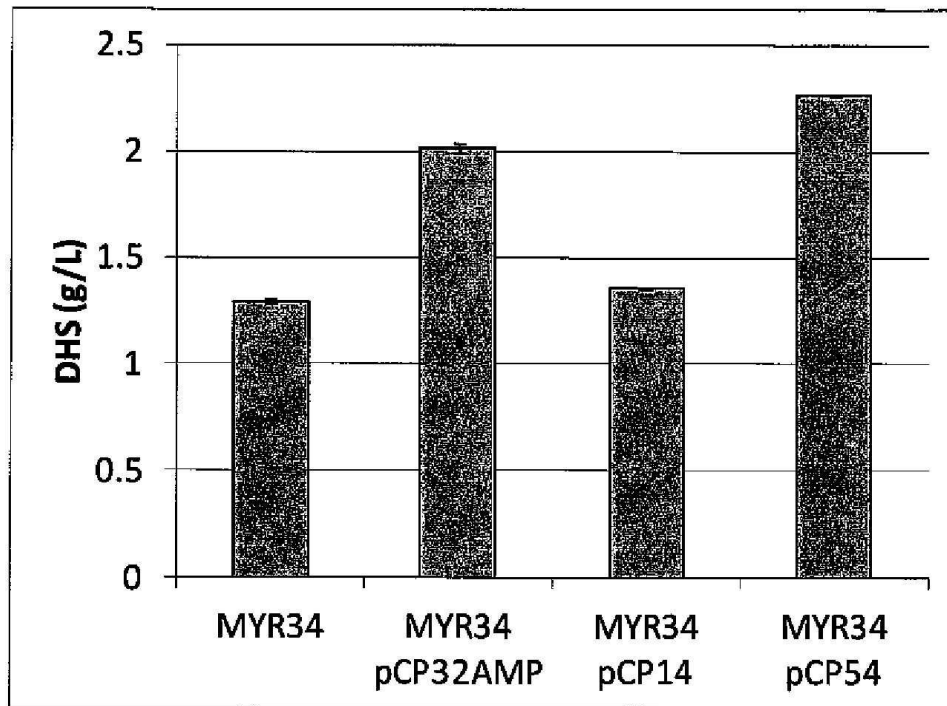
도면4



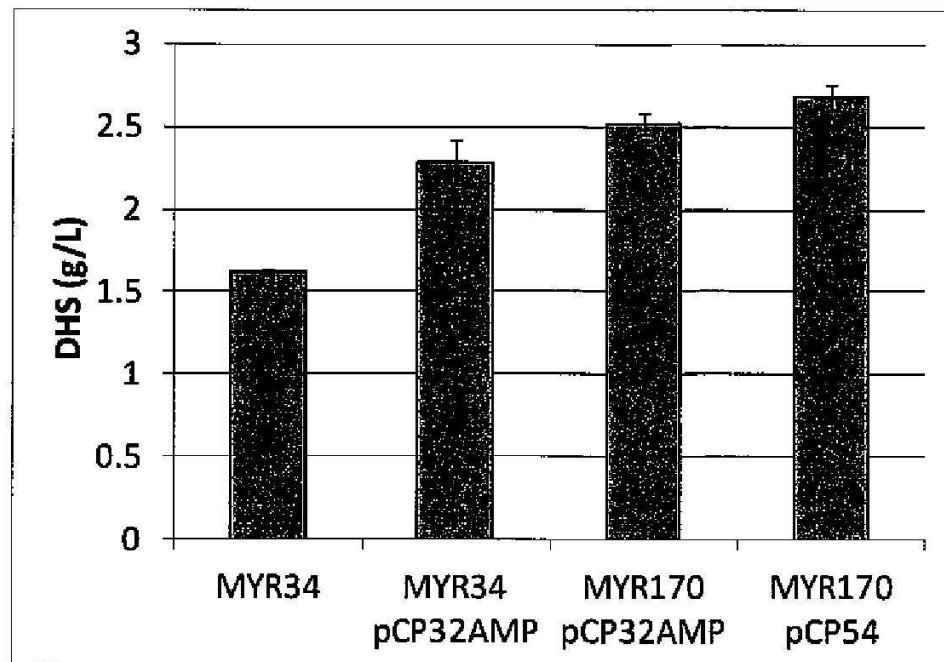
도면5



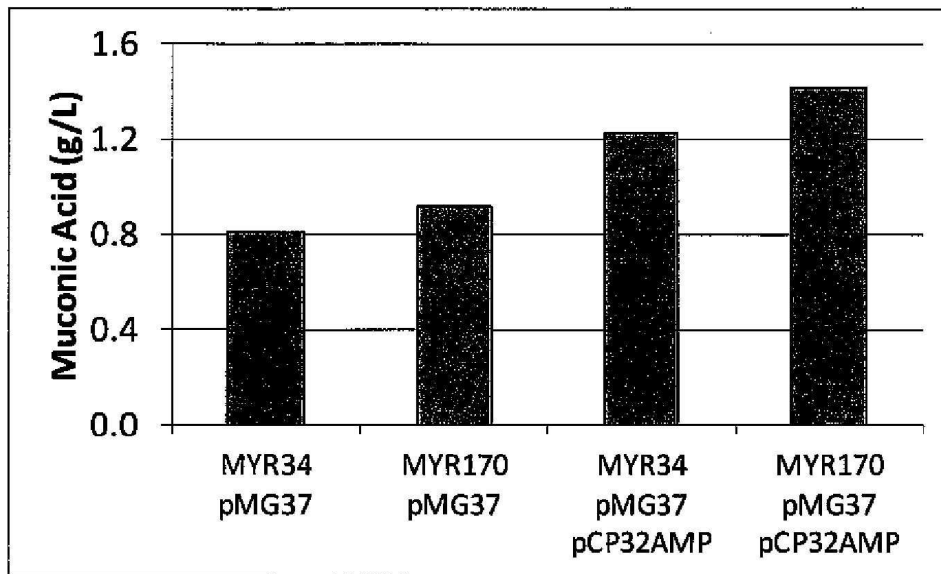
도면6



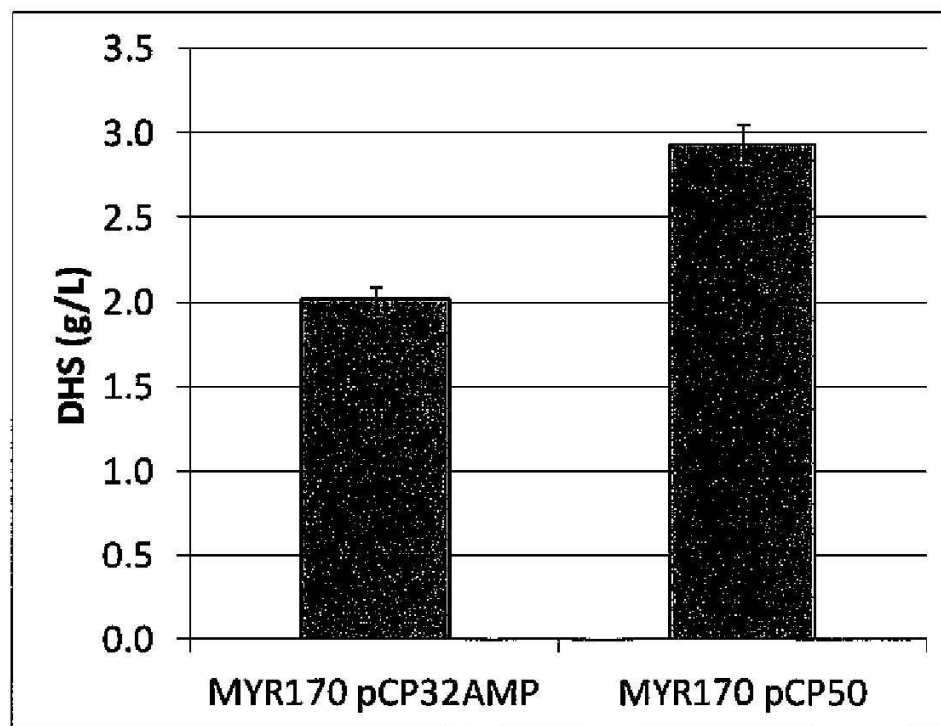
도면7



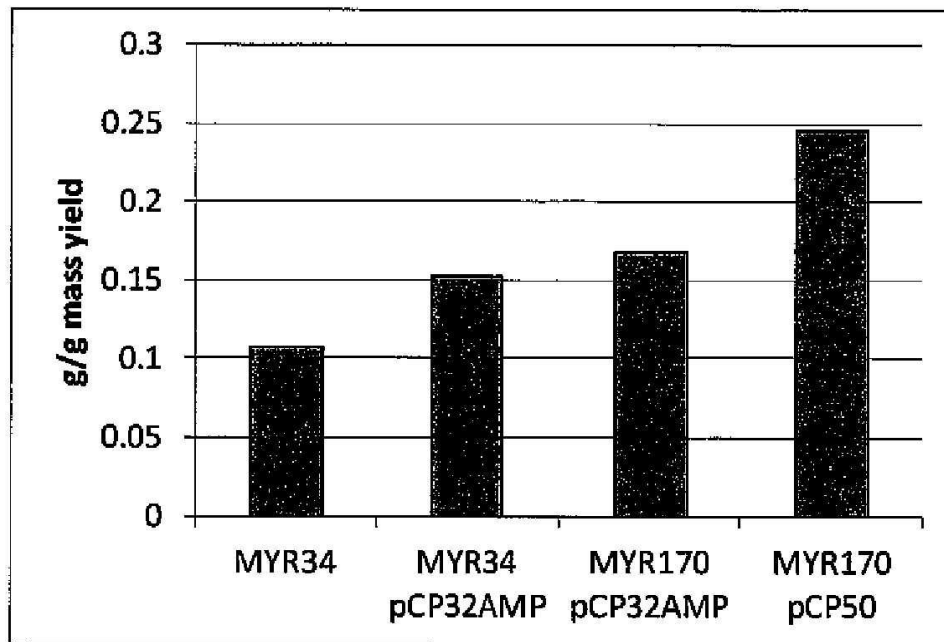
도면8



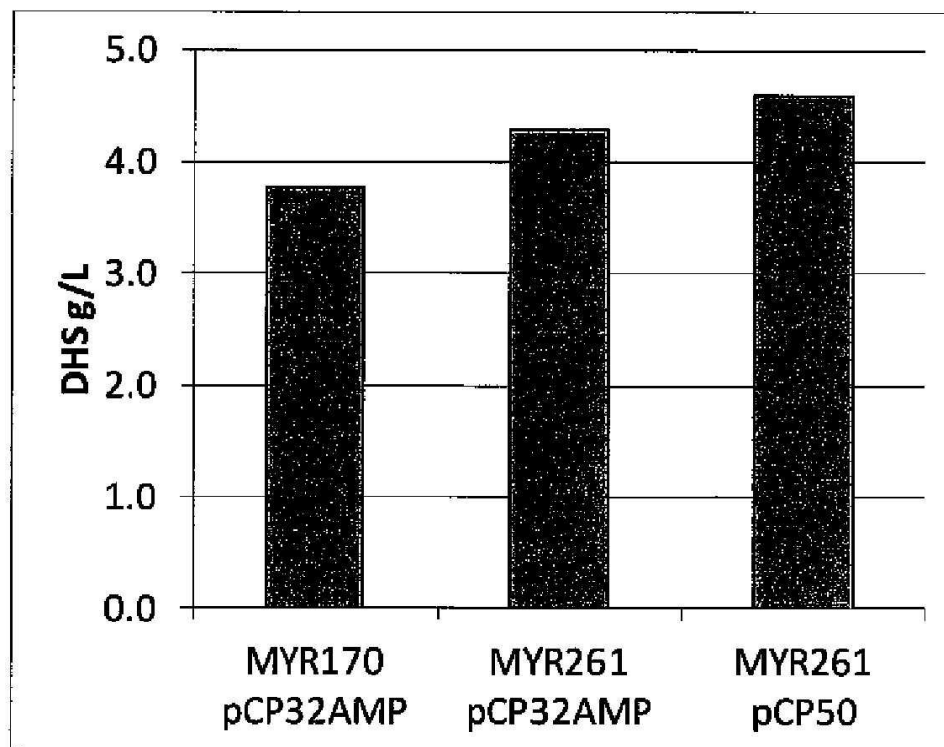
도면9



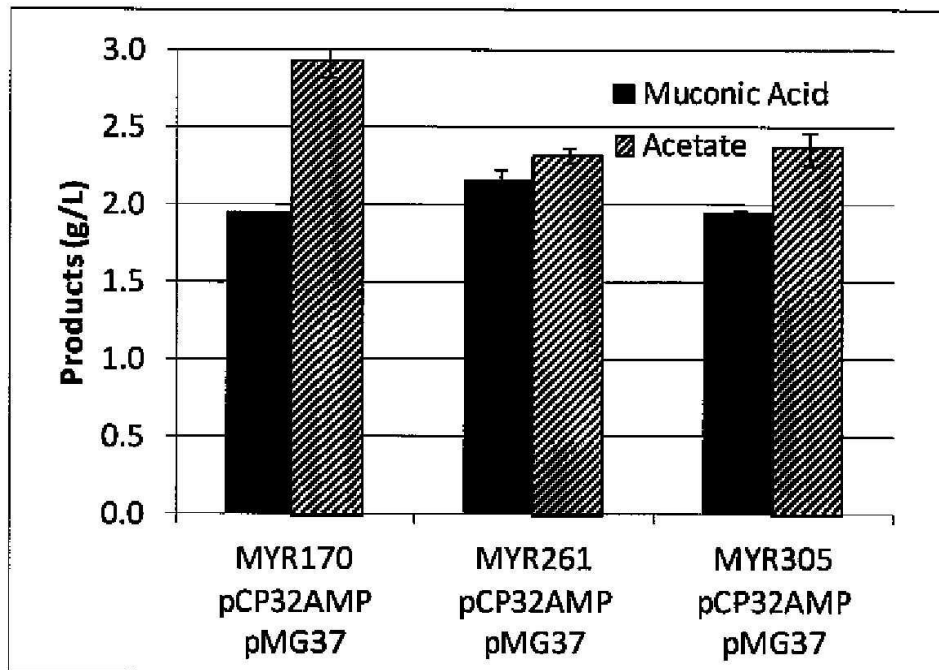
도면10



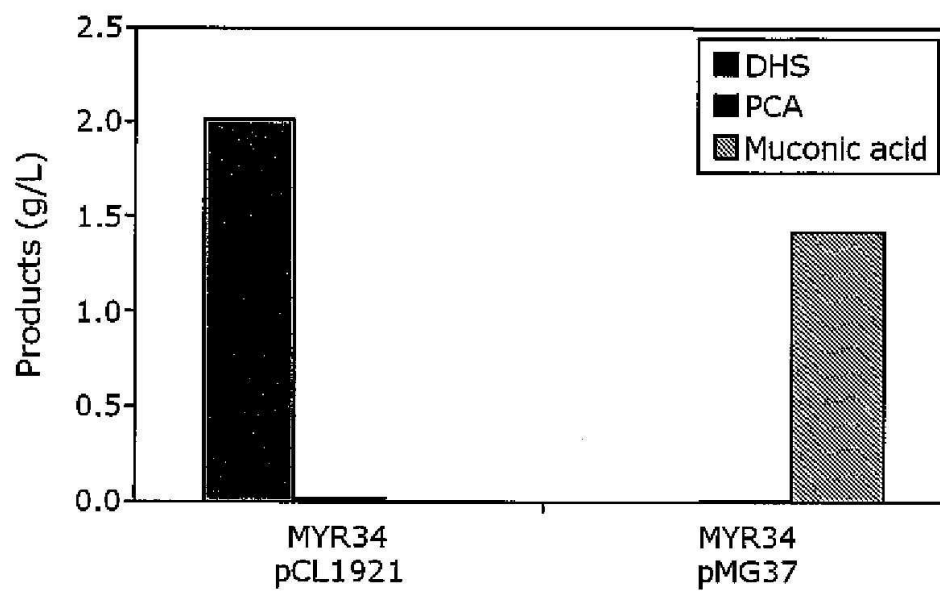
도면11



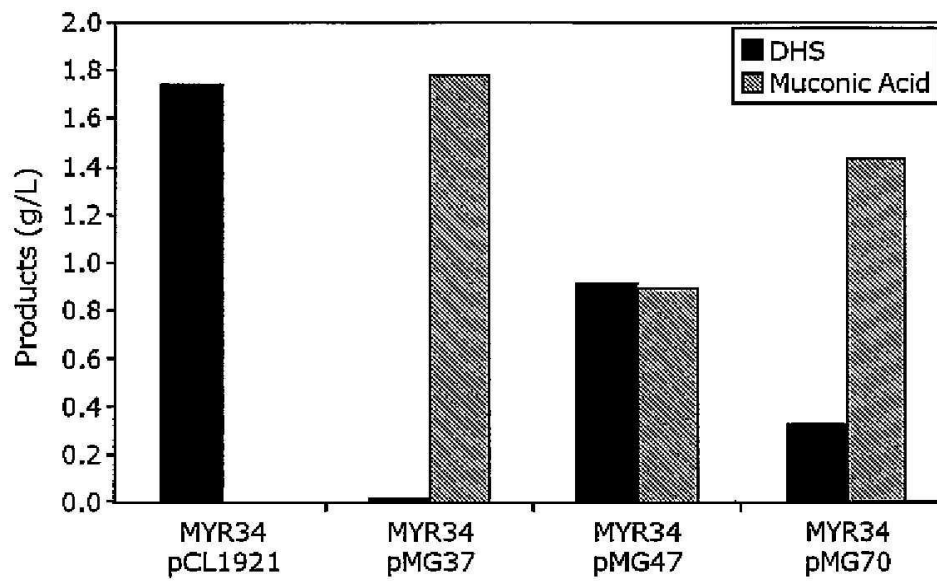
도면12



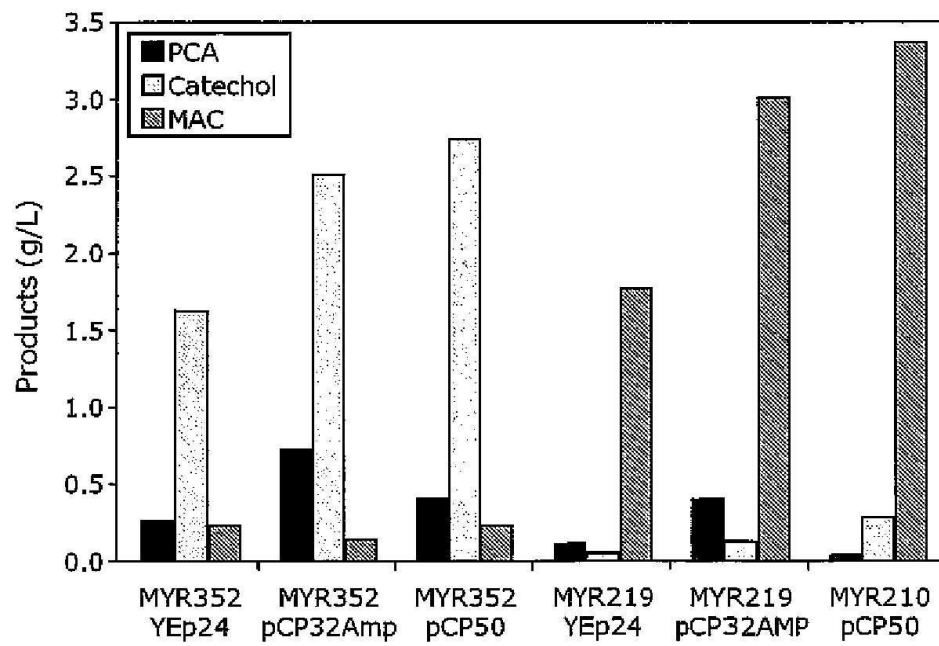
도면13



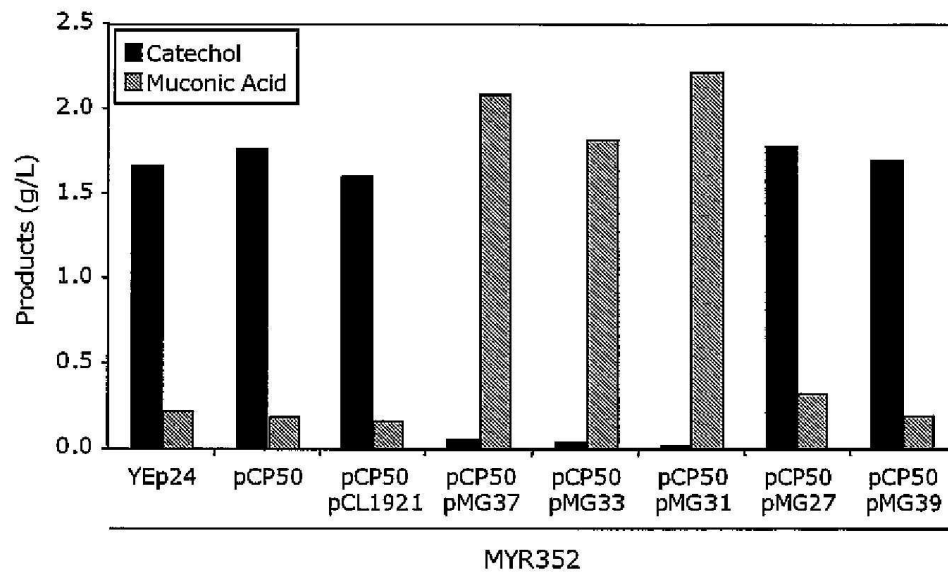
도면14



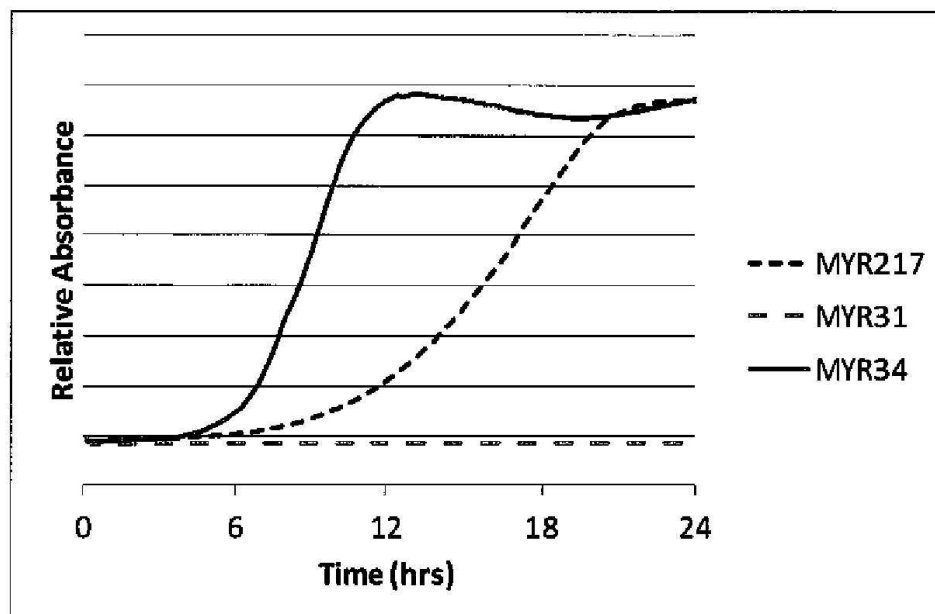
도면15



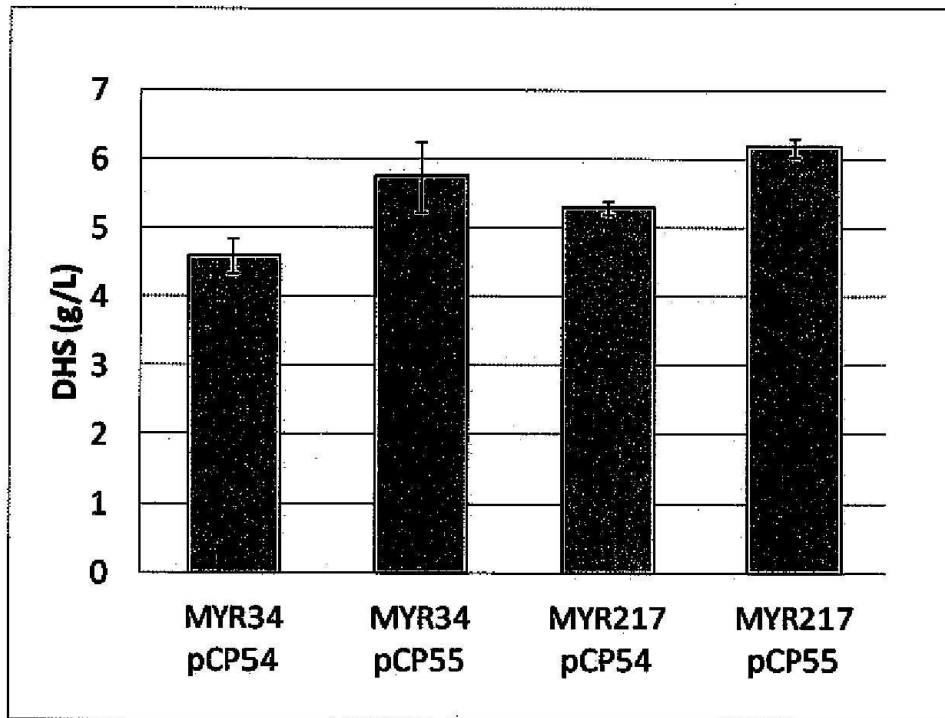
도면16



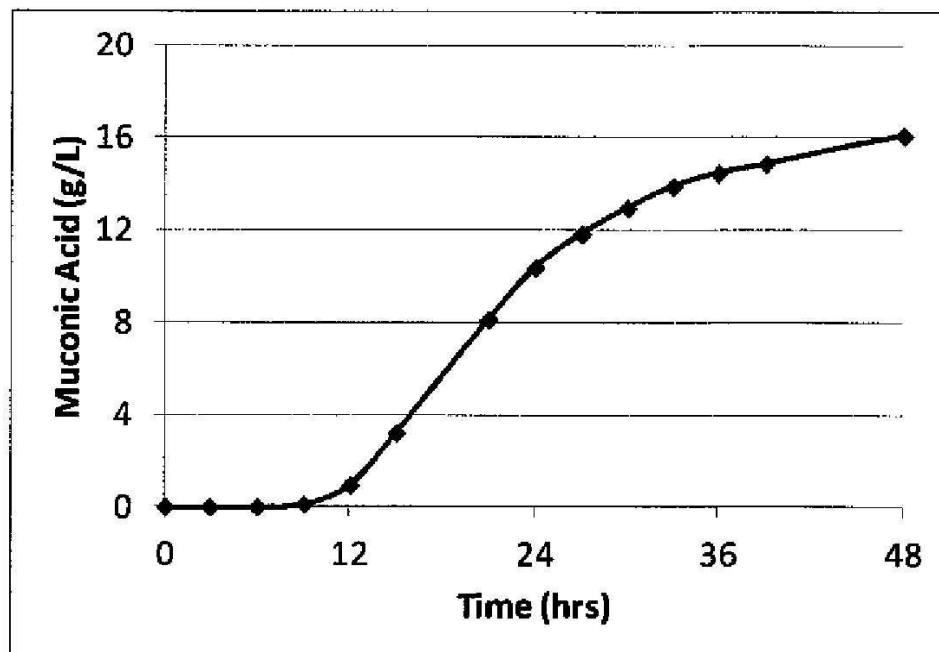
도면17



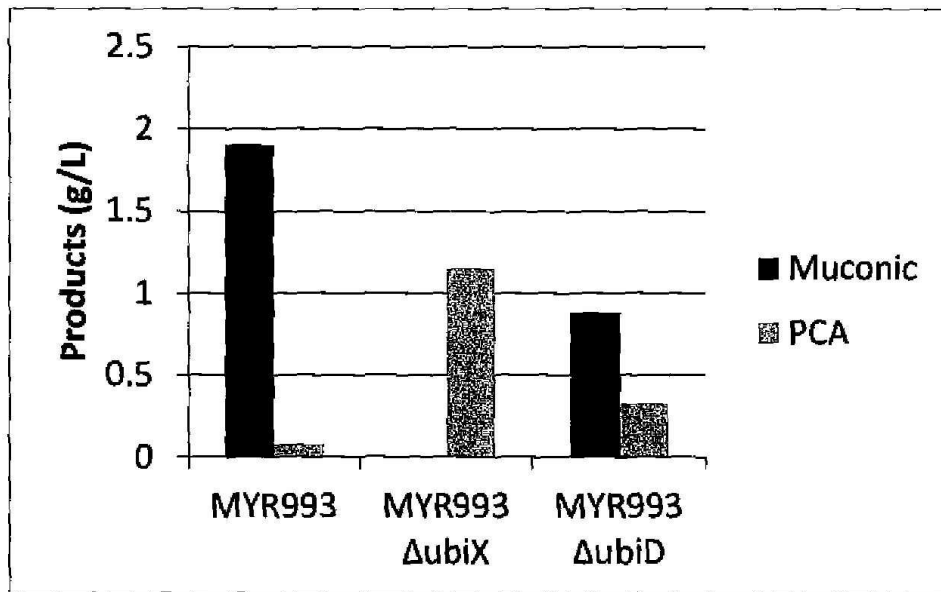
도면18



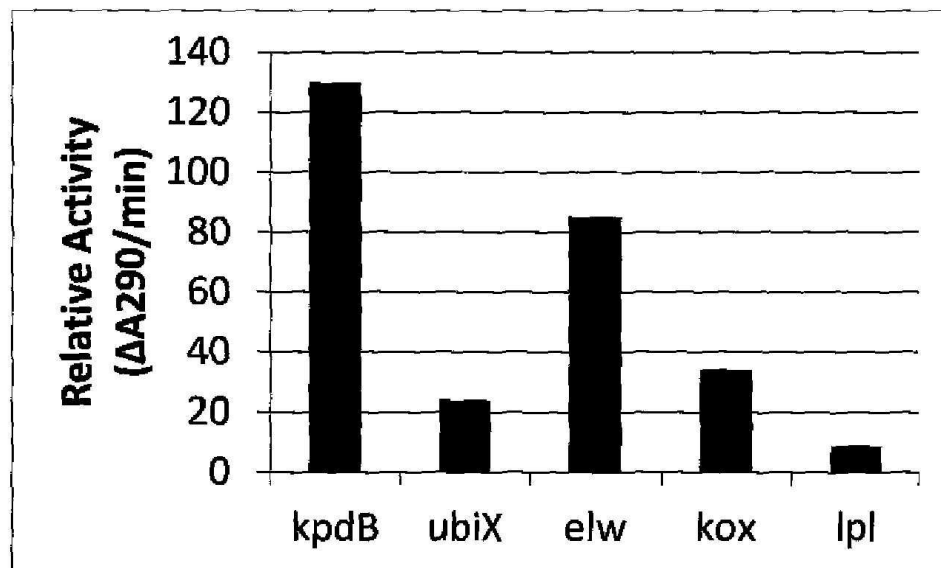
도면19



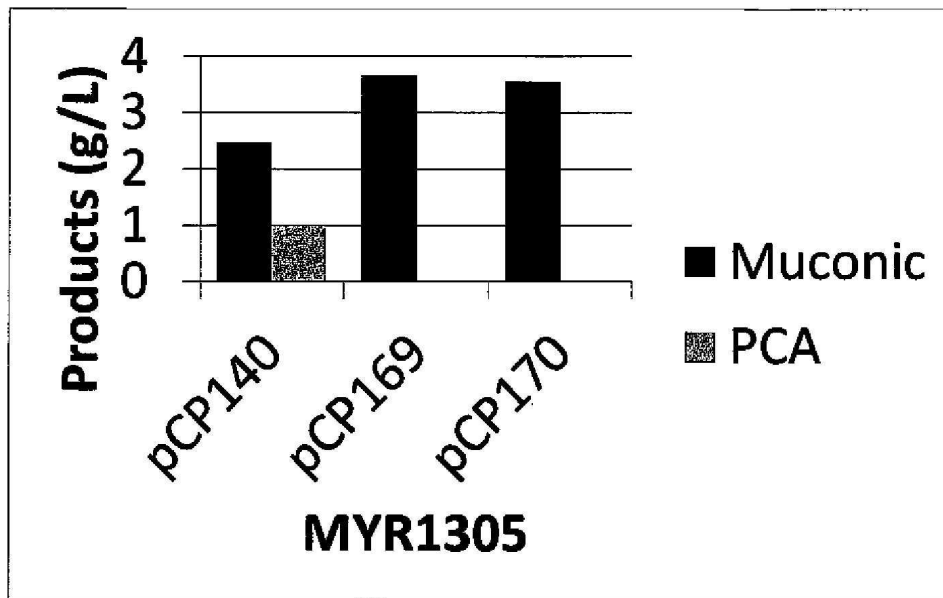
도면20



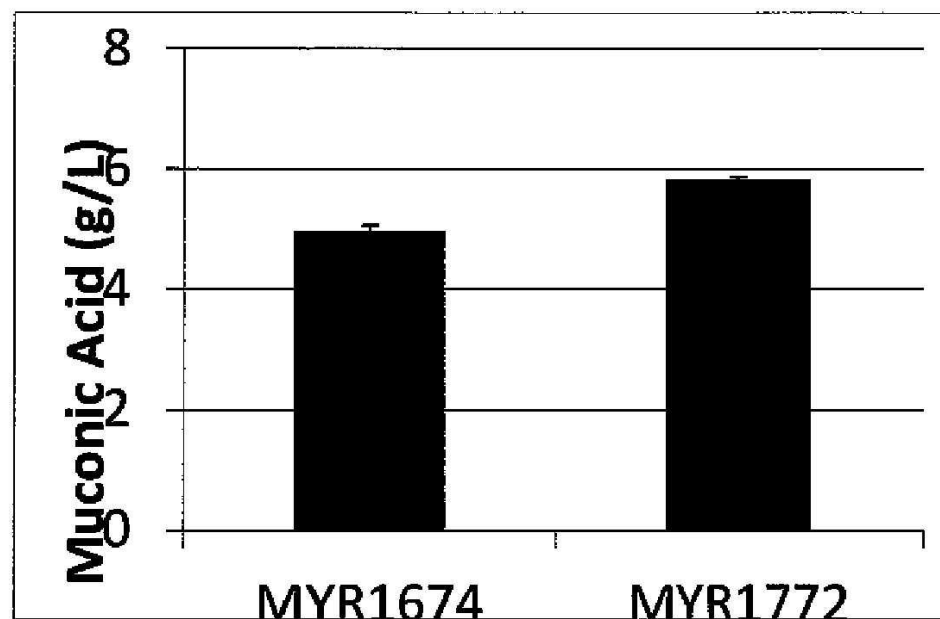
도면21



도면22



도면23



서열목록

<110> Myriant Corporation

Sillers, Ryan

Spencer, Michelle

Yocum, R. Rogers

Hermann, Theron

<120> IMPROVED MUCONIC ACID PRODUCTION FROM GENETICALLY ENGINEERED
MICROORGANISMS

<130> MC2017-01PCT
 <150> US 62/302558
 <151> 2016-03-02
 <160> 94
 <170> PatentIn version 3.5
 <210> 1
 <211> 195
 <212> DNA
 <213> Bacillus subtilis
 <220><221> promoter
 <222> (1)..(195)
 <223> The P15 promoter from Bacillus subtilis phage SP01, with a stem

and loop added just downstream from the transcription start site.

<400> 1
 gctattgacg acagctatgg ttcactgtcc accaaccaaa actgtgctca gtaccgccaa 60
 tatttctccc ttgaggggta caaagagggtg tccctagaag agatccacgc tgtgtaaaaa 120
 ttttacaaaa aggtattgac tttccctaca gggtgtgtaa taatttaatt acaggcgggg 180
 gcaaccccg cgtgtt 195

<210> 2
 <211> 164
 <212> DNA
 <213> Bacillus subtilis
 <220><221> promoter
 <222> (1)..(164)

<223> The P26 promoter from Bacillus subtilis phage SP01

<400> 2
 gcctacctag cttccaagaa agatatacta acagcacaag agcggaaaga tgttttgttc 60
 tacatccaga acaacctctg ctaaaattcc tgaaaaattt tgcaaaaagt tgttgacttt 120
 atctacaagg tgtggtataa taatcttaac aacagcagga cgct 164

<210> 3
 <211> 91
 <212> DNA
 <213> Escherichia coli

<220><221> promoter

<222> (1)..(91)

<223> The PR promoter from Escherichia coli phage

<400> 3

acgttaaadc tatcaccgca agggataaat atctaaccac gtgcgtgttg actatatttac 60

ctctggcggg gataatgggt gcatgtacaa g 91

<210> 4

<211> 359

<212> PRT

<213> Neurospora crassa

<220><221> PEPTIDE

<222> (1)..(359)

<223> Protein sequence of 3-dehydroshikimate dehydratase from
Neurospora crassa encoded by the qa-4 gene.

<400> 4

Met Pro Ser Lys Leu Ala Ile Ser Ser Met Ser Leu Gly Arg Cys Phe

1 5 10 15

Ala Gly His Ser Leu Asp Ser Lys Leu Asp Ala Ala Gln Arg Tyr Gly

20 25 30

Tyr Leu Gly Ile Glu Leu Phe Tyr Glu Asp Leu Val Asp Val Ala Glu

35 40 45

His Leu Ser Asn Glu Arg Pro Ser Pro Glu Gly Pro Phe Val Glu Ala

50 55 60

Gln Ile Ala Ala Ala Arg His Ile Leu Gln Met Cys Gln Ala Arg Gly

65 70 75 80

Leu Glu Val Val Cys Leu Gln Pro Phe Met His Tyr Asp Gly Leu Asn

85 90 95

Asp Arg Ala Glu His Glu Arg Arg Leu Glu Lys Leu Ala Leu Trp Ile

100 105 110

Glu Leu Ala His Glu Leu His Thr Asp Ile Ile Gln Ile Pro Ala Asn

115 120 125

Phe Leu Pro Ala Asn Gln Val Ser Asp Asn Leu Asp Leu Ile Val Ser

130 135 140
 Asp Leu Cys Lys Val Ala Asp Ile Gly Ala Gln Ala Leu Pro Pro Ile
 145 150 155 160
 Arg Phe Ala Tyr Glu Ser Leu Cys Trp Ser Thr Arg Val Asp Leu Trp

 165 170 175
 Glu Arg Cys Trp Asp Ile Val Gln Arg Val Asp Arg Pro Asn Phe Gly
 180 185 190
 Ile Cys Leu Asp Thr Phe Asn Ile Leu Gly Arg Ile Tyr Ala Asp Pro
 195 200 205
 Thr Ser Pro Ser Gly Arg Thr Pro Asn Ala Lys Glu Ala Val Arg Lys
 210 215 220
 Ser Ile Ala Asn Leu Val Ser Arg Val Asp Val Ser Lys Val Phe Tyr
 225 230 235 240

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

 305 310 315 320
 Glu Val Pro Glu Glu Leu Ala Met Arg Gly Ala Ile Ser Trp Ala Lys
 325 330 335
 Leu Val Gln Asp Leu Arg Ile Pro Val Glu Gly Pro Leu Val Thr Met
 340 345 350
 Pro Arg Val Ser Ala Ser Leu
 355
 <210> 5
 <211> 2160
 <212> DNA

<213> Neurospora crassa

<220><221> gene

<222> (1)..(2160)

<223> Genomic DNA sequence of the qa-4 gene from Neurospora crassa plus

surrounding sequences.

<400> 5

gaattcggga aatggaatct tacctgggaa ccgaaatcac agtccgggta ggttatagag	60
catatagtga actgtcaaag ttctagacct ggaccagcca cttggagtcg ttgttttagt	120
tatactaca ttactcact gttgactttc aatcatactt acttagacgg agcaacgcgc	180
cagaatcaa attgttgcag agttgcggta tcaccaagtg gcttcccata atagtgtgcc	240
attcgatgag acagctaact ggaagaccgg tactcgcagg ttgcacgatt acacggaagg	300
attcgggtatt ccgtgtttca tctgtcaaag tccctttcca tatgaatccg aggtactatg	360
actggatctc gatacaagct ggccagcgag gtgcctgcct tgacaggctg tcaactgcgg	420
gacggccggc taagtgttta acacgaagg gtggaagatg tctcgtcccg tcatccaaga	480
ccgtcaacat tcgaggccat ctgatcgttg aagagatgct aaatcttgtg aaacgctcat	540
aggtcgctta ctttcggccc acccgtaa gctttattcc gctgagcaaa cttcggcttc	600
catcccggcg ttacacgttt acatcactta tcgttgcggg tattggccga ttcttcgcaa	660
accgaaacga tgacatcccg aatatctgca atacaccgcc acggccggcg tctttatcac	720
acctcctatg ggagacgaaa gtgccttgat acccctagtc atttgaagat tcaggatggg	780
agacggctgg ccgcttgcgg agttacgttc gagtcttggg cgcaggaacg cttgccgtat	840
tgaatgagac cccgagaagg tcaaatcaaa tcttgggaaga cccaactgc ttcctcattg	900
ccttcactcc ccatatcaat ggggcacatc ctgtgactac cttggtgctt tatttcctca	960
ccatttggcg atacaagctc aaggacaccg aggtgatata cagtcttca aggacactat	1020
ctcacctcaa tatcaagaac cagtctcacc atctcttatt tctccaggat cccccacca	1080
acaacatcgg cttttttttt tccctattc tcaagacca tcaagacgt cacttcgctg	1140
agcctttcgc catgccgtca aagctagcca ttagttccat gtcctaggg cgctgctttg	1200
ccggccactc tctggacagc aagcttgatg ccgctcaacg atacggctat cttggtatcg	1260
agctttttta tgaggatctg gtcgacgttg cagagcattt gtcgaacgag cgtccctctc	1320
ccgaaggccc ttttgtcgaa gctcagatag ccgccgtcg tcatattctc cagatgtgtc	1380
aagccagggg gcttgaggtc gtctgcctcc agcctttcat gcactacgac ggccttaacg	1440
acagggcaga acatgagcgt cgtctggaga agctagcact atggattgag ctcgctcatg	1500

agcttcacac cgacatcatt cagatcccag ccaacttcct ccttgccaac caagtcatg 1560
 acaacctga cctgattgtc tcagatcttt gcaaggtggc cgatattgga gctcaagctt 1620
 tgccccctat ccgctttgcc tacgagagtc tttgctggag caccctgtgc gacctctggg 1680
 agcgctgctg ggacatcgta caacgcgttg accgccccaa ctttggcatt tgccttgaca 1740
 ccttcaacat cctcgccgcg atctatgccg accctacatc tcctagcggg aggacacca 1800
 acgcaaaaga ggcagtcagg aagtccatcg ccaacttggg ctgcgcgtg gatgtctcca 1860
 aagtcttcta cgtccaggtg gttgacgccg agaggctgag caagccacta ctgcccggtc 1920
 acccgtaata caatccagag cagccggcga ggatgagctg gtcgcgcaat ttagactgt 1980
 tctacggcga aacagaatat ggtgcgtatc ttcccgtgaa ggaggttgct cgagcccttt 2040
 tccacggcat tggtttcgag ggctgggtca gtttggagct tttcaaccgc agaattgtctg 2100
 aggagggacc tgaagtgcg gaggaacttg ccatgagagg cgctatctcg tgggccaagt 2160
 2160

<210> 6
 <211> 348
 <212> PRT
 <213> Aspergillus niger
 <220><221> PEPTIDE
 <222> (1)..(340)
 <223> Protein sequence of 3dehydroshikimate dehydratase from
 Aspergillus nidulans. encoded by the qutC gene
 <400> 6

Met Pro Ala Asn Leu Lys Ile Gly Ile Pro Thr Val Ser Leu Ser Lys

1	5	10	15
Pro Gly Leu His Ser Leu Asp His Lys Leu Arg Ser Ala Ala His Gly			
20	25	30	
Phe Ala Gly Ile Glu Leu Phe Ile Asp Asp Leu Ser His Phe Ala Ser			
35	40	45	
Ser Ser Phe Asn Gly Ser Leu Thr Gln Ala Ala Lys Tyr Ile Ser Ser			
50	55	60	
Leu Ala Lys Gln Leu Asn Leu Thr Phe Ile Cys Leu Gln Pro Phe Gly			
65	70	75	80

Phe Tyr Glu Gly Leu Val Asp Thr Asn Gln Ser Thr Tyr Leu Leu Thr
 85 90 95
 Glu Lys Leu Pro Leu Trp Phe Ala Ile Ala Arg Ile Ile Gly Thr Asp
 100 105 110
 Leu Ile Gln Ile Pro Ala Asn Phe Leu Gln Asn Asp Pro Val Thr Gly
 115 120 125
 Ala Ala Arg Thr Ser Gly Asp Ile Arg Leu Ile Val Ser Asp Leu Gln
 130 135 140
 Thr Ile Ala Asp Ile Gly Val Lys Gln Gly Phe Arg Phe Val Tyr Glu
 145 150 155 160
 Ala Leu Cys Trp Ser Thr His Val Asp Thr Trp Glu Ala Ala Trp Asn
 165 170 175
 Val Val Lys Leu Val Asp Arg Glu Asn Phe Gly Ile Cys Leu Asp Ser
 180 185 190
 Phe Asn Thr Arg Thr Pro Leu Pro Ser Leu Gly Arg Arg Arg Met Leu
 195 200 205
 Ser Lys Pro Trp Pro Ser Pro Trp Arg Arg Ser Val Leu Ser Ser Pro
 210 215 220
 Val Glu Asn Trp Thr Ser Gly Lys Ser Ser Thr Ser Ser Leu Ser Met
 225 230 235 240
 Ala Ser Gly Cys Arg Arg Arg Trp Thr Arg Ser Thr Pro Phe Met Trp
 245 250 255
 Arg Ala Asn Pro Arg Arg Met Ser Trp Ser Arg Asn Ala Arg Leu Phe
 260 265 270
 Pro Cys Glu Glu Glu Arg Gly Gly Tyr Leu Pro Val Leu Glu Ile Ala
 275 280 285
 Arg Ala Phe Phe Glu Ile Gly Phe Glu Gly Trp Val Ser Leu Glu Leu
 290 295 300
 Phe Ser Arg Thr Cys Asn Asp Pro Asp Val Asn Thr Val Gly Glu His
 305 310 315 320
 Ala Arg Arg Gly Met Asp Arg Arg Arg Arg Val Val Ala Ala Leu Gly
 325 330 335

Leu Asp Val Glu Val Pro Ala Arg Asn Cys Glu Cys

340

345

<210> 7

<211> 3298

<212> DNA

<213> *Aspergillus niger*

<220><221> gene

<222> (1)..(3298)

<223> Genomic DNA sequence of the *qutC* gene from *Aspergillus nidulans*
plus surrounding sequences

<400> 7

```

aagcttggtt tcaagtgatg atatatagtt atgaggatat aatatgaacc gaaagacgat      60
gtttcttgtg aatatttacg tgatagttgt ctgtctaata tggtagacga gtagaacaac      120
tacatacggg cactacttac agccctagtc attccctccc tcgattgcct accatttata      180
cactttgaac atccacaggc ttgcctccct ccatactctc ctaaacagct tgaacaactc      240
tgagcgccct caccccatct tcaacgccac agccaactcc tcgctcgcca tctcaccct      300

ttccactaac aacatcaaca aaatatcccc actgcgcata aaacggcctc acatcagcat      360
ccttactga aatctgctgc atcgccaact ccgtattcca gcctttctct tccccctgtc      420
cacaagaaac atggtcatag ctccagcgcg tcatatcagg cacactgaga ctgcctctgg      480
ttccaagaat tcgataacag tcgcttgcac tcttggaagg agcggggcgga atcgtaggat      540
tctcgcccggt tctgtttca aagttaacg gcgaaggcgt cgcgtcgag atgagaaatg      600
tgctactat ccagacgca aagcgacgag tcacagcaca gccttctctg gcggtatgtt      660
ccgggttctg gcgcatgcgc tgcaggagtg taccctccgc gtagacccta ctgacgggcc      720

caaacagaaa ctgcagcacg tcgatatcgt ggataagatt aatccccagc acgccgcct      780
tcttcttate tgcgcgccaa gaaccgagcg gcggcgcgaa gtaagaggcc ggcttcagaa      840
gtgtccagag gccgttact gcaacgacgg tgccgagtga gtctgtctct aacaaagact      900
ttgtggtttg gatgtacgga ttgaagcggc ggtggtggcc gatctggatg ttgatcttcg      960
catecttgcc tctcttctca tctttacatt tctgttctt gacggtagcg aggaggtgct      1020
cggccgactc cagatcgta ctgatcggtt tctcaaggag gatattgcgg attccgttct      1080
ccagcagctg gagcgtgacg tccacgtgag tgtgattggg cgtgctcacg atcgccgcgt      1140

ctggtttccc ggttgtctta ccgacaacgt ctaacataga cgtaatagaa tcatagcaag      1200

```

gaacgccaaa tgattctgcg accgggattg cagaggggtga agggtaaca aaagcgatca	1260
gctgggttcg tgggtgtcgt tgcacggatt gtgcgtgacg gggcccaata agtccggcac	1320
cgacgatgac aatgaggata ttcttgtcct ttctcttgct gcagggcacc attgtgcatg	1380
tcggtggctg gaaataaaca gaacagggat atggtaagt cggagaatcg gtgcaggata	1440
gaccggctac ttgatgtagg acgacagtcg cgatctaccg agagcgtgag attcactgtg	1500
ggactgattt atgtaatttg aggcgcagca gacttaggga cttgaaatgt ggctgtctgt	1560
ggatgcattt gcggggtatg gagtacagag tgcatacagc tgtglatatg gagttcctta	1620
cggagaggggt gacctggatg ggggagaacg ggcaaaatgc tcaccggca acctctcaaa	1680
gcgtttaccc ggtatactcc tctgatatca atatttccaa tcagcaccta tatcatcacg	1740
acgtctcctt gaggattccg tagctaaccg ccttgatcc tacattaata aataagccat	1800
ttgctttttc tgcctgcagt gtgattctca atacgattac gtatcacatg cagattgcct	1860
ttacttcagc tgcatttgat cagccacagc tctaagagca aacataacct acctacctac	1920
ctacttcgcc tagggtacat aatcaccgcc atctctcct cgatcagtct tcaactcaat	1980
cagtcattc attctattct taatataata tataccttta gatctccagc agagaccga	2040
agagtcggca attcaaatg cccgcaaacc tcaaatcgg tatcccaacc gtgtccctgt	2100
caaaaccggg cctgcactct cttgaccata agctccgctc ggccgctcat ggcttcgagg	2160
ggatcgagct gtttattgat gacctctccc atttcgctc atcgatcattc aatggctccc	2220
tcactcaagc ggcaaagtat atctctctgc tcgccaagca acttaacctc acatttatct	2280
gcctgcaacc attcggtttc tacgagggtc tggtaggacac aatcagtcg acgtacctgc	2340
tcactgagaa actcccgtc tggtttgca tcgcccgcgt tataggcaca gatctcatcc	2400
aaatccccgc aaatttcctc cagaatgacc ctgtcaccgg ggctgcacga acaagcggcg	2460
acataaggct tatcgtctca gatctgcaga cgatcgaga tatcgggtga aagcagggt	2520
tccgctttgt gtacgaggcg ctctgctggt cgacgcatgt cgatacatgg gaagcagcgt	2580
ggaatgtcgt caagctgggt gatagagaga atttcgggat ctgcctggat agcttcaaca	2640
cgcggaaccc gcttcgctca ctgggaagac gccggatgct gagcaagccg tggccaagtc	2700
catggagacg ctccgttctc tcgtctccag tggagaactg gacatcagga aaatcttcta	2760
catccagctt gtcgatggcg agcggttgtc ggcgccgtg gacgagaagc acccctttca	2820
tgtagggggc caacccccga agaattgagct ggagtcgcaa tgcgcggtta ttcccctgtg	2880
aagaggagag gggtaggtat cttcctgtgt tggagatcg gagggcgttc ttgaaatcg	2940
ggttcgaggg gtgggtgagt ctagagctgt ttcaaggac gtgtaatgat cccgatgtga	3000
acacggtggg ggagcatgcg agacgtggga tggatagaag gaggagggtt gttgcggcgc	3060

taggactcga tgttgagggtg ccagcacgta actgtgaatg ttagcatgaa cggcaaggag 3120
agggtggagg tgcaggtgca ggaggagctg gctgttcagc atcggctgta ggtagtggtg 3180
tcttgaaagg acgatagggt ttgatctaga gatTTTTATT ttgtctaatt actggtaatg 3240

atggcctcat gcacgctgtt gaacacgctg tacaacatca ctgttgaaga tgatacct 3298

<210> 8

<211> 502

<212> PRT

<213> Klebsiella pneumoniae

<220><221> PEPTIDE

<222> (1)..(502)

<223> Protein sequence of protocatechuate decarboxylase (AroY) from
Klebsiella pneumoniae ATCC25597

<400> 8

Met Thr Ala Pro Ile Gln Asp Leu Arg Asp Ala Ile Ala Leu Leu Gln

1 5 10 15

Gln His Asp Asn Gln Tyr Leu Glu Thr Asp His Pro Val Asp Pro Asn

20 25 30

Ala Glu Leu Ala Gly Val Tyr Arg His Ile Gly Ala Gly Gly Thr Val

35 40 45

Lys Arg Pro Thr Arg Ile Gly Pro Ala Met Met Phe Asn Asn Ile Lys

50 55 60

Gly Tyr Pro His Ser Arg Ile Leu Val Gly Met His Ala Ser Arg Gln

65 70 75 80

Arg Ala Ala Leu Leu Leu Gly Cys Glu Ala Ser Gln Leu Ala Leu Glu

85 90 95

Val Gly Lys Ala Val Lys Lys Pro Val Ala Pro Val Val Val Pro Ala

100 105 110

Ser Ser Ala Pro Cys Gln Glu Gln Ile Phe Leu Ala Asp Asp Pro Asp

115 120 125

Phe Asp Leu Arg Thr Leu Leu Pro Ala His Thr Asn Thr Pro Ile Asp

130 135 140

Ala Gly Pro Phe Phe Cys Leu Gly Leu Ala Leu Ala Ser Asp Pro Val

145 150 155 160
 Asp Ala Ser Leu Thr Asp Val Thr Ile His Arg Leu Cys Val Gln Gly

 165 170 175
 Arg Asp Glu Leu Ser Met Phe Leu Ala Ala Gly Arg His Ile Glu Val
 180 185 190
 Phe Arg Gln Lys Ala Glu Ala Ala Gly Lys Pro Leu Pro Ile Thr Ile
 195 200 205
 Asn Met Gly Leu Asp Pro Ala Ile Tyr Ile Gly Ala Cys Phe Glu Ala
 210 215 220
 Pro Thr Thr Pro Phe Gly Tyr Asn Glu Leu Gly Val Ala Gly Ala Leu
 225 230 235 240

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

 305 310 315 320
 Val Gly Pro Gly Glu Glu His Thr Thr Leu Ala Gly Leu Pro Thr Glu
 325 330 335
 Ala Ser Ile Trp Asn Ala Val Glu Ala Ala Ile Pro Gly Phe Leu Gln
 340 345 350
 Asn Val Tyr Ala His Thr Ala Gly Gly Gly Lys Phe Leu Gly Ile Leu
 355 360 365
 Gln Val Lys Lys Arg Gln Pro Ala Asp Glu Gly Arg Gln Gly Gln Ala
 370 375 380

 Ala Leu Leu Ala Leu Ala Thr Tyr Ser Glu Leu Lys Asn Ile Ile Leu
 385 390 395 400

Val Asp Glu Asp Val Asp Ile Phe Asp Ser Asp Asp Ile Leu Trp Ala
 405 410 415
 Met Thr Thr Arg Met Gln Gly Asp Val Ser Ile Thr Thr Ile Pro Gly
 420 425 430
 Ile Arg Gly His Gln Leu Asp Pro Ser Gln Thr Pro Glu Tyr Ser Pro
 435 440 445
 Ser Ile Arg Gly Asn Gly Ile Ser Cys Lys Thr Ile Phe Asp Cys Thr

450 455 460
 Val Pro Trp Ala Leu Lys Ser His Phe Glu Arg Ala Pro Phe Ala Asp
 465 470 475 480
 Val Asp Pro Arg Pro Phe Ala Pro Glu Tyr Phe Ala Arg Leu Glu Lys
 485 490 495
 Asn Gln Gly Ser Ala Lys

500

<210> 9

<211> 5502

<212> DNA

<213> *Klebsiella pneumoniae*

<220><221> gene

<222> (1)..(5502)

<223> DNA sequence of the *aroY* gene of *Klebsiella pneumoniae* 342 plus 2

kilobases of surrounding DNA sequences

<400> 9

gcgacgccga ctggcgatc cgtgaactgc tggcgcgtat gaccacgcgt ctgcagggt 60
 gtgaaacat agaggatgtg attaaggtgg cggagctgtt cgcgccgaac atgccccga 120
 cgatccccgg taaactgtat attctggata ccgatccatg gcagatgcgc tgcgtggcgc 180
 agtggctgtc gcccgccggg gagacgacgt cctttgctcc cgacgactgc tgggcgatac 240
 ggcggggact cagccatccg ccggtgcagg gtgagccga tatcacctgc tatcatctgc 300
 cggaggcgca cgccggccag tcgctctgcg taccgctcat cgcccagggc gaagcgatcg 360

gtctgtgag ctttcagaac gtcaccgcca gtgacgccc ttccggggt tacctggagc 420
 tgatggccga agcgctgggg ctggcgctcg ccaatcagcg ttacgcagc gccctgctgg 480
 aaaaagcgtt gttcgattcg ctgaccggcc tgcgtaaccg ccatcatctt gatgaagcgc 540

tgcactcgca gatggcgctg gcggtccata cccacacccc gctgagctgc ctgatgatcg	600
acatcgatca cttcaaagcc atcaatgacc gctacggcca tgaagccggg gatctggtga	660
ttaagagcgt cgcgaccatt gtgcagcgcg cgggtgcgga tatcggcatt gctttccgct	720
acggcggcga ggagttttta gtgctgctcc ccgggattga cgaagccggg gcgcaccagt	780
gcgccagcga gatctacacc cagggtgcaca atatgacgct gcgcgatggc ctgacggaga	840
taggccaggt ggatgtgtcg attggcatcg ccagctaccc gcagcacacc caaagcgaca	900
gcctgtctcg cgcagcggac gccgcgctgt accgggcgaa agagctgggc cgttcaagga	960
ttgtcagctt tggccgctg aagaccgct aagcgggatt attgctcagc ggcattaagc	1020
agcgagataa ctttccgcac caccgccgaa cggaaatggc ggtggtaaac catgctcagc	1080
tcactctccg ccagccgac ctgaagatcg atatacaca cgttatcaag gcgtaaggcc	1140
ctcgccgaga ccggcaccag cgccaccccc accccggcgg agaccagct aatcatcgag	1200
gtgacatcgt taattcgctg caccacctgc ggcgtaaaac ccgcgacgcg acaggcgtca	1260
ataaatacct gctccagtc ggtgccctgc ggatcgtcaa gcgagatcca gttgtcagt	1320
cgcaacgagg ccagattgag cgccccacg cctgccagcg gatgctgttg ataaagcgcc	1380
aggcaaagtt tttccgcac aaatggcctg accaccagcg cgtccggcgg tgacgccagc	1440
ggcgcgcgga tgatggcgat atccagacgc agatccagca gcgcttcgta gagcatttgc	1500
acatccccct gcaccagcga cagctcaatc ccgggccagt cagcgcgcag ctgcgcagg	1560
agccccggca gtttctgtc atacatcgca ctggagacat agcccagatg caatcgcccc	1620
tgctcgctc gtgcggtgcg ctgggcgtcc aggaccgct gatcgccat ctccagcgcc	1680
agcccgctct tctgcaagaa ggcctcgccc gcggcggtga gggtcaggcg ccggttagcg	1740
cgggagaaga gcaccagcc caggcgctgc tcgagttgtt taatctgctg gctgaggcg	1800
ggctgggca tatgtaaccg ctctgccgcc cgatgcata gtagttcttc agcaacgacc	1860
acaaaatggc gtaacgctcg caaggacatg gccggactcc gcggagtaa ttgataataa	1920
aaatgttate aataaagcat gaatgatgca attgataacc attagcctgc gagcatactg	1980
tgcgcatcga cagcctaagg agaacatcat gaccgcaccg attcaggatc tgcgcgacgc	2040
tatcgcgctg ctgcaacagc atgacaatca gtacctgaa accgatcatc cggttgaccc	2100
taacgctgag ctggccggcg tctatcgcca catcggcgcg ggccgcaccg tgaagcgccc	2160
cacgcgcatc ggcccgcgca tgatgtttaa caatattaag ggctatccgc actcgcgcat	2220
tctggtgggc atgcacgcca gccgccagcg ggccgcgctg ctgctgggct gcgaagcctc	2280
acagctggcg ctggagtag gcaaagcggg gaaaaaacg gtcgcgccgg tggctgttcc	2340
ggccagcagc gccccctgc aggaacaggt ctttctggcc gacgatccgg attttgattt	2400

gcgcaccctg ctccccggcg ccaccaaacac cccgatcgac gccggtcctt tcttctgcct	2460
gggcctggcg ctggccagcg atccccgacga cgcctcgctc accgacgtca ccatccaccg	2520
cttgtgcgtc cagggccggg atgagctgtc gatgtttctc gccgctggcc gccatatcga	2580
agtgtttcgt cagaaagccg aggccgctgg caaacctgtg ccgataacca tcaatatggg	2640
actcgatccg gctatctata tcggcgcttg ctttgaagcg ccgaccacac gtttgctat	2700
aacgaactgg gcgtcgccgg tgcgtgcgt cagcgtccgg tagagctggt acaggcgctc	2760
agcgtcccgg agaaagccat cgctcgcgcc gagatcgta tcgaaggga actgctgccg	2820
ggggtacgcg tcagagaaga tcagcacaca acagcgcca tgcgatgccg gaatttcctg	2880
gttactgcgg cggcgccaat ccgtcgctgc cggtcattaa agtcaaagcg gtgacatgc	2940
gaaacaatgc gattctgcag acgctggtag ggccggcgga agagcatacc accctcgccg	3000
gattgccaac ggaagccagt atctggaatg ctgtcgaggc tgctatcccg ggctttttac	3060
aaaatgtcta cgccacacc gcggcgcgcg gtaaatcct cgggatcctg caggtgaaaa	3120
aacgccagcc cgccgacgaa gggcgtcagg ggcaggccgc gttgctggcg ctggcgacct	3180
attccgagct gaaaaatata attctggtcg atgaagatgt cgatatcttt gacagcgacg	3240
atatcctgtg ggccatgacc acccgcatgc agggggatgt cagcatcacg acgatcccag	3300
gcattcgtgg tcaccagctg gatccttccc agaccccggc ctacagcccg tcgatccgcg	3360
gagagggtat cagttgcaag acgattttcg attgcacggt gccgtggcg ctaaaatcac	3420
acttcgagcg cgcaccgttt gccgatgtcg atcccgctcc gtttgcgccc gagtattttg	3480
cccggctgga aaaaaaccac ggtcagtaaa atcaggatgat agccgccgga gcacggcggc	3540
atcttccggg ccagcatcac ctgcagcggg tggctgacgc agggttagtt gatcgcgcg	3600
gagaggtctt ttttcacctg ctcacgtgc tcgggggtca acacctggct cacgtcgaag	3660
tagtatttca cagataata ccgcacctgc tggccagct ggccaaaggc ggccagctgc	3720
tgtttgacct tagcgtcatc ccatttccc gagtgaataa cgtctgccag ggcacatcc	3780
tgatagccgc tgatttaate tggcttacat tgttttcgaa tcctgacgc agcgctgga	3840
ttttggcgac ctgctcttca ctcagcttca ggtgctggac gaccgatcc tgcgagacag	3900
acggtatata ggcggaggtc gacgcctggc tggctgccgt aaagcaggtg gtcagcgcaa	3960
tggcgagcag ggtgttacgc aagcgagtat tcacagtga tgatccttca aaaaagaaaa	4020
tgagaggcga ttatcactgc gctaataaag actatctgta acaaagggtt aattttaa	4080
tggataaaaa aaggatggta agaaacagaa atcagatccc gggtcagcag cacagaaaga	4140

tatattcatc cticcagtaa cggccctgtc caatgatatc cccggcggcg ctgattaact 4200
gtttttgctt ttggtttcaa tccctcaac gatcacatgg ctggtcaggg tatgaataga 4260
ttgcaacagc cgggaaaaag cggggtcgtt tttttatcc cagaagtaat ctttatccac 4320
tttgacgcaa tcgaagcgga agcgctccag cagcgggaag gacgtcgtcc gcggccaaaa 4380
tcattccagc agaccgggca aagcgccgcc agcgtgtca gcgccgtcag ctcacgtccg 4440
gcgataaact cgtgaaagtt ctcatattt tccagagcaa tgtgtttaca ggagcgcagg 4500
aatcacaga gatagcgatc cgtcagaata aagtggctta ataatcatc aatattcagc 4560

gatatcggtt tattatcgac cgggcgaaa tcaaatacag aaagtaactc gatttgccga 4620
ataaatagag caagcttacc gcgttcggta agcgtggaga aagaaacgt cgacgcagag 4680
gtggtatgtt gtgcgggtgc tataatatct ttagtaagca actcccacga gtggtagctg 4740
ccgtcatcgc taatagcggg ctccagaacg aagcgataag aggtatgttt cacgttttct 4800
tcaaccattt aaaaaatacc aaaaataaga aagggttaag catgtcatat atttccgcc 4860
aacaaaaata gtttaaagt atcgataata atcattcgat agttaaaac tatcaagata 4920
taatttattg atcggtaaat tgaattaata taaattagcc actgccgtaa ctccctctga 4980

aaagtcaatt aaaatatgt ttcaaacag ccagttacca gagtattctg cgtaaagcct 5040
ggtcgtctca cgctttgtgc tgccaggtaa aaaaagagag gggtataaaa aatgaaaaat 5100
acaagccgcc agttttagtc atatcattat gccgaatatg aataacgtg cgctgaggcg 5160
ccgttcgcc tggcatgcca tgagtcctca acaaaaaagt gtgactcagt cgacaaaacg 5220
tcatttttc ccgtatcct gcagcgaaga agagtgaagt ggatgacagg cagtgaaaaa 5280
aataaacgtg attccgctgg ggctgatgct attcatgctc atcgccagcg catggctggg 5340
ccctgcgcgc cggcacaccg gcagcatgca gtgcgtttgg ttgacgggg caatggtgag 5400

ctgcctgccg aagcaacgac tgggcgaagg ctgccgcat catttactgg tcagacgata 5460
aacgggtact cgccgggtgg tgttgaacag attatcgctg gc 5502

<210> 10

<211> 4629

<212> DNA

<213> Acinetobacter baylyi

<220><221> gene

<222> (1)..(4629)

<223> DNA sequence of the catA gene from Acinetobacter baylyi ADP1,
including 410 bases of upstream sequence and two open reading
frames downstream

<400> 10

atctgctcga ccatagtaat gatcacatta tgagctaaat ttacttttta aaatttaa	60
atattatata tatttgaatt ttattgtttt attttaattt ttagcttaga agtttttatt	120
aagatttatt tttaaattag atgtcgaaaa aattagtata ccaaaaaagc atgaaaacat	180
actctcttag gaattggagt cgccatgagt ttcagatata gttgatcagt atggaaggta	240
tagaaacgac tatcgaaata aataagtttg tgggtgtgta agcaaggtaa agctcaaggc	300
tgaggcaaac caagcaaagg ttaattgaac cgatatgcac aacacattca acgatagcgt	360
cgacagataa gtttatcaaa tgatgttttg gcgatttcaa ggagaaagcc atggaagtta	420
aaatatcaa tactcaggat gtgcaagatt ttttacgtgt tgcaagcgga cttagcaag	480
aaggtggcaa tccgcgtgta aagcagatca tccatcgtgt gctttcagat ttatataaag	540
ccattgaaga tttgaatata acttcagatg aatactgggc aggtgtggca tatttaaatc	600
agctaggtgc caatcaagaa gctggtttac tctcgccagg cttgggtttt gaccattacc	660
tcgatatgcg tatggatgcc gaagatgccg cactaggtat tgaaaatgcg acaccacgta	720
ccattgaagg cccgctatac gtggcagggtg cgcctgaatc ggtaggttat gcgcgcatgg	780
atgacggaag tgatccaaat ggtcatacc tgattctaca tggcacgac tttgatgcag	840
atggaaaacc tttaccaat gccaaagttg aaatctggca tgccaatacc aaaggctttt	900
attcacactt cgacccaaca ggcgagcagc aggcgttcaa tatgcgccgt agtattatta	960
ccgatgaaaa cggtcagtat cgcgttcgta ccattttgcc tgcgggttat ggttgccac	1020
cagaaggtec aacgcaacag ttgctgaatc agttgggccg tcatggtaac cgccctgcgc	1080
acattcacta ttttgtttct gccgatggac accgcaaact aactacgcaa attaatgtgg	1140
ctggcgatcc gtacacctat gacgactttg cttatgcaac ccgtgaagc ttggtggttg	1200
atgcagtgga acacaccgat cctgaagcca ttaaggccaa tgatgttgaa ggcccatcgc	1260
ctgaaatggt tttcgatcta aaattgacgc gtttggttga tgggttagat aaccaagttg	1320
ttgatcgtec acgtctagcg gtgtaataca ccaaaatggt tcaaaattat caggcgagt	1380
atcatgatca ctggcctgtt tttatttcag ggaagggtgg agacaattac gtggacaatc	1440
aaatcattca ggaaacgta gataaaattt taagcgtatt gccgaatcag gctgggcaat	1500
tggcacgctt ggttcgtctg atgcagtttg cttgtgacce caccattacc gtcattggta	1560
aataataatca tggtaaaagc cgactactca atgagctgat cgggacagat attttttctg	1620
ttgccgataa acgagagacg attcaactgg ccgaacataa acaagatcag gtgcgttggt	1680

tggatgcacc cggactcgat gcagatgttg cggcagtgga tgatcgtcat gcttttgaag	1740
cagtctggac acaggcagat attcgccctt ttgtgcattc agtccgagaa ggcgaactcg	1800
atgcaactga gcatcatctt ttacaacaac ttattgaaga tgcagacat agccggcgcc	1860
aaaccatact ggtcttgacc cagatagatc agataccgga tcagacaatt ttaaccaga	1920
ttaaaacctc aattgcacag caggtacca aactcgatat ttgggtgtt tcggccactc	1980
gccaccgtca ggttattgaa aatggaaaaa ccttgctgat cgaaaaaagt ggaatcggcg	2040
cgttacgaca tacacttgag caggcacttg ctacaggtcc atctgcacga acgtatgaaa	2100
agaatagatt gctgtctgac ttgcatcadc aacttaagca gttattactc gatcaaaaac	2160
atgtacttca gcaactacaa cagacacagc agcagcaatt gcatgacttt gatacaggac	2220
tcatcaacat actcgataag attcgagtag atcttgagcc cattgtaaat atagatggtc	2280
aagaccaagc actcaatcca gattcatttg ccacgatgtt taaaaataca gcagccaagc	2340
agcaacgtgc caaagtgcag attgcttact cacgtgcctg tatcgagatc aatagccatc	2400
tcatacgtca cgggtgtgtg ggtttaccg cagagcaaca aaccacgata aaaagtattg	2460
atacggatcat tgttgcggtt ttgtgaattt cagtgaattt tcgcgatcag ctacgtgcat	2520
tgttttatac cgataccgaa cgacaacgct tgcaaagaga gtttcgattt tttttgaaa	2580
agtcagcagg ccgaatgatt ttagctgcc aagattgagca gacaatgcgg cagcaagggt	2640
gtattcaaaa tgcaatgatg gcgttgcaac agatggagag tgcagcatga ccagcgcgcg	2700
acacattcaa ttgtttatcg aacacacccg gcagattgcg actgccaag gggatataca	2760
gttggcattg caatcgatgc agcaatggcg cgaagcattt gctacagcat taaaacaaaa	2820
taccttgat ttaacgggct ggtcacgcga gacaaagatc gccaatcaac tcaagcaatt	2880
taaccataag cttacaacgc atgtatcgaa ttgggatacc gaatggcata cttttagtgc	2940
tgctcaatcg gttgcagaag ttttcatga tcgggtgatg ttgcttgat tcggttaagtt	3000
taatgccgga aagagttcat tgtgtaactt actggccgaa tgctttcggt ctcacgaaca	3060
aaccgtgcaa tttttcatg ttcaaatga acagatattt tataccgaat ctcaattacg	3120
cgaaggtgca accgagacga cagcgcaact acaggcgta tgtctgggtg aaaaacttat	3180
tttgctagat acaccaggtt tgcatcttgg tactcagaaa aatgcagcgc tcacacaaaa	3240
atttatcgac agtcagatg gtgtgctgtg gctcagtagc gcaacttcac cgggtcaggt	3300
gcaagagcta gatgcactgg ggcgcgagtt aaagcgtcat aaacctttat ttctgttat	3360
taccgaagc gattttgtcg aagaagatga aattgatggt gagctatgta cagtgtttg	3420

caataaaaat tcagaacaac gtgcgttgca agagtctgat gtattgatgc gtgcgaaaga 3480
 aaaactgcac atatgcaagt ggatgtgagi ttattaaagc cgcccgtgtc cgtttcaact 3540
 caaatggcgc gtgaagcaga tatgaacca caagccatga acgaggctgg ttttgagcga 3600
 ttatttgcag cacttttggc tcttattgag cctgctttgc gctataagca gcgtaaacct 3660
 gccgaagtat tgttgcatTT tttgcaagaa catatcattg aaggTTtaag gttttacctg 3720
 caaccgcatc tagagcaaat acaacaggac ctcaaacagg ctcaagatga tttacgacag 3780
 ctacacaccg atttagccga ggcagtctgg cgtagcgtat tgcctgagct accacaactt 3840

cttgagcaac atgcaagtac acaaaatatt gatgccgtag tgaacagttt gaacgagtgg 3900
 ataaacgtcg cattcgaaca acagcttgca attcagcttg atgcttatgg tttaaatttg 3960
 gattcgttta gcaagatcga aaaaaccgaa aaaatgcagt atgaacgcat tgcgggaatg 4020
 gtggtgcatg atggcttgta cagcactctc acgcagcaga ttcaacaagc tgtcaaagct 4080
 tctacgagtg aattgattga tcagtgtcag gctcaacttg agcagtcaat caaacatgtt 4140
 caaacactcg atgaaacctt catcgattac agcgcagcac tcgatcaact cagccaagcg 4200
 ctacgcattg aataaagagc agtaaatTTT tcagacatat tttattcgat gagtggcctg 4260

atatggtgcg ttgcaaacac ctctgtaca caggcgagaa ttttaggaat gtaattactg 4320
 tggTccatat ttcgaccgc gagtgaattt gggctatagg catcatcatc taaaattgga 4380
 atataaagta gattcttcac cccaatatcc atggcagacg ccggtacgat gcagacgcct 4440
 tcacctgctg ccaccaagcc gagtgccagi tgaatttctc gaatttcggt gagtttggat 4500
 ggtactaggc ctagttcggt aaagagtac tgaataaagg tcgcaaaatt gggcttttga 4560
 gagactgggt acagcagcat cggttcatca ataatttgag agagatgaac ccctgttgct 4620
 gcaaactga 4629

<210> 11
 <211> 311
 <212> PRT
 <213> Acinetobacter baylyi
 <220><221> PEPTIDE
 <222> (1)..(311)
 <223> Protein sequence of CatA (catechol 1,2-dioxygenase) from
 Acinetobacter baylyi ADP1
 <400> 11

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

1 5 10 15

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

260 265 270
Ala Ile Lys Ala Asn Asp Val Glu Gly Pro Phe Ala Glu Met Val Phe
275 280 285
Asp Leu Lys Leu Thr Arg Leu Val Asp Gly Val Asp Asn Gln Val Val
290 295 300
Asp Arg Pro Arg Leu Ala Val
305 310
<210> 12
<211> 1461
<212> DNA
<213> Acinetobacter sp. ADP1
<220><221> gene

<222> (1)..(1461)
<223> DNA sequence of the quiC (3-dehydroshikimate dehydratase)gene
from Acinetobacter sp. ADP1
<400> 12
atgaaattaa cttcttttacg cgtatcttta ttggcgctgg gcttggtaac atcaggtttt 60
gctgcggcag aaacttatac tgtagatcgt taccagatg atagtgaata aggctctttg 120
cgttgggcaa ttgaacaatc taatgcaaat agcgacacag agaatcagat tctgattcag 180
gctgttggtg aggcacctta tgtgatcaag gtggataaac cgttaccacc gattaaatca 240
tctgtaaaaa ttattggtac agaatgggat aaaacgggag aatttattgc gattgatggt 300

tcaaactata tcaagggcga aggcgaataa gcatgtccag gtgcaaatcc aggacaatat 360
ggtaccaatg ttcgtaccat gactttacca ggtttggttc tacaagatgt caatggtgtg 420
accctgaaag gtcttgatgt tcatcgcttc tgtattggtg tactggtaaa tcgttcaagc 480
aataatttga ttcagcataa ccgtatttca aataattacg gtggcgctgg tgtcatgac 540
acgggtgatg atggtaaagg taaccaacg tctaccacca ccaataacaa caaagtattg 600
gataatgtgt ttattgacaa tggcgatggt ctggaactga cgcgtggagc agcattcaac 660
ctgattgcta acaatctgtt tacatcgacc aaagccaatc cagagccgtc tcaaggcatt 720

gaaattcttt gggggaatga caatgcagtg gtgggtaaca aatttgaaaa ctattcagat 780
ggtctacaaa tcaactgggg taaacgtaat tacatcgctt ataacgaatt gaccaataac 840
tctttgggtt tcaatcttac aggtgatgga aacatcttcg atagtaacaa agtgcattggc 900
aatcgatttg gtatcgcaat tcgttctgaa aaagatgcaa atgcacgtat cacacttacc 960

aaaaatcaga tttaggataa tggtaaagat atcaaacgct gtgaggctgg tggttcatgt 1020
 gttccaaacc aacgttttagg tgcaattgta tttaggtgttc ctgcgcttga gcatgaaggt 1080
 tttagtaggct ctctgtggtgg cgggtgtagtc attgaacctg caaaattaca aaaaacatgt 1140

 acacagccaa atcaacaaaa ctgtaatgcc attccgaacc aaggtattca ggcacctaaa 1200
 ctgactgtca gtaaaaaaca acttacagtt gaagttaaag gaacacaaaa ccagcggttac 1260
 aacgtagaat tttttgaaa tcgtaatgca tcttcttccg aagctgagca atatttaggt 1320
 tcaattgttg tagtgacaga tcatcaaggt ctgcaaaag caaactgggc accaaaagtc 1380
 agcatgccat ctgttactgc gaatgtaact gatcacttgg gcgccacttc agagttaagt 1440
 tctgcagtga aaatgagata a 1461

 <210> 13
 <211> 1461
 <212> DNA

 <213> Acinetobacter sp. arc5
 <220><221> gene
 <222> (1)..(1461)
 <223> Codon-optimized DNA sequence of the quiC (3-dehydroshikimate
 dehydratase)gene from Acinetobacter sp. ADP1

 <400> 13

 atgaaactga ccagcctgcg tgtagcctg ctggcactgg gtctggttac cagcggtttt 60
 gcagcagcag aaacctatac cgttgatcgt tatcaggatg atagcgaaaa aggtagcctg 120
 cgttgggcaa ttgaacagag caatgcaaat agcgacaaag aaaaccagat tctgattcag 180
 gcagtggta aagcaccgta tgttatcaaa gttgataaac cgctgcctcc gattaaaagc 240

 agcgttaaaa tcattggcac cgagtgggat aaaaccggtg aatttattgc aattgatggc 300
 agcaactata tcaaaggcga aggtgaaaaa gcatgtccgg gtgcaaatcc gggtcagtat 360
 ggcaccaatg ttcgtaccat gacctgcct ggtctggttc tgcaagatgt taatggtgtt 420
 accctgaaag gtctggatgt tcatcgtttt tgtattggtg ttctggttaa tcgcagcagc 480
 aataacctga ttcagcataa tcgtatcagc aacaattatg gtggtgccgg tgttatgatt 540
 accggtgatg atggtaaagg taatccgacc agcaccacca ccaataataa caaagttctg 600
 gataacgtgt tcatcgataa tggatgatgt ctggaactga cccgtggtgc agcatttaat 660

 ctgattgcaa ataacctgtt taccagcaca aaagccaatc cggaaccgag ccagggtatt 720
 gaaattctgt ggggtaatga taatgccgtg gtgggtaaca aattcgaaaa ctattcagat 780

ggcctgcaaa tcaattgggg taaacgtaac tatatgcct ataacgaact gaccaataac 840
 agcctgggtt tcaatctgac aggtgatggt aacattttcg acagcaataa agtgcattggt 900
 aaccgtattg gtattgcat tctagttaa aaagatgcca atgcacgtat taccctgacc 960
 aaaaatcaga ttgggataa cggcaaagat atcaaactgt gtgaagccgg tggtagctgt 1020
 gttccgaatc agcgtctggg tgcaattgtt ttgggtgttc cggcactgga acatgaaggt 1080

 ttgttggtgta gccgtggcgg tgggtgtgtt attgaaccgg caaaactgca aaaaacctgc 1140
 acccagccga acccagcaga ttgtaatgca attcctaatac agggatttca ggcaccgaaa 1200
 ctgacagtta gcaaaaaaca gctgaccgtt gaagttaaag gcaccctaa tcagcgttat 1260
 aatgtggaat tttttggcaa tctgaatgcc agcagcagcg aagcagaaca gtatctgggt 1320
 agcattgttg ttgttaccga tcatcagggt ctggcaaaag caaattgggc tccgaaagtt 1380
 agcatgccga gcgttaccgc aaatgtgaca gatcatctgg gtgcgaccag cgaactgagc 1440
 agcgcagtta aaatgcgtta a 1461

<210> 14
 <211> 486
 <212> PRT
 <213> Acinetobacter sp. arc5
 <220><221> PEPTIDE
 <222> (1)..(486)
 <223> Protein sequence of QuiC (3-dehydroshikimate dehydrogenase from
 Acinetobacter sp. ADP1
 <400> 14

Met Lys Leu Thr Ser Leu Arg Val Ser Leu Leu Ala Leu Gly Leu Val
 1 5 10 15
 Thr Ser Gly Phe Ala Ala Ala Glu Thr Tyr Thr Val Asp Arg Tyr Gln
 20 25 30
 Asp Asp Ser Glu Lys Gly Ser Leu Arg Trp Ala Ile Glu Gln Ser Asn
 35 40 45
 Ala Asn Ser Ala Gln Glu Asn Gln Ile Leu Ile Gln Ala Val Gly Lys
 50 55 60
 Ala Pro Tyr Val Ile Lys Val Asp Lys Pro Leu Pro Pro Ile Lys Ser
 65 70 75 80
 Ser Val Lys Ile Ile Gly Thr Glu Trp Asp Lys Thr Gly Glu Phe Ile

85	90	95	
Ala Ile Asp Gly Ser Asn Tyr Ile Lys Gly Glu Gly Glu Lys Ala Cys			
100	105	110	
Pro Gly Ala Asn Pro Gly Gln Tyr Gly Thr Asn Val Arg Thr Met Thr			
115	120	125	
Leu Pro Gly Leu Val Leu Gln Asp Val Asn Gly Val Thr Leu Lys Gly			
130	135	140	
Leu Asp Val His Arg Phe Cys Ile Gly Val Leu Val Asn Arg Ser Ser			
145	150	155	160
Asn Asn Leu Ile Gln His Asn Arg Ile Ser Asn Asn Tyr Gly Gly Ala			
165	170	175	
Gly Val Met Ile Thr Gly Asp Asp Gly Lys Gly Asn Pro Thr Ser Thr			
180	185	190	
Thr Thr Asn Asn Asn Lys Val Leu Asp Asn Val Phe Ile Asp Asn Gly			
195	200	205	
Asp Gly Leu Glu Leu Thr Arg Gly Ala Ala Phe Asn Leu Ile Ala Asn			
210	215	220	
Asn Leu Phe Thr Ser Thr Lys Ala Asn Pro Glu Pro Ser Gln Gly Ile			
225	230	235	240
Glu Ile Leu Trp Gly Asn Asp Asn Ala Val Val Gly Asn Lys Phe Glu			
245	250	255	
Asn Tyr Ser Asp Gly Leu Gln Ile Asn Trp Gly Lys Arg Asn Tyr Ile			
260	265	270	
Ala Tyr Asn Glu Leu Thr Asn Asn Ser Leu Gly Phe Asn Leu Thr Gly			
275	280	285	
Asp Gly Asn Ile Phe Asp Ser Asn Lys Val His Gly Asn Arg Ile Gly			
290	295	300	
Ile Ala Ile Arg Ser Glu Lys Asp Ala Asn Ala Arg Ile Thr Leu Thr			
305	310	315	320
Lys Asn Gln Ile Trp Asp Asn Gly Lys Asp Ile Lys Arg Cys Glu Ala			
325	330	335	

Gly Gly Ser Cys Val Pro Asn Gln Arg Leu Gly Ala Ile Val Phe Gly
 340 345 350
 Val Pro Ala Leu Glu His Glu Gly Phe Val Gly Ser Arg Gly Gly Gly
 355 360 365
 Val Val Ile Glu Pro Ala Lys Leu Gln Lys Thr Cys Thr Gln Pro Asn
 370 375 380
 Gln Gln Asn Cys Asn Ala Ile Pro Asn Gln Gly Ile Gln Ala Pro Lys
 385 390 395 400

Leu Thr Val Ser Lys Lys Gln Leu Thr Val Glu Val Lys Gly Thr Pro
 405 410 415
 Asn Gln Arg Tyr Asn Val Glu Phe Phe Gly Asn Arg Asn Ala Ser Ser
 420 425 430
 Ser Glu Ala Glu Gln Tyr Leu Gly Ser Ile Val Val Val Thr Asp His
 435 440 445
 Gln Gly Leu Ala Lys Ala Asn Trp Ala Pro Lys Val Ser Met Pro Ser
 450 455 460
 Val Thr Ala Asn Val Thr Asp His Leu Gly Ala Thr Ser Glu Leu Ser

465 470 475 480
 Ser Ala Val Lys Met Arg

485

<210> 15

<211> 9462

<212> DNA

<213> Escherichia coli

<220><221> misc_feature

<222> (1)..(9462)

<223> DNA sequence of the plasmid pAC21

<400> 15

gttgacagta agacgggtaa gcctgttgat gataccgctg ccttactggg tgcattagcc 60

agtctgaatg acctgtcacg ggataatccg aagtgggtcag actggaaaat cagagggcag 120

gaactgctga acagcaaaaa gtcagatagc accacatagc agacccgcca taaaacgccc 180

tgagaagccc gtgacgggct tttcttgtat tatgggtagt ttccttgcat gaatccataa 240

aaggcgctg tagtgccatt taccgccatt cactgccaga gccgtgagcg cagcgaactg	300
aatgtcacga aaaagacagc gactcaggtg cctgatggtc ggagacaaaa ggaatattca	360
gcgatttgcc cgagcttgcc aggggtgctac ttaagccttt aggggtttta ggtctgtttt	420
gtagaggagc aaacagcggt tgcgacatcc ttttgtaata ctgcggaact gactaaagta	480
gtgagttata cacagggctg ggatctattc tttttatctt tttttattct tttttattc	540
tataaattat aaccacttga atataaaca aaaaaacaca caaaggctca gcggaattta	600
cagagggtct agcagaattt acaagttttc cagcaaaggc ctagcagaat ttacagatac	660
ccacaactca aaggaaaagg actagtaatt atcattgact agcccatctc aattggtata	720
gtgattaaaa tcacctagac caattgagat gtatgtctga attagtgtt ttcaaagcaa	780
atgaactagc gattagtgc tatgacttaa cggagcatga aaccaagcta attttatgct	840
gtgtggcact actcaacccc acgattgaaa accctacaag gaaagaacgg acggtatcgt	900
tcacttataa ccaatacgtc cagatgatga acatcagtag ggaaaatgct tatgggtgat	960
tagctaaagc aaccagagag ctgatgacga gaactgtgga aatcaggaat cctttggtta	1020
aaggctttga gattttccag tggacaaaact atgccaagtt ctcaagcgaa aaattagaat	1080
tagtttttag tgaagagata ttgccttatt ttttcagtt aaaaaattc ataaaatata	1140
atctggaaca tgtaagtct tttgaaaaca aatactctat gaggatttat gagtgttat	1200
taaaagaact aacacaaaag aaaactcaca aggcaaatat agagattagc cttgatgaat	1260
ttaagttcat gtaaatgctt gaaaataact accatgagtt taaaaggctt aaccaatggg	1320
ttttgaaacc aataagtaaa gatttaaca cttacagcaa tatgaaattg gtggttgata	1380
agcgaggccg cccgactgat acgttgattt tccaagtga actagataga caaatggatc	1440
tcgtaaccga acttgagaac aaccagataa aaatgaatgg tgacaaaata ccaacaacca	1500
ttacatcaga ttctactca cgtaacggac taagaaaaac actacacgat gctttaactg	1560
caaaaattca gctcaccagt tttgaggcaa aatttttgag tgacatgcaa agtaagcatg	1620
atctcaatgg ttcgttctca tggctcacgc aaaaacaacg aaccacacta gagaacatac	1680
tggctaaata cggaaggatc tgaggttctt atggctcttg tatctatcag tgaagcatca	1740
agactaaca acaaaagtag aacaactgtt caccgttaga tatcaaaggg aaaactgtcc	1800
atatgcacag atgaaaacgg tgtaaaaaag atagatacat cagagctttt acgagttttt	1860
ggtgcattta aagctgttca ccatgaacag atcgacaatg taacagatga acagcatgta	1920
acacctaata gaacaggiga aaccagtaaa acaaagcaac tagaacatga aattgaacac	1980
ctgagacaac ttgttacagc tcaacagtca cacatagaca gcctgaaaca ggcatgctg	2040
cttatcgaat caaagctgcc gacaacacgg gagccagtga cgcctcccg ggggaaaaaa	2100

tcattggcaat tctggaagaa atagcgcttt cagccggcaa acctgaagcc ggatctgcga	2160
ttctgataac aaactagcaa caccagaaca gcccgtttgc gggcagcaaa acccgttatg	2220
cttgtaaacc gttttgtgaa aaaattttta aaataaaaaa ggggacctct aggggtcccca	2280
attaattagt aatataatct attaaaggtc attcaaaagg tcatccaccg gatcaattcc	2340
cctgctcgcg caggctgggt gccaaagctct cgggtaacat caaggcccga tccttgagc	2400
ccttgccctc ccgcacgatg atcgtgccgt gatcgaaatc cagatccttg acccgcagtt	2460
gcaaaccctc acgatccgc atgcccgctt catacagaag ctgggcgaac aaacgatgct	2520
cgcttccag aaaaccgagg atgcgaacca ctcatccgg ggtcagcacc accggaagc	2580
gccgcgacgg ccgaggtctt ccgatctcct gaagccaggg cagatccgtg cacagcacct	2640
tgccgtagaa gaacagcaag gccgccaatg cctgacgatg cgtggagacc gaaaccttgc	2700
gctcgttcgc cagccaggac agaaatgcct cgacttcgct gctgccaag gttgccgggt	2760
gacgcacacc gtggaacgg atgaaggcac gaaccagtg gacataagcc tgttcggttc	2820
gtaagctgta atgcaagtag cgtatgcgct cagcgaactg gtccagaacc ttgaccgaac	2880
gcagcggtag taacggcgca gtggcggttt tcatggcttg ttatgactgt tttttgggg	2940
tacagtctat gcctcgga tccaagcagc aagcgctta cgccgtgggt cgatgtttga	3000
tgttatggag cagcaacgat gttacgcagc agggcagtcg ccctaaaaca aagttaaca	3060
tcatgaggga agcgggtgac gccgaagtat cgactcaact atcagaggta gttggcgta	3120
tcgagcgcca tctgaaccg acgttgctgg ccgtacattt gtacggctcc gcagtggatg	3180
gcggcctgaa gccacacagt gatattgatt tgcgtgttac ggtgaccgta aggcttgatg	3240
aaacaacgcg gcgagctttg atcaacgacc ttttgaaac ttcggcttcc cctggagaga	3300
gcgagattct ccgcgtgta gaagtcacca ttgttgtgca cgacgacatc attccgtggc	3360
gttatccagc taagcgcaa ctgcaatttg gagaatggca gcgcaatgac attcttgag	3420
glatcttga gccagccag atcgacattg atctggctat cttgctgaca aaagcaagag	3480
aacatagcgt tgccttgta ggtccagcgg cggaggaact ctttgatccg gttcctgaac	3540
aggatctatt tgaggcgcta aatgaaacct taacgctatg gaactcgccg cccgactggg	3600
ctggcgatga gcgaaatgta gtgcttacgt tgtccgcat ttggtacagc gcagtaaccg	3660
gcaaaatcgc gccgaaggat gtcgtgccg actgggcaat ggagcgctg ccggcccagt	3720
atcagcccgt catacttgaa gctagacagg cttatcttgg acaagaagaa gatcgcttgg	3780
cctcgcgcg agatcagttg gaagaatttg tccactacgt gaaaggcgag atcaccaagg	3840
tagtcggcaa ataattgcta acaattcgtt caagccgacg ccgttcgcg gcgcggctta	3900

actcaagcgt tagatgcact aagcacataa ttgctcacag ccaaactatc aggtcaagtc	3960
tgcttttatt atttttaagc gtgcataata agccctacac aaattgggag atatatcatg	4020
aaagctggc tttttcttgt tatcgcaata gttggcgaag taatcgcaac atccgcatta	4080
aaatctagcg agggctttac taagctgac cggtggatga ccttttgaat gacctttaat	4140
agatttatatt actaattaat tggggaccct agaggctccc ttttttattt taaaaatttt	4200
ttcacaaaac ggtttacaag catacgttgg ccgattcatt aatgcagctg gcacgacagg	4260
tttcccgact ggaaagcggg cagtgagcgc aacgaatta atgtgagta gctcactcat	4320
taggcacccc aggcctttaca ctttatgctt ccggctcgta tgttgtgtgg aattgtgagc	4380
ggataacaat ttcacacagg aaacagctat gaccatgatt acgccaagct tgcatgcctg	4440
caggtcgact ctagaggatc cccccgcgc cgacagagt aataggtttt acttaatagc	4500
tcttctgtgc ctttccaggc agtgatccgc attccgttct catggcgagg caacatttcg	4560
ggatggaaga taatgttctt tgctacagga aaatcaaca tatgctgcacc agatgccact	4620
ggcagccgcc cgtgcgcgt tactaactct ataaatgcag ggatctcatc aatgacaaca	4680
tcctcgggac tgtttctgc cagtcccatg atgatggcga catccgtggc atggcctttg	4740
cccgtcagtg acaacgaccc gtacagatcg accacaatat ggctcgtcgc ggttaataag	4800
ccgtacttt ccagccgac aataaaactt tttccggcat tcattggccc cacggtatgc	4860
gaactggagg gaccaatccc aattttgaaa atatcgaatg cactaatcat gtgacggaag	4920
atcacttcgc agaataaata aatcctgttg tcctgttga taccgggaag ccctgggcca	4980
acttttggcg aaaaatgagac gttgatcggc acgtaagagg ttccaacttt caccataatg	5040
aaataagatc actaccgggc gtattttttg agttatcgag attttcagga gctaaggaag	5100
ctaaaatgga gaaaaaatc actggatata ccaccgttga tataatccaa tggcatcgta	5160
aagaacattt tgaggcattt cagtcagttg ctcaatgtac ctataaccag accgttcagc	5220
tgatattac ggccttttta aagaccgtaa agaaaaataa gcacaagttt tatccggcct	5280
ttattcacat ttttgcgcgc ctgatgaatg ctcatccgga attccgtatg gcaatgaaag	5340
acggtgagct ggigatatgg gatagtgttc acccttgta caccgttttc catgagcaaa	5400
ctgaaacgtt ttcacgctc tggagtgaat accacgacga tttccggcag tttctacaca	5460
tatatcgca agatgtggcg tgttacggtg aaaacctggc ctatttcct aaagggttta	5520
ttgagaatat gtttttcgtc tcagccaatc cctgggtgag tttcaccagt ttgtatttaa	5580
acgtggccaa tatggacaac ttcttcgcc ccgttttcac catgggcaaa tattatacgc	5640

aaggcgacaa ggtgctgatg ccgctggcga ttcaggttca tcatgccgtt tgtgatggct	5700
tccatgtcgg cagaatgctt aatgaattac aacagtactg cgatgagtgg cagggcgggg	5760
cgtaattttt ttaaggcagt tattgggtgcc cttaaacgcc tgggtgctacg cctgaataag	5820
tgataataag cggaatgaatg gcagaaattc gaaagcaaat tcgacccggt cgtcggttca	5880
gggcagggtc gttaaatagc cgcttatgtc tattgctggt ttantcggta cccggggatc	5940
gcggccgcgg accggatccc atcacatata cctgccgttc actattatit agtgaaatga	6000
gatattatga tattttctga attgtgattia aaaaggcaac tttatgccca tgcaacagaa	6060
actataaaaa atacagagaa tgaaaagaaa cagatagatt ttttagttct ttaggcccggt	6120
agttctgcaa tccttttatg attttctatc aaacaaaaga ggaaaataga ccagttgcaa	6180
tccaaacgag agtctaatag aatgaggctc aaaagtaaat cgcgcgggtt tgttactgat	6240
aaagcaggca agacctaaaa tgtgtaaagg gcaaaagtga tactttggcg tcacccctta	6300
catattttag gtcttttttt attgtgcgta actaacttgc catcttcaaa caggagggtc	6360
ggaagaagca gaccgctaac acagtacata aaaaaggaga catgaacgat gaacatcaaa	6420
aagtttgcaa aacaagcaac agtattaacc ttactaccg cactgctggc aggaggcgca	6480
actcaagcgt ttgcgaaaga aacgaaccaa aagccatata aggaacata cggcatttcc	6540
catattacac gccatgatat gctgcaaatc cctgaacagc aaaaaaatga aaaatatcaa	6600
gttcctgaat tcgattcgtc cacaattaaa aatatctctt ctgcaaaagg cctggacgtt	6660
tgggacagct ggccattaca aaacgtgac ggcaactgtc caaactatca cggctaccac	6720
atcgtctttg cattagccgg agatcctaaa aatgcggatg acacatcgat ttacatgttc	6780
tatcaaaaag tcggcgaaac ttctattgac agctggaaaa acgctggccg cgtctttaaa	6840
gacagcgaca aattcgatgc aatgattct atcctaaaag accaaacaca agaatggtca	6900
ggttcagcca catttacatc tgacggaaaa atccgtttat tctacactga tttctccggt	6960
aaacattacg gcaaacaaac actgacaact gcacaagtta acgtatcagc atcagacagc	7020
tctttgaaca tcaacgggtg agaggattat aatcaatct ttgacggtga cggaaaaacg	7080
tatcaaaatg tacagcagtt catcgatgaa ggcaactaca gctcaggcga caaccatacg	7140
ctgagagatc ctactacgt agaagataaa ggccacaaat acttagtatt tgaagcaaac	7200
actggaactg aagatggcta ccaaggcgaa gaatctttat ttaacaaagc atactatggc	7260
aaaagcacat cattcttccg tcaagaaagt caaaaacttc tgcaaagcga taaaaaacgc	7320
acggctgagt tagcaaacgg cgctctcggt atgattgagc taaacgatga ttacacactg	7380
aaaaaagtga tgaaccgct gattgcatct aacacagtaa cagatgaaat tgaacgcgcg	7440
aacgtcttta aaatgaacgg caaatgggtac ctgttctactg actccgcgg atcaaaaatg	7500

acgattgacg gcattacgtc taacgatatt tacatgcttg gttatgtttc taattcttta	7560
actggcccat acaagccgct gaacaaaact ggccttgtgt taaaaatgga tcttgatcct	7620
aacgatgtaa cctttactta ctcacacttc gctgtacctc aagcgaaagg aaacaatgtc	7680
gtgattacaa gctatatgac aaacagagga ttctacgcag acaacaatc aacgtttgcg	7740
ccgagcttcc tgctgaacat caaaggcaag aaaacatctg ttgtcaaaga cagcatcctt	7800
gaacaaggac aattaacagt taacaataa aaacgcaaaa gaaaatgcca atatcctatt	7860
ggcattttct tttatttctt ccatttaa at ggatgcatgc gctagcggag tgtatactgg	7920
cttactatgt tggcactgat gaggtgtca gtgaagtgtc tcagcctcgt gagcgggacg	7980
gtcgtaaagt cgttccgtc cacttcactg aacggcaatc cgagggtgtg gatccaatta	8040
aggccacgct gtcatttaaa ttccgttttt ccagttcaaa tgcaattgcc ttcaatgcac	8100
cttcgtagct gtggtgagcc agcgggtgtg gctctcccc atttacggat aagaatgcat	8160
tttccgagtt aataccgtcg gcaatacctg acattaatac ttcacagtcg ctggcatcga	8220
gtacggaaaa cttaatcgaa gacgaaccac agttaataac caaaacaacc ggaaattcat	8280
tcactctttt tctcatcctg agttacggat taaacagtt tgtatacgat gttcaggatg	8340
gtcagcagac caatcacggt aacaacacg ttatccagac gaccacggta tttcgccaga	8400
gacggcgctt tacggatggc atacatcggc aacaggcaca gcagggatgc gataatcgg	8460
gcgccatgg cttaatcag gtcgaggatg ttccgggttg cgtaggcaac aaccaggtg	8520
gagcccatga tgaagatcat gctgagagta ttcagtttac ccagcgacac tttggttttg	8580
tcacctttat aaccgaactt cagaatcaga ccattcaagc cttccagcgt cccagatag	8640
tgaccgaaga aagatttgaa gatagccacg agtgcgatga tggaagccgc atattccagt	8700
gtaatcgca acgttgtttt ggtaccggtc atggacgcaa agtggttagc cagataagaa	8760
agcactggaa tattctgcgc tttggcttcc gccatgttgg ccggagacag agtaaacagg	8820
cagctaaagg caaagaacat caccactgca accatcagca tgctggcacg agaaatgatt	8880
tgggaacatt tacgttcggt gaagtcgca ccgaagtctt tctcactc ttcacgttta	8940
gaaaccacga aggaagagac gattggcgag aagttaaagg agaaaacat gatggaaatc	9000
cccagccaga cagtgatcag gataccgtca tgaccgtta acgacagcga accgaggtca	9060
acctggctga taactgcaga gtccagtaa gggatcagcg acaaaagaat cagcaccagg	9120
ctggcgataa acggccatc caggtagctc agggatccga gctcgaattc actggccgtc	9180
gttttacaac gtcgtgactg ggaaaacct ggctttacc aacttaatcg ccttgcagca	9240
catccccctt tcgccagctg gcgtaatagc gaagaggccc gcaccgatcg ccttcccaa	9300
cagttgcgca gcctgaatgg cgaatggcgc ctgatgcggt attttctcct tacgcatctg	9360

tgcggtatatt cacaccgcat atgggtgact ctacgtacaa tctgtcttga tgccgcatag 9420

ttaagccagc cccgacaccc gccaacaccc gctgacgaat tc 9462

<210> 16

<211> 9430

<212> DNA

<213> Escherichia coli

<220><221> misc_feature

<222> (1)..(9430)

<223> DNA sequence of the plasmid pAC19

<400> 16

gttgacagta agacgggtaa gcctgttgat gataccgtg ccttactggg tgcattagcc 60

agtctgaatg acctgtcacg ggataatccg aagtggtcag actggaaaat cagagggcag 120

gaactgctga acagcaaaaa gtcagatagc accacatagc agaccgcga taaaacgccc 180

tgagaagccc gtgacgggct tttcttgtat tatgggtagt ttccttgcat gaatccataa 240

aaggcgctg tagtgccatt taccgccatt cactgccaga gccgtgagcg cagcgaactg 300

aatgtcacga aaaagacagc gactcaggtg cctgatggtc ggagacaaaa ggaatattca 360

gcgatttgcc cgagcttgcc aggggtgctac ttaagccttt aggggtttta ggtctgtttt 420

gtagaggagc aaacagcgtt tgcgacatcc ttttgtaata ctgcggaact gactaaagta 480

gtgagttata cacagggctg ggatctattc tttttatctt tttttattct ttctttattc 540

tataaattat aaccacttga atataaaca aaaaaacaca caaaggtcta gcggaattta 600

cagaggtctc agcagaattt acaagttttc cagcaaaggt ctacgagaat ttacagatac 660

ccacaactca aaggaaaagg actagtaatt atcattgact agcccatctc aatttggtata 720

gtgattaaaa tcacctagac caattgagat gtatgtctga attagtgtt ttcaaagcaa 780

atgaactagc gattagtgcg tatgacttaa cggagcatga aaccaagcta attttatgct 840

gtgtggcact actcaacccc acgattgaaa accctacaag gaaagaacgg acggtatcgt 900

tcacttataa ccaatacgtc cagatgatga acatcagtag ggaaaatgct tatgggtgat 960

tagctaaagc aaccagagag ctgatgacga gaactgtgga aatcaggaat cctttggtta 1020

aaggctttga gattttccag tggacaaaact atgccaagtt ctcaagcgaa aaattagaat 1080

tagtttttag tgaagagata ttgccttatt ttttcagtt aaaaaattc ataaaatata 1140

atctggaaca tgtaagtct tttgaaaaca aatactctat gaggatttat gagtggttat 1200

taaaagaact aacacaaaag aaaactcaca aggcaaatat agagattagc cttgatgaat 1260

ttaagttcat gttaatgctt gaaaataact accatgagtt taaaaggctt aaccaatggg	1320
ttttgaaacc aataagtaaa gatttaaaca cttacagcaa tatgaaattg gtggttgata	1380
agcgaggccg cccgactgat acgttgattt tccaagtga actagataga caaatggatc	1440
tcgtaaccga acttgagaac aaccagataa aatgaatgg tgacaaaata ccaacaacca	1500
ttacatcaga ttctaccta cgtaacggac taagaaaaac actacacgat gctttaactg	1560
caaaaattca gtcaccagt tttagggcaa aatttttgag tgacatgcaa agtaagcatg	1620
atctcaatgg ttcgtttctca tggctcacgc aaaaacaacg aaccacacta gagaacatac	1680
tggctaaata cggaaggatc tgaggttctt atggctcttg tatctatcag tgaagcatca	1740
agactaacia acaaaagtag aacaactgtt caccgttaga tatcaaaggg aaaactgtcc	1800
atatgcacag atgaaaacgg tgtaaaaaag atagatacat cagagctttt acgagttttt	1860
ggtgcattta aagctgttca ccatgaacag atcgacaatg taacagatga acagcatgta	1920
acacctaata gaacaggatga aaccagtaaa acaaagcaac tagaacatga aattgaacac	1980
ctgagacaac ttgttacagc tcaacagtca cacatagaca gcctgaaaca ggcgatgctg	2040
cttatcgaat caaagctgcc gacaacacgg gagccagtga cgctcccggt ggggaaaaaa	2100
tcatggcaat tciggaagaa atagcgcttt cagccggcaa acctgaagcc ggatctcgca	2160
ttctgataac aaactagcaa caccagaaca gcccgtttgc gggcagcaaa acccgttatg	2220
cttgtaaacc gttttgtgaa aaaattttta aaataaaaaa ggggacctct aggggtccca	2280
attaattagt aatataatct attaaaggtc attcaaaagg tcatccaccg gatcaattcc	2340
cctgctcgcg caggctgggt gccaaagctct cgggtaacat caaggcccga tccttggagc	2400
ccttgcctc ccgcacgatg atcgtgccgt gatcgaaac cagatccttg acccgcagtt	2460
gcaaaccctc actgatccgc atgcccgctt catacagaag ctgggcgaac aaacgatgct	2520
cgcttccag aaaaccgagg atgcgaacca ctctatccgg ggtcagcacc accggcaagc	2580
gccgcgacgg ccgaggtctt ccgatctcct gaagccaggg cagatccgtg cacagcacct	2640
tgccgtagaa gaacagcaag gccccaatg cctgacgatg cgtggagacc gaaaccttgc	2700
gctcgttcgc cagccaggac agaaatgcct cgacttcgct gctgccaag gttgccgggt	2760
gacgcacacc gtggaaacgg atgaaggcac gaaccagtg gacataagcc tgttcggttc	2820
gtaagctgta atgcaagtag cgtatgcgct cacgcaactg gtccagaacc ttgaccgaac	2880
gcagcggtag taacggcgca gtggcggttt tcatggcttg ttatgactgt tttttgggg	2940
tacagtctat gcctcgggca tccaagcagc aagcgcgtta cgccgtgggt cgatgtttga	3000
tgttatggag cagcaacgat gttacgcagc agggcagtcg ccctaaaaca aagttaaaca	3060
tcatgaggga agcgggtgatc gccgaagtat cgactcaact atcagaggta gttggcgtca	3120

tcgagcgcca tctcgaaccg acgttgctgg ccgtacattt gtacggctcc gcagtggatg	3180
gcggcctgaa gccacacagt gatattgatt tgctggttac ggtgaccgta aggcttgatg	3240
aaacaacgcg gcgagctttg atcaacgacc ttttggaac ttcggcttcc cctggagaga	3300
gcgagattct ccgcgctgta gaagtcacca ttgttggtga cgacgacatc attccgtggc	3360
gttatccagc taagcgcgaa ctgcaatttg gagaatggca gcgcaatgac attcttgtag	3420
gtatcttcca gccagccacg atcgacattg atctggctat cttgtcgaca aaagcaagag	3480
aacatagcgt tgccttggtg ggtccagcgg cggaggaact ctttgatccg gttcctgaac	3540
aggatctatt tgaggcgcta aatgaaacct taacgctatg gaactcgccg cccgactggg	3600
ctggcgatga gcgaaatgta gtgcttacgt tgtcccgcac ttggtacagc gcagtaaccg	3660
gcaaaatcgc gccgaaggat gtcgctgccg actgggcaat ggagcgcctg ccggcccagt	3720
atcagcccgat catacttgaa gctagacagg cttatcttgg acaagaagaa gatcgcttgg	3780
cctcgcgcgc agatcagttg gaagaatttg tccactacgt gaaaggcgag atcaccaagg	3840
tagtcggcaa ataattgcta acaattcgtt caagccgacg ccgcttcgcg gcgcggctta	3900
actcaagcgt tagatgcact aagcacataa ttgctcacag ccaaactatc aggtcaagtc	3960
tgcttttatt atttttaagc gtgcataata agccctacac aaattgggag atatatcatg	4020
aaagctggc tttttcttgt tatcgcaata gttggcgaag taatcgcaac atccgcatta	4080
aaatctagcg agggctttac taagctgac cggtggatga ctttttgaat gacctttaat	4140
agatttatatt actaattaat tggggaccct agagggtcccc ttttttattt taaaaatttt	4200
ttcacaaaac gggttacaaag catacgttgg ccgattcatt aatgcagctg gcacgacagg	4260
tttcccgact ggaaagcggg cagttagcgc aacgcaatta atgtgagtta gtcactcat	4320
taggcacccc aggctttaca ctttatgctt ccggctcgta tgttgtgtgg aattgtgagc	4380
ggataacaat ttcacacagg aaacagctat gaccatgatt acgccaagct tgcattgcctg	4440
caggtcgact ctagaggatc cccccgccg ccgacagagt aataggtttt acttaatagc	4500
tcttcctgtc cttccagge agtgatccgc attccgttct catggcgagg caacatttcg	4560
ggatggaaga taatgttctt tgctacagga aaatcaacaa tatgcgcacc agatgccact	4620
ggcagccgcc cgtcgcgct tactaactct ataaatgcag ggatctcatc aatgacaaca	4680
tcttgccgac tgtttctgc cagtcccatg atgatggcga catccgtggc atggcctttg	4740
cccgtcagtg acaacgacce gtacagatcg accacaatat ggctcgtcgc ggttaataag	4800
ccgctacttt ccagccgac aataaaactt tttccggcat tcattggccc cacggtatgc	4860

gaactggagg gaccaatccc aatTTtgaaa atatcgaaTg cactaatcat atccacaccc	4920
tcggaTtgcc gTtcagTgaa gTggagcgga acgaccttac gaccgtcccg cTcacgaggc	4980
TTtacgcact acgtactgcg atggcttcaa TTtccagcgg gagggcggaT ccactaatac	5040
aaaatatatc aaaagTtaat aataatatTa TTcttactTa agactTTTTt gTcttcatTT	5100
TTtagtaaaa aatataaaaa aggccacctc ccgattTTat cggaaggcag cctcttaaat	5160
tcaTtTcata atattaaaa atattattca acttcagaat attgtTggc ataggcagct	5220
gccgcaccca acagtccagg ctgcggataa gTaatcaact taaccggaat cTtggaatg	5280
acgcgttcaa agcgtctctt tgaacaaaag cgctgacgga aaccagattc tggcaaatgg	5340
gaagcgatac gaagaccgac accaccgcca ataacaacac tggTtgacc ctgtgccaaa	5400
gcaagatcac cagcgatagc gccaggctc aagcagaagc gatccaaagc ggcttcagca	5460
aggTtgtctt taccttccaa agccatctgc cataattTaa tatcatccag caagctgaac	5520
ggaacgcctt caatggcagc cagtgcTtcg tagatattac caagaccgg gccagaaata	5580
atgcgttTca tagaaacgcg gcggaacgt TcacgTaaac gtgccagaat TTtgtctTca	5640
agtctgtcaa gcggagcaaa gtcgatatga ccgccttcag TTtcgatgac gaaataacgg	5700
cttTcagTcc gcaacagatg ggcaacaccc aagcccgTtc ccggaccaag aatagTgata	5760
acaccatcgc taggaagcgc Ttcatcagga ccacaaatat gatccagata agaagaatcc	5820
atatgcgcaa ccgcgtgggc aaccgcgccg aagTcattga Tcagaacatg cgtatcgatg	5880
TccagctTTt catTcagagt agctggTctt aatacccaag ggTtatTggT aagTTTtaa	5940
actTcaccat gaaccgggcc agcccatgca atagctgcgg cacgtggcag aggacgaccc	6000
agTTTTTcac cgaacgTtc ccaagTaaTc Tgcaagctag catgtTctgc cgtTTTaaaa	6060
gtTgtTctt ctccaagaga aagaaccgca ccattgtTta cTtccgcaat agagaaacgc	6120
gcagTcgTtc caccgatgTc aatcgcaaca attTccataa taattcctt ctgaaatcag	6180
aaggctaccc aacagTtaa ataagTccgc ccgctTtata ccatcgTtgt aaacaaaaag	6240
tataattggT taagacttat ctaaaaaaga caaaaggatt cagccaaagc aagTTtaact	6300
actTctggga gcgccacatc TctcgaTtt catccaggct ccgacctTtg gTtTccgga	6360
cgaagcgagc aacaatcaag ccacctaaTg tacttaatgc Tgcgaaaacg agataggaga	6420
aaccgtggTt gaaagTctga Ttcaatgctg gagaaccatc ggcaacctTa aacaggaagt	6480
taaccaagat attagTaaTc cattgtccgg taacagcgat aggcatagct gcgcctTga	6540
Tggaactcgg gaacattTct gacagaacaa ccagcagac agggcccat gacataccaa	6600
agactgcaat ataaagaagc acagaagcca aaggcaaaac accaccgact TTgaaccaga	6660
aacagcagcc taaaacagcc atcattgcag ccataccgag agcaccccaa ataagcagag	6720

gtttacggcc gaagcgggtca acaacacggg aagcaatcat ggtgaagatg aagttcacia	6780
caccgataga gatggtctgc aataatgccg tatcagctcc aaaacctaaa ttctggaaca	6840
tctgcggtgc ataatacagc acggcggttaa taccgactaa ctgctggaag gcagcaacgg	6900
atacaccggc aaaaacaacg gtgatacaca aagcaacaca acctgcgctg cttttgtcca	6960
tggttttatt aaagccagct ttaatctttt gaatcgtcag attaggatcg gcttgcggtt	7020
ccagacgagc aaggattttg ctagcctcgg aatgacgtcc cttcatcacc aaccaatgcg	7080
gcgtatccgg tgcggttaac agcagcaata agaaggcaat accgatcagg ctttctgaag	7140
ccggagacca gcaccaacca ctggcattaa cccaatcgat agaaccgaaa tgagccagta	7200
accaggtaaa gatataaccg gttaaagcac ccgtcacat ggccatctgc tgaccagaaa	7260
ccatctgacc acgtttgtct ggccgagcaa ttctagcaat ataggttggg gtcaaggttg	7320
aaacgacacc gatacctaaa ccggcaagaa accggaaaaa gcaaaaaatt tgtaaagccg	7380
aaccaccggt tccaaataat ttttcggtta acgcagcacc aaaaccggcg gcgacgaaac	7440
aatggaact catcaacaat ccgccgcgac gaccgaagcg aataccaatc cagccagaca	7500
gcaaagaacc ggtaacacaa ccgacaaaaa cagcaacaac gaccatcca gaaaggaag	7560
ccgcagccgt agcagacagg tgacgagggg caataaaatg gatatcaacc ggtgtaccga	7620
ttgcagcgat aaccgctgaa tcgtaaccga aaagcaagcc gcctatagca gcgattaggg	7680
ctagtcgctg gactagacc tgactacttt cagaactcat ggcgattcct ctccctctag	7740
agcgtcctgc tgttggttaag attattatac cacacctgt agataaagtc aacaactttt	7800
tgcaaaattt ttcaggaatt ttagcagagg ttgttctgga tgtagaaca aacatctttc	7860
cgctcttgtg ctgttaggat atctttcttg gaagctaggt aggcctcgag ttatggcagt	7920
tggttaaaag gaaacaaaaa gaccgttttc acacaaaacg gtctttttcg atttctttt	7980
acagtcacag ccaactttgc accaattaag gccacgtgt catttaaact ccgtttttcc	8040
agttcaaatg caattgcctt caatgcacct tcgtagctgt ggtgagccag cgggtgctggc	8100
tctcccccattacggataa gaatgcattt tccgagttaa taccgtcggc aatacctgac	8160
attaatactt cacagtcgct ggcatcgagt acggaaaact taatcgaaga cgaaccacag	8220
ttataacca aaacaaccgg aaattcattc atctcttttc tcacctgag ttacggatta	8280
aaacagtttg tatacagatg tcaggatggt cagcagacca atcacggtta caaacagtt	8340
atccagacga ccacggtatt tcgccagaga cggcgcttta cggtatggcat acatcggcaa	8400
caggcacagc agggatgcga taatcggtgc gcccatggct tcaatcaggt cgaggatgtt	8460
cgggttggcg taggcaacaa ccaggtgga gcccatgatg aagatcatgc tgagagtatt	8520
cagtttacc agcgacactt tggttttgtc acctttataa ccgaacttca gaatcagacc	8580

attcaagcct tccagcgicc ccagatagtg accgaagaaa gatttgaaga tagccacgag 8640

tgcatgatg gaagccgcat attccagtgt aatcggaac gttgttttgg taccggtcat 8700

ggacgcaaag tggtagcca gataagaaag cactggaata ttctgcgctt tggcttccgc 8760

catgttggcc ggagacagag taaacaggca gctaaaggca aagaacatca ccactgcaac 8820

catcagcatg ctggcacgag aaatgatttg ggaacattta cgttcggtga agtcgcgacc 8880

gaagtctttc tcatactctt cacgtttaga aaccacgaag gaagagacga ttggcgagaa 8940

gttaaaggag aaaacatga tggaaatccc cagccagaca gtgatcagga taccgtcatg 9000

accggttaac gacagcgaac cgaggtcaac ctggtcgata actgcagagt tccagtaagg 9060

gatcagcgac aaagaaatca gcaccaggct ggcgataaac ggccatacca ggtagctcag 9120

ggtaccgagc tcgaattcac tggccgtcgt tttaacaagt cgtgactggg aaaaccctgg 9180

cgttacccaa cttaatcgcc ttgcagcaca tcccccttcc gccagctggc gtaatagcga 9240

agaggcccgcc accgatcgcc cttcccaaca gttgcgcagc ctgaatggcg aatggcgcc 9300

gatgcggtat tttctcctta cgcactcttg cggtatttca caccgcataat ggtgcactct 9360

cagtacaate tgcctctgat cgcgatagtt aagccagccc cgacacccgc caacacccgc 9420

tgacgaattc 9430

<210> 17

<211> 5768

<212> DNA

<213> Escherichia coli

<220><221> misc_feature

<222> (1)..(5768)

<223> DNA sequence of the plasmid pMH17F

<400> 17

gttgacagta agacgggtaa gcctgttgat gataccgctg ccttactggg tgcattagcc 60

agtctgaatg acctgtcacg ggataatccg aagtggtcag actggaaaat cagagggcag 120

gaactgctga acagcaaaaa gtcagatagc accacatagc agacccgcca taaaacgccc 180

tgagaagccc gtgacgggct tttcttctat tatgggtagt ttcttgcgcat gaatccataa 240

aaggcgccctg tagtgccatt tacccccatt cactgccaga gccgtgagcg cagcgaactg 300

aatgtcacga aaaagacagc gactcagggtg cctgatggtc ggagacaaaa ggaatattca 360

gcgatttgcc cgagcttgcg aggggtgctac ttaagccttt aggggtttta ggtctgtttt 420

glagaggagc aaacagcggtt tgcgacatcc ttttgtaata ctgcggaact gactaaagta 480

gtgagttata cacagggtcg ggatctattc tttttatctt tttttattct ttctttattc	540
tataaattat aaccacttga atataaaca aaaaaacaca caaaggctca gcggaattta	600
cagagggtct agcagaattt acaagttttc cagcaaaggt ctagcagaat ttacagatac	660
ccacaactca aaggaaaagg actagtaatt atcattgact agcccatctc aattggtata	720
gtgattaaaa tcacctagac caattgagat gtatgtctga attagttgtt ttcaaagcaa	780
atgaactagc gattagtgcg tatgacttaa cggagcatga aaccaagcta attttatgct	840
gtgtggcact actcaacccc acgattgaaa accctacaag gaaagaacgg acggtatcgt	900
tcacttataa ccaatacgtc cagatgatga acatcagtag ggaaaatgct tatggtgtat	960
tagctaaagc aaccagagag ctgatgacga gaactgtgga aatcaggaat cctttggtta	1020
aaggctttga gattttccag tggacaaact atgccaagtt ctcaagcgaa aaattagaat	1080
tagtttttag tgaagagata ttgccttattc ttttccagtt aaaaaattc ataaaatata	1140
atctggaaca tgtaagtct tttgaaaaca aatactctat gaggatttat gagtggttat	1200
taaaagaact aacacaaaag aaaactcaca aggcaaatat agagattagc ctgatgaat	1260
ttaagtcat gttaatgctt gaaaataact accatgagtt taaaaggctt aaccaatggg	1320
ttttgaaacc aataagtaaa gatttaaaca cttacagcaa tatgaaattg gtggttgata	1380
agcgaggccg cccgactgat acgttgattt tccaagtga actagataga caaatggatc	1440
tcgtaaccga acttgagaac aaccagataa aaatgaatgg tgacaaaata ccaacaacca	1500
ttacatcaga ttctaccta cgtaacggac taagaaaac actacacgat gctttaactg	1560
caaaaattca gtcaccagt tttgaggcaa aatttttgag tgacatgcaa agtaagcatg	1620
atctcaatgg ttctgttctc tggctcacgc aaaaacaacg aaccacacta gagaacatac	1680
tggctaaata cggaaggatc tgaggttctt atggctcttg tatctatcag tgaagcatca	1740
agactaaca acaaaagtag aacaactgtt caccgttaga tatcaaaggg aaaactgtcc	1800
atatgcacag atgaaaacgg tgtaaaaaag atagatacat cagagctttt acgagttttt	1860
ggtgcattta aagctgttca ccatgaacag atcgacaatg taacagatga acagcatgta	1920
acacctaata gaacaggatga aaccagtaaa acaagcaac tagaacatga aattgaacac	1980
ctgagacaac ttgttacagc tcaacagtca cacatagaca gcctgaaaca ggcgatgctg	2040
cttatcgaat caaagctgcc gacaacacgg gagccagtga cgctcccggt ggggaaaaaa	2100
tcattggcaat tciggaagaa atagcgcttt cagccggcaa acctgaagcc ggatctgcga	2160
ttctgataac aaactagcaa caccagaaca gcccgtttgc gggcagcaaa acccgttatg	2220
cttgtaaac gttttgtgaa aaaattttta aaataaaaaa ggggacctct aggggtccca	2280
attaattagt aatataatct attaaaggtc attcaaaagg tcattccaccg gatcaattcc	2340

cctgctcgcg caggctgggt gccaaagtct cgggtaacat caaggcccga tccttgagc	2400
ccttgccctc ccgcacgatg atcgtgccgt gatcgaaatc cagatccttg acccgagtt	2460
gcaaaccctc actgatccgc atgcccgttc catacagaag ctggcggaac aaacgatgct	2520
cgccctccag aaaaccgagg atgcgaacca ctcatccgg ggtcagcacc accggcaagc	2580
gccgcgacgg ccgaggtctt ccgatctcct gaagccaggg cagatccgtg cacagcacct	2640
tgccgtagaa gaacagcaag gccgccaatg cctgacgatg cgtggagacc gaaaccttgc	2700
gctcgttcgc cagccaggac agaaatgcct cgacttcgct gctgccaag gttgccgggt	2760
gacgcacacc gtggaaacgg atgaaggcac gaaccagtg gacataagcc tgttcggttc	2820
gtaagctgta atgcaagtag cgtatgcgct cacgcaactg gtccagaacc ttgaccgaac	2880
gcagcggtag taacggcgca gtggcggttt tcatggcttg ttatgactgt tttttgggg	2940
tacagtctat gccctgggca tccaagcagc aagcgctta cgccgtgggt cgatgtttga	3000
tgttatggag cagcaacgat gttacgcagc agggcagtcg ccctaaaaca aagttaaaca	3060
tcatgaggga agcggtagtc gccgaagtat cgactcaact atcagaggta gttggcgta	3120
tcgagcgcca tctgaaccg acgttgcttg ccgtacattt gtacggctcc gcagtggatg	3180
gcggcctgaa gccacacagt gatattgatt tgcctgttac ggtgaccgta aggcttgatg	3240
aaacaacgcg gcgagctttg atcaacgacc ttttggaac ttcggcttcc cctggagaga	3300
gcgagattct ccgcgtgta gaagtcacca ttgttgtgca cgacgacatc attccgtggc	3360
gttatccagc taagcgcaa ctgcaatttg gagaatggca gcgcaatgac attcttcgag	3420
gtatcttcga gccagccagc atcgacattg atctggctat cttgtcgaca aaagcaagag	3480
aacatagcgt tgccttggtg ggtccagcgg cggaggaact cttgatccg gttcctgaac	3540
aggatctatt tgaggcgcta aatgaaacct taacgctatg gaactcgccg cccgactggg	3600
ctggcgatga gcgaaatgta gtgcttacgt tgtcccgcac ttggtacagc gcagtaaccg	3660
gcaaaatcgc gccgaaggat gtcgtgccg actgggcaat ggagcgctg ccggcccagt	3720
atcagcccgt catacttgaa gctagacagg ctatcttgg acaagaagaa gatcgcttgg	3780
cctcgcgcg agatcagttg gaagaatttg tccactacgt gaaaggcgag atcaccaagg	3840
tagtcggcaa ataattgcta acaattcgtt caagccgacg ccgcttcgcg gcgcggctta	3900
actcaagcgt tagatgcact aagcacataa ttgctcacag ccaaactatc aggtcaagtc	3960
tgcttttatt attttaagc gtgcataata agccctacac aaattgggag atatatcatg	4020
aaaggctggc ttttcttgt tatcgcaata gttggcgaag taatcgcaac atccgcatta	4080

aaatctagcg agggctttac taagctgac cggtggatga ccttttgaat gacctttaat 4140
 agattatatt actaattaat tggggaccct agaggtcccc ttttttattt taaaaatttt 4200
 ttcacaaaac ggtttacaag catacgttgg ccgattcatt aatgcagctg gcacgacagg 4260
 tttcccgact ggaaagcggg cagttagcgc aacgaatta atgtgagtta gtcactcat 4320
 taggcacccc aggcctttaca ctttatgctt ccggctcgta tgttgtgtgg aattgtgagt 4380
 tacctagaga gggtagaagt tgccgaacat ggcataagt tccccggaca gatttcaggt 4440
 ggtcagcagc aacgcgttgc cattgcgcgt tcgctgtgta tgaagccgaa aattatgttg 4500

 ttgatgagc caacgtcggc gctcgatcct gagatggtga aagaggtgct ggatacgatg 4560
 attgggctgg cgcagtcggg tatgacaatg ttgtgtgtaa cacatgagat ggggtttgca 4620
 cgaaccgtcg ctgaccgggt aatttttatg gatcgtgggg aaatagtga gcaagctgca 4680
 cctgatgaat tttttgcga tcctaaatca gagcgtacga gggcattttt atcgcaggta 4740
 atccattaat tgaatgttag ttgaaaaagc aaaaaggcca tcctttcgga tggcctttcg 4800
 cttgatttga tgtctggcag tttatggcgg gcgtcctgcc cgccaccctc cgggccgttg 4860
 cttcgcaacg ttcaaatccg ctcccggcgg atttgtccta ctcgggagag ttttcaccga 4920

 caaacaacag ataaaaacaa agggccagtc ttcgactga gccttttggt ttatttgatg 4980
 tctggcagtt cctactctc gcatggggag accccacact accatcggcg ctacggcgg 5040
 ttcacttctg agttcgcat ggggtcaggt gggaccaccg cgtactgcc gccagacaaa 5100
 ttcttttcta atctgccga ttttaacct aaaaagtgtg ctgatacca gagtgaact 5160
 ggggacctca ccttaccaa ggggtcgtc taccaactga gccatatcag cacgctaaat 5220
 ttgatgcctg gcagttcct actctcgcat ggggagacc cactacca tcggcgctac 5280
 ggcgtttcac ttctgagttc ggcatggggt caggtgggac caccgcgcta cggccgccag 5340

 gcaaattctg ttttatcaga ccgcttctgc gttctgattt aatctgtatc aggtgaaaa 5400
 tcttctctca tccggataac aatttcacac aggaacacgc tatgacctg attacgcaa 5460
 gctcgagctc gaattcactg gccgtcgtt tacaacgtc tgactgggaa aaccctggcg 5520
 ttaccaact taatgcctt gcagcacatc cccctttcgc cagctggcgt aatagcgaag 5580
 agggccgcac cgatgcctt tccaacagt tgcgcagcct gaatggcgaa tggcgctga 5640
 tgcggtattt tctccttacg catctgtcgc gtatttcaca ccgcatatgg tgcactctca 5700
 gtacaatctg ctctgatgcc gcatagttaa gccagccccg acaccgcca acaccgctg 5760

 acgaattc 5768

<210> 18

<211> 1053

<212> DNA

<213> Escherichia coli

<220><221> gene

<222> (1)..(1053)

<223> DNA sequence of the coding region of the wild type aroG gene

<400> 18

atgaattatc agaacgacga ttacgcatac aaagaaatca aagagttact tcctcctgtc 60

gcattgctgg aaaaattccc cgctactgaa aatgccgcga atacggttgc ccatgcccga 120

aaagcgatcc ataagatcct gaaaggtaat gatgatcgcc tgttggttgt gattggccca 180

tgctcaattc atgacctgt cgcgcaaaa gagtatgcca ctgccttgct ggcgctgcgt 240

gaagagctga aagatgagct ggaaatcgta atgcgcgtct atttgaaaa gccgcgtacc 300

acggtgggct ggaaagggt gattaacgat ccgcatatgg ataatactt ccagatcaac 360

gacggtctgc gtatagcccg taaattgtg ctgatatta acgacagcgg tctgccagcg 420

gcaggtgagt ttctcgatat gatcaccca caatatctcg ctgacctgat gagctggggc 480

gcaattggcg cagctaccac cgaatcgag gtgcaccgag aactggcatc agggctttct 540

tgtccggtcg gttcaaaaa tggcaccgac ggtacgatta aagtggctat cgatgccatt 600

aatgccgcg gtgcgccga ctgcttctg tccgtaacga aatgggggca ttcggcgatt 660

gtgaatacca gcgtaacgg cgattgccat atcattctgc gcggcggtaa agagcctaac 720

tacagcgca agcacgttg tgaagtgaag gaagggtga acaaagcagg cctgccagca 780

caggtgatga tcgatttcag ccatgctaac tcgtccaaac aattcaaaaa gcagatggat 840

gtttgtgctg acgtttgcc gcagattgcc ggtggcgaaa aggccattat tggcgtgatg 900

gtggaaagcc atctggtgga aggcaatcag agcctcgaga gcggggagcc gctggcctac 960

ggtaagagca tcaccgatgc ctgcatcggc tgggaagata ccgatgctct gttacgtcaa 1020

ctggcgaatg cagtaaaagc gcgtcgcggtg taa 1053

<210> 19

<211> 8820

<212> DNA

<213> Escherichia coli

<220><221> misc_feature

<222> (1)..(8820)

<223> DNA sequence of the plasmid pMH28F

<400> 19

gttgacagta agacgggtaa gcctgttgat gataccgctg ccttactggg tgcattagcc	60
agtctgaatg acctgtcacg ggataatccg aagtggtcag actggaaaat cagagggcag	120
gaactgctga acagcaaaaa gtcagatagc accacatagc agaccgcga taaaacgccc	180
tgagaagccc gtgacgggct tttcttgtat tatgggtagt ttccttgcat gaatccataa	240
aaggcgctg tagtgccatt taccgccatt cactgccaga gccgtgagcg cagcgaactg	300
aatgtcacga aaaagacagc gactcagggtg cctgatggtc ggagacaaaa ggaatattca	360
gcgatttgcc cgagcttgcc aggggtgctac ttaagccttt aggggtttta ggtctgtttt	420
gtagaggagc aaacagcgtt tgcgacatcc ttttgtaata ctgcggaact gactaaagta	480
gtgagttata cacagggctg ggatctattc tttttatctt tttttattct ttcctttattc	540
tataaattat aaccacttga atataaaca aaaaaacaca caaaggctca gcggaattta	600
cagagggctc agcagaattt acaagttttc cagcaaaggt ctagcagaat ttacagatac	660
ccacaactca aaggaaaagg actagtaatt atcattgact agcccatctc aattgggtata	720
gtgattaaaa tcacctagac caattgagat gtatgtctga attagtgtt ttcaaagcaa	780
atgaactagc gattagtgcg tatgacttaa cggagcatga aaccaagcta attttatgct	840
gtgtggcact actcaacccc acgattgaaa accctacaag gaaagaacgg acggtatcgt	900
tcacttataa ccaatacgtc cagatgatga acatcagtag ggaaaatgct tatggtgtat	960
tagctaaagc aaccagagag ctgatgacga gaactgtgga aatcaggaat cctttggtta	1020
aaggccttga gattttccag tggacaaact atgccaagtt ctcaagcgaa aaattagaat	1080
tagtttttag tgaagagata ttgccttata ttttccagtt aaaaaattc ataaaatata	1140
atctggaaca tgtaagtct tttgaaaaca aatactctat gaggatttat gagtggttat	1200
taaaagaact aacacaaaag aaaactcaca aggcgaatat agagattagc cttgatgaat	1260
ttaagttcat gttaatgctt gaaaataact accatgagtt taaaaggctt aaccaatggg	1320
ttttgaaacc aataagtaaa gatttaaaaca cttacagcaa tatgaaattg gtggttgata	1380
agcgaggccg cccgactgat acgttgattt tccaagtga actagataga caaatggatc	1440
tcgtaaccga acttgagaac aaccagataa aatgaatgg tgacaaaata ccaacaacca	1500
ttacatcaga ttctaccta cgtaacggac taagaaaaac actacacgat gctttaactg	1560
caaaaattca gtcaccagt tttgaggcaa aatttttgag tgacatgcaa agtaagcatg	1620
atctcaatgg ttcgtttctca tggctcacgc aaaaacaacg aaccacacta gagaacatac	1680
tggctaaata cggaaggatc tgaggttctt atggctcttg tatctatcag tgaagcatca	1740
agactaaca acaaaagtag aacaactgtt caccgttaga tatcaaaggg aaaactgtcc	1800
atatgcacag atgaaaacgg tgtaaaaaag atagatacat cagagctttt acgagttttt	1860

gggtgcattta aagctgttca ccatgaacag atcgacaatg taacagatga acagcatgta	1920
acacctaata gaacaggiga aaccagtaaa acaaagcaac tagaacatga aattgaacac	1980
ctgagacaac ttgttacagc tcaacagtca cacatagaca gcctgaaaca ggcgatgctg	2040
cttatcgaat caaagctgcc gacaacacgg gagccagtga cgcctcccgt ggggaaaaaa	2100
tcatggcaat tctggaagaa atagcgcttt cagccggcaa acctgaagcc ggatctgcga	2160
ttctgataac aaactagcaa caccagaaca gcccgtttgc gggcagcaaa acccgttatg	2220
cttgtaaacc gttttgtgaa aaaattttta aaataaaaaa ggggacctct aggggtccca	2280
attaattagt aatataatct attaaaggtc attcaaaagg tcatccaccg gatcaattcc	2340
cctgctcgcg caggctgggt gccaaagctct cgggtaacat caaggcccga tccttgagc	2400
ccttgccctc ccgcacgatg atcgtgccgt gatcgaaatc cagatccttg acccgagtt	2460
gcaaaccctc actgatccgc atgcccgttc catacagaag ctggcggaac aaacgatgct	2520
cgccttcag aaaaccgagg atgcgaacca ctcatccgg ggtcagcacc accggcaagc	2580
gccgcgacgg ccgaggtctt ccgatctcct gaagccaggg cagatccgtg cacagcacct	2640
tgccgtagaa gaacagcaag gccgccaatg cctgacgatg cgtggagacc gaaaccttgc	2700
gctcgttcgc cagccaggac agaaatgcct cgacttcgct gctgccaag gttgccgggt	2760
gagcacacc gtggaaacgg atgaaggcac gaaccagtg gacataagcc tgttcggttc	2820
gtaagctgta atgcaagtag cgtatgcgct cagcgaactg gtccagaacc ttgaccgaac	2880
gcagcggtgg taacggcgca gtggcggttt tcatggcttg ttatgactgt tttttgggg	2940
tacagtctat gccctgggca tccaagcagc aagcgcgtta cgccgtgggt cgatgtttga	3000
tgttatggag cagcaacgat gttacgcagc agggcagtcg ccctaaaaca aagttaaaca	3060
tcatgagga agcgggtatc gccgaagtat cgactcaact atcagaggta gttggcgta	3120
tcgagcgcca tctcgaaccg acgttgctgg ccgtacattt gtacggctcc gcagtggatg	3180
gcggcctgaa gccacacagt gatattgatt tgctggttac ggtgaccgta aggcttgatg	3240
aaacaacgcg gcgagctttg atcaacgacc ttttgaaac ttcggcttcc cctggagaga	3300
gcgagattct ccgcgctgta gaagtcacca ttgttgigca cgacgacatc attccgtggc	3360
gttatccagc taagcgcgaa ctgcaatttg gagaatggca gcgcaatgac attcttgag	3420
gtatcttcga gccagccagc atcgacattg atctggctat cttgctgaca aaagcaagag	3480
aacatagcgt tgccttggtg ggtccagcgg cggaggaact cttgatccg gttcctgaac	3540
aggatctatt tgaggcgcta aatgaaacct taacgctatg gaactcgccg cccgactggg	3600

ctggcgatga gcgaaatgta gtgcttacgt tgtcccgcac ttggtacagc gcagtaaccg	3660
gcaaaatcgc gccgaaggat gtgctgccc acgtgggcaat ggagcgcctg cgggccagc	3720
atcagcccgt catacttgaa gctagacagg cttatcttgg acaagaagaa gatcgcttgg	3780
cctcgcgcgc agatcagttg gaagaatttg tccactacgt gaaaggcgag atcaccaagg	3840
tagtcggcaa ataattgcta acaattcgtt caagccgacg ccgcttcgcg gcgcggctta	3900
actcaagcgt tagatgcaat aagcacataa ttgctcacag ccaaactatc aggtcaagtc	3960
tgcttttatt atttttaagc gtgcataata agccctacac aaattgggag atatatcatg	4020
aaaggctggc tttttcttgt tatcgcaata gttggcgaag taatcgcaac atccgcatta	4080
aaatctagcg agggctttac taagctgac ccgtggatga ccttttgaat gacctttaat	4140
agattatatt actaattaat tggggaccct agaggcccc tttttatatt taaaaatatt	4200
ttcacaaaac ggtttacaag catacgttgg ccgattcatt aatgcagctg gcacgacagg	4260
tttcccgcact ggaaagcggg cagtgcgcgc aacgcaatta atgtgagtta gctcactcat	4320
taggcacccc aggcctttaca ctttatgctt ccggctcgta tgttgtgtgg aattgtgagt	4380
tacctagaga ggggtgagaat tgccgaacat gcgcataagt ttcccggaca gatttcaggt	4440
ggtcagcagc aacgcgttgc cattgcgcgt tcgctgtgta tgaagccgaa aattatgttg	4500
tttgatgagc caacgtcggc gctcgatcct gagatggtga aagaggctgt ggatacgatg	4560
attgggctgg cgcagtcggg tatgacaatg ttgtgtgtaa cacatgagat ggggtttgca	4620
cgaaccgtcg ctgaccgggt aatttttatg gatcgctggg aaatagtga gcaagctgca	4680
cctgatgaat tttttgcgca tccataatca gagcgtacga gggcattttt atcgcaggta	4740
atccattaat tgaatgttag ttcgaaaagc aaaaaggcca tcccttcgga tggcctttcg	4800
cttgatttga tgtctggcag tttatggcgg gcgtcctgcc cgccaccctc cgggccgttg	4860
cttcgcaacg ttcaaatccg gtgacggaag atcacttcgc agaataaata aatcctggtg	4920
tccctgttga taccgggaag ccctgggcca acttttggcg aaaatgagac gttgatcggc	4980
acgtaagagg ttccaacttt caccataatg aaataagatc actaccgggc gtattttttg	5040
agttatcgag attttcagga gctaaggaag ctaaaatgga gaaaaaatc actggatata	5100
ccaccgttga tatatccaa tggcatcgta aagaacattt tgaggcattt cagtcagttg	5160
ctcaatgtac ctataaccag accgttcagc tggatattac ggccttttta aagaccgtaa	5220
agaaaaataa gcacaagttt tatccggcct ttattcacat tcttccccgc ctgatgaatg	5280
ctcatccgga attccgtatg gcaatgaaag acggtgagct ggtgatatgg gatagtgttc	5340
acccttgtta caccgttttc catgagcaaa ctgaaacgtt ttcatcgctc tggagtgaat	5400
accacgacga ttccggcag tttctacaca tatattcgca agatgtggcg tgttacgggtg	5460

aaaacctggc ctatttcctt aaagggttta ttgagaatat gtttttcgtc tcagccaatc	5520
cctgggtgag tttcaccagt tttgatttaa acgtggccaa tatggacaac ttcttcgccc	5580
ccgttttcac catgggcaaa tattatacgc aaggcgacaa ggtgctgatg ccgctggcga	5640
ttcaggttca tcatgccgtt tgtgatggct tccatgtcgg cagaatgctt aatgaattac	5700
aacagtactg cgatgagtgg cagggcgggg cgtaattttt ttaaggcagt tattggtgcc	5760
cttaaagcc tggtgctacg cctgaataag tgataataag cggatgaatg gcagaaattc	5820
gaaagcaaat tcgaccgggt cgtcggttca gggcagggtc gttaaatagc cgcttatgtc	5880
tattgctggt ttantcggta cccggggatc gcggccgagg accggatccc atcacatata	5940
cctgcccgtt actattattt agtgaatatg gatattatga tattttctga attgtgatta	6000
aaaaggcaac tttatgccca tgcaacagaa actataaaaa atacagagaa tgaaaagaaa	6060
cagatagatt ttttagttct ttaggcccgt agtctgcaaa tccttttatg attttctatc	6120
aaacaaaaga ggaaaataga ccagttgcaa tccaaacgag agtctaatag aatgaggtcg	6180
aaaagtaaat cgcgcggtt tggtactgat aaagcaggca agacctaaaa tgtgtaaagg	6240
gcaaagtgtg tactttggcg tccccctta catatttttag gtcttttttt attgtgcgta	6300
actaacttgc catcttcaaa caggagggct ggaagaagca gaccgctaac acagtacata	6360
aaaaaggaga catgaacgat gaacatcaaa aagtttgcaa aacaagcaac agtattaacc	6420
tttactaccg cactgctggc aggaggcgca actcaagcgt ttgcgaaaga aacgaaccaa	6480
aagccatata aggaacata cggcatttcc catattacac gccatgatat gctgcaaatc	6540
cctgaacagc aaaaaaatga aaaatatcaa gtctctgaat tcgattcgtc cacaattaaa	6600
aatatctctt ctgcaaagg cctggacgtt tgggacagct ggccattaca aaacgctgac	6660
ggcactgtcg caaactatca cggctaccac atcgtctttg cattagccgg agatcctaaa	6720
aatgcggatg acacatcgat ttacatgttc tatcaaaaag tcggcgaaac ttctattgac	6780
agctggaaaa acgctggccg cgtcttttaa gacagcgaca aattcgatgc aaatgattct	6840
atcctaaaag accaaacaca agaatgttca ggttcagcca catttacatc tgacggaaaa	6900
atccgtttat tctacactga tttctccggt aaacattacg gcaaacaaac actgacaact	6960
gcacaagtta acgtatcagc atcagacagc tctttgaaca tcaacggtgt agaggattat	7020
aaatcaatct ttgacgttga cggaaaaacg tatcaaaatg tacagcagtt catcgatgaa	7080
ggcaactaca gtcaggcga caaccatacg ctgagagatc ctactacgt agaagataaa	7140
ggccacaaat acttagtatt tgaagcaaac actggaactg aagatggcta ccaaggcgaa	7200
gaatctttat ttaacaaagc atactatggc aaaagcacat cattcttccg tcaagaaagt	7260
caaaaacttc tgcaaagcga taaaaaacgc acggctgagt tagcaaacgg cgctctcgtt	7320

atgattgagc taaacgatga ttacacactg aaaaaagtga tgaaaccgct gattgcatct 7380

aacacagtaa cagatgaaat tgaacgcgcg aacgtcttta aaatgaacgg caaatggtac 7440

ctgttcactg actccccggg atcaaaaatg acgattgacg gcattacgtc taacgatatt 7500

tacatgcttg gttatgtttc taattcttta actggcccat acaagccgct gaacaaaact 7560

ggccttgtgt taaaatgga tcttgatcct aacgatgtaa cctttactta ctcacacttc 7620

gctgtacctc aagcgaaagg aaacaatgtc gtgattacaa gctatatgac aaacagagga 7680

ttctacgcag acaacaatc aacgtttgcg ccgagcttcc tgctgaacat caaaggcaag 7740

aaaacatctg ttgtcaaaga cagcatcctt gaacaaggac aattaacagt taacaaataa 7800

aaacgcaaaa gaaaatgcca atatcctatt ggcatcttct tttatttctt ccatttaaatt 7860

ggatgcatgc gctagcggag tgtatactgg ctactatgt tggcactgat gaggggtgtca 7920

gtgaagtgtc tctccccgcg ggatttgtcc tactcgggag agtgttcacc gacaaacaac 7980

agataaaaca aaaggcccgact tcttcgact gagccttttg ttttatttga tgtctggcag 8040

ttccctactc tcgcatgggg agaccccaca ctaccatcgg cgctacggcg gtttcacttc 8100

tgagttcggc atggggtcag gtgggaccac cgcgctactg ccgccagaca aattcttttc 8160

taatctgcgc aactttaacc taaaagtgg tgctgatacc cagagtcgaa ctggggacct 8220

cacccttacc aagggtgcgc tctaccaact gagccatata agcacgctaa atttgatgcc 8280

tggcagttcc ctactctcgc atggggagac cccacactac catcggcgct acggcgtttc 8340

acttctgagt tcggcatggg gtcaggtggg accaccgcgc tacggccgcc aggcaaatc 8400

tgttttatca gaccgcttct gcgttctgat ttaatctgta tcaggctgaa aatcttctct 8460

catccggata acaatttcac acaggaaaca gctatgacca tgattacgcc aagctcgagc 8520

tcgaattcac tgccctcgt tttacaacgt cgtgactggg aaaaccctgg cgttacccaa 8580

cttaatcgcc ttgcagcaca tcccccttc gccagctggc gtaatagcga agaggccgc 8640

accgatgcc ctccccaa gttgcgcagc ctgaatggcg aatgggcct gatgcggtat 8700

tttctcetta cgcactgtg cggtatttca caccgcatat ggtgcactct cagtacaatc 8760

tgctctgatg ccgcatagtt aagccagccc cgacaccgc caacaccgc tgacgaattc 8820

8820

<210> 20

<211> 4774

<212> DNA

<213> Escherichia coli

<220><221> misc_feature

<222> (1)..(4774)

<400> 20

gttgacagta agacgggtaa gcctgttgat gataccgctg ccttactggg tgcattagcc 60

agtctgaatg acctgtcacg ggataatccg aagtggtcag actggaaaat cagagggcag 120

gaactgctga acagcaaaaa gtcagatagc accacatagc agaccgcga taaaacgccc 180

tgagaagccc gtgacgggct tttcttgtat tatgggtagt ttccttgcat gaatccataa 240

aaggcgcctg tagtgccatt taccgccatt cactgccaga gccgtgagcg cagcgaactg 300

aatgtcacga aaaagacagc gactcaggtg cctgatggtc ggagacaaaa ggaatattca 360

gcgatttgcc cgagcttgcc aggggtgctac ttaagccttt aggggtttta ggtctgtttt 420

gtagaggagc aaacagcggt tgcgacatcc ttttgtaata ctgcggaact gactaaagta 480

gtgagttata cacagggtcg ggatctattc tttttatctt tttttattct tttttattc 540

tataaattat aaccacttga atataaaca aaaaaacaca caaaggctca gcggaattta 600

cagagggtct agcagaattt acaagttttc cagcaaaggt ctagcagaat ttacagatac 660

ccacaactca aaggaaaagg actagtaatt atcattgact ageccatctc aatttggtata 720

gtgattaaaa tcacctagac caattgagat gtatgtctga attagtgttt tcaaaagcaa 780

atgaactagc gattagtgcg tatgacttaa cggagcatga aaccaagcta attttatgct 840

gtgtggcact actcaacccc acgattgaaa accctacaag gaaagaacgg acggtatcgt 900

tcacttataa ccaatacgtc cagatgatga acatcagtag ggaaaatgct tatggtgtat 960

tagctaaagc aaccagagag ctgatgacga gaactgtgga aatcaggaat cctttggtta 1020

aaggctttga gattttccag tggacaaact atgccaagtt ctcaagcgaa aaattagaat 1080

tagtttttag tgaagagata ttgccttatt ttttcagtt aaaaaattc ataaaatata 1140

atctggaaca tgtaagtct tttgaaaaca aatactctat gaggatttat gagtgttat 1200

taaaagaact aacacaaaag aaaactcaca aggcaaatat agagattagc cttgatgaat 1260

ttaagttcat gttaatgctt gaaaataact accatgagtt taaaaggctt aaccaatggg 1320

ttttgaaacc aataagtaaa gatttaaaca cttacagcaa tatgaaattg gtggttgata 1380

agcgaggccg cccgactgat acgttgattt tccaagtga actagataga caaatggatc 1440

tcgtaaccga acttgagaac aaccagataa aaatgaatgg tgacaaaata ccaacaacca 1500

ttacatcaga ttctaccta cgtaacggac taagaaaaac actacacgat gctttaactg 1560

caaaaattca gctcaccagt tttgaggcaa aatttttgag tgacatgcaa agtaagcatg 1620

atctcaatgg ttcgtttctc tggctcacgc aaaaacaacg aaccacacta gagaacatac 1680

tggttaaata cggaaggatc tgaggttctt atggctcttg tatctatcag tgaagcatca	1740
agactaacia acaaaagtag aacaactgtt caccgttaga tatcaaaggg aaaactgtcc	1800
atatgcacag atgaaaacgg tgtaaaaaag atagatacat cagagctttt acgagttttt	1860
ggtgcattta aagctgttca ccatgaacag atcgacaatg taacagatga acagcatgta	1920
acacctaata gaacagggtga aaccagtaaa acaaagcaac tagaacatga aattgaacac	1980
ctgagacaac ttgttacagc tcaacagtca cacatagaca gcctgaaaca ggcgatgctg	2040
cttatcgaat caaagctgcc gacaacacgg gagccagtga cgcctcccgt ggggaaaaaa	2100
tcatggcaat tctggaagaa atagcgcttt cagccggcaa acctgaagcc ggatctgcga	2160
ttctgataac aaactagcaa caccagaaca gcccgtttgc gggcagcaaa acccgttatg	2220
cttgtaaacc gttttgtgaa aaaattttta aaataaaaaa ggggacctct aggggtcccca	2280
attaattagt aatataatct attaaaggtc attcaaaagg tcatccaccg gatcaattcc	2340
cctgctcgcg caggctgggt gccaaagctct cgggtaacat caaggcccga tccttggagc	2400
ccttgccctc ccgcacgatg atcgtgccgt gatcgaaatc cagatccttg acccgagtt	2460
gcaaaccctc actgatccgc atgcccgttc catacagaag ctggcgcaac aaacgatgct	2520
cgccttcag aaaaccgagg atgcgaacca ctcatccgg ggtcagcacc accggcaagc	2580
gccgcgacgg ccgaggtctt ccgatctctt gaagccaggg cagatccgtg cacagcacct	2640
tgccgtagaa gaacagcaag gccgccaatg cctgacgatg cgtggagacc gaaaccttgc	2700
gctcgttcgc cagccaggac agaaatgcct cgacttcgtt gctgccaag gttgccgggt	2760
gacgcacacc gtggaaacgg atgaaggcac gaacccagtg gacataagcc gtctcggttc	2820
gtaagctgta atgcaagtag cgtatgcgtt cagcgaactg gtccagaacc ttgaccgaac	2880
gcagcggtgg taacggcgca gtggcggttt tcatggcttg ttatgactgt ttttttgggg	2940
tacagtctat gcctcgggca tccaagcagc aagcgcgtta cgccgtgggt cgatgtttga	3000
tgttatggag cagcaacgat gttacgcagc agggcagtcg ccctaaaaca aagttaaaca	3060
tcatgagga agcgggtgat gccgaagtat cgactcaact atcagaggta gttggcgta	3120
tcgagcgcca tctcgaaccg acgttgcttg ccgtacattt gtacggctcc gcagtggatg	3180
gcggcctgaa gccacacagt gatattgatt tgctggttac ggtgaccgta aggcttgatg	3240
aaacaacgcg gcgagctttg atcaacgacc ttttggaac ttcggttcc cctggagaga	3300
gcgagattct ccgcgctgta gaagtcacca ttgttggtga cgacgacatc attccgtggc	3360
gttatccagc taagcgcgaa ctgcaatttg gagaatggca gcgcaatgac attcttcgag	3420

gtatcttcga gccagccacg atcgacattg atctggctat cttgctgaca aaagcaagag 3480
aacaatagcgt tgccttggta ggtccagcgg cggaggaact ctttgatccg gttcctgaac 3540
aggatctatt tgaggcgcta aatgaaacct taacgctatg gaactcgccg cccgactggg 3600
ctggcgatga gcgaaatgta gtgcttacgt tgtcccgcat ttggtacagc gcagtaaccg 3660
gcaaaatcgc gccgaaggat gtcgctgccg actgggcaat ggagcgcctg cgggccagct 3720
atcagcccgt catacttgaa gctagacagg cttatcttgg acaagaagaa gatcgcttgg 3780
cctcgcgcg c agatcagttg gaagaatttg tccactacgt gaaaggcgag atcaccaagg 3840

tagtcggcaa ataatgtcta acaattcgtt caagccgacg ccgcttcgcg gcgcggctta 3900
actcaagcgt tagatgcaact aagcacataa ttgctcacag ccaaactatc aggtcaagtc 3960
tgcttttatt atttttaagc gtgcataata agccctacac aaattgggag atatatcatg 4020
aaaggctggc tttttcttgt tatcgcaata gttggcgaag taatcgcaac atccgcatta 4080
aaatctagcg agggcctttac taagctgac cgggtgatga ccttttgaat gacctttaat 4140
agattatatt actaattaat tggggaccct agagggtccc tttttatatt taaaaatttt 4200
ttcacaaaac ggtttacaag catacgttgg ccgattcatt aatgcagctg gcacgacagg 4260

tttcccgact ggaaagcggg cagtgagcgc aacgcaatta atgtgagtta gctcactcat 4320
taggcacccc aggcctttaca ctttatgctt ccggctcgta tgttgtgtgg aattgtgagc 4380
ggataacaat ttacacacagg aaacagctat gaccatgatt acgccaagct tgcatgcctg 4440
caggtcgact ctagaggatc cccgggtacc gagctcgaat tcaactggccg tcgttttaca 4500
acgtcgtgac tgggaaaacc ctggcggttac ccaacttaat cgccttgacg cacatcccc 4560
tttcgccagc tggcgtaata gcgaagaggc ccgcaccgat cgccttccc aacagttgcg 4620
cagcctgaat ggcgaaatggc gcctgatgcg gtattttctc cttacgcac tgtgcggtat 4680

ttcacaccgc atatggtgca ctctcagtac aatctgctct gatgccgcat agttaagcca 4740
gccccgacac ccgccaacac ccgctgacga attc 4774

<210> 21
<211> 6432
<212> DNA
<213> Escherichia coli
<220><221> misc_feature
<222> (1)..(6432)
<223> DNA sequence of the plasmid pMG27
<400> 21

gttgacagta agacgggtaa gcctgttgat gataccgctg ccttactggg tgcattagcc 60

agtctgaatg acctgtcacg ggataatccg aagtggtcag actggaaaat cagagggcag	120
gaactgctga acagcaaaaa gtcagatagc accacatagc agaccgcga taaaacgccc	180
tgagaagccc gtgacgggct tttcttgtat tatgggtagt ttccttgcat gaatccataa	240
aaggcgcttg tagtgccatt taccgccatt cactgccaga gccgtgagcg cagcgaactg	300
aatgtcacga aaaagacagc gactcaggtg cctgatggtc ggagacaaaa ggaatattca	360
gcgatttgcc cgagcttgcc aggggtgctac ttaagccttt agggttttaa ggtctgtttt	420
glagaggagc aaacagcgtt tgcgacatcc ttttgtaata ctgcggaact gactaaagta	480
gtgagtata cacagggctg ggatctattc tttttatctt tttttattct tttttattc	540
tataaattat aaccacttga atataaaca aaaaaacaca caaaggtcta gcggaattta	600
cagagggtct agcagaattt acaagttttc cagcaaaggt ctagcagaat ttacagatac	660
ccacaactca aaggaaaagg actagtaatt atcattgact agcccatctc aattggtata	720
gtgattaaaa tcacctagac caattgagat gtatgtctga attagtgtt ttcaaagcaa	780
atgaactagc gattagtgc tatgacttaa cggagcatga aaccaagcta attttatgct	840
gtgtggcact actcaacccc acgattgaaa accctacaag gaaagaacgg acggtatcgt	900
tcacttataa ccaatacgt cagatgatga acatcagtag ggaaaatgct tatgggtgat	960
tagctaaagc aaccagagag ctgatgacga gaactgtgga aatcaggaat cctttggtta	1020
aaggctttga gattttccag tggacaaact atgccaagtt ctcaagcgaa aaattagaat	1080
tagtttttag tgaagagata ttgccttatt tttccagtt aaaaaattc ataaaatata	1140
atctggaaca tgtaagtct tttgaaaaca aatctctat gaggatttat gagtgttat	1200
taaaagaact aacacaaaag aaaactcaca aggcaaatat agagattagc ctgatgaat	1260
ttaagttcat gttaatgctt gaaaataact accatgagtt taaaaggctt aaccaatggg	1320
ttttgaaacc aataagtaaa gatttaaca cttacagcaa tatgaaattg gtggttgata	1380
agcgaggccg cccgactgat acgttgattt tccaagtga actagataga caaatggatc	1440
tcgtaaccga acttgagaac aaccagataa aatgaatgg tgacaaaata ccaacaacca	1500
ttacatcaga ttctaccta cgtaacggac taagaaaaac actacacgat gctttaactg	1560
caaaaattca gctcaccagt tttgaggcaa aatttttgag tgacatgcaa agtaagcatg	1620
atctcaatgg ttcgttctca tggctcacgc aaaaacaacg aaccacacta gagaacatac	1680
tggctaaata cggaaggatc tgaggttctt atggctcttg tatctatcag tgaagcatca	1740
agactaaca acaaaagtag aacaactgtt caccgttaga tatcaaaggg aaaactgtcc	1800

atatgcacag atgaaaacgg tgtaaaaaag atagatacat cagagctttt acgagttttt	1860
ggtgcattta aagctgttca ccatgaacag atcgacaatg taacagatga acagcatgta	1920
acacctaata gaacaggatga aaccagtaaa acaaagcaac tagaacatga aattgaacac	1980
ctgagacaac ttgttacagc tcaacagtca cacatagaca gccigaaaca ggcatgctg	2040
cttatcgaat caaagctgcc gacaacacgg gagccagtga cgcctcccggt ggggaaaaaa	2100
tcatggcaat tctggaagaa atagcgcttt cagccggcaa acctgaagcc ggatctgcga	2160
ttctgataac aaactagcaa caccagaaca gcccgtttgc gggcagcaaa acccgttatg	2220
cttgtaaacc gttttgtgaa aaaattttta aaataaaaaa ggggacctct aggggtccca	2280
attaattagt aatataatct attaaaggic attcaaaagg tcatccaccg gatcaattcc	2340
cctgctcgcg caggctgggt gccaaagctct cgggtaacat caaggcccga tccttggagc	2400
ccttgcctc ccgcacgatg atcgtgccgt gatcgaaatc cagatccttg acccgagtt	2460
gcaaaccctc acgatccgc atgcccgttc cacaagaag ctgggcgaac aaacgatgct	2520
cgccttcag aaaaccgagg atgcgaacca ctcatccgg ggtcagcacc accggcaagc	2580
gccgcgacgg ccgaggtctt ccgatctcct gaagccaggg cagatccgtg cacagcacct	2640
tgccgtagaa gaacagcaag gccgccaatg cctgacgatg cgtggagacc gaaaccttgc	2700
gctcgttcgc cagccaggac agaaatgcct cgacttcgct gctgccaag gttgccgggt	2760
gacgcacacc gtggaaacgg atgaaggcac gaaccagtg gacataagcc tgttcggttc	2820
gtaagctgta atgcaagtag cgtatgcgct cagcgaactg gtccagaacc ttgaccgaac	2880
gcagcggtag taacggcgca gtggcggttt tcatggcttg ttatgactgt tttttgggg	2940
tacagtctat gcctcgggca tccaagcagc aagcgcgtta cgccgtgggt cgatgtttga	3000
tgttatggag cagcaacgat gttacgcagc agggcagtcg ccctaaaaca aagttaaaca	3060
tcatgaggga agcgggtgac gccgaagtat cgactcaact atcagaggta gttggcgtea	3120
tcgagcgcca tctgaaccg acgttgctgg ccgtacattt gtacggctcc gcagtggatg	3180
gcggcctgaa gccacacagt gatattgatt tgcgtgttac ggtgaccgta aggcttgatg	3240
aaacaacgcg gcgagctttg atcaacgacc ttttggaac ttcggcttcc cctggagaga	3300
gcgagattct ccgcgtgta gaagtcacca ttgttgtgca cgacgacatc attccgtggc	3360
gttatccagc taagcgcgaa ctgcaatttg gagaatggca gcgcaatgac attcttgag	3420
gtatcttcga gccagccacg atcgacattg atctggctat cttgctgaca aaagcaagag	3480
aacatagcgt tgccttggtg ggtccagcgg cggaggaaact ctttgatccg gttcctgaac	3540

aggatctatt tgaggcgcta aatgaaacct taacgctatg gaactcgccg cccgactggg	3600
ctggcgatga gcgaaatgta gtgcttacgt tgtcccgcat ttgttacagc gcagtaaccg	3660
gcaaaatcgc gccgaaggat gtgctgccc actgggcaat ggagcgccgt cgggccagct	3720
atcagcccgct catacttgaa gctagacagg cttatcttgg acaagaagaa gatcgcttgg	3780
cctcgcgcg agatcagttg gaagaatttg tccactacgt gaaaggcgag atcaccaagg	3840
tagtcggcaa ataattgtcta acaattcgtt caagccgacg ccgcttcgcg gcgcggctta	3900
actcaagcgt tagatgcact aagcacataa ttgctcacag ccaactatc aggtcaagtc	3960
tgcttttatt atttttaagc gtgcataata agccctacac aaattgggag atatatcatg	4020
aaaggctggc tttttcttgt tatcgcaata gttggcgaag taatcgcaac atccgcatta	4080
aatcttagcg agggctttac taagctgac cgttgatga cttttgaat gaccttaat	4140
agattatatt actaattaat tggggaccct agaggctccc ttttttattt taaaaatttt	4200
ttcacaaaac ggittacaag catacgttgg ccgattcatt aatgcagctg gcacgacagg	4260
tttcccgact ggaaagcggg cagtgagcgc aacgaatta atgtgagta gtcactcat	4320
taggcacccc aggctttaca ctttatgctt ccggctcgta tgttgtgtgg aattgtgagc	4380
ggataacaat ttcacacagg aaacagctat gaccatgatt acgccaagct tgcatgcctg	4440
caggctgacc gttaaatcta tcaccgcaag ggataaatat ctaacaccgt gcgtgttgac	4500
tattttacct ctggcgggtga taatggttgc atgtactaat ctagataagg aatatagcca	4560
tgaccgcacc gattcaggat ctgctgatg caattgccct gctgcaacag catgataatc	4620
agtatctgga aaccgatcat ccggttgatc cgaatgcaga actggcaggc gtttatcgtc	4680
atattggtgc cgggtggcacc gttaaacgtc cgacccgtat tgggccggca atgatgttta	4740
ataacattaa aggttatccg cacagccgta ttctggttgg tatgcatgca agccgtcagc	4800
gtgcagcact gctgctgggt tgtgaagcaa gtcagctggc actggaagtt ggtaaagcag	4860
ttaaaaaacc gggtgcaccg gtgggtgttc cggcaagcag cgcaccgtgt caagagcaga	4920
tttttctggc agatgatccg gattttgatc tgcgtaccct gctgcctgca cataccaata	4980
ccccgattga tgcaggctccg tttttttgtc tgggtctggc cctggcaagc gatccggtgg	5040
atgcaagcct gaccgatgtt accattcatc gtctgtgtgt tcagggtcgt gatgaactga	5100
gcatgttctt ggcagcaggt cgccatattg aagtttttcg tcagaaagca gaagcagcag	5160
gtaaaccgct gccgattacc attaatatgg gtctggaccc agcaatctat attggcgcat	5220
gttttgaagc accgaccacc ccgtttggtt ataatgaact gggtgttgcc ggtgcactgc	5280
gtcagcgtcc gggtgaactg gttcaggggtg ttagcgttcc ggaaaaagca attgcacgtg	5340
ccgaaattgt tattgaaggt gaactgctgc ctggtgttcg tgttcgtgaa gatcagcata	5400

ccaattcagg tcatgcaatg ccggaatttc cgggttattg tgggtggtgca aatccgagcc 5460
tgccggttat taaagttaaa gccgttacca tgcgcaataa cgcaattctg caaaccttgg 5520
ttggtccggg tgaagaacat accaccctgg caggtctgcc gaccgaagca agcatttga 5580
atgcagtga agcagcaatt ccgggttttc tgcaaaatgt ttatgcccat accgcaggcg 5640

gtggtaaatt tctgggtatt ctgcaagtga aaaaacgtca gcctgccgat gaagtcgtc 5700
agggtcaggc agccctgctg gcgctggcaa cctatagcga actgaaaaat atcattctgg 5760
tggatgagga tgtggacatt tttgatagt atgatatctt gtgggcaatg accaccgta 5820
tgcagggtga tgttagcatt accaccattc cgggtattcg cggtcacag ctggaccga 5880
gccagacacc ggaatattca ccgagcattc gtggtaatgg tattagctgc aaaaccatct 5940
ttgattgtac cgttccgtgg gcactgaaaa gccattttga acgtgcaccg tttgcagatg 6000
ttgatccgcg tccgtttgca cctgaatatt ttgcacgtct ggaaaaaat cagggcagcg 6060

caaaataagc taataacagg cctgctggta atcgcaggaa tttttatttg gatggatccc 6120
cgggtaccga gctcgaattc actggccgtc gttttacaac gtcgtgactg ggaaaaccct 6180
ggcgttacc aacttaatcg ccttcagca catccccctt tcgccagctg gcgtaatagc 6240
gaagaggccc gcaccgatcg ccttcccaa cagttgcga gcctgaatgg cgaatggcgc 6300
ctgatcggt attttctct tacgcattc tgcggtattt cacaccgat atggtgcact 6360
ctcagtaaa tctgctctga tgccgatag ttaagccage cccgacacc gccaacacc 6420
gctgacgaat tc 6432

<210> 22
<211> 7294
<212> DNA
<213> Escherichia coli
<220><221> misc_feature
<222> (1)..(7294)
<223> DNA sequence of the plasmid pMG31
<400> 22

gttgacagta agacgggtaa gcctgttgat gataccgtg cttactggg tgcattagcc 60
agtctgaatg acctgtcacg ggataatccg aagtggtcag actggaaaat cagagggcag 120
gaactgctga acagcaaaaa gtcagatagc accacatagc agaccgcca taaaacccc 180
tgagaagccc gtgacgggt tttcttgat tatgggtagt ttccttgcat gaatccataa 240
aaggcgctg tagtgccatt taccgccatt cactgccaga gccgtgagcg cagcgaactg 300

aatgtcacga aaaagacagc gactcaggtg cctgatggtc ggagacaaaa ggaatattca	360
gcgatttgcc cgagcttgcc aggggtgctac ttaagccttt agggttttta ggtctgtttt	420
gtagaggagc aaacagcgtt tgcgacatcc ttttgtaata ctgcggaact gactaaagta	480
gtgagttata cacagggctg ggatctattc tttttatctt tttttattct ttttttattc	540
tataaattat aaccacttga atataaaca aaaaaacaca caaaggtcta gcggaattta	600
cagagggctc agcagaattt acaagttttc cagcaaaggt ctagcagaat ttacagatac	660
ccacaactca aaggaaaagg actagtaatt atcattgact agcccatctc aattggtata	720
gtgattaaaa tcacctagac caattgagat gtatgtctga attagtgtt ttcaaagcaa	780
atgaactagc gattagtgcg tatgacttaa cggagcatga aaccaagcta attttatgct	840
gtgtggcact actcaacccc acgattgaaa accctacaag gaaagaacgg acggtatcgt	900
tcacttataa ccaatacgtc cagatgatga acatcagtag ggaaaatgct tatggtgtat	960
tagctaaagc aaccagagag ctgatgacga gaactgtgga aatcaggaat cctttggtta	1020
aaggctttga gattttccag tggacaaact atgccaagtt ctcaagcgaa aaattagaat	1080
tagtttttag tgaagagata ttgccttattc ttttccagtt aaaaaaatc ataaaatata	1140
atctggaaca tgtaagtc tttgaaaaca aatactctat gaggatttat gagtggttat	1200
taaaagaact aacacaaaag aaaactcaca aggcaaatat agagattagc ctgatgaat	1260
ttaagtcat gttaatgctt gaaaataact accatgagtt taaaaggctt aaccaatggg	1320
ttttgaaacc aataagtaaa gatttaaaca cttacagcaa tatgaaattg gtggttgata	1380
agcgaggccg cccgactgat acgttgattt tccaagtga actagataga caaatggatc	1440
tcgtaaccga acttgagaac aaccagataa aaatgaatgg tgacaaaata ccaacaacca	1500
ttacatcaga ttcttaccta cgtaacggac taagaaaaac actacacgat gctttaactg	1560
caaaaattca gtcaccagt tttgaggcaa aatttttgag tgacatgcaa agtaagcatg	1620
atctcaatgg ttcgtttctca tggctcacgc aaaaacaacg aaccacacta gagaacatac	1680
tggctaaata cggaaggatc tgaggttctt atggctcttg tatctatcag tgaagcatca	1740
agactaaca acaaaagtag aacaactgtt caccgttaga tatcaaaggg aaaactgtcc	1800
atatgcacag atgaaaacgg tgtaaaaaag atagatacat cagagctttt acgagttttt	1860
ggtgcattta aagctgttca ccatgaacag atcgacaatg taacagatga acagcatgta	1920
acacctaata gaacagggtga aaccagtaaa acaaagcaac tagaacatga aattgaacac	1980
ctgagacaac ttgttacagc tcaacagtca cacatagaca gcctgaaaca ggcgatgctg	2040
cttatcgaat caaagctgcc gacaacacgg gagccagtga cgctcccggt ggggaaaaaa	2100
tcatggcaat tciggaagaa atagcgcttt cagccggcaa acctgaagcc ggatctcgca	2160

ttctgataac aaactagcaa caccagaaca gcccgtttgc gggcagcaaa acccgttatg	2220
cttgtaaacc gttttgtgaa aaaattttta aaataaaaaa ggggacctct aggggtcccca	2280
attaattagt aatataatct attaaaggtc attcaaaagg tcatccaccg gatcaattcc	2340
cctgctcgcg caggctgggt gccaaagtct cgggtaacat caaggcccga tccttggagc	2400
ccttgccctc ccgcacgatg atcgtgccgt gatcgaaatc cagatccttg acccgcagtt	2460
gcaaaccctc actgatccgc atgcccggtc catacagaag ctggcggaac aaacgatgct	2520
cgccctccag aaaaccgagg atgcgaacca ctcatccgg ggtcagcacc accggcaagc	2580
gccgcgacgg ccgaggtctt ccgatctct gaagccaggg cagatccgtg cacagcacct	2640
tgccgtagaa gaacagcaag gccgccaatg cctgacgatg cgtggagacc gaaaccttgc	2700
gctcgttcgc cagccaggac agaaatgcct cgacttcgt gctgccaag gttgccgggt	2760
gacgcacacc gtggaaacgg atgaaggcac gaaccagtg gacataagcc tgttcggttc	2820
glaagctgta atgcaagtag cgtatgcgt cacgcaactg gtccagaacc ttgaccgaac	2880
gcagcggtgg taacggcgca gtggcggttt tcatggcttg ttatgactgt tttttgggg	2940
tacagtctat gctcgggca tccaagcagc aagcgcgtta cgccgtgggt cgatgtttga	3000
tgttatggag cagcaacgat gttacgcagc agggcagtcg ccctaaaaca aagttaaaca	3060
tcatgaggga agcgggtgac gccgaagtat cgactcaact atcagaggta gttggcgta	3120
tcgagcgcca tctgaaccg acgttgctgg ccgtacattt gtacggctcc gcagtggatg	3180
gcggcctgaa gccacacagt gatattgatt tgctggttac ggtgaccgta aggcttgatg	3240
aaacaacgcg gcgagctttg atcaacgacc ttttggaaac ttcggcttcc cctggagaga	3300
gcgagattct ccgcgctgta gaagtcacca ttgttgtgca cgacgacatc attccgtggc	3360
gttatccagc taagcgcgaa ctgcaatttg gagaatggca gcgcaatgac attcttcgag	3420
gtatcttcga gccagccagc atcgacattg atctggtat cttgctgaca aaagcaagag	3480
aacatagcgt tgccttggtg ggtccagcgg cggaggaaact ctttgatccg gttcctgaac	3540
aggatctatt tgaggcgcta aatgaaacct taacgctatg gaactcgccg cccgactggg	3600
ctggcgatga gcgaaatgta gtgcttacgt tgtcccgcat ttggtacagc gcagtaaccg	3660
gcaaaatcgc gccgaaggat gtcgtgccg actgggcaat ggagcgctg ccggcccagt	3720
atcagcccg cactattgaa gctagacagg ctatcttgg acaagaagaa gatcgcttgg	3780
cctcgcgcgc agatcagttg gaagaatttg tccactacgt gaaaggcgag atcaccaagg	3840
tagtcggcaa ataattgcta acaattcgtt caagccgacg ccgcttcgcg gcgcggctta	3900
actcaagcgt tagatgcact aagcacataa ttgctcacag ccaaactatc aggtcaagtc	3960
tgcttttatt atttttaagc gtgcataata agccctacac aaattgggag atatatcatg	4020

aaagctggc ttttcttgt tatcgcaata gttggcgaag taatcgcaac atccgcatta	4080
aaatctagcg agggctttac taagctgac cggtggatga ccttttgaat gacctttaat	4140
agattatatt actaattaat tggggaccct agaggteccc ttttttattt taaaaatttt	4200
ttcacaaaac ggtttacaag catacgttgg ccgattcatt aatgcagctg gcacgacagg	4260
tttcccgaact ggaagcgagg cagtgcgcgc aacgaatta atgtgagtt gctcactcat	4320
taggcacccc aggttttaca ctttatgctt ccggctcgta tgttgtgtgg aattgtgagc	4380
ggataacaat ttcacacagg aaacagctat gaccatgatt acgccaagct tgctattgac	4440
gacagctatg gttcactgac caccaaccaa aactgtgctc agtaccgcca atatttctcc	4500
cttgaggggt acaaagagggt gtccctagaa gagatccacg ctgtgtaaaa attttcaaaa	4560
aaggtattga ctttcctac aggggtgtga ataatttaac tacaggcggg ggcaaccccc	4620
cctgtttctg agaggaggaa tatagccatg gaagtgaata tcttcaacac ccaggatgtt	4680
caggattttc tgcgtgttgc aagcggctcg gaacaagagg gtggtaatcc gcgtgtttaa	4740
caaattattc atcgtgttct gagcgacctg tataaagcaa ttgaagatct gaatatcacc	4800
agcgacgaat attgggcagg cgttgcatat ctgaatcagc tgggtgcaaa tcaagaagca	4860
ggtctgctga gtccgggtct gggttttgat cattatctgg atatgcgtat ggatgcagaa	4920
gatgcagcac tgggtattga aaatgcaaca ccgcgtacca ttgaaggtcc gctgtatgtt	4980
gcgggtgcac cggaagcgt tggttatgca cgcatggatg atggtagcga tccgaatggt	5040
catacctga ttttgcattg caccattttt gatgcagatg gtaaacgct gccgaatgca	5100
aaagtgaata tttggcatgc aaacacaaaa ggtttttata gccattttga tccgaccggt	5160
gaacagcagg cttttaatat gcgtcgtagc attattaccg atgagaatgg tcagtatcgt	5220
gttcgtacca ttttgcctgc cggttatggt tgcctccgg aaggccgac ccagcaactg	5280
ctgaaccaac tgggtcgtca tggtaatcgt ccggcacata ttcattattt tgttagcgca	5340
gatgttcacc gtaaacgac caccagatt aatgttgccg gtgatccgta tacctatgat	5400
gattttgcat atgccaccg tgaaggtctg gttgttgatg cagttgaaca taccgatccg	5460
gaagcaatta aagccaatga tgtggaaggt ccttttgccg aaatgggtgtt tgatctgaaa	5520
ctgaccgctc tggttgatgg tgttgataat caggttgtgg atcgtccgct tctggcagtt	5580
taatacacca aaatggttca aaattatcag gcgagtgatc atgacactg gcctgttttt	5640
atttcaggga aggttgaga caattacgtg gataatcaga tcatccaaga aaccgtggat	5700
aaaattctga gcgttctgcc gaatcaggca ggtcagctgg cacgtctggt gcgtctgatg	5760

caatttgcac gcgatccgac cattaccgtt attggcaaat ataaccatgg taaaagccgt 5820
 ctgctgaatg aactgattgg caccgatatc tttagcgttg cagataaacg tgaaaccatt 5880
 cagctggccg aacataaaca ggatcaggtt cgttggcttg atgcacctgg tctggatgcc 5940
 gatgttgcag cagttagatg tctgcatgca tttgaagcag tttggaccca ggcagatatt 6000
 cgtctgtttg ttcatacgct tctgaaggt gaactggatg caaccgaaca ccatctgctg 6060
 caacagctga ttgaagatgc cgatcatagc cgtcgtcaga ccattctggt tctgaccag 6120
 attgatcaga ttccggatca gaccatcctg acacagatta aaaccagcat tgcacagcag 6180

gttccgaaac tggatatttg ggcagttagc gcaaccgctc atcgtcaggg cattgaaaac 6240
 ggtaaaaccc tgctgatcga aaaaagcggg attgggtcac tgcgccatac cctggaacag 6300
 gcactggcac aggtgccgag cgcacgtacc tatgaaaaa atcgtctgct gtcagatctg 6360
 caccatcagc tgaacaact gctgctggat cagaaacatg tttgcaaca actgcaacag 6420
 acacagcaac agcagctgca tgattttgat accggtctga ttaacattct ggacaaaatt 6480
 cgtgttgatc tggaaccgat tgtgaatatt gatggtcagg atcaagcact gaatccggat 6540
 agctttgcaa ccatgtttaa aaacaccgca gcaaacagc agcgtgccaa agttcagatt 6600

gcataatgcc gtgcatgcat tgaaatcaac agccatctga ttcgccatgg tgttgttggt 6660
 ctgcctgcgg aacagcagac caccattaaa agcattgata ccgtgattgt tgccgtgttt 6720
 ggtatcagcg ttaaatttcg tgatcagctg cgtgccctgt tttataccga taccgaacgt 6780
 cagcgtctgc aactggaatt tctgtttctat tttgaaaaa gtgccggtcg catgattctg 6840
 gcagcaaaaa ttgaacagac catgcgtcag cagggctgta ttcagaatgc catgatggca 6900
 ctgaacaaa tggaaagcgc agcataaaaa cacggacgcc gcaaacggcg tccgaatttc 6960
 ttggtcgact ctagaggatc cccgggtacc gagctcgaat tcactggccg tctgtttaca 7020

acgtcgtgac tgggaaaacc ctggcgttac ccaacttaat cgcttgagc caccatcccc 7080
 tttcgccagc tggcgtaata gcgaagaggc ccgcaccgat cgcccttccc aacagttgcg 7140
 cagcctgaat ggcgaatggc gcctgatgag gtattttctc cttacgcac tctgcggtat 7200
 ttcacaccgc atatggtgca ctctcagtag aatctgctct gatgccgcat agttaagcca 7260
 gccccgacac ccgccaacac ccgctgacga attc 7294

<210> 23

<211> 8952

<212> DNA

<213> Escherichia coli

<220><221> misc_feature

<222> (1)..(8952)

<223> DNA sequence of the plasmid pMG33

<400> 23

gttgacagta agacgggtaa gcctgttgat gataccgctg ccttactggg tgcattagcc	60
agtctgaatg acctgtcacg ggataatccg aagtggtcag actggaaaat cagagggcag	120
gaactgctga acagcaaaaa gtcagatagc accacatagc agacccgcca taaaacgccc	180
tgagaagccc gtgacgggct tttcttgtat tatgggtagt ttccttgcat gaatccataa	240
aaggcgctg tagtgccatt taccgccatt cactgccaga gccgtgagcg cagcgaactg	300
aatgtcacga aaaagacagc gactcagggtg cctgatggtc ggagacaaaa ggaatattca	360
gcgatttgcc cgagcttgcg aggggtgctac ttaagccttt aggggtttta ggtctgtttt	420
gtagaggagc aaacagcggt tgcgacatcc ttttgtaata ctgcggaact gactaaagta	480
gtgagttata cacagggctg ggatctattc tttttatctt tttttattct ttcctttattc	540
tataaattat aaccacttga atataaaca aaaaaacaca caaaggctca gcggaattta	600
cagagggctc agcagaattt acaagttttc cagcaaaggc ctacgagaat ttacagatac	660
ccacaactca aaggaaaagg actagtaatt atcattgact agcccatctc aattggtata	720
gtgattaaaa tcacctagac caattgagat gtatgtctga attagttgtt ttcaaagcaa	780
atgaactagc gattagtgcg tatgacttaa cggagcatga aaccaagcta attttatgct	840
gtgtggcact actcaacccc acgattgaaa acctacaag gaaagaacgg acggtatcgt	900
tcacttataa ccaatacgct cagatgatga acatcagtag ggaaaatgct tatggtgtat	960
tagctaaagc aaccagagag ctgatgacga gaactgtgga aatcaggaat cctttggtta	1020
aaggctttga gattttccag tggacaaact atgccaaagt ctcaagcgaa aaattagaat	1080
tagtttttag tgaagagata ttgccttatt tttccagtt aaaaaattc ataaaatata	1140
atctggaaca tgtaagtct tttgaaaaca aatactctat gaggatttat gagtggttat	1200
taaaagaact aacacaaaag aaaactcaca aggcataat agagattagc cttgatgaat	1260
ttaagttcat gttaatgctt gaaaataact accatgagtt taaaaggctt aaccaatggg	1320
ttttgaaacc aataagtaaa gatttaacaa cttacagcaa tatgaaattg gtggttgata	1380
agcgaggccg cccgactgat acgttgattt tccaagtga actagataga caaatggatc	1440
tcgtaaccga acttgagaac aaccagataa aaatgaatgg tgacaaaata ccaacaacca	1500
ttacatcaga ttcttaccta cgtaacggac taagaaaaac actacacgat gctttaactg	1560

caaaaattca gctcaccagt tttagggcaa aatttttgag tgacatgcaa agtaagcatg	1620
atctcaatgg ttcgtttctca tggctcacgc aaaaacaacg aaccacacta gagaacatac	1680
tggttaata cggaaggatc tgaggttctt atggctcttg tatctatcag tgaagcatca	1740
agactaaca acaaaagtag aacaactgtt caccgttaga tatcaaaggg aaaactgtcc	1800
atatgcacag atgaaaacgg tgtaaaaaag atagatacat cagagctttt acgagttttt	1860
ggtgcattta aagctgttca ccatgaacag atcgacaatg taacagatga acagcatgta	1920
acacctaata gaacagggtga aaccagtaaa acaaagcaac tagaacatga aattgaacac	1980
ctgagacaac ttgttacagc tcaacagtca cacatagaca gcctgaaaca ggcgatgctg	2040
cttatcgaat caaagctgcc gacaacacgg gagccagtga cgcctcccggt ggggaaaaaa	2100
tcatggcaat tctggaagaa atagcgcttt cagccggcaa acctgaagcc ggatctgcga	2160
ttctgataac aaactagcaa caccagaaca gcccgtttgc gggcagcaaa acccgttatg	2220
cttgtaaac gttttgtgaa aaaattttta aaataaaaaa ggggacctct aggggtccca	2280
attaattagt aatataatct attaaaggtc attcaaaagg tcatccaccg gatcaattcc	2340
cctgctcgcg caggctgggt gccaaagctct cgggtaacat caaggcccgga tccttggagc	2400
ccttgccctc ccgcacgatg atcgtgccgt gatcgaaatc cagatccttg acccgcagtt	2460
gcaaaccctc actgatccgc atgcccgttc catacagaag ctgggcgaac aaacgatgct	2520
cgccttcag aaaaccgagg atgcgaacca ctcatccgg ggtcagcacc accggcaagc	2580
gccgcgacgg ccgaggctct ccatctcct gaagccaggg cagatccgtg cacagcacct	2640
tgccgtagaa gaacagcaag gccgccaatg cctgacgatg cgtggagacc gaaaccttgc	2700
gctcgttcgc cagccaggac agaaatgcct cgacttcgct gctgccaag gttgccgggt	2760
gacgcacacc gtggaaacgg atgaaggcac gaaccagtg gacataagcc tgttcggttc	2820
gtaagctgta atgcaagtag cgtatgcgt cagcaactg gtcagaacc ttgaccgaac	2880
gcagcggtgg taacggcgca gtggcggttt tcatggcttg ttatgactgt tttttgggg	2940
tacagtctat gcctcgggca tccaagcagc aagcgcgtta cgccgtgggt cgatgtttga	3000
tgttatggag cagcaacgat gttacgcagc agggcagtcg ccctaaaaca aagttaaaca	3060
tcatgaggga agcgggtgac gccgaagtat cgactcaact atcagaggta gttggcgta	3120
tcgagcgcca tctgaaccg acgttgcttg ccgtacattt gtacggctcc gcagtggatg	3180
gcggcctgaa gccacacagt gatattgatt tgctggttac ggtgaccgta aggcttgatg	3240
aaacaacgcg gcgagctttg atcaacgacc ttttggaac ttcggcttcc cctggagaga	3300

gcgagattct ccgcgctgta gaagtcacca ttgtttgtga cgacgacatc attccgtggc	3360
gttatccagc taagcgcgaa ctgcaatttg gagaatggca gcgcaatgac attcttgag	3420
gtatcttcga gccagccacg atcgacattg atctggctat cttgctgaca aaagcaagag	3480
aacatagcgt tgccttggtg ggtccagcgg cggaggaact ctttgatccg gttcctgaac	3540
aggatctatt tgaggcgcta aatgaaacct taacgctatg gaactcgccg cccgactggg	3600
ctggcgatga gcgaaatgta gtgcttacgt tgtcccgcac ttggtacagc gcagtaaccg	3660
gcaaaatcgc gccgaaggat gtcgctgccg actgggcaat ggagcgcctg ccggcccagt	3720
atcagcccgt catacttgaa gctagacagg cttatcttgg acaagaagaa gatcgcttgg	3780
cctcgcgcgc agatcagttg gaagaatttg tccactacgt gaaaggcgag atcaccaagg	3840
tagtcggcaa ataattgtcta acaattcgtt caagccgacg ccgcttcgcg gcgcggctta	3900
actcaagcgt tagatgcact aagcacataa ttgctcacag ccaaactatc aggtcaagtc	3960
tgcttttatt atttttaagc gtgcataata agccctacac aaattgggag atatatcatg	4020
aaaggctggc tttttcttgt tategcaata gttggcgaag taatcgcaac atccgcatta	4080
aaatctagcg agggcctttac taagctgacg cgggtggatga ccttttgaat gacctttaat	4140
agatttatatt actaattaat tggggaccct agaggcccc ttttttattt taaaaatttt	4200
ttcacaaaac ggtttacaag catacgttgg ccgattcatt aatgcagctg gcacgacagg	4260
tttcccgaact ggaaagcggg cagtgagcgc aacgcaatta atgtgagtta gtcactcat	4320
taggcacccc aggcctttaca ctttatgctt ccggctcgta tgttgtgtgg aattgtgagc	4380
ggataacaat ttcacacagg aaacagctat gaccatgatt acgccaagct tgctattgac	4440
gacagctatg gttcactgtc caccaaccaa aactgtgctc agtaccgcca atatttctcc	4500
cttgaggggt acaaagaggt gtccttagaa gagatccacg ctgtgtaaaa attttacaaa	4560
aaggatttga ctttccctac aggggtgtgta ataatttaat tacaggcggg ggcaaccccg	4620
cctgtttctgc agaggaggaa tatagccatg gaagtgaata tcttcaacac ccaggatgtt	4680
caggattttc tgcgtgttgc aagcggctctg gaacaagagg gtggtaatcc gcgtgttaaa	4740
caaattatc atcgtgttct gagcgacctg tataaagcaa ttgaagatct gaatatcacc	4800
agcgacgaat attgggcagg cgttgcatat ctgaatcagc tgggtgcaaa tcaagaagca	4860
ggctctgtga gtccgggtct gggttttgat cattatctgg atatcgctat ggatgcagaa	4920
gatgcagcac tgggtattga aaatgcaaca ccgcgtacca ttgaaggtcc gctgtatgtt	4980
gcgggtgcac cgaaaagcgt tggttatgca cgcatggatg atggtagcga tccgaatggt	5040
cataccctga ttctgcatgg caccatTTTT gatgcagatg gtaaacgct gccgaatgca	5100
aaagtTgaaa ttTggcatgc aaacacaaaa ggcttttata gccattttga tccgaccggt	5160

gaacagcagg cctttaatat gcgtcgtagc attattaccg atgagaatgg tcagtatcgt	5220
gttcgtacca ttctgcctgc cggttatggt tgcctccgg aaggtccgac ccagcaactg	5280
ctgaaccaac tgggtcgtca tggtaatcgt ccggcacata ttcattatit tgttagcgca	5340
gatggtcacc gtaaacgtgac caccagatt aatgttgccg gtgatccgta tacctatgat	5400
gattttgcat atgccaccg tgaaggtctg gttgttgatg cagttgaaca taccgatccg	5460
gaagcaatta aagccaatga tgtggaaggt ccttttgccg aaatggtgtt tgatctgaaa	5520
ctgaccgcgc tggttgatgg tgttgataat caggttgtgg atcgtccgcg tctggcagtt	5580
taatacacca aaatggttca aaattatcag gcgagtgatc atgatcactg gcctgttttt	5640
atttcaggga aggttgaga caattacgtg gataatcaga tcatccaaga aaccgtggat	5700
aaaattctga gcgttctgcc gaatcaggca ggtcagctgg cacgtctggt gcgtctgatg	5760
caatttgcat gcgatccgac cattaccgtt attggcaaat ataaccatgg taaaagccgt	5820
ctgctgaatg aactgattgg caccgataac tttagcgttg cagataaacg tgaaaccatt	5880
cagctggccg aacataaaca ggatcaggtt cgttggtctg atgcacctgg tctggatgcc	5940
gatgttcgag cagttgatga tcgtcatgca tttgaagcag tttggacca ggcagatatt	6000
cgtctgtttg ttcatacgct tcgtgaaggt gaactggatg caaccgaaca ccatctgctg	6060
caacagctga ttgaagatgc cgatcatagc cgtcgtcaga ccattctggt tctgaccag	6120
attgatcaga ttccggatca gaccatctg acacagatta aaaccagcat tgcacagcag	6180
gttccgaaac tggatatttg ggcagttagc gcaaccgctc atcgtcaggg cattgaaaac	6240
ggtaaaacc tgcgatcga aaaaagcgggt attggtgcac tgcgccatac cctggaacag	6300
gcactggcac aggtgccgag cgcacgtacc tatgaaaaa atcgtctgct gtcagatctg	6360
caccatcagc tgaacaact gctgctggat cagaaacatg ttctgcaaca actgcaacag	6420
acacagcaac agcagctgca tgattttgat accggtctga ttaacattct ggacaaaatt	6480
cgtgttgatc tggaaaccgat tgtgaatatt gatggtcagg atcaagcact gaatccggat	6540
agctttgcaa ccatgtttaa aaacaccgca gcaaaacagc agcgtgccaa agttcagatt	6600
gcatatagcc gtgcatgcat tgaatacaac agccatctga ttcgccatgg tgttgttggt	6660
ctgcctgcgg aacagcagac caccattaaa agcattgata ccgtgattgt tgccgtgttt	6720
ggtatcagcg ttaaatttcg tgatcagctg cgtgccctgt tttataccga taccgaacgt	6780
cagcgtctgc aacgtgaatt tcgtttctat tttgaaaaa gtgccggtcg catgattctg	6840
gcagcaaaaa ttgaacagac catgcgtcag cagggtctga ttcagaatgc catgatggca	6900
ctgcaacaaa tggaaagcgc agcataaaaa cacggacgcc gcaaacggcg tccgaatttc	6960
ttggtcgacc gttaaactca tcaccgcaag ggataaatat ctaacaccgt gcgtgttgac	7020

tattttacct ctggcgggtga taatggttgc atgtactaat ctagataagg aatatagcca	7080
tgaccgcacc gattcaggat ctgcgtgatg caattgccct gctgcaacag catgataatc	7140
agtatctgga aaccgatcat ccggttgatc cgaatgcaga actggcaggc gtttatcgtc	7200
atattggtgc cggtaggcacc gttaaacgtc cgaccctgat tggtagcgca atgatgttta	7260
ataacattaa aggttatccg cacagccgta tttcgttgg tatgcatgca agccgtcagc	7320
gtgcagcact gctgctgggt tgtgaagcaa gtcagctggc actggaagtt ggtaaagcag	7380
ttaaaaaacc ggttgaccg gtggttggtc cggcaagcag cgcaccgtgt caagagcaga	7440
tttttctggc agatgatccg gattttgatc tgcgtaccct gctgcctgca cataccaata	7500
ccccgattga tgcaggctccg tttttttgtc tgggtctggc cctggcaagc gatccggtgg	7560
atgcaagcct gaccgatgtt accattcatc gtctgtgtgt tcagggtcgt gatgaactga	7620
gcatgttctt ggcagcaggc cgccatattg aagtttttcg tcagaaaagca gaagcagcag	7680
glaaaccgct gccgattacc attaatatgg gtctggacc agcaatctat attggcgcat	7740
gttttgaagc accgaccacc ccgtttggtt ataatgaact gggtgttgcc ggtgcactgc	7800
gtcagcgtcc ggttgaactg gttcagggtg ttagcgttcc ggaaaaagca attgcacgtg	7860
ccgaaattgt tattgaaggt gaactgctgc ctggtgttcg tgttcgtgaa gatcagcata	7920
ccaattcagg tcatgcaatg ccggaatttc cgggttattg tggtagtgca aatccgagcc	7980
tgccggttat taaagttaaa gccgttacca tgcgaataa cgcaattctg caaaccttgg	8040
ttggtccggg tgaagaacat accaccctgg caggctctgcc gaccgaagca agcatttga	8100
atgcagtga agcagcaatt ccgggttttc tgcaaatgt ttatgcccat accgcaggcg	8160
gtggtaaatt tctgggtatt ctgcaagtga aaaaacgtca gcctgccgat gaaggctcgtc	8220
agggtcaggc agccctgctg gcgctggcaa cctatagcga actgaaaaat atcattctgg	8280
tggatgagga tgtggacatt tttgatagt atgatattct gtgggcaatg accaccgta	8340
tgcagggtga tgtagcatt accaccattc cgggtattcg cggatcatcag ctggaccgca	8400
gccagacacc ggaatattca ccgagcattc gtgtaatgg tattagctgc aaaaccatct	8460
ttgattgtac cgttccgtgg gcactgaaaa gccattttga acgtgcaccg tttgcagatg	8520
ttgatccgcg tccgtttgca cctgaatatt ttgcacgtct ggaaaaaat cagggcagcg	8580
caaaataagc taataacagg cctgctggta atgcaggaa tttttatttg gatggatccc	8640
cgggtaccga gctcgaattc actggccgtc gttttacaac gtcgtgactg ggaaaaccct	8700
ggcgttacc aacttaatcg ccttgacgca catccccctt tcgccagctg gcgtaatagc	8760

gaagaggccc gcaccgatcg cccttcccaa cagttgcgca gcctgaatgg cgaatggcgc 8820
ctgatcggtt attttctcct tacgcatctg tgcggtatct cacaccgcat atggtgcact 8880
ctcagtacaa tctgctctga tgccgcatag ttaagccagc cccgacaccc gccaacaccc 8940
gctgacgaat tc 8952

<210> 24
<211> 10630
<212> DNA
<213> Escherichia coli
<220><221> misc_feature
<222> (1)..(10630)
<223> DNA sequence of the plasmid pMG37
<400> 24

gttgacagta agacgggtaa gcctgttgat gataccgctg cttactggg tgcattagcc 60

agttctgaatg acctgtcacg ggataatccg aagtgggtcag actggaaaat cagagggcag 120
gaactgctga acagcaaaaa gtcagatagc accacatagc agacccgcca taaaacgccc 180
tgagaagccc gtgacgggct tttcttgtat tatgggtagt ttccttgcat gaatccataa 240
aaggcgctgt tagtgccatt taccgccatt cactgccaga gccgtgagcg cagcgaactg 300
aatgtcacga aaaagacagc gactcagggt cctgatggtc ggagacaaaa ggaatattca 360
gcgatttgcc cgagcttgcc aggggtgctac ttaagccttt aggggtttta ggtctgtttt 420
gtagaggagc aaacagcggt tgccgacatcc ttttgtaata ctgcggaact gactaaagta 480

gtgagttata cacagggtct ggatctattc tttttatctt tttttattct ttctttattc 540
tataaattat aaccacttga atataaaca aaaaaacaca caaagggtcta gcggaattta 600
cagaggggtct agcagaattt acaagttttc cagcaaaggc ctagcagaat ttacagatac 660
ccacaactca aaggaaaagg actagtaatt atcattgact agcccatctc aattgggtata 720
gtgattaaaa tcacctagac caattgagat gtatgtctga attagttgtt ttcaaagcaa 780
atgaactagc gattagtcgc tatgacttaa cggagcatga aaccaagcta attttatgct 840
gtgtggcact actcaacccc acgattgaaa accctacaag gaaagaacgg acggtatcgt 900

tcacttataa ccaatacgtc cagatgatga acatcagtag ggaaaatgct tatgggtgat 960
tagctaaagc aaccagagag ctgatgacga gaactgtgga aatcaggaat cctttgggtta 1020
aaggctttga gatgttccag tggacaaact atgccaagtt ctcaagcgaa aaattagaat 1080
tagtttttag tgaagagata ttgccttatt tttccagtt aaaaaattc ataaaatata 1140
atctggaaca tgtaagtct tttgaaaaca aatactctat gaggatttat gagtgggtat 1200

taaaagaact aacacaaaag aaaactcaca aggcaaatat agagattagc cttgatgaat	1260
ttaagttcat gttaatgctt gaaaataact accatgagtt taaaaggctt aaccaatggg	1320
ttttgaaacc aataagtaaa gatttaaaca cttacagcaa tatgaaattg gtggttgata	1380
agcgaggccg cccgactgat acgttgattt tccaagtga actagataga caaatggatc	1440
tcgtaaccga acttgagaac aaccagataa aatgaatgg tgacaaaata ccaacaacca	1500
ttacatcaga ttctaccta cgtaacggac taagaaaaac actacacgat gctttaactg	1560
caaaaattca gctcaccagt tttagggcaa aatttttgag tgacatgcaa agtaagcatg	1620
atctcaatgg ttctgttctca tggctcacgc aaaaacaacg aaccacacta gagaacatac	1680
tggctaaata cggaaggatc tgaggttctt atggctcttg tatctatcag tgaagcatca	1740
agactaacia acaaaagtag aacaactgtt caccgttaga tatcaaaggg aaaactgtcc	1800
atatgcacag atgaaaacgg tgtaaaaaag atagatacat cagagctttt acgagttttt	1860
ggtgcattta aagctgttca ccatgaacag atcgacaatg taacagatga acagcatgta	1920
acacctaata gaacaggatga aaccagtaaa acaaagcaac tagaacatga aattgaacac	1980
ctgagacaac ttgttacagc tcaacagtca cacatagaca gcctgaaaca ggcgatgctg	2040
cttatcgaat caaagctgcc gacaacacgg gagccagtga cgcctcccggt ggggaaaaaa	2100
tcatggcaat tctggaagaa atagcgcttt cagccggcaa acctgaagcc ggatctgcga	2160
ttctgataac aaactagcaa caccagaaca gcccgtttgc gggcagcaaa acccgttatg	2220
cttgtaaacc gttttgtgaa aaaattttta aaataaaaaa ggggacctct aggggtccca	2280
attaattagt aatataatct attaaaggtc attcaaaagg tcatccaccg gatcaattcc	2340
cctgctcgcg caggctgggt gccaaagctct cgggtaacat caaggcccga tccttgagac	2400
ccttgccctc ccgcacgatg atcgtgccgt gatcgaaac cagatccttg acccgagtt	2460
gcaaaccctc actgatccgc atgcccgttc catacagaag ctgggcgaac aaacgatgct	2520
cgctttccag aaaaccgagg atgcgaacca ctcatccgg ggtcagcacc accggcaagc	2580
gccgcgacgg ccgaggtctt ccgatctcct gaagccaggg cagatccgtg cacagcacct	2640
tgccgtagaa gaacagcaag gccgccaatg cctgacgatg cgtggagacc gaaaccttgc	2700
gctcgttcgc cagccaggac agaaatgcct cgacttcgct gctgcccgaag gttgccgggt	2760
gacgcacacc gtggaaacgg atgaaggcac gaaccagtg gacataagcc tgttcggttc	2820
gtaagctgta atgcaagtag cgtaatgcgt cagcaactg gtccagaacc ttgaccgaac	2880
gcagcggtgg taacggcgca gtggcggttt tcatggcttg ttatgactgt ttttttgggg	2940

tacagtctat gcctcgggca tccaagcagc aagcgcgtta cgccgtgggt cgatgtttga	3000
tgttatggag cagcaacgat gttacgcagc agggcagtcg ccctaaaaca aagttaaaca	3060
tcatgaggga agcgggtgac gccgaagtat cgactcaact atcagaggta gttggcgtca	3120
tcgagcgcca tctcgaaccg acgttgctgg ccgtacattt gtacggctcc gcagtggatg	3180
gcggcctgaa gccacacagt gatattgatt tgctggttac ggtgaccgta aggcttgatg	3240
aaacaacgcg gcgagctttg atcaacgacc ttttggaac ttcggcttcc cctggagaga	3300
gcgagattct ccgcgtgta gaagtcacca ttgttgtgca cgacgacatc attccgtggc	3360
gttatccagc taagcgcgaa ctgcaatttg gagaatggca gcgcaatgac attcttgag	3420
glatcttga gccagccacg atcgacattg atctggctat cttgctgaca aaagcaagag	3480
aacatagcgt tgccttggtg ggtccagcgg cggaggaact ctttgatccg gttcctgaac	3540
aggatctatt tgaggcgcta aatgaaacct taacgctatg gaactcgccg cccgactggg	3600
ctggcgatga gcgaaatgta gtgcttacgt tgtcccgcat ttgttacagc gcagtaaccg	3660
gcaaaatcgc gccgaaggat gtcgtgccc actgggcaat ggagcgctg ccggcccagt	3720
atcagcccg cactattgaa gctagacagg ctatcttgg acaagaagaa gatcgcttgg	3780
cctcgcgcgc agatcagttg gaagaatttg tccactacgt gaaaggcgag atcaccaagg	3840
tagtcggcaa ataattgcta acaattcgtt caagccgacg ccgcttcgcg gcgcggctta	3900
actcaagcgt tagatgcact aagcacataa ttgctcacag ccaactatc aggtcaagtc	3960
tgcttttatt atttttaagc gtgcataata agccctacac aaattgggag atatatcatg	4020
aaagctggc tttttcttgt tatcgcaata gttggcgaag taatcgcaac atccgcatta	4080
aatctagcg agggctttac taagctgac cggtggatga cttttgaat gacctttaat	4140
agattatatt actaattaat tggggaccct agagggtcccc tttttatatt taaaaatatt	4200
ttcacaaaac ggtttacaag catacgttgg ccgattcatt aatgcagctg gcacgacagg	4260
tttcccgact ggaaagcggg cagtgagcgc aacgcaatta atgtgagtta gtcactcat	4320
taggcacccc aggttttaca ctttatgctt ccggctcgta tgttgtgtgg aattgtgagc	4380
ggataacaat ttcacacagg aaacagctat gaccatgatt acgccaagct tgctattgac	4440
gacagctatg gttcactgtc caccaaccaa aactgtgctc agtaccgcca atatttctcc	4500
cttgaggggt acaaagaggt gtcctagaa gagatccagc ctgtgtaaaa attttataaa	4560
aaggtattga ctttcctac aggggtgtgta ataatttaat tacaggcggg ggcaaccccg	4620
cctgttctgc agaggaggaa tatagccatg gaagtgaaaa tcttcaacac ccaggatgtt	4680

caggattttc tgcgtgttgc aagcggctctg gaacaagagg gtggtaatcc gcgtgtttaa	4740
caaattattc atcgtgttct gagegacctg tataaagcaa ttgaagatct gaatatcacc	4800
agcgacgaat attgggcagg cgttgcatat ctgaatcagc tgggtgcaaa tcaagaagca	4860
ggtctgctga gtccgggtct gggttttgat cattatctgg atatgcgtat ggatgcagaa	4920
gatgcagcac tgggtattga aaatgcaaca ccgcgtacca ttgaaggtcc gctgtatgtt	4980
gcgggtgcac cggaaagcgt tggttatgca cgcatggatg atggtagcga tccgaatggt	5040
cataccctga ttctgcatgg caccattttt gatgcagatg gtaaacgct gccgaatgca	5100
aaagtgtaaa tttggcatgc aaacacaaaa ggcttttata gccattttga tccgaccggt	5160
gaacagcagg cctttaatat gcgtcgtagc attattaccg atgagaatgg tcagtatcgt	5220
gttcgtacca ttctgcctgc cggttatggt tgcctccgg aaggtccgac ccagcaactg	5280
ctgaaccaac tgggtcgtca tggtaatcgt ccggcacata ttcattatit tgttagcgca	5340
gatgttcacc gtaaacgtac caccagatt aatgttccg gtgatccgta tacctatgat	5400
gattttgcat atgccaccg tgaaggtctg gttgttgatg cagtgaaca taccgatccg	5460
gaagcaatta aagccaatga tgtggaaggt cttttgccc aaatggtgtt tgatctgaaa	5520
ctgaccgctc tggttgatgg tgttgataat caggttgtgg atcgtccgcg tctggcagtt	5580
taatacacca aaatggttca aaattatcag gcgagtgatc atgatcactg gcctgttttt	5640
atttcaggga aggggtggaga caattacgtg gataatcaga tcatccaaga aaccgtggat	5700
aaaattctga gcgttctgcc gaatcaggca ggtcagctgg cacgtctggt gcgtctgatg	5760
caatttgcat gcgatccgac cattaccgtt attggcaaat ataaccatgg taaaagccgt	5820
ctgctgaatg aactgattgg caccgatac tttagcgttg cagataaacg tgaaaccatt	5880
cagctggccg aacataaaca ggatcaggtt cgttggctgg atgcacctgg tctggatgcc	5940
gatgttgacg cagttagatg tgcgtatgca tttgaagcag tttggacca ggcagatatt	6000
cgtctgtttt ttcatagcgt tegtgaaggt gaactggatg caaccgaaca ccatctgctg	6060
caacagctga ttgaagatgc cgatcatagc cgtcgtcaga ccattctggt tctgaccag	6120
attgatcaga ttccggatca gaccatcctg acacagatta aaaccagcat tgcacagcag	6180
gttccgaaac tggatatttg ggcagttagc gcaaccgctc atcgtcaggg cattgaaaac	6240
ggtaaaacc tgcgtatcga aaaaagcgggt attggtgcac tgcgccatac cctggaacag	6300
gcactggcac aggtgccgag cgcacgtacc tatgaaaaa atcgtctgct gtcagatctg	6360
caccatcagc tgaacaact gctgctggat cagaaacatg ttctgcaaca actgcaacag	6420
acacagcaac agcagctgca tgattttgat accggtctga ttaacattct ggacaaaatt	6480
cgtgttgatc tggaaccgat tgtgaatatt gatggtcagg atcaagcact gaatccgat	6540

agctttgcaa ccatgtttaa aaacaccgca gcaaaacagc agcgtgcaa agttcagatt	6600
gcataatagcc gtgcatgcat tgaaatcaac agccatctga ttcgcatgg tttgttgg	6660
ctgcctgcgg aacagcagac caccattaaa agcattgata ccgtgattgt tgccgtgttt	6720
ggatatcagcg ttaaatttcg tgatcagctg cgtgccctgt tttataccga taccgaacgt	6780
cagcgtctgc aacgtgaatt tcgtttctat ttgaaaaa gtgccggtcg catgattctg	6840
gcagcaaaaa ttgaacagac catgcgtcag cagggctgta ttcagaatgc catgatggca	6900
ctgcaacaaa tggaaagcgc agcataaaaa cacggacgcc gcaaacggcg tccgaatttc	6960
ttggtcgacc gttaaactta tcaccgcaag ggataaatat ctaacaccgt gcgtgttgac	7020
tattttacct ctggcgggta taatggttgc atgtactaat ctagataagg aatatagcca	7080
tgaccgcacc gattcaggat ctgcgtgatg caattgccct gctgcaacag catgataatc	7140
agtatctgga aaccgatcat ccggttgatc cgaatgcaga actggcaggc gtttatcgtc	7200
atattgggtgc cggtaggcacc gttaaactgc cgaccctgat tggtaggcga atgatgttta	7260
ataacattaa aggttatccg cacagccgta ttctggttgg tatgcatgca agccgtcagc	7320
gtgcagcact gctgctgggt tgtgaagcaa gtcagctggc actggaagtt ggtaaagcag	7380
ttaaaaaacc ggttgcaccg gtggttgttc cggcaagcag cgcaccgtgt caagagcaga	7440
tttttctggc agatgatccg gattttgatc tgcgtaccct gctgcctgca cataccaata	7500
ccccgattga tgcaggctccg tttttttgtc tgggtctggc cctggcaagc gatccggtgg	7560
atgcaagcct gaccgatgtt accattcatc gtctgtgtgt tcagggtcgt gatgaactga	7620
gcatgttccct ggcagcaggt cgccatattg aagtttttcg tcagaaagca gaagcagcag	7680
gtaaaccgct gccgattacc attaatatgg gtctggaccc agcaatctat attggcgcac	7740
gttttgaagc accgaccacc ccgtttggtt ataatagaact ggggtgttgc ggtgcactgc	7800
gtcagcgtcc ggttgaactg gttcagggtg ttagcgttcc ggaaaaagca attgcacgtg	7860
ccgaaattgt tattgaaggt gaactgctgc ctggtgttcg tgttcgtgaa gatcagcata	7920
ccaattcagg tcatgcaatg ccggaatttc cgggttattg tgggtgtgca aatccgagcc	7980
tgccggttat taaagttaaa gccgttacca tgcgcaataa cgcaattctg caaacctgg	8040
ttggtccggg tgaagaacat accaccctgg caggtctgcc gaccgaagca agcatttga	8100
atgcagtga agcagcaatt ccgggttttc tgcaaatgt ttatgcccat accgcaggcg	8160
gtggtaaatt tcigggattt ctgcaagtga aaaaacgtca gcctgccgat gaaggtcgtc	8220
agggtcaggc agccctgctg gcgctggcaa cctatagcga actgaaaaat atcattctgg	8280
tggatgagga tgtggacatt tttgatagt atgatattct gtgggcaatg accaccgta	8340
tgcagggtga tgttagcatt accaccattc cgggtatttc cggtcacag ctggaccga	8400

gccagacacc ggaatattca ccgagcattc gtggtaatgg tattagctgc aaaaccatct	8460
ttgattgtac cgttccgtgg gcactgaaaa gccattttga acgtgcaccg ttgacgatg	8520
ttgatccgcg tccgttttga cctgaatatt ttgcacgtct ggaaaaaaat cagggcagcg	8580
caaaataagc taataacagg cctgctggta atcgacaggaa tttttatttg gatggatccg	8640
cctacctagc ttccaagaaa gatataccta cagcacaga gcggaaagat gttttgttct	8700
acatccagaa caacctctgc taaaattcct gaaaaatttt gcaaaaagtt gttgacttta	8760
tctacaaggt gtggtataat aatcttaaca acagcaggac gctcccgggt tgaggaaaac	8820
ctaatagaaac tgaccagcct gcgtgttagc ctgctggcac tgggtctggt taccagcgg	8880
ttgacgacag cagaaaccta taccgttgat cgttatcagg atgatagcga aaaaggtagc	8940
ctgcgttggg caattgaaca gagcaatgca aatagcgac aagaaaacca gattctgatt	9000
caggcagttg gtaaagcacc gtatgttacc aaagttagata aaccgctgcc tccgattaaa	9060
agcagcgtta aaatcattgg caccgagtgg gataaaaccg gtgaatttat tgcaattgat	9120
ggcagcaact atatcaaagg cgaaggtaga aaagcatgac cgggtgcaaa tccgggtcag	9180
tatggcacca atgttctac catgacctg cctggtctgg ttctgcaaga tgttaatggt	9240
gttacctga aaggtctgga tgttcatcgt tttgtattg gtgttctggt taatcgacg	9300
agcaataacc tgattcagca taatcgtatc agcaacaatt atggtggtgc cgggtttatg	9360
attaccggtg atgatggtaa aggtaatccg accagcacca ccaccaataa taacaaagtt	9420
ctggataacg tgttcatcga taatggtagt ggtctggaac tgaccctgg tgcagcattt	9480
aatctgattg caaataacct gtttaccagc acaaaagcca atccggaacc gagccagggt	9540
attgaaattc tgtggggtaa tgataatgcc gtggtgggta acaaatcga aaactattca	9600
gatggcctgc aaatcaattg gggtaaacgt aactatatcg cctataacga actgaccaat	9660
aacagcctgg gtttcaatct gacaggtgat ggtaacattt tcgacagcaa taaagtgc	9720
ggtaaccgta ttggtattgc cattcgtagt gaaaaagatg ccaatgcacg tattaccctg	9780
acaaaaaac agatttgga taacggcaaa gatataaac gttgtgaagc cgggtgtagc	9840
tgtgttccga atcagcgtct gggtagcaatt gtttttggtg ttccggcact ggaacatgaa	9900
ggttttgttg gtagccgtgg cggtaggtgtt gttattgaac cggcaaaact gcaaaaaacc	9960
tgacccagc cgaaccagca gaattgtaat gcaattccta atcagggtat tcaggcacccg	10020
aaactgacag ttagcaaaaa acagctgacc gttgaagtta aaggcacccc taatcagcgt	10080
tataatgtgg aattttttgg caatcgtaat gccagcagca gcgaagcaga acagtatctg	10140

ggtagcattg ttgttggtac cgatcatcag ggctctggcaa aagcaaattg ggctccgaaa 10200
 gtttagcatgc cgagcggttac cgcaaatgtg acagatcatc tgggtgacgac cagcgaactg 10260
 agcagcgcag ttaaaatgcg ttaaatgcat gcgcgccgcg ttcgcgcggc gctttttttt 10320
 ggtaccgagc tcgaattcac tggccgtcgt tttaacaacgt cgtgactggg aaaaccctgg 10380
 cgttacccaa cttaatgcgc ttgcagcaca tcccccttc gccagctggc gtaatagcga 10440
 agaggccgc accgatgcgc ctcccaaca gttgcgcagc ctgaatggcg aatggcgctt 10500
 gatgcggtat ttctcctta cgcactgtg cggtatttca caccgcatat ggtgcactct 10560

 cagtacaatc tgctctgatg ccgcatagt t aagccagccc cgacaccgc caacaccgc 10620
 tgacgaattc 10630
 <210> 25
 <211> 6452
 <212> DNA
 <213> Escherichia coli
 <220><221> misc_feature
 <222> (1)..(6452)
 <223> DNA sequence of the plasmid pMG39
 <400> 25

 gttgacagta agacgggtaa gcctgttgat gataccgtg ccttactggg tgcattagcc 60
 agtctgaatg acctgtcacg ggataatccg aagtggtcag actggaaaat cagagggcag 120
 gaactgctga acagcaaaaa gtcagatagc accacatagc agaccgcga taaaacgccc 180

 tgagaagccc gtgacgggct tttcttgat tatgggtagt ttccttgcat gaatccataa 240
 aaggcgctg tagtgccatt taccgccatt cactgccaga gccgtgagcg cagcgaactg 300
 aatgtcacga aaaagacagc gactcagggt cctgatggc ggagacaaaa ggaatattca 360
 gcgatttgcc cgagcttgcc aggggtgctac ttaagccttt aggggtttta ggtctgtttt 420
 gtagaggagc aaacagcgtt tgcgacatcc ttttgtaata ctgcggaact gactaaagta 480
 gtgagttata cacagggtg ggatctattc tttttatctt tttttattct tttttattc 540
 tataaattat aaccattga atataaaca aaaaaacaca caaaggtcta gcggaattta 600

 cagagggtct agcagaattt acaagttttc cagcaaaggc ctagcagaat ttacagatac 660
 ccacaactca aaggaaaagg actagtaatt atcattgact agcccatctc aattggtata 720
 gtgattaaaa tcacctagac caattgagat gtatgtctga attagttgtt ttcaaagcaa 780
 atgaactagc gattagtcgc tatgacttaa cggagcatga aaccaagcta attttatgct 840
 gtgtggcact actcaacccc acgattgaaa acctacaag gaaagaacgg acggtatcgt 900

tcacttataa ccaatacgtc cagatgatga acatcagtag ggaaaatgct tatgggtgat	960
tagctaaagc aaccagagag ctgatgacga gaactgtgga aatcaggaat cctttggtta	1020
aaggctttga gattttccag tggacaaaact atgccaagtt ctcaagcgaa aaattagaat	1080
tagtttttag tgaagagata ttgccttatac tttccagtt aaaaaattc ataaaatata	1140
atctggaaca tgtaagtct tttgaaaaca aatactctat gaggatttat gagtggttat	1200
taaaagaact aacacaaaag aaaactcaca aggcaaatat agagattagc cttgatgaat	1260
ttaagttcat gttaatgctt gaaaataact accatgagtt taaaaggctt aaccaatggg	1320
ttttgaaacc aataagtaaa gatttaaaaca cttacagcaa tatgaaattg gtggttgata	1380
agcgaggccg cccgactgat acgttgattt tccaagtga actagataga caaatggatc	1440
tcgtaaccga acttgagaac aaccagataa aaatgaatgg tgacaaaata ccaacaacca	1500
ttacatcaga ttctaccta cgtaacggac taagaaaaac actacacgat gctttaactg	1560
caaaaattca gctcaccagt tttgaggcaa aatttttgag tgacatgcaa agtaagcatg	1620
atctcaatgg ttctttctca tggctcacgc aaaaacaacg aaccacacta gagaacatac	1680
tggctaaata cggaaggatc tgaggttctt atggctcttg tatctatcag tgaagcatca	1740
agactaaca acaaaagtag aacaactgtt caccgttaga tatcaaaggg aaaactgtcc	1800
atatgcacag atgaaaacgg tgtaaaaaag atagatacat cagagctttt acgagttttt	1860
ggtgcattta aagctgttca ccatgaacag atcgacaatg taacagatga acagcatgta	1920
acacctaata gaacaggta aaccagtaaa acaagcaac tagaacatga aattgaacac	1980
ctgagacaac ttgttacagc tcaacagtca cacatagaca gcctgaaaca ggcgatgctg	2040
cttatcgaat caaagctgcc gacaacacgg gagccagtga cgcctcccgt ggggaaaaaa	2100
tcatggcaat tctggaagaa atagcgcttt cagccggcaa acctgaagcc ggatctgcga	2160
ttctgataac aaactagcaa caccagaaca gcccgtttgc gggcagcaaa acccgttatg	2220
cttgtaaacc gttttgtgaa aaaattttta aaataaaaaa ggggacctct aggggtccca	2280
attaattagt aatataatct attaaagtc attcaaaagg tcatccaccg gatcaattcc	2340
cctgctcgcg caggctgggt gccaaagctct cgggtaacat caaggcccga tccttgagc	2400
ccttgccctc ccgcacgatg atcgtgccgt gatcgaaac cagatccttg acccgagtt	2460
gcaaaccctc actgatccgc atgcccgttc catacagaag ctgggcgaac aaacgatgct	2520
cgctttccag aaaaccgagg atgcgaacca cttcatccgg ggtcagcacc accggcaagc	2580
gccgcgacgg ccgaggtctt ccgatctcct gaagccaggg cagatccgtg cacagcacct	2640

tgccgtagaa gaacagcaag gccgccaatg cctgacgatg cgtggagacc gaaaccttgc	2700
gctcgttcgc cagccaggac agaaatgcct cgacttcgct gctgccaag gttgccgggt	2760
gacgcacacc gtggaaacgg atgaaggcac gaacccagtg gacataagcc tgttcggttc	2820
gtaagctgta atgcaagtag cgtatgcgct cagcgaactg gtccagaacc ttgaccgaac	2880
gcagcggtag taacggcgca gtggcggttt tcatggcttg ttatgactgt tttttgggg	2940
tacagtctat gcctcgggca tccaagcagc aagcgcgtta cgccgtgggt cgatgtttga	3000
tgttatggag cagcaacgat gttacgcagc agggcagtcg ccctaaaaca aagttaaaca	3060
tcatgagga agcgggtgac gccgaagtat cgactcaact atcagaggta gttggcgta	3120
tcgagcgcca tctcgaaccg acgttgctgg ccgtacattt gtacggctcc gcagtggatg	3180
gcggcctgaa gccacacagt gatattgatt tgcgtgttac ggtgaccgta aggcttgatg	3240
aaacaacgcg gcgagctttg atcaacgacc ttttgaaac ttcggcttcc cctggagaga	3300
gcgagattct ccgcgctgta gaagtcacca ttgttgtgca cgacgacatc attccgtggc	3360
gttatccagc taagcgcgaa ctgcaatttg gagaatggca gcgcaatgac attcttcag	3420
gtatcttga gccagccagc atcgacattg atctggtat cttgtcgaca aaagcaagag	3480
aacatagcgt tgccttggtg ggtccagcgg cggaggaact cttgatccg gttcctgaac	3540
aggatctatt ttaggcgcta aatgaaacct taacgctatg gaactcgccg cccgactggg	3600
ctggcgatga gcgaaatgta gtgcttacgt tgtccgcat ttggtacagc gcagtaaccg	3660
gcaaaatcgc gccgaaggat gtgcgtgccg actgggcaat ggagcgctg ccggcccagt	3720
atcagcccgt catacttgaa gctagacagg cttatcttgg acaagaagaa gatcgcttgg	3780
cctcgcgcg agatcagttg gaagaatttg tcactacgt gaaaggcgag atcaccaagg	3840
tagtcggcaa ataattgcta acaattcgtt caagccgacg ccgcttcgcg gcgcggctta	3900
actcaagcgt tagatgcact aagcacataa ttgctcacag ccaaactatc aggtcaagtc	3960
tgcttttatt atttttaagc gtgcataata agccctacac aaattgggag atatatcatg	4020
aaagctggc tttttcttgt tatcgcaata gttggcgaag taatcgcaac atccgcatta	4080
aaatctagcg agggctttac taagctgac cggtggatga cttttgaat gacctttaat	4140
agatttatatt actaattaat tggggaccct agaggtcccc tttttatatt taaaaatatt	4200
ttcacaaaac ggtttacaag catacgttgg ccgattcatt aatgcagctg gcacgacagg	4260
tttcccgaact ggaaagcggg cagtgcgcgc aacgcaatta atgtgagttg gtcactcat	4320
taggcacccc aggcctttaca ctttatgctt ccggctcgta tgttgtgtgg aattgtgagc	4380

ggataacaat ttcacacagg aaacagctat gaccatgatt acgccaagct tgcctgcctg	4440
caggtcgact ctagaggatc cgcctacctt gcttccaaga aagatctcct aacagcacia	4500
gagcggaag atgttttgtt ctacatccag aacaacctct gctaaaattc ctgaaaaatt	4560
ttgcaaaaag ttgttgactt tatctacaag gtgtgggtata ataattctaa caacagcagg	4620
acgtccccg gttaggagaa acctaatgaa actgaccagc ctgcgtgtta gcctgctggc	4680
actgggtctg gttaccagcg gttttgcagc agcagaaacc tataccgttg atcggtatca	4740
ggatgatagc gaaaaaggta gcctgcgttg ggcaattgaa cagagcaatg caaatagcgc	4800
acaagaaaac cagattctga ttacaggcagt tggtaaagca ccgtatgtta tcaaagtga	4860
taaaccgctg cctccgatta aaagcagcgt taaaatcatt ggccaccagt gggataaac	4920
cggatgaattt attgcaattg atggcagcaa ctatatcaa ggccaagggtg aaaaagcatg	4980
tccgggtgca aatccgggtc agtatggcac caatgttcgt accatgaccc tgcctggtct	5040
ggttctgcaa gatgttaatg gtgttaccct gaaaggctcg gatgttcac gttttgtat	5100
tgggtgtctg gtaaatcgca gcagcaataa cctgattcag cataatcgta tcagcaacaa	5160
ttatgggtgt gccggtgtta tgattaccgg tgatgatgtt aaaggtaatc cgaccagcac	5220
caccaccaat aataacaaag ttctggataa cgtgttcac gataatgggt atggctctga	5280
actgaccctg ggtgcagcat ttaatctgat tgcaataaac ctgtttacca gcacaaaagc	5340
caatccggaa ccgagccagg gtattgaaat tctgtgggtt aatgataatg ccgtggtggg	5400
taacaaattc gaaaactatt cagatggcct gcaaatcaat tggggttaac gtaactatat	5460
cgctataac gaactgacca ataacagcct gggtttcaat ctgacaggtg atggtaacat	5520
tttcgacagc aataaagtgc atggtaaccg tattgggtatt gccattcgta gtgaaaaaga	5580
tgccaatgca cgtattacc tgaccaaaaa tcagatttgg gataacggca aagatatcaa	5640
acgttgtgaa gccggtggta gctgtgttcc gaatcagcgt ctgggtgcaa ttgtttttgg	5700
tgttccggca ctggaacatg aaggttttgt tggtagccgt ggccggtgtg ttgttatga	5760
accggcaaaa ctgcaaaaaa cctgcacca gccgaaccag cagaattgta atgcaattcc	5820
taatcagggt attcaggcac cgaaactgac agttagcaaa aaacagctga ccgttgaagt	5880
taaaggcacc cctaactcag gttataatgt ggaatttttt ggcaatcgta atgccagcag	5940
cagcgaagca gaacagtatc tgggtagcat tgttgttgt accgatcatc agggctctggc	6000
aaaagcaaat tgggtctcga aagttagcat gccgagcgtt accgcaaatg tgacagatca	6060
tctgggtgca accagcgaac tgagcagcgc agttaaaatg cgttaaatgc atgcgcgccg	6120
cgttcgcgcg gcgttttttt ttggtagcga gctcgaattc actggcgcgc gttttacaac	6180
gtcgtgactg ggaaaacctt ggcgttacc aacttaatcg ccttcagca catccccctt	6240

tcgccagctg gcgtaatagc gaagaggccc gcaccgatcg cccttcccaa cagttgcgca	6300
gcctgaatgg cgaatggcgc ctgatgcggt attttctcct tacgcatctg tgcggtat	6360
cacaccgcat atggtgcaact ctcagtacaa tctgctctga tgccgcatag ttaagccagc	6420
cccgcacccc gccaacaccc gctgacgaat tc	6452

<210> 26
 <211> 10012
 <212> DNA
 <213> Escherichia coli
 <220><221> misc_feature
 <222> (1)..(10012)
 <223> DNA sequence of the plasmid pMG47
 <400> 26

gttgacagta agacgggtaa gcctgttgat gataccgctg cttactggg tgcattagcc	60
agtctgaatg acctgtcacg ggataatccg aagtgggtcag actggaaaat cagagggcag	120
gaactgctga acagcaaaaa gtcagatagc accacatagc agaccgcga taaaacgccc	180
tgagaagccc gtgacgggct tttctgtat tatgggtagt ttccttgcag gaatccataa	240
aaggcgctg tagtgccatt taccgccatt cactgccaga gccgtgagcg cagcgaactg	300

aatgtcacga aaaagacagc gactcaggtg cctgatggc ggagacaaaa ggaatattca	360
gcgatttgcc cgagcttgcc aggggtgtac ttaagccttt aggggtttta ggtctgtttt	420
gtagaggagc aaacagcggt tgcgacatcc ttttgtaata ctgcggaact gactaaagta	480
gtgagttata cacagggctg ggatctattc tttttatct tttttattct tttttattc	540
tataaattat aaccacttga atataaaca aaaaaacaca caaaggtcta gcggaattta	600
cagagggctc agcagaattt acaagttttc cagcaaaggt ctagcagaat ttacagatac	660
ccacaactca aaggaaaagg actagtaatt atcattgact agcccatctc aatttgtata	720

gtgattaaaa tcacctagac caattgagat gtatgtctga attagtgtt ttcaaagcaa	780
atgaactagc gattagtgc tatgacttaa cggagcatga aaccaagcta attttatgct	840
gtgtggcact actcaacccc acgattgaaa accctacaag gaaagaacgg acggtatcgt	900
tcacttataa ccaatacgt cagatgatga acatcagtag ggaaaatgct tatggtgtat	960
tagctaaagc aaccagagag ctgatgacga gaactgtgga aatcaggaat cctttggtta	1020
aaggctttga gatgttccag tggacaaact atgccaagtt ctcaagcgaa aaattagaat	1080
tagtttttag tgaagagata ttgccttatt tttccagtt aaaaaattc ataaaatata	1140

atctggaaca tgtaagtct tttgaaaaca aatactctat gaggatttat gagtggttat	1200
taaaagaact aacacaaaaa aaaactcaca aggcaaatat agagatttagc cttgatgaat	1260
ttaagttcat gttaatgctt gaaaataact accatgagtt taaaaggctt aaccaatggg	1320
ttttgaaacc aataagtaaa gattttaaaca cttacagcaa tatgaaattg gtgggtgata	1380
agcgaggccg cccgactgat acgttgattt tccaagtga actagataga caaatggatc	1440
tcgtaaccga acttgagaac aaccagataa aatgaatgg tgacaaaata ccaacaacca	1500
ttacatcaga ttcttaccta cgtaacggac taagaaaaac actacacgat gctttaactg	1560
caaaaattca gctcaccagt tttgaggcaa aatTTTTgag tgacatgcaa agtaagcatg	1620
atctcaatgg ttctgttctca tggctcacgc aaaaacaacg aaccacacta gagaacatac	1680
tggctaaata cggaaggatc tgaggttctt atggctcttg tatctatcag tgaagcatca	1740
agactaaca acaaaagtag aacaactgtt caccgttaga tatcaaaggg aaaactgtcc	1800
atatgcacag atgaaaacgg tgtaaaaaag atagatacat cagagctttt acgagttttt	1860
ggtgcattta aagctgttca ccatgaacag atcgacaatg taacagatga acagcatgta	1920
acacctaata gaacaggtga aaccagtaaa acaaagcaac tagaacatga aattgaacac	1980
ctgagacaac ttgttacagc tcaacagtca cacatagaca gcctgaaaca ggcatgctg	2040
cttatcgaat caaagctgcc gacaacacgg gagccagtga cgcctcccggt ggggaaaaaa	2100
tcatggcaat tctggaagaa atagcgcttt cagccggcaa acctgaagcc ggatctgcga	2160
ttctgataac aaactagcaa caccagaaca gcccgtttgc gggcagcaaa acccgttatg	2220
cttgtaaacc gttttgtgaa aaaattttta aaataaaaaa ggggacctct agggccccca	2280
attaattagt aatataatct attaaaggtc attcaaaagg tcatccaccg gatcaattcc	2340
cctgctcgcg caggctgggt gccaaactct cgggtaacat caaggcccgga tccttgagc	2400
ccttgccctc ccgcacgatg atcgtgccgt gatcgaaac cagatccttg acccgcagtt	2460
gcaaaccctc actgatccgc atgccgttc catacagaag ctgggcgaac aaacgatgct	2520
cgccttcag aaaaccgagg atgcgaacca ctcatccgg ggtcagcacc accggaagc	2580
gccgcgacgg ccgaggtctt ccgatctctt gaagccaggg cagatccgtg cacagcacct	2640
tgccgtagaa gaacagcaag gccccaatg cctgacgatg cgtggagacc gaaaccttgc	2700
gctcgttcgc cagccaggac agaaatgcct cgacttcgct gctgccaag gttgccgggt	2760
gacgcacacc gtggaaacgg atgaaggcac gaaccagtg gacataagcc tgttcggttc	2820
gtaagctgta atgcaagtag cgtatgcgt cagcgaactg gtccagaacc ttgaccgaac	2880
gcagcggtgg taacggcgca gtggcggttt tcatggcttg ttatgactgt tttttgggg	2940
tacagtctat gcctcgggca tccaagcagc aagcgcttga cgccgtgggt cgatgtttga	3000

tgttatggag cagcaacgat gttacgcagc agggcagtcg ccctaaaaca aagttaaaca	3060
tcatgagggg agcgggtgat gccgaagtat cgactcaact atcagaggta gttggcgta	3120
tcgagcgcca tctcgaaccg acgttgctgg ccgtacattt gtacggctcc gcagtggatg	3180
gcggcctgaa gccacacagt gatattgatt tgcctggttac ggtgaccgta aggcttgatg	3240
aaacaacgcg gcgagctttg atcaacgacc ttttggaac ttcggcttcc cctggagaga	3300
gcgagattct ccgcgtgta gaagtcacca ttgttgtgca cgacgacatc attccgtggc	3360
gttatccagc taagcgcgaa ctgcaatttg gagaatggca gcgcaatgac attcttgag	3420
gtatcttcga gccagccagc atcgacattg atctggctat cttgctgaca aaagcaagag	3480
aacatagcgt tgccttggtg ggtccagcgg cggaggaact cttgatccg gttcctgaac	3540
aggatctatt tgaggcgcta aatgaaacct taacgctatg gaactcgccg cccgactggg	3600
ctggcgatga gcgaaatgta gtgcttacgt tgtcccgcac ttggtacagc gcagtaaccg	3660
gcaaaatcgc gccgaaggat gtcgtgccg actgggcaat ggagcgcctg ccggcccagt	3720
atcagcccgt catacttgaa gctagacagg cttatcttgg acaagaagaa gatcgcttgg	3780
cctcgcgcgc agatcagttg gaagaatttg tccactacgt gaaaggcgag atcaccaagg	3840
tagtcggcaa ataattgtcta acaattcgtt caagccgacg ccgcttcgcg gcgcggctta	3900
actcaagcgt tagatgcaat aagcacataa ttgctcacag ccaaactatc aggtcaagtc	3960
tgcttttatt atttttaagc gtgcataata agccctacac aaattgggag atatatcatg	4020
aaaggctggc tttttcttgt tatcgcaata gttggcgaag taatcgcaac atccgcatta	4080
aaatctagcg agggctttac taagctgac cgttgatga ccttttgaat gaccttaat	4140
agattatatt actaatfaat tggggaccct agaggctccc tttttatatt taaaaatttt	4200
ttcacaaaac ggtttacaag catacgttgg ccgattcatt aatgcagctg gcacgacagg	4260
tttcccgact ggaaagcggg cagtgagcgc aacgcaatta atgtgagttg gctcactcat	4320
taggcacccc aggcctttaca ctttatgctt ccggctcgta tgttgtgtgg aattgtgagc	4380
ggataacaat ttacacaggg aaacagctat gaccatgatt acgccaagct tgctattgac	4440
gacagctatg gttcactgtc caccaaccaa aactgtgctc agtaccgcca atatttctcc	4500
cttgaggggt acaaagaggt gtcctagaa gagatccacg ctgtgtaaaa attttacaaa	4560
aaggtattga ctttccctac aggggtgtgta ataatttaat tacaggcggg ggcaaccccg	4620
cctgttctgc agaggaggaa tatagccatg gaagtgtaaa tcttcaacac ccaggatgtt	4680
caggattttc tgcgtgttgc aagcggctctg gaacaagagg gtggtaatcc gcgtgttaaa	4740
caaattatc atcgtgttct gagcgacctg tataaagcaa ttgaagatct gaatatcacc	4800
agcgacgaat attgggcagg cgttgcatac ctgaatcagc tgggtgcaaa tcaagaagca	4860

ggctctgctga gtccgggtct gggttttgat cattatctgg atatgcgtat ggatgcagaa	4920
gatgcagcac tgggtattga aaatgcaaca ccgcgtacca ttgaaggtcc gctgtatgtt	4980
gcgggtgcac cggaagcgt tggttatgca cgcattgatg atggtagcga tccgaatggt	5040
cataacctga ttctgcatgg caccattttt gatgcagatg gtaaacgcgt gccgaatgca	5100
aaagttgaaa ttggcatgc aaacacaaaa ggcttttata gccattttga tccgaccggt	5160
gaacagcagg cctttaatat gcgtcgtagc attattaccg atgagaatgg tcagtatcgt	5220
gttcgtacca ttctgcctgc cggttatggt tgcctccgg aaggltccgac ccagcaactg	5280
ctgaaccaac tgggtcgtca tggtaatcgt ccggcacata ttcattatit tgttagcgca	5340
gatggtcacc gtaaacigac caccagatt aatgttccg gtgatccgta tacctatgat	5400
gattttgcat atgccaccg tgaaggtctg gttgttgatg cagtgaaca taccgatccg	5460
gaagcaatta aagccaatga tgtggaaggt ctttttccg aaatggtgtt tgatctgaaa	5520
ctgaccctgc tggttgatgg tgttgataat caggttgtgg atcgtccgctg tctggcagtt	5580
taatacacca aaatggttca aaattatcag gcgagtgatc atgatcactg gcctgttttt	5640
atttcaggga aggggtggaga caattacgtg gataatcaga tcatccaaga aaccgtggat	5700
aaaattctga gcgttctgcc gaatcaggca ggtcagctgg cagctctggt gcgtctgatg	5760
caatttgcac gcgatccgac cattaccgtt attggcaaat ataaccatgg taaaagccgt	5820
ctgctgaatg aactgattgg caccgataac tttagcgttg cagataaacg tgaaaccatt	5880
cagctggccg aacataaaca ggatcaggtt cgttggctgg atgcacctgg tctggatgcc	5940
gatgttgcat cagttagatg tcgtcatgca tttgaagcag tttggacca ggcagatatt	6000
cgtctgtttg ttcatagcgt tcgtgaaggt gaactggatg caaccgaaca ccatctgctg	6060
caacagctga ttgaagatgc cgatcatagc cgtcgtcaga ccattctggt tctgaccag	6120
attgatcaga ttccggatca gaccatcctg acacagatta aaaccagcat tgcacagcag	6180
gttccgaaac tggatatttg ggcagttagc gcaaccctgc atcgtcaggg cattgaaaac	6240
ggtaaaaccc tgctgatcga aaaaagcgggt attggtgcac tgcgccatac cctggaacag	6300
gcactggcac aggtgccgag cgcacgtacc tatgaaaaaa atcgtctgct gtcagatctg	6360
caccatcagc tgaacaact gctgctggat cagaaacatg ttctgcaaca actgcaacag	6420
acacagcaac agcagctgca tgattttgat accggtctga ttaacattct ggacaaaatt	6480
cgtgttgatc tggaaccgat tgtgaatatt gatggtcagg atcaagcact gaatccgat	6540
agctttgcaa ccatgtttaa aaacaccgca gcaaaacagc agcgtgccaa agttcagatt	6600

gcataatagcc gtgcatgcat tgaaatcaac agccatctga ttcgccatgg tgttgttgg	6660
ctgcctgcgg aacagcagac caccattaaa agcattgata ccgtgattgt tgccgtgttt	6720
ggatcagcg ttaaatttcg tgatcagctg cgtgccctgt tttataccga taccgaacgt	6780
cagcgtctgc aacgtgaatt tcgtttctat ttgaaaaaa gtgccggtcg catgattctg	6840
gcagcaaaaa ttgaacagac catgcgtcag cagggtgta ttcagaatgc catgatggca	6900
ctgcaacaaa tggaaagcg agcataaaaa cacggacgcc gcaaaccggc tccgaatttc	6960
ttggtcgacc gttaaatcta tcaccgcaag ggataaatat ctaacaccgt gcgtgttgac	7020
tattttacct ctggcgggta taatggttgc atgtactaat ctagataagg aatatagcca	7080
tgaccgcacc gattcaggat ctgcgtgatg caattgccct gctgcaacag catgataatc	7140
agtatctgga aaccgatcat ccggttgatc cgaatgcaga actggcaggc gtttatcgtc	7200
atattggtgc cgggtggcacc gttaaacgtc cgaccgtat tgggtccggca atgatgttta	7260
ataacattaa aggttatccg cacagccgta tttcgttgg tatgcatgca agccgtcagc	7320
gtgcagcact gctgctgggt tgtgaagcaa gtcagctggc actggaagt ggtaaagcag	7380
ttaaaaaac gggtgcaccg gtggttgttc cggcaagcag cgcaccgtgt caagagcaga	7440
tttttctggc agatgatccg gatattgatc tgcgtaccct gctgcctgca cataccaata	7500
ccccgattga tgcaggctccg ttttttctg tgggtctggc cctggcaagc gatccggtgg	7560
atgcaagcct gaccgatgtt accattcatc gtctgtgtgt tcagggtcgt gatgaactga	7620
gcatgttctt ggcagcaggt cgccatattg aagtttttcg tcagaaagca gaagcagcag	7680
gtaaaccgct gccgattacc attaatatgg gtctggacc agcaatctat attggcgcgt	7740
gttttgaagc accgaccacc ccgtttggtt ataatgaact ggggtgttgc ggtgcactgc	7800
gtcagcgtcc ggttgaactg gttcagggtg ttagcgttcc ggaaaaagca attgcacgtg	7860
ccgaaattgt tattgaaggt gaactgctgc ctggtgttcg tgttcgtgaa gatcagcata	7920
ccaattcagg tcatgcaatg ccggaatttc cgggttatgt tgggtgtgca aatccgagcc	7980
tgccggttat taaagttaaa gccgttacca tgcgcaataa cgcaattctg caaaccttg	8040
ttggtccggg tgaagaacat accaccctgg cagggtctgc gaccgaagca agcatttga	8100
atgcagttag agcagcaatt ccgggttttc tgcaaatgt ttatgccat accgcaggcg	8160
gtggtaaatt tctgggtatt ctgcaagtga aaaaacgtca gcctgccgat gaaggtcgtc	8220
agggtcaggc agccctgctg gcgctggcaa cctatagcga actgaaaaat atcattctgg	8280
tggatgagga tgtggacatt tttgatagt atgatattct gtgggcaatg accaccgta	8340
tgcagggtga tgttagcatt accaccattc cgggtattcg cggtcacag ctggaccga	8400
gccagacacc ggaatatca ccgagcattc gtggtaatgg tattagctgc aaaaccatct	8460

ttgattgtac cgttccgtgg gcactgaaaa gccattttga acgtgcaccg tttgcagatg	8520
ttgatccgcg tccgttttga cctgaatatt ttgcacgtct ggaaaaaaat cagggcagcg	8580
caaaataagc taataacagg cctgctggta atcgcaggaa tttttatttg gatggatccg	8640
cctacctagc ttccaagaaa gatacctaa cagcacaaga gcggaaagat gttttgttct	8700
acatccagaa caacctctgc taaaattcct gaaaaatfff gcaaaaagtt gttgacttta	8760
tctacaaggt gtggtataat aatcttaaca acagcaggac gctcccgggt tgaggaaaac	8820
ctaataaat atagcctgtg caccattagc ttctgtcacc agctgattag ctttaccgat	8880
attgttcagt ttgcctatga aaacggcttt gaaggatttg aactgtgggg cacccatgca	8940
cagaatctgt atatgcaaga atatgaaacc accgaacgtg aactgaattg cctgaaagat	9000
aaaaccctgg aaattacat gatcagcgat tatctggata ttagcctgag cgcagatfff	9060
gaaaaaacca tcgaaaaatg tgaacagctg gcaattctgg ccaattggtt taaaacgaac	9120
aaaattcgta cctttgccgg tcagaaaggt agtgcagatt ttagccagca agaacgtcaa	9180
gagtatgtga atcgtattcg catgatttgt gaactgtttg cccagcataa tatgtatgtt	9240
ctgctggaaa cccatccgaa tacctgacc gatacctgc cgagcaccct ggaactgctg	9300
ggtgaagttg atcatccgaa tctgaaaatc aacctggatt ttctgcatat ctgggaaagc	9360
ggtgcagatc cggttgatag ctttcagcag ctgcgtccgt ggattcagca ttatcacttt	9420
aaaaacatta gcagcgaga ctatctgcat gtgtttgaac cgaataatgt ttatgcagca	9480
gcaggtaatc gtaccggtat ggttccgctg tttgaaggca ttgttaacta tgatgaaatc	9540
atccaagaag tgcgcgatac cgatcatfff gcaagcctgg aatggtttgg tcataacgca	9600
aaagatatte tgaaagccga aatgaaagt ctgaccaatc gtaatctgga agttgttacc	9660
agctaaatgc atgcgcgccg cgttcgcgcg gcgctttttt ttggtaccga gctcgaatc	9720
actggccgtc gttttacaac gtctgtactg ggaaaaccct ggcgttacc aacttaatcg	9780
ccttgacga catccccctt tcgccagctg gcgtaatagc gaagaggccc gcaccgatcg	9840
cccttcccaa cagttgcgca gcctgaatgg cgaatggcgc ctgatgcggt attttctcct	9900
tacgcatctg tgcggtatft cacaccgat atggtgcact ctcagtacaa tetgctctga	9960
tgccgcatag ttaagccagc cccgacacc gccaacacc gctgacgaat tc	10012
<210> 27	
<211> 10249	
<212> DNA	
<213> Escherichia coli	
<220><221> misc_feature	

<222> (1)..(10249)

<223> DNA sequence of the plasmid pMG70

<400> 27

gttgacagta agacgggtaa gcctgttgat gataccgctg ccttactggg tgcattagcc	60
agtctgaatg acctgtcacg ggataatccg aagtggctcag actggaaaat cagagggcag	120
gaactgctga acagcaaaaa gtcagatagc accacatagc agacccgcca taaaacgccc	180
tgagaagccc gtgacgggct tttcttgtat tatgggtagt ttccttgcat gaatccataa	240
aaggcgctg tagtgccatt taccgccatt cactgccaga gccgtgagcg cagcgaactg	300
aatgtcacga aaaagacagc gactcagggt cctgatggtc ggagacaaaa ggaatattca	360
gcgatttgcc cgagcttgcc aggggtgctac ttaagccttt aggggttttaa ggtctgtttt	420
gtagaggagc aaacagcggt tgcgacatcc ttttgtaata ctgcggaact gactaaagta	480
gtgagttata cacagggctg ggatctattc tttttatctt tttttattct ttcctttattc	540
tataaattat aaccacttga atataaaca aaaaaacaca caaaggtcta gcggaattta	600
cagaggggtct agcagaattt acaagttttc cagcaaaggt ctagcagaat ttacagatac	660
ccacaactca aaggaaaagg actagtaatt atcattgact agcccatctc aattggtata	720
gtgattaaaa tcacctagac caattgagat gtatgtctga attagtgtt ttcaaagcaa	780
atgaactagc gattagtgcg tatgacttaa cggagcatga aaccaagcta attttatgct	840
gtgtggcact actcaacccc acgattgaaa accctacaag gaaagaacgg acggtatcgt	900
tcacttataa ccaatacgtc cagatgatga acatcagtag ggaaaatgct tatggtgtat	960
tagctaaagc aaccagagag ctgatgacga gaactgtgga aatcaggaat cctttggtta	1020
aaggctttga gattttccag tggacaaact atgccaagtt ctcaagcgaa aaattagaat	1080
tagtttttag tgaagagata ttgccttatt ttttccagtt aaaaaattc ataaaatata	1140
atctggaaca tgtaagtct tttgaaaaca aatactctat gaggatttat gagtggttat	1200
taaaagaact aacacaaaag aaaactcaca aggcaaatat agagattagc ctgtatgaat	1260
ttaagttcat gttaatgctt gaaaataact accatgagtt taaaaggctt aaccaatggg	1320
ttttgaaacc aataagtaaa gatttaaaca cttacagcaa tatgaaattg gtggttgata	1380
agcgaggccg cccgactgat acgttgattt tccaagtga actagataga caaatggatc	1440
tcgtaaccga acttgagaac aaccagataa aaatgaatgg tgacaaaata ccaacaacca	1500
ttacatcaga ttcttaccta cgtaacggac taagaaaaac actacacgat gctttaactg	1560
caaaaattca gctcaccagt tttgaggcaa aatttttgag tgacatgcaa agtaagcatg	1620
atctcaatgg ttcgtttctca tggctcacgc aaaaacaacg aaccacacta gagaacatac	1680

tggctaaata cggaaggatc tgaggttctt atggctcttg tatctatcag tgaagcatca	1740
agactaaca acaaaagtag aacaactgtt caccgttaga tatcaaaggg aaaactgtcc	1800
atatgcacag atgaaaacgg tgtaaaaaag atagatacat cagagctttt acgagttttt	1860
ggtgcattta aagctgttca ccatgaacag atcgacaatg taacagatga acagcatgta	1920
acacctaata gaacaggatga aaccagtaaa acaaagcaac tagaacatga aattgaacac	1980
ctgagacaac ttgttacagc tcaacagtca cacatagaca gcctgaaaca ggcgatgctg	2040
cttatcgaat caaagctgcc gacaacacgg gagccagtga cgcctcccggt ggggaaaaaa	2100
tcatggcaat tctggaagaa atagcgcttt cagccggcaa acctgaagcc ggatctgcga	2160
ttctgataac aaactagcaa caccagaaca gcccgtttgc gggcagcaaa acccgttatg	2220
cttgtaaacc gttttgtgaa aaaattttta aaataaaaaa ggggacctct aggggtccca	2280
attaattagt aatataatct attaaaggtc attcaaaagg tcatccaccg gatcaattcc	2340
ctgtctcgcg caggctgggt gccaaagtct cgggtaacat caaggcccga tccttggagc	2400
ccttgccctc ccgcacgatg atcgtgccgt gatcgaaatc cagatccttg acccgcagtt	2460
gcaaaccctc actgatccgc atgcccgttc catacagaag ctggcgcaac aaacgatgct	2520
cgccttcag aaaaccgagg atgcgaacca ctcatccgg ggtcagcacc accggcaagc	2580
gccgcgacgg ccgaggtctt ccgatctect gaagccaggg cagatccgtg cacagcacct	2640
tgccgtagaa gaacagcaag gccgccaatg cctgacgatg cgtggagacc gaaaccttgc	2700
gctcgttcgc cagccaggac agaaatgcct cgacttcgct gctgcccag gttgccgggt	2760
gacgcacacc gtggaaacgg atgaaggcac gaacccagtg gacataagcc tgttcggttc	2820
gtaagctgta atgcaagtag cgtatgcgct cagcgaactg gtccagaacc ttgaccgaac	2880
gcagcggtgg taacggcgca gtggcggttt tcatggcttg ttatgactgt ttttttgggg	2940
tacagtctat gcctcgggca tccaagcagc aagcgcgtta cgccgtgggt cgatgtttga	3000
tgttatggag cagcaacgat gttacgcagc agggcagtcg ccctaaaaca aagttaaaca	3060
tcatgaggga agcgggtgac gccgaagtat cgactcaact atcagaggta gttggcgta	3120
tcgagcgcca tctcgaaccg acgttgctgg ccgtacattt gtacggctcc gcagtggatg	3180
gcggcctgaa gccacacagt gatattgatt tgctggttac ggtgaccgta aggcttgatg	3240
aaacaacgcg gcgagctttg atcaacgacc ttttggaac ttcggcttcc cctggagaga	3300
gcgagattct ccgcgctgta gaagtcacca ttgttgta cgcagacatc attccgtggc	3360
gttatccagc taagcgcgaa ctgcaatttg gagaatggca gcgcaatgac attcttcgag	3420
gtatcttcga gccagccagc atcgacattg atctggctat cttgtgaca aaagcaagag	3480
aacatagcgt tgccttggtg ggtccagcgg cggaggaact ctttgatccg gttcctgaac	3540

aggatctatt tgaggcgcta aatgaaacct taacgtatg gaactcgccg cccgactggg	3600
ctggcgatga gcgaaatgta gtgcttacgt tgtcccgcat ttggtacagc gcagtaaccg	3660
gcaaaatcgc gccgaaggat gtcgtgccg actgggcaat ggagcgctg cgggccagt	3720
atcagcccgat catacttgaa gctagacagg cttatcttgg acaagaagaa gatcgcttgg	3780
cctcgcgcg agatcagttg gaagaatttg tccactacgt gaaaggcgag atcaccaagg	3840
tagtcggcaa ataattgtcta acaattcgtt caagccgacg ccgcttcgcg gcgcggctta	3900
actcaagcgt tagatgcaat aagcacataa ttgctcacag ccaactatc aggtcaagtc	3960
tgcttttatt atttttaagc gtgcataata agccctacac aaattgggag atatatcatg	4020
aaaggctggc tttttcttgt tatcgcaata gttggcgaag taatcgcaac atccgcatta	4080
aaatctagcg agggctttac taagctgac cgttgatga cttttgaat gacctttaat	4140
agattatatt actaattaat tggggacct agaggctccc tttttatatt taaaaatttt	4200
ttcacaaaac gggttacaag catacgttgg ccgattcatt aatgcagctg gcacgacagg	4260
tttcccgact ggaaagcggg cagtgcgcg aacgaatta atgtgagta gtcactcat	4320
taggcacccc aggttttaca ctttatgctt ccggctcgta tgttgtgtgg aattgtgagc	4380
ggataacaat ttcacacagg aaacagctat gaccatgatt acgccaagct tgctattgac	4440
gacagctatg gttcactgtc caccaaccaa aactgtgctc agtaccgcca atatttctcc	4500
cttgaggggt acaaagaggt gtccttagaa gagatccacg ctgtgtaaaa attttcaaaa	4560
aaggtattga ctttccctac aggggtgtgta ataatttaac tacaggcgagg ggcaaccccg	4620
cctgttctgc agaggaggaa tatagccatg gaagtgaata tcttcaacac ccaggatgtt	4680
caggattttc tgcgtgttgc aagcggctcg gaacaagagg gtggtaatcc gcgtgttaaa	4740
caaattatc atcgtgttct gacgacacgt tataaagcaa ttgaagatct gaatatcacc	4800
agcgacgaat attgggcagg cggttcatac ctgaatcagc tgggtgcaaa tcaagaagca	4860
ggtctgctga gtccgggtct gggttttgat cattatctgg atatgcgtat ggatgcagaa	4920
gatgcagcac tgggtattga aaatgaaca ccgcgtacca ttgaaggtcc gctgtatgtt	4980
gcgggtgcac cgaaagcgt tggttatgca cgcattgatg atggtagcga tccgaatggt	5040
cataacctga ttctgcatgg caccattttt gatgcagatg gtaaaccgct gccgaatgca	5100
aaagtgaata tttggcatgc aaacacaaaa ggcttttata gccattttga tccgaccggt	5160
gaacagcagg cctttaatat gcgtcgtagc attattaccg atgagaatgg tcagtatcgt	5220
gttcgtacca ttctgcctgc cggttatggt tgcctccgg aaggtccgac ccagcaactg	5280

ctgaaccaac tgggtcgtca tggtaatcgt ccggcacata ttcattatit tgttagcgca	5340
gatggtcacc gtaaactgac caccagatt aatgttgccg gtgatccgta tacctatgat	5400
gattttgcat atgccaccgg tgaaggtctg gttgttgatg cagttgaaca taccgatccg	5460
gaagcaatta aagccaatga tgtggaaggt ccttttgccg aaatgggtgt tgcctgaaa	5520
ctgaccgcgc tgggtgatgg tgttgataat caggttggtg atcgcccgcg tctggcagtt	5580
taatacacca aaatggttca aaattatcag gcgagtgtc atgatcactg gcctgttttt	5640
atctcaggga aggggtggaga caattacgtg gataatcaga tcatccaaga aaccgtggat	5700
aaaattctga gcgttctgcc gaatcaggca ggtcagctgg cacgtctggt gcgtctgatg	5760
caatttgcat gcgatccgac cattaccgtt attggcaaat ataaccatgg taaaagccgt	5820
ctgctgaatg aactgattgg caccgataat tttagcgttg cagataaacg tgaaaccatt	5880
cagctggccg aacataaaca ggatcaggtt cgttggtctg atgcacctgg tctggatgcc	5940
gatgttgcat cagttgatga tctgtatgca tttgaagcag tttggacca ggcagatatt	6000
cgtctgtttg ttcatacgct tctgaaggt gaactggatg caaccgaaca ccatctgtg	6060
caacagctga ttgaagatgc cgatcatagc cgtctgcaga ccattctggt tctgaccag	6120
attgatcaga ttccggatca gaccatcctg acacagatta aaaccagcat tgcacagcag	6180
gttccgaaac tggatatttg ggcagttagc gcaaccgcgc atcgccaggc cattgaaaac	6240
ggtaaaaccg tctgtatcga aaaaagcggg attggtgcac tgcgccatcc cctggaacag	6300
gcactggcac aggtgccgag cgcacgtacc tatgaaaaaa atcgctctgt gtcagatctg	6360
caccatcagc tgaacaact gctgtctggat cagaaacatg tttgcaaca actgcaacag	6420
acacagcaac agcagctgca tgattttgat accggtctga ttaacattct ggacaaaatt	6480
cgtgttgatc tggaaaccgat tgtgaatatt gatggtcagg atcaagcact gaatccggat	6540
agctttgcaa ccatgtttta aaacaccgca gcaaaacagc agcgtgcaa agttcagatt	6600
gcataatgcc gtgcatgcat tgaaatcaac agccatctga ttcgccatgg tgttgttggt	6660
ctgcctgcgg aacagcagac caccattaaa agcattgata ccgtgattgt tgcctgtttt	6720
ggatcagcg ttaaatttcg tgatcagctg cgtgccctgt tttataccga taccgaacgt	6780
cagcgtctgc aacgtgaatt tcttttctat tttgaaaaaa gtgccggtcg catgattctg	6840
gcagcaaaaa ttgaacagac catgcgtcag cagggtctga ttcagaatgc catgatggca	6900
ctgcaacaaa tggaaagcgc agcataaaaa cacggacgcc gcaaacggcg tccgaatttc	6960
ttggtcgacc gttaaatcta tcaccgcaag ggataaatat ctaacaccgt gcgtgttgac	7020
tattttacct ctggcggatg taatggttgc atgtactaat ctagataagg aatatagcca	7080
tgaccgcacc gattcaggat ctgcgtgatg caattgccct gctgcaacag catgataatc	7140

agtatctgga aaccgatcat ccggttgatc cgaatgcaga actggcaggc gtttatcgtc	7200
atattggtgc cggtaggcacc gttaaacgtc cgacccgtat tggtagcgca atgatgttta	7260
ataacattaa aggttatccg cacagccgta ttctggttgg tatgcatgca agccgtcagc	7320
gtgcagcact gctgctgggt tgtgaagcaa gtcagctggc actggaagtt ggtaaagcag	7380
ttaaaaaacc ggttgaccg gtggttgttc cggcaagcag cgcaccgtgt caagagcaga	7440
tttttctggc agatgatccg gattttgatc tgcgtaccct gctgcctgca cataccaata	7500
ccccgattga tgcaggtccg ttttttgtc tgggtctggc cctggcaagc gatccggtgg	7560
atgcaagcct gaccgatgtt accattcatt gtcgtgtgtt tcagggtcgt gatgaactga	7620
gcatgttcct ggcagcaggt cgccatattg aagtttttcg tcagaaagca gaagcagcag	7680
gtaaaccgct gccgattacc attaatatgg gtcgtggacc agcaatctat attggcgcct	7740
gttttgaagc accgaccacc ccgtttggtt ataatagaact ggggtgttgc ggtgcactgc	7800
gtcagcgtcc ggttgaactg gttcagggtg ttagcgttcc ggaaaaagca attgcacgtg	7860
ccgaaattgt tattgaaggt gaactgctgc ctggtgttcg tgttcgtgaa gatcagcata	7920
ccaattcagg tcatgcaatg ccggaatttc cgggttattg tggtagtgca aatccgagcc	7980
tgccggttat taaagttaaa gccgttacca tgcgcaataa cgcaattctg caaaccttgg	8040
ttggtccggg tgaagaacat accaccctgg caggtctgcc gaccgaagca agcatttga	8100
atgcagtga agcagcaatt ccgggttttc tgcaaatgt ttatgcccat accgcaggcg	8160
gtggtaaatt tctgggtatt ctgcaagtga aaaaacgtca gcctgccgat gaaggtcgtc	8220
agggtcaggc agccctgctg gcgctggcaa cctatagcga actgaaaaat atcattctgg	8280
tggatgagga tgtggacatt ttgatagt atgatattct gtgggcaatg accaccgta	8340
tgcagggtga tgttagcatt accaccattc cgggtattcg cggatcatcag ctggaccga	8400
gccagacacc ggaatattca ccgagcattc gtggtaatgg tattagctgc aaaaccatct	8460
ttgattgtac cgttccgtgg gcaatgaaaa gccattttga acgtgcaccg ttgcagatg	8520
ttgatcccg tccgtttgca cctgaatatt ttgcacgtct ggaaaaaat cagggcagcg	8580
caaaataagc taataacagg cctgctggta atcgaggaa tttttatttg gatggatccg	8640
cctacctagc ttccaagaaa gatacctaa cagcacaaga gcgaaagat gttttgttct	8700
acatccagaa caacctctgc taaaattcct gaaaaatttt gcaaaaagtt gttgacttta	8760
tctacaaggt gtggtataat aatcttaaca acagcaggac gctcccggtg tgaggaaaac	8820
ctaagccga gcaaaactggc aattagcagc atgagcctgg gtcgttgttt tgcaggtcat	8880
agcctggata gtaaaactgga tgcagcacag cgttatggtt atctgggtat tgaactgttt	8940
tatgaggatc tggttgatgt tgcagaacat ctgagcaatg aacgtccgag tccggaaggt	9000

ccgtttgttg aagcacagat tgcagcagca cgtcatattc tgcaaatgtg tcaggcacgt 9060

ggctctggaag ttgtttgtct gcaaccgttt atgcattatg atggctctgaa tgatcgtgcc 9120

gaacatgaac gtcgtctgga aaaactggca ctgtggattg aactggcaca tgaactgcat 9180

accgatatta ttcagattcc ggcaaatatt ctgcctgcaa atcaggtttag cgataatctg 9240

gatctgattg ttagcgatct gtgtaaagt gcatatattg gtgcacagc actgcctccg 9300

attcgttttg catatgaaag cctgtgttgg agcacccgtg ttgatctgtg ggaacgttgt 9360

tgggatattg ttcagcgtgt ggatcgtccg aattttggtt tttgtctgga tacctttaac 9420

atcctgggtc gcatttatgc agatccgacc agcccgagcg gtcgtacccc gaatgcaaaa 9480

gaagcagttc gtaaaagcat tgccaatctg gttagccgtg tggatgttag caaagttttt 9540

tatgttcagg ttgtggatgc cgaacgtctg agtaaacgcg tgctgcctgg tcatccgtat 9600

tataaccggg aacagcctgc acgtatgagc tggtcacgta attgtcgtct gttctatggt 9660

gaaaccgaat atggtgcata tctgccgttt aaagaattg cacgcgcact gtttcatggt 9720

attggttttg aaggttgggt tagcctggaa ctgtttaatc gtcgtatgag cgaagaaggt 9780

ccggaagtgc ctgaagaact ggccatgcgt ggtgcaatta gctgggcaaa actggttcag 9840

gatctgcgta ttccggttga aggtccgctg gttaccatgc ctctgtttag cgcaagcctg 9900

taaatgcatg cgcgcgcgt tgcgcggcg ctttttttg gtaccgagct cgaattcact 9960

ggccgtcgtt ttacaacgtc gtgactggga aaacctggc gttaccaac ttaatgcct 10020

tgcagcatat cccctttcg ccagctggcg taatagcgaa gagggccgca ccgatcgccc 10080

ttccaacag ttgcgcagcc tgaatggcga atggcgctg atcggtatt ttctcttac 10140

gcatctgtgc ggtatttcac accgcatatg gtgcactctc agtacaatct gctctgatgc 10200

cgcatagtta agccagcccc gacacccgcc aacaccgct gacgaattc 10249

<210> 28

<211> 7623

<212> DNA

<213> Escherichia coli

<220><221> misc_feature

<222> (1)..(7623)

<223> DNA sequence of the plasmid pCP32AMP

<400> 28

acatgaatgg tcttcggttt ccgtgtttcg taaagtctgg aaacgcggaa gtcagcgccc 60

tgcaccatta tgttccggat ctgcatcgca ggaatgtgct ggctaccctg tggaacacct 120

acatctgtat taacgaagcg ctggcattga ccttgagtga tttttctctg gtcccgccgc	180
atccataccg ccagttgttt accctcaciaa cgttcagta accgggcatg ttcacatca	240
gtaacccgta tcgtgagcat cctctctctg ttcacggtg tcattacccc catgaacaga	300
aattccccct tacacggagg catcaagtga ccaaacagga aaaaaccgcc cttaacatgg	360
cccgctttat cagaagccag acattaacgc ttttgagaa actcaacgag ctggacgcgg	420
atgaacaggc agacatctgt gaatcgcttc acgaccacgc tgatgagctt taccgcagct	480
gcctcgcgcg tttcggatg gacggtgaaa acctctgaca catgcagctc ccggagacgg	540
tcacagcttg tctgtaagcg gatgccggga gcagacaagc ccgtcagggc gcgtcagcgg	600
gtgttggcgg gtgtcggggc gcagccatga cccagtcacg tagcgatagc ggagtgtata	660
ctggcttaac tatgcggcat cagagcagat tgtactgaga gtgcaccata tgcggtgtga	720
aataccgcac agatgcgtaa ggagaaaata ccgcatcagg cgctcttccg cttctctgct	780
cactgactcg ctgcgctcgg tcgttcggct gcggcgagcg gtatcagctc actcaaaggc	840
ggtaatacgg ttatccacag aatcagggga taacgcagga aagaacatgt gagcaaaagg	900
ccagcaaaag gccaggaacc gtaaaaaggc cgcgttctg gcgtttttcc ataggtccg	960
ccccctgac gagcatcaca aaaatcgacg ctcaagtcag aggtggcgaa acccgacagg	1020
actataaaga taccaggcgt tccccctgg aagctccctc gtgcgctctc ctgttccgac	1080
cctgccgctt accggatacc tgtecgctt tctccctcg ggaagcgtgg cgctttctca	1140
tagctcacgc ttaggtatc tcagttcggg ttaggtcgtt cgctccaagc tgggtgtgt	1200
gcacgaaccc cccgttcagc ccgaccgtg cgccttatcc ggtaactatc gtcttgagtc	1260
caaccggta agacacgact tatcgccact ggcagcagcc actggtaaca ggattagcag	1320
agcgaggtat gtaggcggtg ctacagagtt ctgaagtgg tggcctaact acggtacac	1380
tagaaggaca gtatttggtg tctgcgctct gctgaagcca gttaccttcg gaaaaagagt	1440
tggtagctct tgatccggca aacaaccac cgtggttagc ggtggttttt ttgtttgcaa	1500
gcagcagatt acgcgcagaa aaaaaggatc tcaagaagat cctttgatct tttctacggg	1560
gtctgacgct cagtggacg aaaactcacg ttaagggatt ttggtcatga gattatcaaa	1620
aaggatcttc acctagatcc ttttaatta aaaatgaagt tttaatcaa tctaaagtat	1680
atatagtaa acttggctcg acagttacca atgcttaac agtgaggcac ctatctcagc	1740
gatctgtcta tttcgttcac ccatagttgc ctgactcccc gtcgtgtaga taactacgat	1800
acgggagggc ttaccatctg gccccagtgc tgcaatgata ccgcgagacc cacgctcacc	1860
ggctccagat ttatcagcaa taaaccagcc agccggaagg gccgagcgca gaagtggtec	1920

tgcaacttta tccgcctcca tccagtctat taattgttgc cggaagcta gagtaagtag	1980
ttcgccagtt aatagtttgc gcaacgttgt tgccattgct gcaggcatcg tgggtgcacg	2040
ctcgtcgttt ggtatggctt cattcagctc cggttcccaa cgatcaaggc gagttacatg	2100
atcccccatg ttgtgcaaaa aagcggtttag ctcttcgggt cctccgatcg ttgtcagaag	2160
taagtggcc gcagtgttat cactcatggt tatggcagca ctgcataatt ctcttactgt	2220
catgccatcc gtaagatgct tttctgtgac tggtagtac tcaaccaagt cattctgaga	2280
atagtgtatg cggcgaccga gttgtcttgc cccggcgctc acacgggata ataccgcgcc	2340
acatagcaga actttaaaag tgctcatcat tggaaaacgt tcttcggggc gaaaactctc	2400
aaggatctta ccgtgttga gatccagttc gatgtaaccc actcgtgcac ccaactgac	2460
ttcagcatct tttactttca ccagcgttgc tgggtgagca aaaacaggaa ggcaaaatgc	2520
cgcaaaaaag ggaataaggc gcacacggaa atgttgaata ctcatactct tcctttttca	2580
atattattga agcatttacc aggggtattg tctcatgagc ggatacatat ttgaatgtat	2640
ttagaaaaat aaacaaatag ggggtccgcg cacatttccc cgaaaagtgc cacctgacgt	2700
ctaagaaacc attattatca tgacattaac ctataaaaat aggcgtatca cgaggccctt	2760
tcgtcttcaa gaattctgaa ccagtcctaa aacgagtaaa taggaccggc aattcttcaa	2820
gcaataaaca ggaataccaa ttattaaaag ataacttagt cagatcgtac aataaagctt	2880
tgaagaaaaa tgcgccttat tcaatctttg ctataaaaaa tggcccaaaa tctcacattg	2940
gaagacattt gatgacctca tttctttcaa tgaagggcct aacggagtgt actaatgttg	3000
tgggaaattg gagcgataag cgtgtcttgc cgtggccag gacaacgtat actcatcaga	3060
taacagcaat acctgatcac tacttcgcac tagtttctcg gtactatgca tatgatccaa	3120
tatcaaagga aatgatagca ttgaaggatg agactaatcc aattgaggag tggcagcata	3180
tagaacagct aaagggtagt gctgaaggaa gcatacgata ccccgcatgg aatgggataa	3240
tatcacagga ggtactagac tacctttcat cctacataaa tagacgcata taagtacgca	3300
tttaagcata aacacgcact atgccgttct tctcatgtat atatataac aggcaacacg	3360
cagatatagg tgcgacgtga acagttagct gtatgtgcgc agctcgcgtt gcattttcgg	3420
aagcgtcgt tttcggaac gctttgaagt tcctattccg aagttcctat tctctagaaa	3480
gtataggaac ttcagagcgc ttttgaaaac caaaagcgt ctgaagacgc actttcaaaa	3540
aacaaaaaac gcaccggact gtaacgagct actaaaatat tgcgaatacc gcttcacaa	3600
acattgtcga aaagtatctc tttgtatata atctctgtgc tatacccta tataacctac	3660

ccatccacct ttcgctcctt gaacttgcac ctaaactcga cctctacatt ttttatgttt	3720
atctctagta ttactcttta gacaaaaaaa ttgtagtaag aactattcat agagtgaatc	3780
gaaaacaata cgaaaatgta aacatttcct atacgtagta tatagagaca aaatagaaga	3840
aaccgttcat aattttctga ccaatgaaga atcatcaacg ctatcacttt ctgttcacaa	3900
agtatgcga atccacatcg gtatagaata taatcgggga tgcctttatc ttgaaaaaat	3960
gcaccgcgag cttcgctagt aatcagtaaa cgcgggaagt ggagtcaggc ttttttatg	4020
gaagagaaaa tagacaccaa agtagccttc ttctaacctt aacggaccta cagtgcacaa	4080
agttatcaag agactgcatt atagagcgca caaaggagaa aaaaagtaat ctaagatgct	4140
ttgttagaaa aatagcgctc tcgggatgca tttttgtaga acaaaaaaga agtatagatt	4200
ctttgttggc aaaatagcgc tctcgcttg catttctgtt ctgtaaaaat gcagctcaga	4260
ttctttgttt gaaaaattag cgctctcgcg ttgcatTTTT gttttacaaa aatgaagcac	4320
agattcttcg ttggtaaaaat agcgctttcg cgttgcatTT ctgttctgta aaaatgcagc	4380
tcagattctt tgtttgaaaa attagcgctc tcgcgttgca tttttgttct acaaaatgaa	4440
gcacagatgc ttcgttaaca aagatatgct attgaagtgc aagatggaaa cgcagaaaaat	4500
gaaccgggga tgcgacgtgc aagattacct atgcaataga tgcaatagtt tctccaggaa	4560
ccgaaataca tacattgtct tccgtaaagc gctagactat atattattat acaggttcaa	4620
atatactatc tgtttcaggg aaaactccca ggttcggatg ttcaaaattc aatgatgggt	4680
aacaagtacg atcgtaaate tgtaaaacag tttgtcgat attaggtgt atctcctcaa	4740
agcgtattcg aatatcattg agaagctgca gcgtcacatc ggataataat gatggcagcc	4800
attgtagaag tgccttttgc atttctagtc tctttctcgg tctagctagt ttactacat	4860
cgcgaagata gaatcttaga tcacactgcc tttgctgagc tggatcaata gagnaacaaa	4920
agagtggtaa ggcctcgtaa aaggacaagg acctgagcgg aagtgtatcg tacagtagac	4980
ggagtatact agtatagctc atagtccttg gaattctcat gtttgacagc ttatcatcga	5040
taagcttttc aattcaattc atcatTTTT ttttattctt tttttgatt tcggtttctt	5100
tgaaatTTTT ttgattcggc aatctccgaa cagaaggaag aacgaaggaa ggagcacaga	5160
cttagattgg tatatatag catatgtatg gttgaagaaa catgaaattg ccagtatcc	5220
ttaacccaac tgcacagaac aaaaacctgc aggaacgaa gataaatcat gtcgaaagct	5280
acataaagg aacgtgctgc tactcatcct agtctgttg ctgcaagct atttaatatc	5340
atgcacgaaa agcaaacaaa cttgtgtgct tcattggatg ttcgtaccac caaggaatta	5400
ctggagttag ttgaagcatt aggtcccaaa atttgtttac taaaaacaca tgtggatatac	5460
ttgactgatt ttccatgga gggcacagtt aagccgctaa aggcattatc cgccaagtac	5520

aatTTTTtac tcttcgaaga cagaaaattt gctgacattg gtaatacagt caaattgcag	5580
tactctgcgg gtgtatacag aatagcagaa tgggcagaca ttacgaatgc acacggtgtg	5640
gtgggccag gtattgttag cggtttgaag caggcggcag aagaagtaac aaaggaacct	5700
agaggccttt tgatgttagc agaattgtca tgcaagggtc ccctatctac tggagaatat	5760
actaagggtg ctgttgacat tgcgaagagc gacaaagatt ttgttatcgg ctttattgct	5820
caaagagaca tgggtggaag agatgaaggt tacgattggt tgattatgac acccggtgtg	5880
ggttttagatg acaagggaga cgcatgtggg caacagtata gaaccgtgga tgatgtggtc	5940
tctacaggat ctgacattat tattgttgga agaggactat ttgcaaaggg aagggtatgct	6000
aaggtagagg gtgaacgtta cagaaaagca ggctgggaag catatttgag aagatgcggc	6060
cagcaaaact aaaaaactgt attataagta aatgcatgta tactaaactc acaaattaga	6120
gcttcaattt aattatatca gttattacc gggaatctcg gtcgtaatga cttgaaataa	6180
ttaacaaca aaggagtac agttagaaat ttaggagag atctcgtttt tcgcgacaat	6240
ctggcgtttt tcttgctaata tccaggatta atccgttcac agtgtaaaac cccgtttaca	6300
cattctgacg gaagatatag attggaagta ttgcattcac taagataagt atggcaacac	6360
tggaacagac atgaattatc agaacgacga ttacgcac aaagaaatca aagagttact	6420
tcctcctgtc gcattgtcgg aaaaattccc cgctactgaa aatgcccgga atacggttgc	6480
ccatgcccga aaagcgatcc ataagatcct gaaaggtaat gatgatgcc tgttggttgt	6540
gattggccca tgctcaattc atgatcctgt cgcggcaaaa gagtatgcca ctgcttgct	6600
ggcgctgctg gaagagctga aagatgagct ggaaatcgta atgcgctct attttgaana	6660
gccgctgacc acggtgggct ggaaagggt gattaacgat ccgcatatgg ataatagctt	6720
ccagatcaac gacggtctgc gtatagcccg taaattgtcg cttgatatta acgacagcgg	6780
tctgccagcg gcaggtgagt ttctcgatat gatcaccca caatatctcg ctgacctgat	6840
gagctggggc gcaattggcg cacgtaccac cgaatcgag gtgcaccg aactggcatc	6900
agggtttct tgtccggtcg gcttcaaaaa tggcaccgac ggtacgatta aagtggctat	6960
cgatgccatt aatgccgccc gtgcgccgca ctgcttctg tccgtaacga aatgggggca	7020
ttcggcgatt gtgaatacca gcggtaacgg cgattgccat atcattctgc gcggcggtaa	7080
agagcctaac tacagcgca agcacgttgc tgaagtga aaagggtga acaaagcagg	7140
cctgccagca caggtgatga tcgatttcag ccatgctaac tcgtccaaac aattcaaaaa	7200
gcagatggat gtttgtgtc acgtttgcc gcagattgcc ggtggcgaaa aggccattat	7260
tggcgtgatg tggaagacc atctggtgga aggcaatcag agcctcgaga gcggggagcc	7320
gctggcctac ggtaagagca tcaccgatgc ctgcatcggc tgggaagata ccgatgctct	7380

gttacgtcaa ctggcgaatg cagtaaaagc gcgtcgcggg taaggtttaa ttgtcggatg	7440
cgccgtcaga gtggcgtatc cgatgaatca ccacaggcct gataagtcgc gcagcgtcgc	7500
atcaggcaat gtgtccatt gttagcaaca aaaaagccga ctacttgca gtcggctttc	7560
tcattttaaa cgaatgacgt ttacttcgct ttaccctggg ttgcaaccgc cgctgctttc	7620
gct	7623
<210> 29	
<211> 7630	
<212> DNA	
<213> Escherichia coli	
<220><221> misc_feature	
<222> (1)..(7630)	
<223> DNA sequence of the plasmid pCP14	
<400> 29	
ctcagggcta ttgacgacag ctatggttca ctgtccacca accaaaactg tgctcagtac	60
cgccaatatt tctcccttga ggggtacaaa gaggtgtccc tagaagagat ccacgtgtg	120
taaaaatfff aaaaaagggt attgactttc cctacagggt gtgtaataat ttaattacag	180
gcggggggcaa ccccgctgt tctagaggag gaggaatcgc catggagagg attgtcgta	240
ctctcgggga acgtagtac ccaattacca tcgcatctgg ttgttttaac gaaccagctt	300
cattcttacc gctgaaatcg ggcgagcagg tcatgttgg caccaacgaa accctggctc	360
ctctgtatct cgataaggtc cgcgcggtac ttgaacaggc ggggtttaac gtcgatagcg	420
ttatcctccc tgacggcgag cagtataaaa gcctggctgt actcgatacc gtctttacgg	480
cgttgttaca aaagccgcat ggtcgcgata ctacgtgggt ggcgcttggc ggcggcgtag	540
tgggcgatct gaccggcttc gcggcggcga gttatcagcg cgggtttcgt ttattcaag	600
tcccgcacac gttactgtcg caggctgatt cctccgttgg cggcaaaact gcggtcaacc	660
atccctcggg taaaaacatg attggcgcgt tctaccagcc tgcttcagtg gtggtggatc	720
tcgactgtct gaaaacgctt ccccgcgctg agttagcgtc ggggctggca gaagtcatca	780
aatacggcat tattcttgac ggtgcgtttt tcaactggct ggaagagaat ctggatgcgt	840
tgttgcgtct ggacgggccg gcaatggcgt actgtattcg ccgttggtgt gaactgaagg	900
cagaagtgtg cgccgccgac gagcgcgaaa ccgggttacg tgctttactg aatctgggac	960
acacctttgg tcatgccatt gaagctgaaa tggggtatgg caattggta catggtgaag	1020
cggtcgtcgc ggglatgggt atggcggcgc ggacgtcgga acgtctcggg cagtttagtt	1080

ctgccgaaac gcagcgtatt ataaccctgc tcacgcgggc tgggttacg gtcaatgggc	1140
cgcgcgaaat gtccgcgcag gcgtatttac cgcataatgct gcgtgacaag aaagtccttg	1200
cgggagagat gcgcttaatt cttccgttgg caattggtaa gagtgaagtt cgcagcggcg	1260
tttcgcacga gcttgttctt aacgccattg ccgattgtca atcagcgtaa tcatcgttca	1320
tgctgatgc cgctatgtag gccggataag gcgttcacgc cgcattccgc aaccgatgcc	1380
tgatgcgacg cggtcgcgtc ttatcaggcc tacaggtcga tgccgatatg tacatcgtat	1440
tcggcaatta atacatagca acatgaatgg tcttcggttt ccgtgtttcg taaagtcagg	1500
aaacgcggaa gtcagcgccc tgcaccatta tgttcggat ctgcatcgca ggatgctgct	1560
ggctaccctg tggaacacct acatctgtat taacgaagcg ctggcattga ccctgagtga	1620
tttttctctg gtccccccgc atccataaccg ccagttgttt accctcaca cgttcagta	1680
accgggcatg ttcatcatca gtaaccgta tcgtgagcat cctctctcgt ttcacggta	1740
tcattacccc catgaacaga aattccccct tacacggagg catcaagtga ccaaacagga	1800
aaaaaccgcc cttaacatgg cccgctttat cagaagccag acattaacgc ttctggagaa	1860
actcaacgag ctggacgcgg atgaacaggc agacatctgt gaatcgcttc acgaccacgc	1920
tgatgagctt taccgcagct gcctcgcgcg ttccggtgat gacggtgaaa acctctgaca	1980
catgcagctc cgggagacgg tcacagcttg tctgtaagcg gatgccggga gcagacaagc	2040
ccgtcagggc gcgtcagcgg gtgttggcgg gtgtcggggc gcagccatga cccagtcacg	2100
tagcgatagc ggagtgtata ctggcttaac tatgcggcat cagagcagat tgtactgaga	2160
gtgcaccata tgcggtgtga aataccgcac agatgcgtaa ggagaaaata ccgcatcagg	2220
cgctcttccg cticctcgt cactgactcg ctgcgctcgg tcgttcggct gcggcgagcg	2280
gtatcagctc actcaaaggc ggtaatacgg ttatccacag aatcagggga taacgcagga	2340
aagaacatgt gagcaaaagg ccagcaaaag gccaggaacc gtaaaaaggc cgcgttgctg	2400
gcgtttttcc ataggctccg ccccttgac gagcatcaca aaaatcgac ctcaagtcag	2460
aggtggcgaa acccgacagg actataaaga taccaggcgt ttccccctgg aagctccctc	2520
gtgcgctctc ctgttcgac cctgccgtt accggatacc tgtccgcctt tctccctcg	2580
ggaagcgtgg cgcttttca tagctcacgc ttaggtatc tcagttcggg ttaggtcgtt	2640
cgctccaagc tgggctgtgt gcacgaaccc cccgttcage ccgaccgtg cgccttatcc	2700
ggtaactatc gtcttgagtc caaccggta agacacgact tatcgccact ggcagcagcc	2760
actggtaaca ggattagcag agcgaggtat gtaggcggtg ctacagagtt cttgaagtgg	2820
tggcctaact acggctacac tagaaggaca gtatttggta tctgcgctct gctgaagcca	2880
gttaccttcg gaaaagagt tggtagctct tgatccggca aacaaccac cgctggtagc	2940

ggtaggttttt ttgtttgcaa gcagcagatt acgcgcagaa aaaaaggatc tcaagaagat	3000
cctttgatct tttctacggg gtctgacgct cagtggacg aaaactcacg ttaagggatt	3060
ttggtcatga gattatcaaa aaggatcttc acctagatcc ttttaaatta aaaatgaagt	3120
tttaaataca tctaaaglat atatgagtaa acttggctcg acagttacca atgcttaatc	3180
agttaggcac ctatctcagc gatctgtcta tttcgttcat ccatagttgc ctgactcccc	3240
gtcgtgtaga taactacgat acgggagggc ttaccatctg gccccagtc tgcaatgata	3300
ccgcgagacc cacgctcacc ggctccagat ttatcagcaa taaaccagcc agccggaagg	3360
gccgagcgca gaagtgggcc tgcaacttta tccgcctcca tccagtctat taattgttgc	3420
cgggaagcta gtagaagtag ttcgccagtt aatagtttgc gcaacgttgt tgccattgct	3480
gcaggcatcg tgggtgcacg ctctgtcttt ggtatggctt cattcagctc cggttcccaa	3540
cgatcaaggc gaggtaacat gatccccatg ttgtgcaaaa aagcggttag ctcttccgt	3600
cctccgatcg ttgtcagaag taagtggcc gcagtgttat cactcatggt tatggcagca	3660
ctgcataatt ctcttactgt catgccatcc gtaagatgct tttctgtgac tggtagtac	3720
tcaaccaagt cattctgaga atagtgtatg cggcgaccga gttgtcttgc cccggcgtea	3780
acacgggata ataccgcgcc acatagcaga actttaaaag tgctcatcat tggaaaacgt	3840
tcttcggggc gaaaactctc aaggatctta ccgctgttga gatccagttc gatgtaacct	3900
actcgtgcac ccaactgac ttcagcatct tttactttca ccagcgtttc tgggtgagca	3960
aaaacaggaa ggcaaaatgc cgcaaaaaag ggaataaggg cgacacggaa atgttgaata	4020
ctcactactc tctttttca atattattga agcatttacc agggttattg tctcatgagc	4080
ggatacatat ttgaatgtat ttagaaaaat aaacaaatag gggttccgcg cacatttccc	4140
cgaaggtgc cacctgacgt ctaagaaacc attattatca tgacattaac ctataaaaat	4200
aggcgtatca cgagccctt tcgtcttcaa gaattctgaa ccagtcctaa aacgagtaaa	4260
taggaccggc aattcttcaa gcaataaaca ggaataccaa ttattaaaag ataacttagt	4320
cagatcgtac aataaagctt tgaagaaaaa tgcgccttat tcaatctttg ctataaaaaa	4380
tggcccaaaa tctcacattg gaagacattt gatgacctca tttctttcaa tgaaggccct	4440
aacggagttg actaatgttg tgggaaattg gagcgataag cgtgcttctg ccgtggccag	4500
gacaacgtat actcatcaga taacagcaat acctgatcac tacttcgcac tagttttctg	4560
gtactatgca tatgatccaa tatcaaagga aatgatagca ttgaaggatg agactaatcc	4620
aattgaggag tggcagcata tagaacagct aaagggtagt gctgaaggaa gcatacgata	4680

ccccgcatgg aatgggataa tatcacagga ggtactagac tacctttcat cctacataaa	4740
tagacgcata taagtacga ttttaagcata aacacgcact atgccgttct tctcatgtat	4800
atataatatac aggcaacacg cagatatagg tgcgacgtga acagtgagct gtatgtgcgc	4860
agctcgcgtt gcattttcgg aagcgtcgtt tttcggaac gctttgaagt tcctattccg	4920
aagttcctat tctctagaaa gtataggaac ttcagagcgc ttttgaac caaaagcgt	4980
ctgaagacgc actttcaaaa aaccaaaaac gcaccggact gtaacgagct actaaaatat	5040
tgcgaatacc gcttcacaa acattgctca aaagtatctc tttgctatat atctctgtgc	5100
tatatcccta tataacctac ccatccacct ttcgtctctt gaacttgcac cttaactcga	5160
cctctacatt ttttatgttt atctctagta ttactcttta gacaaaaaaa ttgtagtaag	5220
aactattcat agagtgaatc gaaaacaata cgaaaatgta aacatttcct atacgtagta	5280
tatagagaca aaatagaaga aaccgttcat aattttctga ccaatgaaga atcatcaacg	5340
ctatcacttt ctgttcacaa agtatgcgca atccacatcg gtatagaata taatcgggga	5400
tgcctttatc ttgaaaaat gcaccgcag cttcgctagt aatcagtaa cgcgggaagt	5460
ggagtcaggc ttttttatg gaagagaaaa tagacaccaa agtagccttc ttctaacctt	5520
aacggaccta cagtgcacaa agttatcaag agactgcatt atagagcgca caaaggagaa	5580
aaaaagtaat ctaagatgct ttgttagaaa aatagcgctc tcgggatgca tttttgtaga	5640
acaaaaaaga agtatagatt ctttttgggt aaaatagcgc tctcgcgttg catttctgtt	5700
ctgtaaaaat gcagctcaga ttctttgttt gaaaaattag cgctctcgcg ttgcattttt	5760
gttttacaaa aatgaagcac agattcttcg ttggtaaaat agcgctttcg cgttgcattt	5820
ctgttctgta aaaatgcagc tcagattctt tgtttgaaaa attagcgctc tcgcgttgca	5880
tttttgttct acaaaatgaa gcacagatgc ttcgttaaca aagatatgct attgaagtgc	5940
aagatggaaa cgcagaaaaat gaaccgggga tgcgacgtgc aagattacct atgcaataga	6000
tgcaatagtt tctccaggaa ccgaataca tacattgtct tccgtaaagc gctagactat	6060
atattattat acaggttcaa atatactatc tgtttcaggg aaaactcca ggttcggatg	6120
ttcaaaaatc aatgatgggt aacaagtacg atcgtaaate tgtaaacag tttgtcggat	6180
attagctgt atctctcaa agcgtattcg aatatcattg agaagctgca gcgtcacatc	6240
ggataataat gatggcagcc attgtagaag tgccttttgc atttctagtc tctttctcgg	6300
tctagctagt ttactacat cgcaagata gaatcttaga tcacactgcc tttgctgagc	6360
tggatcaata gagtaacaaa agagtggtaa ggccctcgta aaggacaagg acctgagcgg	6420
aagtgtatcg tacagtagac ggagttact agtatagtct atagtcctg gaattctcat	6480
gtttgacagc ttatcatcga taagcttttc aattcaatc atcatttttt ttttattctt	6540

ttttttgatt tcggtttctt tgaaatTTTT ttgattcggT aatctccgaa cagaaggaag	6600
aacgaaggaa ggagcacaga cttagattgg tatatatag catatgtagt gttgaagaaa	6660
catgaaattg cccagtattc ttaacccaac tgcacagaac aaaaacctgc aggaaacgaa	6720
gataaatcat gtcgaaagct acatataagg aacgtgctgc tactcatcct agtcctgttg	6780
ctgccaagct atttaatatc atgcacgaaa agcaaacaaa cttgtgtgct tcattggatg	6840
ttcgtaccac caaggaatta ctggagttag ttgaagcatt aggtcccaaa atttgtttac	6900
taaaaacaca tgtggatatc ttgactgatt ttccatgga gggcacagtt aagccgctaa	6960
aggcattatc cgccaagtac aattttttac tcttcgaaga cagaaaattt gctgacattg	7020
gtaatacagt caaattgcag tactctgcgg gtgtatacag aatagcagaa tgggcagaca	7080
ttacgaatgc acacggtgtg gtgggccag gtattgttag cggtttgaag caggcggcag	7140
aagaagtaac aaaggaacct agaggccttt tgatgttagc agaattgtca tgcaagggct	7200
ccctatctac tggagaatat actaagggtA ctgttgacat tgcgaagagc gacaaagatt	7260
ttgttatcgg ctttattgct caaagagaca tgggtggaag agatgaaggt tacgattggt	7320
tgattatgac acccgggtgtg gggttagatg acaagggaga cgcatgggt caacagtata	7380
gaaccgtgga tgatgtggct tctacaggat ctgacattat tattgttgga agaggactat	7440
ttgcaaaggg aagggtatgct aaggtagagg gtgaacgtta cagaaaagca ggctgggaag	7500
catatttgag aagatgcggc cagcaaaact aaaaaactgt attataagta aatgcatgta	7560
tactaaactc acaaataga gcttcaattt aattatatca gttattacc gggaatctcg	7620
gtcgtaatga	7630
<210> 30	
<211> 10015	
<212> DNA	
<213> Escherichia coli	
<220><221> misc_feature	
<222> (1)..(10015)	
<223> DNA sequence of the plasmid pCP50	
<400> 30	
cttgaaataa ttaacaaaca aaggagttac agttagaaat tgtaggagag atctcgtttt	60
tcgcgacaat ctggcgtttt tcttgctaata tccaggatta atccgttcat agtgtaaaac	120
cccgtttaca cattctgacg gaagatatag attggaagta ttgcattcac taagataagt	180
atggcaacac tggaacagac atgaattatc agaacgacga ttacgcatac aaagaaatca	240

aagagttact tctcctgtc gcattgctgg aaaaattccc cgctactgaa aatgccgcga	300
atacggttgc ccatgccga aaagcgatcc ataagatcct gaaaggtaat gatgatcgcc	360
tgttggttgt gattggccca tgcctcaattc atgatcctgt cgcggcaaaa gagtatgcca	420
ctcgcttgct ggcgctgcgt gaagagctga aagatgagct ggaaatcgta atgcgcgtct	480
atTTTgaaaa gccgcgtacc acggtgggct ggaaagggct gattaacgat ccgcataatg	540
ataatagctt ccagatcaac gacggtctgc gtatagcccg taaattgctg ctgatatta	600
acgacagcgg tctgccagcg gcaggtgagt ttctcgatat gatcaccca caatatctcg	660
ctgacctgat gagctggggc gcaattggcg cacgtaccac cgaatcgag gtgcaccgcg	720
aactggcatc agggctttct tgtccggtcg gcttcaaaaa tggcacccgac ggtacgatta	780
aagtggctat cgatgccatt aatgccgcg gtgcgccga ctgcttcctg tccgtaacga	840
aatgggggca ttcggcgatt gtgaatacca gcgtaacgg cgattgccat atcattctgc	900
gcggcggtaa agagcctaac tacagcgca agcacgttg tgaagtgaag gaagggtga	960
acaaagcagg cctgccagca caggtgatga tcgatttcag ccatgctaac tcgtccaaac	1020
aattcaaaaa gcagatggat gtttgtgctg acgtttgcca gcagattgcc ggtggcgaaa	1080
aggccattat tggcgtgatg gtggaagcc atctggtgga aggcaatcag agcctcgaga	1140
gcggggagcc gctggcctac ggtaagagca tcaccgatgc ctgcatcggc tgggaagata	1200
ccgatgctct gttacgtcaa ctggcgaatg cagtaaaagc gcgtcgcggt taaggtttaa	1260
ttgtcggatg cgcgcgcaga gtggcgatc cgatgaatca ccacaggcct gataagtgc	1320
gcagcgtgc atcaggcaat gtgctccatt gttagcaaca aaaaagccga ctcaattgca	1380
gtcggctttc tcatTTTaaa cgaatgacgt ttacttcgct ttaccctggt ttgcaaccgc	1440
cgctgctttc gctacatgaa tggctcttcgg ttccggtgt tcgtaaagtc tggaaacgcg	1500
gaagtacgcg cctgcacca ttatgttccg gatctgcatc gcaggatgct gctggctacc	1560
ctgtggaaca cctacatctg tattaacgaa gcgctggcat tgaccctgag tgatttttct	1620
ctggtccgc cgcattcata ccgccattg ttaccctca caacgttcca gtaaccgggc	1680
atgttcatca tcagtaacce gtatcgtgag catcctctct cgtttcatcg gtatcattac	1740
ccccatgaac agaaattccc cttacacgg aggcattcaag tgaccaaaca ggaaaaaac	1800
gcccttaaca tggcccgctt tatcagaagc cagacattaa cgcttctgga gaaactcaac	1860
gagctggacg cggatgaaca ggcagacatc tgtgaatcgc ttacgacca cgctgatgag	1920
ctttaccgca gctgcctcgc gcgtttcggg gatgacgggtg aaaacctctg acacatgcag	1980
ctcccgaga cggctacagc ttgtctgtaa gcggatgccg ggagcagaca agcccgctag	2040
ggcgcgtcag cgggtgttgg cgggtgtcgg ggcgcagcca tgaccagtc acgtagcgat	2100

agcggagtgt atactggctt aactatgcgg catcagagca gattgtactg agagtgcacc	2160
atatgcggtg tgaataaccg cacagatgcg taaggagaaa ataccgcac aggcgctctt	2220
ccgcttcttc gctcactgac tcgctgcgct cggctcgttcg gctgcggcga gcggtatcag	2280
ctcactcaaa ggcggttaata cggttatcca cagaatcagg ggataacgca ggaaagaaca	2340
tgtagcaaaa aggccagcaa aaggccagga accgtaaaaa ggccgcgttg ctggcgtttt	2400
tccataggct cggccccct gacgagcac acaaaaatcg acgtcaagt cagaggtggc	2460
gaaacccgac aggactataa agataccagg cgtttcccc tggaagctcc ctctgcgct	2520
ctctgttcc gaccctgcg cttaccgat acctgtccgc ctttctccct tcgggaagcg	2580
tggcgctttc tcatagctca cgctgtagg atctcagttc ggtgtaggtc gttcgtcca	2640
agctgggctg tgtgcacgaa cccccgttc agcccgaccg ctgcgcctta tccgtaact	2700
atcgtcttga gtccaaccg gtaagacacg acttatcgcc actggcagca gccactggta	2760
acaggattag cagagcgagg tatgtaggcg gtgctacaga gttcttgaag tggtagccta	2820
actacggcta cactagaagg acagtatttg gtatctgcgc tctgctgaag ccagttacct	2880
tcggaaaaag agttggtagc tcttgatccg gcaaacaaac caccgctggt agcggtggtt	2940
tttttgttt caagcagcag attacgcgca gaaaaaagg atctcaagaa gatcctttga	3000
tcttttctac ggggtctgac gtcagtgga acgaaaactc acgttaaggg attttggtea	3060
tgagattatc aaaaaggatc ttacactaga tctttttaa ttaaaaatga agttttaaat	3120
caatctaag tatatatgag taaacttggc ctgacagtta ccaatgctta atcagtagg	3180
cacctatctc agcgatctgt ctatttcgtt catccatagt tgctgactc cccgtcgtgt	3240
agataactac gatacgggag ggcttaccat ctggccccag tgctgcaatg ataccgcgag	3300
accacgctc accggtcca gatttatcag caataaacca gccagccgga agggccgagc	3360
gcagaagtgg tctgcaact ttatccgct ccatccagtc tattaattgt tgccgggaag	3420
ctagagtaag tagttcgcca gttaatagtt tgcgcaacgt tgttgccatt gctgcaggca	3480
tcgtggtgtc acgctcgtcg tttggtatgg ctctattcag ctccggttcc caacgatcaa	3540
ggcgagttac atgatecccc atgttgtgca aaaaagcgg tagctccttc ggtcctccga	3600
tcgttgtcag aagtaagttg gccgcagtgt tatcactcat ggttatggca gcaactgcata	3660
attctcttac tgtcatgcca tccgtaagat gcttttctgt gactggtgag tactcaacca	3720
agtcattctg agaatagtgt atgcggcgac cgagttgctc ttgcccggcg tcaacacggg	3780
ataataccgc gccacatagc agaactttaa aagtgtcat cattggaaaa cgttcttcgg	3840
ggcgaaaact ctcaaggatc ttaccgctgt tgagatccag ttcatgtaa cccactcgtg	3900
caccaactg atcttcagca tcttttactt tcaccagcgt tctgggtga gcaaaaacag	3960

gaaggcaaaa tgccgcaaaa aagggaataa gggcgacacg gaaatgttga atactcatac	4020
tcttcctttt tcaatattat tgaagcattt atcagggtta ttgtctcatg agcggataca	4080
tatttgaatg tatttagaaa aataaacaaa taggggttcc ggcacattt ccccgaaaag	4140
tgccacctga cgictaagaa accattatta tcatgacatt aacctataaa aataggcgta	4200
tcacgaggcc ctttcgtctt caagaattct gaaccagtcc taaaacgagt aaataggacc	4260
ggcaattctt caagcaataa acaggaatac caattattaa aagataactt agtcagatcg	4320
tacaataaag ctttgaagaa aaatgcgcct tattcaatct ttgctataaa aaatggccca	4380
aaatctcaca ttggaagaca ttgatgacc tcatttcttt caatgaaggg cctaacggag	4440
ttgactaatg ttgtgggaaa ttggagcgat aagcgtgctt ctgccgtggc caggacaacg	4500
tatactcatc agataacagc aatactgat cactacttcg cactagtttc tcggtactat	4560
gcatatgac caatatcaaa ggaaatgata gcattgaagg atgagactaa tccaattgag	4620
gagtggcagc atatagaaca gctaaagggt agtgcgaag gaagcatacg ataccccgca	4680
tggaatggga taatatcaca ggaggtacta gactacctt catcctacat aaatagacgc	4740
atataagtac gcatttaagc ataaacacgc actatgccgt tcttctcatg tatatatata	4800
tacaggcaac acgcagatat aggtgcgacg tgaacagtga gctglatgtg cgcagctcgc	4860
gttgcatttt cggaagcgtt cgttttcgga aacgctttga agttcctatt ccgaagticc	4920
tattctctag aaagtatagg aacttcagag cgcttttgaa aaccaaaagc gctctgaaga	4980
cgcactttca aaaaacaaa aacgcaccgg actgtaacga gctactaaaa tattgcgaat	5040
accgcttcca caaacattgc tcaaaagtat ctctttgcta tatactcttg tgctatatec	5100
ctatataacc tacccatcca cctttcgtc cttgaacttg catctaaact cgaccttac	5160
atTTTTatg tttatctcta gtattactct ttagacaaaa aaattgtagt aagaactatt	5220
catagagtga atcgaaaaca atacgaaaat gtaaacattt cctatacgta gtatatagag	5280
acaaaataga agaaaccgtt cataattttc tgaccaatga agaactatca acgctatcac	5340
tttctgttca caaagtatgc gcaatccaca tcggtataga atataatcgg ggatgccttt	5400
atcttgaaaa aatgcacccg cagcttcgct agtaatcagt aaacgcggga agtggagtca	5460
ggcttttttt atggaagaga aaatagacac caaagtagcc ttcttctaac cttaacggac	5520
ctacagtga aaaaattatc aagagactgc attatagagc gcacaaagga gaaaaaaagt	5580
aatctaagat gctttgttag aaaaatagcg ctctcgggat gcatttttgt agaacaaaaa	5640
agaagtatag attctttgtt ggtaaaatag cgctctcgcg ttgcatttct gttctgtaaa	5700

aatgcagctc agattctttg ttgaaaaat tagcgctctc gcgttgcat tttgttttac	5760
aaaaatgaag cacagattct tcgttggtta aatagcgctt tcgcgttgca tttctgttct	5820
gtaaaaatgc agctcagatt ctttgtttga aaaattagcg ctctcgctt gcatttttgt	5880
tctacaaaat gaagcacaga tgcttcgtta acaaagatat gctattgaag tgcaagatgg	5940
aaacgcagaa aatgaaccgg ggatgcgacg tgcaagatta cctatgcaat agatgcaata	6000
gtttctccag gaaccgaaat acatacattg tcttccgtaa agcgctagac tatatattat	6060
tatacagggt caaatafact atctgtttca gggaaaactc ccaggttcgg atgttcaaaa	6120
ttcaatgatg ggtaacaagt acgatcgtaa atctgtaaaa cagtttgtcg gatattaggc	6180
tgtatctcct caaagcgtat tcgaatatca ttgagaagct gcagcgtcac atcgataat	6240
aatgatggca gccattgtag aagtgccttt tgcatctcta gtctctttct cggcttagct	6300
agttttacta catcgcaag atagaatctt agatcacact gcctttgctg agctggatca	6360
atagagtaac aaaagagtgg taaggcctcg ttaaaggaca aggacctgag cggaagtgt	6420
tcgtacagta gacggagtat actagtatag tctatagtcc gtggaattct catgtttgac	6480
agcttatcat cgataagctt ttcaattcaa ttcattcttt tttttttatt cttttttttg	6540
atttcggttt ctttgaaatt tttttgattc ggtaatctcc gaacagaagg aagaacgaag	6600
gaaggagcac agacttagat tggatatata acgcatatgt agtgttgaag aaacatgaaa	6660
ttgccagta ttcttaaccc aactgcacag acaaaaaacc tgcaggaaac gaagataaat	6720
catgtcgaaa gctacatata aggaacgtgc tgctactcat cctagtcttg ttgctgcaa	6780
gctatttaat atcatgcacg aaaagcaaac aaacttgtgt gcttcattgg atgttcgtac	6840
caccaaggaa ttactggagt tagttgaagc attaggtccc aaaatttgtt tactaaaaac	6900
acatgtggat atcttgactg atttttccat ggagggcaca gttaagccgc taaaggcatt	6960
atccgccaag tacaattttt tactcttcga agacagaaaa tttgctgaca ttggtaatac	7020
agtcaaattg cagtactctg cgggtgtata cagaatagca gaatgggcag acattacgaa	7080
tgcacacggt gtggtgggcc caggtattgt tagcggtttg aagcaggcgg cagaagaagt	7140
aacaaaggaa cctagaggcc ttttgatgtt agcagaattg tcatgcaagg gctccctatc	7200
tactggagaa tatactaagg gtactgttga cattgcgaag agcgacaaag atttgttat	7260
cggctttatt gctcaaagag acatgggtgg aagagatgaa gggtacgatt ggttgattat	7320
gacacccggt gtgggtttag atgacaaggg agacgcattg ggtcaacagt atagaaccgt	7380
ggatgatgtg gtctctacag gatctgacat tattattgtt ggaagaggac tatttgcaaa	7440
gggaagggat gctaaggtag aggggtgaacg ttacagaaaa gcaggctggg aagcatattt	7500
gagaagatgc ggccagcaaa actaaaaaac tgtattataa gtaaatgcat gtatactaaa	7560

ctcacaaatt agagcttcaa ttttaattata tcagttatta cccgggaatc tcggtcgtaa	7620
tgaaggaaa agcgcaacgg acgggcgagt agattgcga acatgcgagc atgatccaga	7680
gatttctgaa gcagcaaaag gatgttccat gtacatgacg cgcggcttgc ggtaaattgt	7740
tggcaaatTT tccggcgtag cccaaaacgc gctgtcgtca agtcgttaag ggcgtgccct	7800
tcacatccg atctggagtc aaaatgtcct cacgtaaaga gcttgccaat gctattcgtg	7860
cgctgagcat ggacgcagta cagaaagcca aatccggtca cccgggtgcc cctatgggta	7920
tggctgacat tgccgaagtc ctgtggcgtg atttctgaa acacaaccg cagaatccgt	7980
cctgggctga ccgtgaccgc ttctgtctgt ccaacggcca cggctccatg ctgatctaca	8040
gcctgtctga cctcaccggt tacgatctgc cgatggaaga actgaaaaac ttccgtcagc	8100
tgcactctaa aactccgggc caccggaag taggttatac cgctgggtgtg gaaaccacca	8160
ccggtccgct gggtcagggt attgccaacg cagtcggtat ggcgattgca gaaaaaacgc	8220
tggcggcgca gttaaccgt ccaggtcacg acattgtcga ccactacacc tacgccttca	8280
tgggcgacgg ctgcatgatg gaagcatct cccacgaagt ttgctctctg gcgggtacgc	8340
tgaagctggg taaactgatt gcgttctacg atgacaacgg tatctcaatc gatggtcacg	8400
ttgaaggctg gttcactgac gacaccgcaa tgcgtttcga agcttacggc tggcacgtta	8460
ttcgcgacat cgacggtcat gacgcggcat ccatcaaacg cgcagtagaa gaagcgcgcg	8520
cagtactga caaacgtcc ctgctgatgt gcaaaacat catcggtttc ggttccccga	8580
acaaagccgg taccacgac tcccacggtg cgcgctggg cgacgtgaa attgccctga	8640
cccgcgaaca gctgggctgg aaatacgcgc cgttcgaaat cccgtctgaa atctatgctc	8700
agtgggatgc gaaagaagca ggccaggcga aagaatctgc atggaatgag aagtttgcgg	8760
cttacgcgaa agcttatccg caggaagcgg ctgaatttac ccgccgtatg aaaggcgaaa	8820
tgcctctga cttcgacgcc aaagcgaaag agtttatcgc taaactgcag gctaatccgg	8880
cgaaaatcg cagccgtaaa gcgtcgcaga atgctatcga agcgttcggc ccgtgttgc	8940
ctgaattcct cggcggctct gctgacctgg caccgtctaa cctgaccctg tggctctggtt	9000
ctaaagcaat caacgaagat gctgcaggta actacatcca ctacggtgtt cgcgagttcg	9060
glatgaccgc gattgctaac ggtatctccc tgcacggtgg tttctgccg tacacctcca	9120
ccttctgat gttcgtggaa tacgcagta acgccgtacg tatggctgcg ctgatgaaac	9180
agcgtcaggt gatggtttac acccagcact ccatcggtct gggcgaagat ggcccgactc	9240
accagccggt ttagcaggtc gcttctctgc gcgtgacccc gaacatgtct acatggcgtc	9300
cgtgtgacca ggttgaatcc cgggtcgcgt ggaaatcagg cgttgagcgt caggacggcc	9360
cgactgcgct tatcctctcc cgtcagaacc tggcgcagca ggaacgaact gaagagcaac	9420

tggaacat cgcgcgcggt gggttatgtgc tgaaagactg cgccgggtcag ccggaactga	9480
ttttcatcgc taccggttca gaagtgaac tggtgtgtgc tgcctacgaa aaactgactg	9540
ccgaaggcgt gaaagcgcg cggtgttcca tgcgtctac cgacgcatTT gacaagcagg	9600
atgctgctta ccgtgaatcc gtactgccga aagcggttac tgcacgcgtt gctgtagaag	9660
cgggtattgc tgactactgg tacaagtatg ttggcctgaa cggtgtctatc gtcggtatga	9720
ccaccttcgg tgaatctgct ccggcagagc tgcgttttga agagttcggc ttcactgttg	9780
ataacgttgt tgcgaaagca aaagaactgc tgtaattagc atttcgggta aaaaggtcgc	9840
ttcggcgacc ttttttatta ctttgatatg tccgtttgcg gacaagcaat agataaagcg	9900
tggttagat cacaaatatt tataatgcaat aaatatcaat tatgtaatat gcatcacgat	9960
atgcgtattg acatttgttg ttataactat aactcaatgt tatataagaa attaa	10015
<210> 31	
<211> 9065	
<212> DNA	
<213> Escherichia coli	
<220><221> misc_feature	
<222> (1)..(9065)	
<223> DNA sequence of the plasmid pCP54	
<400> 31	
acatgaatgg tcttcggttt ccgtgtttcg taaagtctgg aaacgcggaa gtcagcgccc	60
tgcaccatta tgttccggat ctgcatcgca ggatgctgct ggctaccctg tggaacacct	120
acatctgtat taacgaagcg ctggcattga ccttgagtga tttttctctg gtcccgccgc	180
atccataccg ccagttgttt accctcaciaa cgttccagta accgggcatg ttcacatca	240
gtaaccgcta tctgtagcat cctctctctg ttcacggtg tcatcaccac catgaacaga	300
aattccccct tacacggagg catcaagtga ccaaacagga aaaaaccgcc cttaacatgg	360
cccgttttat cagaagccag acattaacgc tttctggagaa actcaacgag ctggacgcgg	420
atgaacaggc agacatctgt gaatcgcttc acgaccacgc tgatgagctt taccgcagct	480
gcctcgcgcg tttcggtgat gacggtgaaa acctctgaca catgcagctc ccggagacgg	540
tcacagcttg tctgtaagcg gatgccggga gcagacaagc ccgtcagggc gcgtcagcgg	600
gigtgtggcgg gtgtcggggc gcagccatga ccagtcacg tagcgatagc ggagtgtata	660
ctggcttaac tatgcggcat cagagcagat tgtactgaga gtgcaccata tgcggtgtga	720
aataaccgcac agatgcgtaa ggagaaaata ccgcatcagg cgctcttccg cttcctcgct	780

cactgactcg ctgcgctcgg tcgttcggct gcggcgagcg gtatcagctc actcaaaggc	840
ggtaatacgg ttatccacag aatcagggga taacgcagga aagaacatgt gagcaaaagg	900
ccagcaaaag gccaggaacc gtaaaaagcg cgcgttgctg gcgtttttcc ataggctccg	960
ccccctgac gagcatcaca aaaatcgacg ctcaagtcag aggtggcgaa acccgacagg	1020
actataaaga taccaggcgt tccccctgg aagctccctc gtgcgctctc ctgttcggac	1080
cctgcgctt accggatacc tgtecgctt tctcccttcg ggaagcgtgg cgcttttctca	1140
tagctcacgc ttaggtatc tcagttcggf taggtcgtt cgctccaagc tgggctgtgt	1200
gcacgaacc cccgttcagc ccgaccgtg cgccttatcc ggtaactatc gtcttgagtc	1260
caaccggta agacacgact tatcgccact ggcagcagcc actggtaaca ggattagcag	1320
agcgaggat tagggcggtg ctacagagtt ctgaagtgg tggcctaact acggctacac	1380
tagaaggaca gtatttggtg tctgcgctct gctgaagcca gttaccttcg gaaaaagagt	1440
tggtagctct tgatccggca aacaaccac cgctggtagc ggtggttttt ttgtttgcaa	1500
gcagcagatt acgcgcagaa aaaaaggatc tcaagaagat cctttgatct tttctacggg	1560
gtctgacgct cagtggaacg aaaactcacg ttaagggtt ttggtcatga gattatcaaa	1620
aaggatcttc acctagatcc ttttaatta aaaatgaagt tttaatcaa tctaaagtat	1680
atatgagtaa acttggtctg agtggcggtt ttcattgctt gttatgactg ttttttggg	1740
gtacagtcta tgctcgggc atccaagcag caagcgcgtt acgccgtggg tcgatgtttg	1800
atgttatgga gcagcaacga tgttacgcag cagggcagtc gccctaaaac aaagttaaac	1860
atcatgaggg aagcgtgat cgccgaagta tcgactcaac tatcagaggt agttggcgtc	1920
atcgagcgcc atctgaacc gacgttgctg gccgtacatt tgtacggctc cgcagtggat	1980
ggcggcctga agccacacag tgatattgat ttgctggtta cggtgaccgt aaggcttgat	2040
gaaacaacgc ggcgagcttt gatcaacgac cttttggaaa cttcggttc ccctggagag	2100
agcgagattc tccgcgctgt agaagtcacc attgttgtgc acgacgacat cattccgtgg	2160
cgttatccag ctaagcgca actgcaattt ggagaatggc agcgcaatga cattcttgca	2220
ggtatcttcg agccagccac gatcgacatt gatctggcta tcttgctgac aaaagcaaga	2280
gaacatagcg ttgccttggg aggtccagcg gcggaggaac tctttgatcc ggttcctgaa	2340
caggatctat ttgaggcgtt aaatgaaacc ttaacgtat ggaactcgcc gcccgactgg	2400
gctggcgatg agcgaaatgt agtgcttacg ttgtcccgca tttggtacag cgcagtaacc	2460
ggcaaaatcg cgccgaagga tgtegtgcc gactgggcaa tggagcgctt gccggcccag	2520
tatcagcccc tcatacttga agctagacag gcttatcttg gacaagaaga agatcgcttg	2580
gcctcgcg cgatcagtt ggaagaattt gtccactacg tgaaaggcga gatcaccaag	2640

gtagtcggca aataatgtct aacaattcgt tcaagccgac gccgcttcgc ggcgcggtt	2700
aactcaagcg ttagatgcac taagcacata attgtcaca gccaaactat cagaattctg	2760
aaccagtcct aaaacgagta aataggaccg gcaattcttc aagcaataaa caggaatacc	2820
aattattaaa agataactta gtcagatcgt acaataaagc tttgaagaaa aatgcgcctt	2880
attcaatctt tgctataaaa aatggcccaa aatctcacat tggaagacat ttgatgacct	2940
catttctttc aatgaagggc ctaacggagt tgactaatgt tgtgggaaat tggagcgata	3000
agcgtgcttc tgccgtggcc aggacaacgt atactcatca gataacagca atacctgac	3060
actacttcgc actagtttct cggtactatg catatgatcc aatatcaaag gaaatgatag	3120
cattgaagga tgagactaat ccaattgagg agtggcagca tatagaacag ctaaagggt	3180
gtgtgaagg aagcatagca taccgccat ggaatgggat aatatcacag gaggtactag	3240
actaccttc atctacata aatagacgca tataagtacg catttaagca taaacacgca	3300
ctatgccgtt ctctcatgt atatatatat acaggcaaca cgcagatata ggtgcgacgt	3360
gaacagttag ctgtatgtgc gcagctcgcg ttgcattttc ggaagcgctc gttttcgaa	3420
acgttttgaa gttcctatc cgaagtacct attctctaga aagtatagga acttcagagc	3480
gcttttgaaa accaaaagcg ctctgaagac gcactttcaa aaaacaaaa acgcaccgga	3540
ctgtaacgag ctactaaaat attgcgaata ccgcttcac aaacattgct caaaagtatc	3600
tctttgctat atatctctgt gctatatccc tatataacct acccatccac ctttcgctcc	3660
ttgaacttgc atctaaactc gacctctaca tttttatgt ttatctctag tattactctt	3720
tagacaaaa aattgtagta agaactattc atagagtga tcgaaaacaa tacgaaaatg	3780
taaacatttc ctatactag tatatagaga caaaatagaa gaaaccgttc ataattttct	3840
gaccaatgaa gaatcatcaa cgctatcact ttctgttcac aaagtatgcg caatccacat	3900
cggtatagaa tataatcggg gatgccttta tcttgaaaa atgcacccgc agcttcgcta	3960
glaatcagta aacgcgggaa gtggagtcag gcttttttta tggaagagaa aatagacacc	4020
aaagtagcct tcttctaacc ttaacggacc tacagtcaa aaagtatca agagactgca	4080
ttatagagcg cacaaaggag aaaaaagta atctaagatg ctttgttaga aaaatagcgc	4140
tctcgggatg catttttgta gaacaaaaaa gaagtataga ttctttgttg gtaaaatagc	4200
gctctcgcgt tgcatctctg ttctgtaaaa atgcagctca gattctttgt ttgaaaaatt	4260
agcgtctctg cgttgcatct ttgttttaca aaaatgaagc acagattctt cgttggtaaa	4320
atagcgcttt cgcttgcat ttctgttctg taaaaatgca gctcagattc tttgtttgaa	4380

aaattagcgc tctcgcgttg cttttttgtt ctacaaaatg aagcacagat gcttcgttaa	4440
caaagatatg ctattgaagt gcaagatgga aacgcagaaa atgaaccggg gatgcgacgt	4500
gcaagattac ctatgcaata gatgcaatag tttctccagg aaccgaaata catacattgt	4560
cttccgtaaa gcgctagact atatatattt atacaggttc aaatatacta tctgtttcag	4620
ggaaaactcc caggttcggg tgttcaaaat tcaatgatgg gtaacaagta cgatcgtaaa	4680
tctgtaaaac agtttgcgg atattaggct gtatctcttc aaagcgtatt cgaatatcat	4740
tgagaagctg cagcgtcaca tcggataata atgatggcag ccattgtaga agtgcctttt	4800
gcatttctag tctctttctc ggtctagcta gttttactac atcgcaaga tagaatctta	4860
gatcacactg ctttgcgtga gctggatcaa tagagtaaca aaagagtggg aaggcctcgt	4920
taaaggacaa ggacctgagc ggaagtgtat cgtacagtag acggagtata ctagtatagt	4980
ctatagtccg tggaattctc atgtttgaca gcttatcctc gataagcttt tcaattcaat	5040
tcatcatttt tttttttctc ttttttttga tttcggtttc tttgaaattt ttttgattcg	5100
glaatctccg aacagaagga agaacgaagg aaggagcaca gacttagatt ggtatatata	5160
cgcataatgta gtgttgaaga aacatgaaat tgcccagtat tcttaacca actgcacaga	5220
acaaaaacct gcaggaaacg aagataaatc atgtcgaaag ctacatataa ggaacgtgct	5280
gctactcctc ctagtctgtg tgctgccaag ctatttaata tcatgcacga aaagcaaaca	5340
aacttgtgtg cttcattgga tgttcgtacc accaaggaat tactggagtt agttgaagca	5400
ttaggtccca aaatttgttt actaaaaaca catgtggata tcttgactga tttttccatg	5460
gagggcacag ttaagccgct aaaggcatta tccgccaagt acaatttttt actcttcgaa	5520
gacagaaaat ttgctgacat tggtaataca gtcaaattgc agtactctgc ggggtgtatac	5580
agaatagcag aatgggcaga cattacgaat gcacacggtg tgggtggccc aggtattgtt	5640
agcggtttga agcaggcggc agaagaagta acaaaggaac ctagaggcct tttgatgtta	5700
gcagaattgt catgcaaggc ctcctatct actggagaat atactaaggg tactgttgac	5760
attgcgaaga gcgacaaaga ttttgttctc ggctttattg ctcaaagaga catgggtgga	5820
agagatgaag gttacgattg gttgattatg acacccggtg tgggtttaga tgacaaggga	5880
gacgcattgg gtcaacagta tagaaccgtg gatgatgtgg tctctacagg atctgacatt	5940
attattgttg gaagaggact atttgcaaag ggaagggatg ctaaggtaga ggggtgaacgt	6000
tacagaaaag caggctggga agcatatttg agaagatgcg gccagcaaaa ctaaaaaact	6060
gtattataag taaatgcatg tatactaaac tcacaaatta gagcttcaat ttaattatat	6120
cagttattac ccgggaatct cggtcgtaat gacttgaaat aattaacaaa caaaggagtt	6180
acagttagaa attgtaggag agatctcggt tttcgcgaca atctggcggt tttcttgcta	6240

attccaggat taatccgttc atagtgtaaa accccgttta cacattctga cggaagatat	6300
agattggaag tattgcattc actaagataa gtatggcaac actggaacag acatgaatta	6360
tcagaacgac gatttacgca tcaaagaaat caaagagtta cttcctcctg tcgcattgct	6420
ggaaaaattc cccgctactg aaaatgccgc gaatacgggt gcccatgccc gaaaagcgat	6480
ccataagatc ctgaaaggta atgatgatcg cctgttggtt gtgattggcc catgctcaat	6540
tcatgatcct gtcgcggaag aagagtatgc cactcgcttg ctggcgctgc gtgaagagct	6600
gaaagatgag ctggaaatcg taatgcgcgt ctattttgaa aagccgcgta ccacggtggg	6660
ctggaaaggg ctgattaacg atccgcataat ggataatagc ttccagatca acgacggtct	6720
gcgtatagcc cgtaaattgc tgcttgatat taacgacagc ggtctgccag cggcaggtga	6780
gtttctcgat atgatcacc cacaatatct cgctgacctg atgagctggg gcgcaattgg	6840
cgcacgtacc accgaatcgc aggtgcaccg cgaactggca tcagggtctt ctgtccggt	6900
cggtttcaaa aatggcaccg acggtacgat taaagtggct atcgatgcca ttaatgccgc	6960
cggtgcgccg cactgcttcc tgctcgtaac gaaatggggg cattcggcga ttgtgaatac	7020
cagcggtaac ggcgattgcc atatcattct gcgcggcggt aaagagccta actacagcgc	7080
gaagcacgtt gctgaagtga aagaagggtt gaacaaagca ggcctgccag cacaggtgat	7140
gatcgatttc agccatgcta actcgtccaa acaattcaaa aagcagatgg atgtttgtgc	7200
tgacgtttgc cagcagattg ccggtggcga aaaggccatt attggcgtga tgggtgaaag	7260
ccatctgggtg gaaggcaatc agagcctcga gagcggggag ccgctggcct acggtaaag	7320
catcaccgat gctgcatcg gctgggaaga taccgatgct ctgttacgtc aactggcgaa	7380
tgcagtaaaa gcgcgtcgcg ggtaagggtt aattgtcgga tgcgccgtca gagtggcgta	7440
tccgatgaat caccacagcg ctgataagtc gcgcagcgtc gcatcaggca atgtgtcca	7500
ttgttagcaa caaaaaagcc gactcacttg cagtcggctt tctcatitta aacgaatgac	7560
gtttacttcg ctttacctg gtttgcaacc gccgtgctt tcgctctcga ggctattgac	7620
gacagctatg gttcactgtc caccaaccaa aactgtgtc agtaccgcca atatttctcc	7680
cttgaggggt acaaagaggt gtcctagaa gagatccacg ctgtgtaaaa attttcaaaa	7740
aaggtattga ctttccttac aggggtgtgt ataatttaac tacaggcggg ggcaaccccg	7800
cctgttctag aggaggagga atcgccatgg agaggattgt cgttactctc ggggaacgta	7860
gttacccaat taccatcgca tctggtttgt ttaatgaacc agcttcattc ttaccgctga	7920
aatcgggcga gcaggtcatg ttggtcacca acgaaacct ggctcctctg tatctcgata	7980
aggtccgcgg cgtacttgaa caggcgggtg ttaacgtcga tagcgttacc ctccctgacg	8040
gcgagcagta taaaagcctg gctgtactcg ataccgtctt tacggcgttg ttacaaaagc	8100

cgcatggtcg cgatactacg ctggcggcgc ttggcggcgg cgtagtgggc gatctgaccg 8160

gcttcgcggc ggcgagttat cagcgcggtg ttcgtttcat tcaagtcccg acgacgttac 8220

tgctgcaggt cgattccctcc gttggcggca aaactgcggg caaccatccc ctcggtaaaa 8280

acatgattgg cgcgtttctac cagcctgctt cagtgggtgg ggatctcgac tgtctgaaaa 8340

cgttccccc gcgtgagtta gcgtcggggc tggcagaagt catcaaatac ggcattattc 8400

ttgacggcgc gtttttcaac tggctggaag agaatctgga tgcgttgttg cgtctggacg 8460

gtccggcaat ggcgtaactgt attcgcggtt gttgtgaact gaaggcagaa gttgtcgccg 8520

ccgacgagcg cgaaaccggg ttacgtgctt tactgaatct gggacacacc ttgtgtcatg 8580

ccattgaagc tgaaatgggg tatggcaatt ggttacatgg tgaagcggtc gctgcgggta 8640

tggatgatgg ggcgcggacg tcggaacgtc tcgggcagtt tagttctgcc gaaacgcagc 8700

gtattataac cctgctcacg cgggctgggt taccgggtcaa tgggccgcgc gaaatgtccg 8760

cgcaggcgta ttaccgcat atgctgcgtg acaagaaagt ccttgcggga gagatgcgct 8820

taattcttcc gttggcaatt ggtaagagtg aagttcgag cggcgtttcg cagcagcttg 8880

ttcttaacgc cattgccgat tgtcaatcag cgtaatcacc gttcatgcct gatgccgcta 8940

tgtaggccgg ataaggcgtt cagcccgcat ccggcaaccg atgcctgatg cgacgcggtc 9000

gcgtcttacc aggcctacag gtcatgccg atatgtacat cgtattcggc aattaataca 9060

tagca 9065

<210> 32

<211> 11475

<212> DNA

<213> Escherichia coli

<220><221> misc_feature

<222> (1)..(11475)

<223> DNA sequence of the plasmid pCP55

<400> 32

acatgaatgg tcttcggttt ccgtgttttcg taaagtctgg aaacgcggaa gtcagcgccc 60

tgcaccatta tgttcggat ctgcatcgca ggatgctgct ggctaccctg tggaacacct 120

acatctgtat taacgaagcg ctggcattga cctgagtgga tttttctctg gtcccgccgc 180

atccataacc ccagttgttt accctcaca cgttcagta accgggcatg ttcacatca 240

gtaaccgta tcgtgagcat cctctctcgt ttcacggta tcattacccc catgaacaga 300

aattccccct tacacggagg catcaagtga ccaaacagga aaaaaccgcc cttaacatgg 360

cccgccttat cagaagccag acattaacgc ttctggagaa actcaacgag ctggacgcgg	420
atgaacaggc agacatctgt gaatcgcttc acgaccacgc tgatgagctt taccgcagct	480
gcctcgcgcg ttctggatgat gacggtgaaa acctctgaca catgcagctc ccggagacgg	540
tcacagcttg tctgtaagcg gatgccggga gcagacaagc ccgtcagggc gcgtcagcgg	600
gtgttggcgg gtgtcggggc gcagccatga cccagtcacg tagcgatagc ggagtgtata	660
ctggcttaac tatgcggcat cagagcagat tgtactgaga gtgcaccata tgcgggtgtga	720
aataccgcac agatgcgtaa ggagaaaata ccgcatcagg cgctcttcgc ctccctcgt	780
cactgactcg ctgcgctcgg tcgttcggct gcggcgagcg gtatcagctc actcaaaggc	840
ggtaatacgg ttatccacag aatcagggga taacgcagga aagaacatgt gagcaaaagg	900
ccagcaaaag gccaggaacc gtaaaaaggc cgcgttgctg gcgtttttcc ataggctccg	960
ccccctgac gagcatcaca aaaatcgacg ctcaagtcag aggtggcgaa acccgacagg	1020
actataaaga taccaggcgt tccccctgg aagctccctc gtgcgctctc ctgttccgac	1080
cctgccgctt accggatacc tgctccctt tctcccttcg ggaagcgtgg cgctttctca	1140
tagctcacgc ttaggtatc tcagttcggt gtaggtcggt cgctccaagc tgggctgtgt	1200
gcacgaaccc cccgttcagc ccgaccgtg cgccttatcc ggtaactatc gtcttgagtc	1260
caaccggta agacacgact tatgccact ggcagcagcc actggtaaca ggattagcag	1320
agcgaggat gtaggcggtg ctacagagtt ctgaagtgg tggcctaact acggtacac	1380
tagaaggaca gtatttggtg tctgcgctct gctgaagcca gttaccttcg gaaaaagagt	1440
tggtagctct tgatccggca aacaaaccac cgtcggtagc ggtggttttt ttgtttgcaa	1500
gcagcagatt acgcgcagaa aaaaaggatc tcaagaagat cctttgatct tttctacggg	1560
gtctgacgct cagtggaaac aaaactcacg ttaagggtatt ttggatcatga gattatcaaa	1620
aaggatcttc acctagatcc ttttaaatga aaaatgaagt tttaaatcaa tctaaagtat	1680
atatgagtaa acttggtctg acagttacca atgcttaatc agtgaggcac ctatctcagc	1740
gatctgtcta tttcgttcat ccatagttgc ctgactcccc gtcgtgtaga taactacgat	1800
acgggaggggc ttaccatctg gccccagtgc tgcaatgata ccgcgagacc cacgctcacc	1860
ggctccagat ttatcagcaa taaaccagcc agccggaagg gccgagcgca gaagtgggcc	1920
tgcaacttta tccgctcca tccagtctat taattgttgc cgggaagcta gagtaagtag	1980
ttcgccagtt aatagtttgc gcaacgttgt tgccattgct gcaggcatcg tgggtgcacg	2040
ctcgtcgttt ggtatggctt cattcagctc cggttcccaa cgatcaaggc gagttacatg	2100
atcccccatg ttgtgcaaaa aagcgggttag ctcccttcggt cctccgatcg ttgtcagaag	2160
taagtggcc gcagtgttat cactcatggt tatggcagca ctgcataatt ctcttactgt	2220

catgccatcc gtaagatgct tttctgtgac tggtagtac tcaaccaagt cattctgaga	2280
atagtgtatg cggcgaccga gttgctcttg cccggcgta acacgggata ataccgcgcc	2340
acatagcaga actttaaaag tgctcatcat tggaaaacgt tcttcggggc gaaaactctc	2400
aaggatctta ccgctgttga gatccagttc gatgtaacc actcgtgcac ccaactgac	2460
ttcagcatct tttactttca ccagcgtttc tgggtgagca aaaacaggaa ggcaaaatgc	2520
cgcaaaaaag ggaataaggc cgacacggaa atgttgaata ctcatactct tcctttttca	2580
atattattga agcatttacc aggggtattg tctcatgagc ggatacatat ttgaatgtat	2640
ttagaaaaat aaacaaatag gggttccgcg cacatttccc cgaaaagtgc cacctgacgt	2700
ctaagaaacc attattatca tgacattaac ctataaaaat aggcgtatca cgaggccctt	2760
tcgtcttcaa gaattctgaa ccagtcctaa aacgagtaaa taggaccggc aattcttcaa	2820
gcaataaaca ggaataccaa ttattaaaag ataacttagt cagatcgtac aataaagctt	2880
tgaagaaaaa tgcgccttat tcaatctttg ctataaaaaa tggcccaaaa tctcacattg	2940
gaagacattt gatgacctca tttctttcaa tgaaggcctt aacggagtgt actaatgttg	3000
tgggaaattg gagcgataag cgtgtctctg ccgtggccag gacaacgtat actcatcaga	3060
taacagcaat acctgatcac tacttcgcac tagtttctcg gtactatgca tatgatccaa	3120
tatcaaagga aatgatagca ttgaaggatg agactaatcc aattgaggag tggcagcata	3180
tagaacagct aaagggtagt gctgaaggaa gcatacgata ccccgcatgg aatgggataa	3240
tatcacagga ggtactagac tacctttcat cctacataaa tagacgcata taagtacgca	3300
tttaagcata aacacgcaat atgccgttct tctcatgtat atatataac aggcaacacg	3360
cagatatagg tgcgacgtga acagttagct gtatgtgcgc agctcgcgtt gcattttcgg	3420
aagcgctcgt tttcggaac gctttgaagt tcctattccg aagttcctat tctctagaaa	3480
gtataggaac ttcagagcgc ttttgaaaac caaaagcgt ctgaagacgc actttcaaaa	3540
aacaaaaac gcaccggact gtaacgagct actaaaatat tgcaataacc gttccacaa	3600
acattgctca aaagtatctc tttgctatat atctctgtgc tatatcccta tataacctac	3660
ccatccacct ttcgctcctt gaacttgcat ctaaaactga cctctacatt ttttatgttt	3720
atctctagta ttactcttta gacaaaaaaa ttgtagtaag aactattcat agagtgaatc	3780
gaaaacaata cgaaaatgta aacatttctt atacgtagta tatagagaca aaatagaaga	3840
aaccgttcat aattttctga ccaatgaaga atcatcaacg ctatcacttt ctgttcacaa	3900
agtatgcgca atccacatcg gtatagaata taatcgggga tgcctttatc ttgaaaaaat	3960

gcacccgcag cttcgctagt aatcagtaaa cgcgggaagt ggagtcaggc ttttttatg	4020
gaagagaaaa tagacaccaa agtagccttc ttctaacctt aacggaccta cagtgcacaaa	4080
agttatcaag agactgcatt atagagcgca caaaggagaa aaaaagtaat ctaagatgct	4140
ttgttagaaa aatagcgctc tcgggatgca tttttgtaga acaaaaaaga agtatagatt	4200
ctttgttggg aaaaatagcg tctcgcttg cttttctgtt ctgtaaaaat gcagctcaga	4260
ttctttgttt gaaaaattag cgctctcgcg ttgcattttt gttttacaaa aatgaagcac	4320
agattcttcg ttggtaaaat agcgctttcg cgttgcatth ctgttctgta aaaatgcagc	4380
tcagattctt tgtttgaaaa attagcgctc tcgcgttgca tttttgttct acaaaatgaa	4440
gcacagatgc ttcgttaaca aagatatgct attgaagtgc aagatggaaa cgcagaaaaat	4500
gaaccgggga tgcgacgtgc aagattacct atgcaataga tgcaatagtt tctccaggaa	4560
ccgaaataca tacattgtct tccgtaaacg gctagactat atattattat acaggttcaa	4620
atatactatc tgtttcaggg aaaactccca ggttcggatg ttcaaaattc aatgatgggt	4680
aacaagtacg atcgtaaatc tgtaaacag tttgtcggat attaggctgt atctcctcaa	4740
agcgtattcg aatatcattg agaagctgca gcgtcacatc ggataataat gatggcagcc	4800
attgtagaag tgccttttgc atttctagtc tctttctcgg tctagctagt tttactacat	4860
cgcgaagata gaatcttaga tcacactgcc tttgtgagc tggatcaata gagtaacaaa	4920
agagtggtaa ggctcgtta aaggacaagg acctgagcgg aagtgtatcg tacagtagac	4980
ggagtatact agtatagct atagtcctg gaattctcat gtttgacagc ttatcatcga	5040
taagcttttc aattcaattc atcatttttt ttttattctt ttttttgatt tcggtttctt	5100
tgaaattttt ttgattcggg aatctccgaa cagaaggaag aacgaaggaa ggagcacaga	5160
cttagattgg tatatatagc catatgtagt gttgaagaaa catgaaattg cccagtattc	5220
ttaacceaac tgcacagAAC aaaaacctgc aggaacgaa gataaatcat gtcgaaagct	5280
acataaagg aacgtgctgc tactcatcct agtcctgttg ctgccaagct atttaatatc	5340
atgcacgaaa agcaacaaa cttgtgtgct tcattggatg ttcgtaccac caaggaatta	5400
ctggagttag ttgaagcatt aggtcccaaa atttgtttac taaaaacaca tgtggatatac	5460
ttgactgatt ttccatgga gggcacagtt aagccgctaa aggcattatc cgccaagtac	5520
aattttttac tcttcgaaga cagaaaattt gctgacattg gtaatacagt caaattgcag	5580
tactctcggg gtgtatacag aatagcagaa tgggcagaca ttacgaatgc acacggtgtg	5640
gigggccag gtattgttag cggtttgaag caggcggcag aagaagtaac aaaggaacct	5700
agaggccttt tgatgttagc agaattgtca tgcaagggt ccctatctac tggagaatat	5760
actaagggtg ctgttgacat tgcgaagagc gacaaagatt ttgttatcgg ctttattgct	5820

caaagagaca tgggtggaag agatgaaggt tacgattggt tgattatgac acccggtgtg	5880
ggttttagatg acaaggga cgcattgggt caacagtata gaaccgtgga tgatgtggtc	5940
tctacaggat ctgacattat tattgttga agaggactat ttgcaaagg aagggatgct	6000
aaggtagagg gtgaacgtta cagaaaagca ggctgggaag catatttgag aagatgcggc	6060
cagcaaaact aaaaaactgt attataagta aatgcatgta tactaaactc acaaattaga	6120
gcttcaattt aattatatca gttattacc gggaatctcg gtcgtaatga aaggaaaagc	6180
gcaacggacg ggcgagtaga ttgcgcaaca tgcgagcatg atccagagat ttctgaagca	6240
gcaaaaggat gttccatgta catgacgcgc ggcttgcggg aaattgttgg caaattttcc	6300
ggcgtagccc aaaacgcgct gtcgtcaagt cgttaagggc gtgcccttca tcatccgatc	6360
tggagtcaaa atgtcctcac gtaaagagct tgccaatgct attcgtgcgc tgagcatgga	6420
cgcagtacag aaagccaaat ccggtcacc gggtgcccct atgggtatgg ctgacattgc	6480
cgaagtcctg tggcgtgatt tctgaaaca caaccgcag aatccgtcct gggtgaccg	6540
tgaccgttc gtgctgtcca acggccacgg ctccatgctg atctacagcc tgctgcacct	6600
caccggttac gatctgccga tggaagaact gaaaaacttc cgtcagctgc actctaaaac	6660
tccgggccac ccggaagtag gttataccgc tgggtgtgga accaccaccg gtccgctggg	6720
tcagggtatt gccaacgcag tcggtatggc gattgcagaa aaaacgctgg cggcgcagtt	6780
taaccgtcca ggtcacgaca ttgtcgacca ctacacctac gccttcatgg gcgacggctg	6840
catgatggaa ggcatctccc acgaagtttg ctctctggcg ggtacgctga agctgggtaa	6900
actgattgcg ttctacgatg acaacggtat ctcaatcgat ggtcacgttg aaggctggtt	6960
cactgacgac accgcaatgc gtttcgaagc tiacggctgg cacgttattc gcgacatcga	7020
cggatcatgac gggcatcca tcaaacgcgc agtagaagaa gcgcgcgcag tgactgacaa	7080
accgtccctg ctgatgtgca aaaccatcat cggtttcggg tcccgaaca aagccgttac	7140
ccacgactcc cacggtgcgc cgctgggcga cgctgaaatt gccctgacct gcgaacagct	7200
gggtggaaa tacgcgccgt tcgaaatccc gtctgaaatc tatgctcagt gggatgcgaa	7260
agaagcaggc caggcgaaag aatctgcatg gaatgagaag tttgcggctt acgcgaaagc	7320
ttatccgcag gaagcggctg aatttaccg ccgtatgaaa ggcgaaatgc cgtctgactt	7380
cgacgcaaaa gcgaaagagt ttatcgctaa actgcaggct aatccggcga aaatcgccag	7440
ccgtaaagcg tcgcagaatg ctatcgaagc gtccggccc ctgttgccctg aattcctcgg	7500
cggctctgct gacctggcac cgtctaacct gacctgtgg tctggttcta aagcaatcaa	7560
cgaagatgct gcagtaact acatccacta cgggtgttcg gagttcggtg tgaccgcgat	7620
tgctaacggt atctccctgc acggtgggtt cctgccgtac acctccacct tctgatgtt	7680

cgtaggaatac gcacgtaacg ccgtacgtat ggctgctgctg atgaaacagc gtcaggtgat	7740
ggtttacacc cagcactcca tcggctctggg cgaagatggc ccgactcacc agcccggttga	7800
gcaggtcgct tctctgctcg tgaccccgaa catgtctaca tggcgtccgt gtgaccaggt	7860
tgaatccgctg gtgcgctgga aatacggcgt tgagcgtcag gacggcccca ctgcgcttat	7920
cctctcccgat cagaacctgg cgcagcagga acgaactgaa gagcaactgg caaacatcgc	7980
gcgcggtggt tatgtgctga aagactgcgc cggtcagccg gaactgattt tcatcgctac	8040
cggttcagaa gttgaactgg ctgttgctgc ctacgaaaaa ctgactgccg aaggcgtgaa	8100
agcgcgctg gtgtccatgc cgtctaccga cgcatttgac aagcaggatg ctgcttaccg	8160
tgaatccgta ctgccgaaag cggttactgc acgcgttgct gtagaagcgg gtattgctga	8220
ctactggtac aagtatgttg gcctgaacgg tgctatcgtc ggatgacca ccttcggtga	8280
atctgctccg gcagagctgc tgtttgaaga gttcggttc actgttgata acgttgctgc	8340
gaaagcaaaa gaactgctgt aattagcatt tcgggtaaaa aggtcgcttc ggcgaccttt	8400
tttattacct tgatatgtcc gtttcggac aagcaataga taaagcgtgt ttagatcac	8460
aaatatattat atgcaataaa tatcaattat gtaatatgca tcacgatatg cgtattgaca	8520
tttgttgta taactataac tcaatgttat ataagaaatt aacttgaaat aattaacaaa	8580
caaaggagtt acagttagaa attgtaggag agatctcggt tttcgcgaca atctggcggt	8640
tttcttgcta attccaggat taatccgttc atagtgtaaa accccgttta cacattctga	8700
cggaagatat agattggaag tattgcattc actaagataa gtatggcaac actggaacag	8760
acatgaatta tcagaacgac gatttacgca tcaaagaaat caaagagtta ctctctctg	8820
tcgcatgtct ggaaaaattc cccgtactg aaaatgccgc gaatcgggt gcccatgcc	8880
gaaaagcgat ccataagatc ctgaaaggta atgatgatc cctgttggt gtgattggcc	8940
catgctcaat tcatgatcct gtgcggcaa aagagtatgc cactcgcttg ctggcgtgc	9000
glgaagagct gaaagatgag ctggaaatcg taatgcgcgt ctattttgaa aagccgcgta	9060
ccacggtggg ctggaaagg ctgattaacg atccgcataat ggataatagc ttccagatca	9120
acgacggtct gcgtatagcc cgtaaatgct tgcttgatat taacgacagc ggtctgccag	9180
cggcaggtga gtttctcgat atgatcacc cacaatatct cgctgacctg atgagctggg	9240
gcgcaattgg cgcacgtacc accgaatcgc aggtgcaccg cgaactggca tcagggtttt	9300
cttgtccggt cggcttcaaa aatggcaccg acggtacgat taaagtggct atcgatgcca	9360
ttaatgccgc cggtagccgc cactgcttc tgtccgtaac gaaatggggg cattcggcga	9420

ttgtgaatac cagcggtaac ggcgattgcc atatcattct gcgcggcggt aaagagccta	9480
actacagcgc gaagcacgtt gctgaagtga aagaagggt gaacaaagca ggcctgccag	9540
cacagtgat gatcgatttc agccatgcta actcgtccaa acaattcaaa aagcagatgg	9600
atgtttgtgc tgacgtttgc cagcagattg ccggtggcga aaaggccatt attggcgtga	9660
tggtagaaag ccatctggtg gaaggcaatc agagcctcga gagcggggag ccgctggcct	9720
acggtaagag catcaccgat gcctgcatcg gctgggaaga taccgatgct ctgttacgtc	9780
aactggcgaa tgcagtaaaa gcgcgtcgcg ggtaaggttt aattgtcgga tgcgccgtca	9840
gagtggcgta tccgatgaat caccacaggc ctgataagtc gcgcagcgtc gcatcaggca	9900
atgtgtccca ttgttagcaa caaaaaagcc gactcacttg cagtcggctt tctcatTTTA	9960
aacgaatgac gtttacttcg ctttaccctg gtttgcaacc gccgtgctt tcgctctcga	10020
ggctattgac gacagctatg gttcactgtc caccaaccaa aactgtgctc agtaccgcca	10080
atatttctcc cttgaggggt acaaagaggt gtccttagaa gagatccacg ctgtgtaaaa	10140
attttacaaa aaggtattga ctttcctac aggggtgtga ataatttaac tacagcgagg	10200
ggcaaccccg cctgttctag aggaggagga atcgccatgg agaggattgt cgttactctc	10260
ggggaacgta gttaccaat taccatcgca tctggtttgt ttaatgaacc agcttcattc	10320
ttaccgctga aatcgggcga gcaggtcatg ttggtcacca acgaaacct ggctcctctg	10380
tatctcgata aggtccgcgg cgtacttgaa caggcgggtg ttaacgtcga tagcgttatc	10440
ctccctgacg gcgagcagta taaaagcctg gctgtactcg ataccgtctt tacggcgttg	10500
ttacaaaagc cgcattggtc cgatactacg ctggtggcgc ttggcggcgg cgtagtgggc	10560
gatctgaccg gcttcgcggc ggcgagttat cagcgcgggtg ttcgtttcat tcaagtcagg	10620
acgacgttac tgtcgagggt cgattcctcc gttggcggca aaactgcggg caaccatccc	10680
ctcggtaaaa acatgattgg cgcgttctac cagcctgctt cagtgggtgt ggatctcgac	10740
tgtctgaaaa cgcttcccc gcgtgagtta gcgtcggggc tggcagaagt catcaatac	10800
ggcattattc ttgacggtgc gtttttcaac tggttggaag agaacttgga tgcgttgttg	10860
cgtctggacg gtccggcaat ggcgtactgt attcgccgtt gttgtgaact gaaggcagaa	10920
gttgtcgccg ccgacgagcg cgaaaccggg ttacgtgctt tactgaatct gggacacacc	10980
tttggtcatg ccattgaagc tgaatgggg tatggcaatt gggtacatgg tgaagcggtc	11040
gctgcgggta tggtagatgg gcgcgggacg tcggaacgtc tcgggcagtt tagttctgcc	11100
gaaacgcagc gtattataac cctgctcacg cgggctgggt taccggtcaa tgggccgcgc	11160
gaaatgtccg cgcaggcgta ttaccgcat atgctgcgtg acaagaaagt ccttcgggga	11220
gagatgcgct taattcttcc gttggcaatt ggtaagagtg aagttcgag cggcgtttcg	11280

cacgagcttg ttcttaacgc cattgccgat tgcatacag cgtaatcatc gttcatgcct 11340
gatgccgcta ttaggccggg ataaggcggt cacgccgat ccggcaaccg atgcctgatg 11400
cgacgcggtc gcgtcttacc aggcctacag gtcgatgccg atatgtacat cgtattcggc 11460
aattaataca tagca 11475

<210> 33

<211> 7769

<212> DNA

<213> Escherichia coli

<220><221> misc_feature

<222> (1)..(7769)

<223> DNA sequence of the plasmid YEP24

<400> 33

gaattctgaa ccagtcctaa aacgagtaaa taggaccggc aattcttcaa gcaataaaca 60
ggaataccaa ttattaaaag ataacttagt cagatcgtac aataaagctt tgaagaaaaa 120
tgcgccttat tcaatctttg ctataaaaaa tggcccaaaa tctcacattg gaagacattt 180
gatgacctca tttctttcaa tgaagggcct aacggagttg actaatgttg tgggaaattg 240
gagcgataag cgtgcttctg ccgtggccag gacaacgtat actcatcaga taacagcaat 300

acctgatcac tacttcgcac tagtttctcg gtactatgca tatgatccaa tatcaaagga 360
aatgatagca ttgaaggatg agactaatcc aattgaggag tggcagcata tagaacagct 420
aaagggtagt gctgaaggaa gcatacgata ccccgcatgg aatgggataa tatcacagga 480
ggtactagac tacctttcat cctacataaa tagacgcata taagtagcga ttttagcata 540
aacacgcaat atgccgttct tctcatgtat atatatatac aggcaacacg cagatatagg 600
tgcgacgtga acagttagct gtatgtgcgc agctcgcgtt gcattttcgg aagcgtctgt 660
tttcggaaac gctttgaagt tcctattccg aagttcctat tctctagaaa gtataggaac 720

ttcagagcgc ttttgaaaac caaaagcgct ctgaagacgc actttcaaaa aacaaaaaac 780
gcaccggact gtaacgagct actaaaatat tgcgaatacc gcttcacaa acattgctca 840
aaagtatctc tttgctatat atctctgtgc tatatcccta tataacctac ccatccacct 900
ttcgctcctt gaacttgcac ctaaactcga cctctacatt ttttatgttt atctctagta 960
ttactcttta gacaaaaaaa ttgtagtaag aactattcat agagtgaatc gaaaacaata 1020
cgaaaatgta aacatttcct atacgtagta tatagagaca aaatagaaga aaccgttcat 1080
aattttctga ccaatgaaga atcatcaacg ctatcacttt ctgttcacaa agtatgcgca 1140

atccacatcg gtatagaata taatcgggga tgcctttatc ttgaaaaaat gcacccgcag	1200
cttcgctagt aatcagtaaa cgcgggaagt ggagtcaggc tttttttatg gaagagaaaa	1260
tagacaccaa agtagccttc ttctaacctt aacggaccta cagtgcacaaa agttatcaag	1320
agactgcatt atagagcgca caaaggagaa aaaaagtaat ctaagatgct ttgttagaaa	1380
aatagcgctc tcgggatgca tttttgtaga acaaaaaaga agtatagatt ctttgttggt	1440
aaaatagcgc tctcgcttg catttctgtt ctgtaaaaat gcagctcaga ttctttgttt	1500
gaaaaattag cgctctcgcg ttgcattttt gttttacaaa aatgaagcac agattcttcg	1560
ttggtaaaaat agcgctttcg cgttgcatth ctgttctgta aaaatgcagc tcagattctt	1620
tgtttgaaaa attagcgctc tcgcttgca tttttgttct acaaaatgaa gcacagatgc	1680
ttcgtaaca aagatatgct attgaagtgc aagatggaaa cgcagaaaat gaaccgggga	1740
tgcgactgc aagattacct atgcaataga tgcaatagtt tctccaggaa ccgaaataca	1800
tacattgtct tccgtaaagc gctagactat atattattat acaggttcaa atatactatc	1860
tgtttcaggg aaaactccca ggttcggatg ttcaaaatc aatgatgggt aacaagtacg	1920
atcgtaaadc tgtaaaacag ttgtcggat attaggctgt atctctcaa agcgtattcg	1980
aatatcattg agaagctgca gcgtcacatc ggataataat gatggcagcc attgtagaag	2040
tgccttttgc atttctagtc tctttctcgg tctagctagt ttactacat cgcgaagata	2100
gaatcttaga tcacactgcc ttgtctgagc tggatcaata gagtaacaaa agagtggtaa	2160
ggcctcgta aaggacaagg acctgagcgg aagtgtatcg tacagtagac ggagtatact	2220
agtatagtct atagtccgtg gaattctcat gtttgacagc ttatcatcga taagcttttc	2280
aattcaatc atcatTTTT ttttattctt ttttttgatt tcggtttctt tgaaattttt	2340
ttgattcggg aatctccgaa cagaaggaag aacgaaggaa ggagcacaga cttagattgg	2400
tatatatacg catatgtagt gttgaagaaa catgaaattg cccagtattc ttaaccaaac	2460
tgcacagaac aaaaacctgc aggaacgaa gataaatcat gtcgaaagct acatataagg	2520
aacgtgctgc tactcatcct agtctgttg ctgccaagct atttaatac atgcacgaaa	2580
agcaacaaa cttgtgtgct tcattggatg ttctgtaccac caaggaatta ctggagttag	2640
ttgaagcatt aggtcccaaa atttgtttac taaaaacaca tgtggatata ttgactgatt	2700
tttccatgga gggcacagtt aagccgctaa aggcattatc cgccaagtac aattttttac	2760
tcttcgaaga cagaaaattt gctgacattg gtaatacagt caaattgcag tactctcgcg	2820
gigtatacag aatagcagaa tgggcagaca ttacgaatgc acacggtgtg gtgggccag	2880
gtattgttag cggtttgaag caggcggcag aagaagtaac aaaggaacct agaggccttt	2940
tgatgttagc agaattgtca tgcaagggtc ccctatctac tggagaatat actaagggtta	3000

ctgttgacat tgcgaagagc gacaaagatt ttgttatcgg ctttattgct caaagagaca	3060
tgggtggaag agatgaaggt tacgattggt tgattatgac acccggtgtg ggtttagatg	3120
acaagggaga cgcattgggt caacagtata gaaccgtgga tgatgtggtc tctacaggat	3180
ctgacattat tattgttga agaggactat ttgcaaaggg aagggatgct aaggtagagg	3240
gtgaacgtta cagaaaagca ggctgggaag catatttgag aagatgcggc cagcaaaact	3300
aaaaaactgt attataagta aatgcatgta tactaaactc acaaattaga gcttcaattt	3360
aattatatca gttattaccc gggaatctcg gtcgtaatga tttttataat gacgaaaaaa	3420
aaaaaattgg aaagaaaaag ctttaatgcg gtagtttata acagttaaat tgctaacgca	3480
gtcaggcacc gtgtatgaaa tctaacaatg cgtcatcgt catcctcggc accgtcacc	3540
tggatgctgt aggcataggc ttggttatgc cggtaactgcc gggcctcttg cgggatatcg	3600
tccattccga cagcatcgcc agtcactatg gcgtgctgct agcgtatat gcgttgatgc	3660
aatttctatg cgcacccgtt ctgggagcac tgcgcaccg ctttggccgc cgcccagtc	3720
tgctcgcttc gctacttga gccactatcg actacgcgat catggcgacc acaccgtcc	3780
tgtggatcct ctacgccga cgcacgtgg ccggcatcac cggcgccaca ggtgcggttg	3840
ctggcgcccta tatcgccgac atcacgatg gggaagatcg ggctcgccac ttcgggctca	3900
tgagcgcttg tttcgcgctg ggtatggtgg caggccccgt ggccggggga ctgttggcg	3960
ccatctcctt gcattgcacca ttcttgcgg cggcgggtgct caacggcctc aacctactac	4020
tgggtgcttt cctaattcag gagtgcata agggagagcg tcgaccgatg cccttgagag	4080
cttcaaccc agtcagctcc ttccggtggg cgcggggcat gactatcgtc gccgcactta	4140
tgactgtctt ctttatcatg caactcgtag gacaggtgcc ggacgcgctc tgggtcattt	4200
tcggcgagga ccgttttcg tggagcgca cgatgatcgg cctgtcgtt gcggtattcg	4260
gaatcttgca cgcctcgt caagccttcg tcaactgtcc cgccacaaa cgtttcgcg	4320
agaagcaggc cattatcgcc ggcatggcgg ccgacgcgt gggctacgtc ttgctggcgt	4380
tcgcgacgcg aggttgatg gccttccca ttatgattct tctcgttcc ggcggcatcg	4440
ggatgccccg gttgcaggcc atgctgtcca ggcaggtaga tgacgacat caggacagc	4500
ttcaaggatc gtcgcggct cttaccagcc taacttcgat cactggaccg ctgatcgta	4560
cggcgattta tgccgctcg gcgagcaca ggaacgggtt ggcatggatt gtaggcgccg	4620
ccctatacct tgctgcctc ccgcgttgc gtcgcggtgc atggagccgg gccacctga	4680
cctgaatgga agccggcgcc acctcgctaa cggattcacc actccaagaa ttggagccaa	4740
tcaattcttg cggagaactg tgaatgcgca aaccaacct tggcagaaca tatccatcg	4800
gtccgccatc tccagagcc gcacgcggcg catctcgggc agcgttgggt cctggccacg	4860

gggtgcgcattg atcgtgctcc tgtcgttgag gacccggcta ggctggcggg gttgccttac	4920
tggttagcag aatgaatcac cgatacgcga gcgaacgtga agcgactgct gctgcaaaac	4980
gtctgcgacc tgagcaacaa catgaatggt ctccggtttc cgtgtttcgt aaagtctgga	5040
aacgcggaag tcagcgcctt gcaccattat gtcccgatc tgcatcgcag gatgctgctg	5100
gctacctgtt ggaacacctt catctgtatt aacgaagcgc tggcattgac cctgagtcat	5160
ttttctctgg tcccgcgcga tccataccgc cagttgttta cctcacaac gttccagtaa	5220
ccgggcatgt tcatcatcag taaccgctat cgtgagcatc ctctctcgtt tcatcggtat	5280
cattaccccc atgaacagaa attccccctt acacggaggc atcaagtgcg caaacaggaa	5340
aaaacgcgcc ttaacatggc ccgctttatc agaagccaga cattaacgtt tctggagaaa	5400
ctcaacgagc tggacgcgga tgaacaggca gacatctgtg aatcgcttca cgaccacgtt	5460
gatgagcttt accgcagctg cctcgcgcgt ttccgtgatg acggtgaaaa cctctgacac	5520
atgcagctcc cggagacggt cacagcttgt ctgtaagcgg atgccgggag cagacaagcc	5580
cgtcagggcg cgtcagcggg tgttggcggg tgtcggggcg cagccatgac ccagtcacgt	5640
agcgatagcg gagtgtatc tggcttaact atgcggcatc agagcagatt gtactgagag	5700
tgcaccatat gcggtgtgaa ataccgcaca gatgcgtaag gagaaaatac cgcacagggc	5760
gctcttccgc ttctctgctc actgactcgc tgcgctcggg cgttcggctg cggcgagcgg	5820
tatcagctca ctcaaaggcg gtaatacggg tatccacaga atcaggggat aacgcaggaa	5880
agaacatgtg agcaaaaggc cagcaaaagg ccaggaaccg taaaaaggcc gcgttgctgg	5940
cgtttttcca taggctccgc cccctgacg agcatcaca aaatcgacgc tcaagtcaga	6000
ggtggcgaaa cccgacagga ctataaagat accagcggtt tccccctgga agctccctcg	6060
tgcgctctcc tgttccgacc ctgccgctta ccgataacct gtccgccttt ctcccttcgg	6120
gaagcgtggc gctttctcat agctcacgtt gtaggtatct cagttcggtg taggtcgttc	6180
gctccaagct gggctgtgtg cacgaacccc ccgttcagcc cgaccgtgc gccttatccg	6240
gtaactatcg tcttgagtc aaccgggtaa gacacgactt atcgccactg gcagcagcca	6300
ctggtaacag gattagcaga gcgaggtatg taggcggtgc tacagagttc ttgaagtggg	6360
ggcctaacta cggctacact agaaggacag tatttggtat ctgcgctctg ctgaagccag	6420
ttaccttcgg aaaaagagtt ggtagctctt gatccggcaa acaaaccacc gctggtagcg	6480
gtgggttttt tgtttgcaag cagcagattt cgcgcagaaa aaaaggatct caagaagatc	6540
ctttgatctt ttctacgggg tctgacgctc agtggaacga aaactcacgt taagggattt	6600

tggtcatgag attatcaaaa aggatcttca cctagatcct tttaaattaa aaatgaagtt 6660
 ttaaatacaat ctaaagtata tatgagtaaa ctggtctga cagttacaa tgcttaataca 6720
 gtgaggcacc tatctcagcg atctgtctat ttcgttcac catagttgcc tgactccccg 6780
 tcgtgtagat aactacgata cgggagggct taccatctgg cccagtgct gcaatgatac 6840
 cgcgagaccc acgtcaccg gctccagatt tatcagcaat aaaccagcca gccggaaggg 6900
 ccgagcgag aagtggctct gcaactttat ccgcctccat ccagtcattt aattgttgcc 6960
 gggaagctag agtaagtagt tcgccagtta atagtttgcg caacgttgtt gccattgctg 7020

caggcatcgt ggtgtcacgc tcgtcgtttg gtatggcttc attcagctcc ggttcccaac 7080
 gatcaaggcg agttacatga tccccatgt tgtgcaaaaa agcggttagc tccttcggtc 7140
 ctccgatcgt tgcagaagt aagtggccg cagtgttacc actcatggtt atggcagcac 7200
 tgcataattc tcttactgtc atgccatccg taagatgctt tctgtgact ggtgagtact 7260
 caaccaagtc attctgagaa tagtgtatgc ggcgaccgag ttgctcttgc ccggcgtaaa 7320
 cacgggataa taccgcgcca catagcagaa ctttaaaagt gctcatcatt ggaaaacgtt 7380
 ctccggggcg aaaactctca aggatcttac cgtgttgag atccagttcg atgtaacca 7440

ctcgtgcacc caactgatct tcagcatctt ttactttcac cagcgtttct gggtgagcaa 7500
 aaacaggaag gcaaaatgcc gcaaaaaagg gaataaggcg gacacggaaa tgttgaatac 7560
 tcatactctt ctttttcaa tattattgaa gcatttatca gggttattgt ctcatgagcg 7620
 gatacatatt tgaatgtatt tagaaaaata acaaatagg ggttccgcgc acatttcccc 7680
 gaaaagtgcc acctgacgtc taagaaacca ttattatcat gacattaacc tataaaaata 7740
 ggcgtatcac gaggcccttt cgtcttcaa 7769

<210> 34
 <211> 2000
 <212> DNA

<213> Escherichia coli
 <220><221> gene
 <222> (1)..(2000)
 <223> DNA sequence of the deleted aroE region
 <400> 34

actacgtccg tcctctgaaa tcttcagcgg atggacatat cgtcaaagtt ctggaggggc 60
 aggtttgccc tgcatgtggc gcaaatctgg tattacgcca gggacgcttt ggtatgttta 120
 ttggttgcat taactacct gaatgcgaac ataccgaact tatcgataaa ccggacgaaa 180
 cagcaattac atgcccccaa tgctggacgg gccatctggt ccagcgccgc tcccgttatg 240

gcaaaacatt tcactcttgt gatcgctacc cggagtgtca atttgccatt aacttcaaac	300
ccatagctgg agaatgccct gagtgtcatt atccgctact catcgaaaag aaaaccgcgc	360
aggggtgtaaa acacttttgt gccagtaaac aatgtggaaa gccggtttcg gcggaataat	420
aacgtgaata ataacctgca aagagacgct atcgacgctg cgatagatgt tctcaatgaa	480
gaacgtgtca tcgcctatcc aacggaagcc gttttcggtg ttgggtgcga tcttgatagc	540
gaaacagcag tgatgcgact gttggagtta aaacagcgtc cggttgataa ggggctgatt	600
ttaatcgag caaattacga gcagcttaaa ccctatattg atgacaccat gttgactgac	660
gtgcagcgtg aaaccatttt ttcccgtgg ccaggtcctg tcacctttgt ctttcccgcg	720
cctgcgacaa caccgcgctg gttgacgggc cgctttgatt cgcttgctgt acgagtcacc	780
gaccatccgt tgggtgttgc tttgtgccag gcttatggta aaccgctggt ttctaccagt	840
gccaaactga gtggattgcc accttgtcga acagtagacg aagttcgcgc acaatttggc	900
gcggcgttcc cggttgtgcc tggtgaaacg ggggggcgtt taaatccttc agaaatccgc	960
gatgccctga cgggtgaact gtttcgacag gggtaacata atcaggccat ccagtttccg	1020
gacaggaag agtgggacga gaataaaaaa tgtgtatgtt ttcccgtct cgtgaatggt	1080
atgcaactga catgcgcgat ctctggcgag agtctggcgt atcgctttac tggagatacg	1140
ccagaacagt ggttagcgag ttttcgtcag catcgctggg acctggaaga agaagcggaa	1200
aacttaattc aggaacaaag tgaagatgat caaggctggg tctggttacc ctgatccaga	1260
tattcgtcct tccatttcac gtaattattc gcggaatagc gtaaccagc cttctcttca	1320
tcacttaacg ggcgcatctg tttagcgggg ctaccgagat acagatatcc gctctccagc	1380
cgtttatatt gtgggaccag actaccgcga ccaatcatca catcatcttc tactattgcg	1440
ccatcaagta aaattgagcc catcccaacc aaaactcgat tgccaatggt gcagccgtgg	1500
agcatcacct tgtgaccaac agtgacatct tcgccaatgg ttaatgggtt gccatctggg	1560
ttgtacgagg atttatgagt gacatgcaac atactgcat cctggatatt ggtgcgtgct	1620
ccgatctgta cataatgtac atctccacga atcacaacga gcggccagat ccccatca	1680
tcagccagac gaacgtcacc aatcacgaca ctgctatcgt cgatcattac gcgctgaccg	1740
atttgtggaa aaagatcgcg gtatgggcgt aaaacatcag acatacttac ctcagcaata	1800
aatgatttac taatgacttt gggggcatta ttggccttgt gcaagtcttt tagtatgcaa	1860
aaaagcaccg ttttgtgtgc gattgcagca aaaagggtga aaaaacaaca aacagaaaaa	1920
aagatcaaaa aaatacttgt gcaaaaaatt gggatcccta taatgcccct ccgttgagac	1980

gacaacgtga aacacttcac 2000

<210> 35

<211> 2460

<212> DNA

<213> Escherichia coli

<220><221> gene

<222> (1)..(2460)

<223> DNA sequence of the integrated cassette ack::P15aroB

<400> 35

gatcggcggc ataaaacgga tcgcataacg cgtcatcttg ataacgcgat tttcgacaaa 60

gaccggggca aggcgttttt ccagcggcca cgtctttgag taatgctgtc cccggcgaaa 120

caagctaaaa aaattaacag aacgattatc cggcgttgac atgcttcacc tcaacttcac 180

atataaagat tcaaaaattt gtgcaaattc acaactcagc gggacaacgt tcaaacatt 240

ttgtcttcca taccactat caggatcct ttagcagcct gaaggcctaa gtagtacata 300

ttcattgagt cgicaaattc atatacatta tgccattggc tgaaaattac gcaaatggc 360

atagactcaa gatatttctt ccatcatgca aaaaaaatt tgcatgcat gatgttaac 420

ataaatgtcg gtgtcatcat gcgctacgt ctatggctcc ctgacgtttt tttagccacg 480

tatcaattat aggtacttcc ctcgaggcta ttgacgacag ctatggttca ctgtccacca 540

accaaaactg tgctcagtac cgccaatatt tctcccttga ggggtacaaa gaggtgtccc 600

tagaagagat ccacgtgtg taaaaatttt acaaaaagggt attgactttc cctacagggt 660

gtgtaataat ttaattacag gcgggggcaa cccgcctgt tctagaggag gaggaatcgc 720

catggagagg attgtcgta ctctcgggga acgtagtac ccaattacca tcgcatctgg 780

tttgtttaat gaaccagctt cattcttacc gctgaaatcg ggcgagcagg tcatgttggt 840

caccaacgaa accctggctc ctctgtatct cgataaggtc cgcggcgtac ttgaacaggc 900

gggtgttaac gtcgatagcg ttatcctccc tgacggcgag cagtataaaa gcctggctgt 960

actcgatacc gtctttacgg cgttgttaca aaagccgcat ggtcgcgata ctacgtggt 1020

ggcgcttggc ggcggcgtag tgggcgatct gaccggcttc gcggcggcga gttatcagcg 1080

cggtgttcgt ttattcaag tcccagcagc gttactgtcg caggtcgatt cctccgttgg 1140

cggcaaaact gcggtcaacc atcccctcgg taaaaacatg attggcgcgt tctaccagcc 1200

tgcttcagtg gtgttgatc tcgactgtct gaaaacgctt cccccgcgtg agttagcgtc 1260

ggggctggca gaagtcatca aatacggcat tattcttgac ggtgcgtttt tcaactggct 1320

ggaagagaat ctggatgcgt tgttgcgtct ggacgggccg gcaatggcgt actgtattcg 1380

ccgttggtgt gaactgaagg cagaagttgt cgccgccgac gagcgcgaaa cggggttacg	1440
tgctttactg aatctgggac acacctttgg tcatgccatt gaagctgaaa tggggatatgg	1500
caattggtta catggtgaag cggtcgctgc gggatatggtg atggcggcgc ggacgtcgga	1560
acgtctcggg cagtttagtt ctgccgaaac gcagcgtatt ataaccctgc tcacgcgggc	1620
tgggttaccg gtcaatgggc cgcgcgaaat gtccgcgcag gcgtatttac cgcatactgt	1680
gcgtgacaag aaagtccttg cgggagagat gcgcttaatt cttccgttgg caattggtaa	1740
gagtgaagtt cgcagcggcg ttctgcacga gcttgcttctt aacgccattg ccgattgtca	1800
atcagcgtaa tcatcgttca tgccgatgac cgctatgtag gccggataag gcgttcacgc	1860
cgcataccggc aaccgatgcc tgatgcgacg cggtcgctgc ttatcaggcc tacaggtcga	1920
tgccgatatg tacatcgtat tcggcaatta atacatagca ttccacaccg ccagctcagc	1980
tggcgggtgct gttttgtaac ccgccaaatc ggccgtaacg aaagaggata aaccgtgtcc	2040
cgtattatta tgctgatccc taccggaacc agcgtcggtc tgaccagcgt cagccttggc	2100
gtgatccgtg caatggaacg caaaggcgtt cgtctgagcg ttttcaaacc tatcgtcag	2160
ccgcgtaccg gtggcgatgc gcccgatcag actacgacta tcgtgcgtgc gaactcttcc	2220
accacgacgg ccgctgaacc gctgaaaatg agctacgttg aaggctctgt ttccagcaat	2280
cagaaagatg tgctgatgga agagatcgtc gcaaaactacc acgctaacac caaagacgt	2340
gaagtcgttc tggttgaagg tctggccccg acacgtaagc accagtttgc ccagctctctg	2400
aactacgaaa tcgctaaaaa gctgaatgcg gaaatcgtct tcgttatgtc tcagggcact	2460
	2460
<210> 36	
<211> 1000	
<212> DNA	
<213> Escherichia coli	
<220><221> gene	
<222> (1)..(1000)	
<223> DNA sequence of the poxB region	
<400> 36	
gcggccccgc tccgtatatg gattgggtag agcaggaagt gaaagcgtc gccgtgacgc	60
gtttctttaa agagaaattc ttcaccccag tagcgggaagc agcgaccagc ggtctgaaat	120
tcaccaaact gcaaccggca cgagaatttt acgccccggt tggcaccacg ctactggagg	180
cgctggaaag caataacgtt ccggttgctc ccgcctgccg tgcgggtgtt tgcggctgct	240

gtaagacaaa agtgggtttcc ggtgaatata cggtagagcag cacaatgacg ctgaccgacg 300
ccgaaatcgc tgaaggttac gtactggcct gctcctgcc tccgcagggg gatttggttc 360
tcgcataatc gccttatgcc cgatgatatt cctttcatcg ggctatttaa ccgttagtgc 420
ctcctttctc tcccatccct tccccctccg tcagatgaac taaacttggt accgttatca 480
cattcaggag atggagaacc aaagggtggc atttcccgtc ataataagga catgccatga 540

ttgatttacg cagtgatacc gttacccgac cgagccgcgc catgctcgaa gcaatgatgg 600
ccgccccggt tggggacgac gtttacggag acgaccctac cgtaaatgct ctgcaggact 660
acgccgcaga gctttccggt aaagaagccg ccatttttct gccgaccgac actcaggcca 720
acctggtcgc tctgctcagt cactgcgaac gtggcggaaga gtatattgtc ggtcaggccg 780
cgcataacta tctgtttgaa gccggtggcg cagcgggtgct gggcagtatt cagccgcaac 840
ccatcgacgc ggctgccgac ggcacgctac cgttgataa agtggcgatg aaaatcaaac 900
ccgacgatat ccatttcgcc cgcaccaaat tactcagtct ggaaaacacc cacaacggca 960

aagtgtgcc gcgtgaatac ctgaaagaag catgggaatt 1000

<210> 37

<211> 3392

<212> DNA

<213> Escherichia coli

<220><221> gene

<222> (1)..(3392)

<223> DNA sequence of the integrated cassette poxB::tktA

<400> 37

gcggccccgc tccgtatatg gattgggtag agcaggaagt gaaagcgctc ggcgtgacgc 60
gtttctttta agagaaattc ttacccccag tagcgggaagc agcgaccagc ggtctgaaat 120
tcaccaaact gcaaccggca cgagaatttt acgccccggt tggcaccacg ctactggagg 180
cgctggaaag caataacgtt ccggttgtcg ccgcctgccg tgcgggtgtt tgcggctgct 240

gtaagacaaa agtgggtttcc ggtgaatata cggtagagcag cacaatgacg ctgaccgacg 300
ccgaaatcgc tgaaggttac gtactggcct gctcctgcc tccgcagggg gatttggttc 360
tcgcataatc gccttatgcc cgatgatatt cctttcatcg ggctatttaa ccgttagtgc 420
ctcctttctc tcccatccct tccccctccg tcagatgaac taaacttggt accgttatca 480
cattcaggag atggagaacc aaggaaaagc gcaacggacg ggcgagtaga ttgcgcaaca 540
tgcgagcatg atccagagat ttctgaagca gcaaaaggat gttccatgta catgacgcgc 600

ggcttgcggt aaattgttgg caaatTTTcc ggcgtagccc aaaacgcgt gtcgtcaagt	660
cgtaaagggc gtgcccttca tcatccgatc tggagtcaaa atgtcctcac gtaaagagct	720
tgccaatgct attcgtgcgc tgagcatgga cgcagtagac aaagccaaat ccggtcaccc	780
gggtgccctt atgggtatgg ctgacattgc cgaagtcttg tggcgtgatt tctgaaaca	840
caaccgcag aatccgtcct gggctgaccg tgaccgttc gtgtgtcca acggccacgg	900
ctccatgctg atctacagcc tgctgcacct caccggttac gatctgccga tggaagaact	960
gaaaaacttc cgtcagctgc actctaaaac tccgggccac ccggaagtag gttataccgc	1020
tgggtgtgaa accaccaccg gtccgtggg tcagggtatt gccaacgcag tcggtatggc	1080
gattgcagaa aaaacgctgg cggcgcagtt taaccgtcca ggtcacgaca ttgtcgacca	1140
ctacacctac gccttcatgg gcgacggctg catgatggaa ggcatctccc acgaagtttg	1200
ctctctggcg ggtacgtga agctgggtaa actgattgcg ttctacgat acaacggtat	1260
ctcaatcgat ggtcacgttg aaggctggtt cactgacgac accgcaatgc gtttcgaagc	1320
ttacggctgg cagttattc gcgacatga cggcatgac gcggcatcca tcaaacgcgc	1380
agtagaagaa gcgcgcgcag tgactgacaa accgtccctg ctgatgtgca aaacatcat	1440
cggtttcggt tccccgaaca aagccggtac ccacgactcc cagggtgcgc cgctgggcga	1500
cgctgaaatt gccctgaccc gcgaacagct gggctggaaa tacgcgccgt tcgaaatccc	1560
gtctgaaatc tatgtcagt gggatgcgaa agaagcaggc caggcgaaag aatctgcatg	1620
gaatgagaag tttagcggtt acgcgaaagc ttatccgcag gaagcggctg aatttaccg	1680
ccgtatgaaa ggcgaaatgc cgtctgactt cgacgcaaaa gcgaaagagt ttatcgctaa	1740
actgcaggct aatccggcga aaatcgccag ccgtaaagcg tcgcagaatg ctatcgaagc	1800
gttcggcccc ctgttgctg aattctcgg cgctctgct gacctggcac cgtctaacct	1860
gacctgtgg tctggttcta aagcaatcaa cgaagatgct gcaggtaact acatccacta	1920
cgggtgtcgc gatttcggtg tgaccgcgat tgctaacggt atctccctgc acggtggttt	1980
cctgccgtac acctccacct tctgatgtt cgtggaatac gcacgtaacg ccgtacgtat	2040
ggctgcgctg atgaaacagc gtcaggatg ggtttacacc cagactcca tcggtctggg	2100
cgaagatggc ccgactcacc agccggttga gcaggctcgt tctctgcgcg tgaccccgaa	2160
catgtctaca tggcgtccgt gtgaccaggt tgaatccgcg gtcgcgtgga aatacggcgt	2220
tgagcgtcag gacggccccg ctgcgttat cctctccctg cagaacctgg cgcagcagga	2280
acgaactgaa gagcaactgg caaacatcgc gcgcgggtgt tatgtgtga aagactgcgc	2340

cggtcagccg gaactgattt tcatcgctac cggttcagaa gttgaactgg ctgttgctgc 2400
ctacgaaaaa ctgactgccg aaggcgtgaa agcgcgcgtg gtgtccatgc cgtctaccga 2460
cgcatctgac aagcaggatg ctgcttaccg tgaatccgta ctgccgaaag cggttactgc 2520
acgcgttgct gtagaagcgg gtattgctga ctactggtac aagtatgttg gcctgaacgg 2580
tgctatcgtc ggtatgacca ctttcgggtga atctgctccg gcagagctgc tgtttgaaga 2640
gttcggcttc actgttgata acgttggtgc gaaagcaaaa gaactgctgt aattagcatt 2700
tcgggtaaaa aggtcgcttc ggcgaccttt ttattacct tgatatgtcc gtttcgggac 2760

aagcaataga taaagcgtgt tgtagatcac aaatatattat atgcaataaa tatcaattat 2820
gtaatatgca tcacgatatg cgtattgaca ttgtttgta taaactataac tcaatgttat 2880
ataagaaatt aaaaagggtg gcatttcccg tcataataag gacatgccat gattgattta 2940
cgcatgata ccgttaccgg accgagccgc gccatgctcg aagcaatgat ggccgccccg 3000
gttggggacg acgtttacgg agacgaccct accgttaatg ctctgcagga ctacgccgca 3060
gagctttccg gtaaagaagc cgccattttt ctgccgaccg gcactcaggc caacctggtc 3120
gctctgctca gtcactgcga acgtggcgaa gagtatatgt tcggtcaggc cgcgcataac 3180

tatctgtttg aagccgggtg cgcagcgggtg ctgggcagta ttcagccgca acccatcgac 3240
gcggctgccg acggcacgct accgctggat aaagtggcga tgaaaatcaa acccgacgat 3300
atccatttcg cccgcaccaa attactcagt ctggaaaaca cccacaacgg caaagtgtctg 3360
ccgcgtgaat acctgaaaga agcatgggaa tt 3392

<210> 38

<211> 1045

<212> DNA

<213> Escherichia coli

<220><221> gene

<222> (1)..(1045)

<223> DNA sequence of the ptsHI region

<400> 38

gaagatgaaa gctttaccaa caagaatatt gtggttattc taccatcatc gggtagcgt 60

tatttaagca ccgcattgtt tgccgatctc ttcactgaga aagaattgca acagtaatgc 120
cagcttgta aaaatgcgta aaaaagcacc ttttaggtg cttttttgtg gcctgcttca 180
aactttgcc ctccttgga ttgattcagc ctgtcggaac tggatattta ccagactaat 240
tattttgatg cgcgaaatta atcgttacag gaaaagccaa agctgaatcg attttatgat 300
ttggttcaat tcttccttta gcggcataat gtttaatgac gtacgaaac tcagcggta 360

acaccgcca gcaatggact gtattgcgt cttcgtgcgt cgcgtctgtt aaaaactggc 420
gctaacaata caggctaaag tcgaaccgcc aggctagact ttagttccac aacactaaac 480

ctataagttg gggaaataca atgttccagc aagaagtac cattaccgct cgcacaatct 540
gctaataccac gagatgcggc ccaatttact gcttaggaga agatcatggg ttgttcgat 600
aaactgaaat ctctggtttc cgacgacaag aaggataccg gaactattga gatcattgct 660
ccgctctctg gcgagatcgt caatatcgaa gacgtgccgg atgtcgtttt tgcggaaaaa 720
atcgttgggt atggtattgc tatcaacca acgggtaaca aaatggtcgc gccagtagac 780
ggcaccattg gtaaaatctt tgaaaccaac cacgcattct ctatcgaatc tgatagcggc 840
gttgaactgt tcgtccactt cggtatcgac accgttgaac tgaaaggcga aggcctcaag 900

cgatttgctg aagaaggica gcgcgtgaaa gtggcgata ctgtcattga atttgatctg 960
ccgctgctgg aagagaaagc caagtctacc ctgactcggg ttgttatctc caacatggac 1020
gaaatcaaag aactgatcaa actgt 1045

<210> 39
<211> 4595
<212> DNA
<213> Escherichia coli
<220><221> gene
<222> (1)..(4595)
<223> DNA sequence of the integrated cassette tdc::glf-glk
<400> 39

ctgatttctt tgcgctgat cccttactgg aactctgcag ttatcgacca ggttgacctc 60
ggttcgtgt cgtaaccgg tcatgacggt atcctgatca ctgtctggct ggggatttcc 120

atcatggttt tctcctttaa cttctcgcca atcgtctctt ccttcgtggt ttctaacgt 180
gaagagtatg agaaagactt cggtcgcgac ttaccgaac gtaaatgttc ccaaatcatt 240
tctcgtgcca gcattctgat ggttgacagt gtgatgttct ttgcctttag ctgcctgttt 300
actctgtctc cggccaacat ggcggaagcc aaagcgaga atattccagt gctttcttat 360
ctggctaacc actttgcgtc catgaccggt accaaaacaa cgttcgcgat tacactggaa 420
tatgcggctt ccatcatcgc actcgtggct atcttcaaat ctttcttcgg tcactatctg 480
gggacgctgg aaggcttgaa tggctctgatt ctgaagttcg gttataaagg tgacaaaacc 540

aaagtgtcgc tgggtaaact gaatactctc agcatgatct tcatcatggg ctccacctgg 600
gttgttcct acgccaaccc gaacatctc gacctgattg aagccatggg cgcaccgatt 660

atcgcatccc tgcgtgacct gttgccgatg tatgcatcc gtaaagcgcc gtctctggcg	720
aaataccgtg gtcgtctgga taacgtgttt gttaccgtga ttggtctgct gaccatcctg	780
aacatcgtat acaaactgtt ttaatccgta actcaggatg agaaaagaga tgaatgaatt	840
tccggttggt ttggttatta actgtgggtc gtcttcgatt aagttttccg tactcgatgc	900
cagcgactgt gaagtattaa tgtcaggatg tgccgacggt attaactcgg aaaatgcatt	960
cttatccgta aatgggggag agccagcacc gctggctcac cacagctacg aagggtcatt	1020
gaaggcaatt gcatttgaac tggaaaaacg gagtttaaat gacagcgtgg ccttaattgg	1080
tgcaaaagtg gctgtgactg taaaaagaaa tcgaaaaaga ccgttttggt tgaacacggt	1140
ctttttgttt ccttttaacc aactgccata actcgaggcc tacctagctt ccaagaaaga	1200
tatcctaaca gcacaagagc ggaaagatgt tttgttctac atccagaaca acctctgcta	1260
aaattcctga aaaaatttgc aaaaagttgt tgactttatc tacaaggtgt ggtataataa	1320
tcttaacaac agcaggacgc tctagaggga gaggaatcgc catgagttct gaaagtagtc	1380
agggtctagt cacgcgacta gccctaatacg ctgctatagg cggcttgctt ttcggttacg	1440
attcagcgtt tatcgtgca atcggtacac cggttgatat ccattttatt gccctctgc	1500
acctgtctgc tacggctcgc gcttccttt ctgggatggt cgttgttgct gttttggctc	1560
gttgtgttac cgtttctttg ctgtctggct ggattggtat tcgttcggt cgtcgcggcg	1620
gattgttgat gattccatt tgtttcgtc ccgccggtt ttggtctgcg ttaaccgaaa	1680
aattatttgg aaccggtggt tcggctttac aaattttttg cttttccgg tttcttccg	1740
gtttaggtat cgggtgctgt tcaaccttga cccaacctt tattgctgaa attgctccgc	1800
cagacaaacg tggtcagatg gtttctggtc agcagatggc catgtgacg ggtgctttaa	1860
ccggttatat ctttaccgtg ttactggctc atttcggttc tatcgattgg gttaatgcca	1920
gtggttggtg ctggtctccg gcttcagaag gctgatcgg tattgccttc ttattgctgc	1980
tgtaaacgc accggatagc ccgcattggt tggatgaa gggacgtcat tccaggcta	2040
gcaaaatcct tgcctgtctg gaaccgcaag ccgacctaa tctgacgatt caaaagatta	2100
aagctggctt tgataaagcc atggacaaaa gcagcgcagg tttgtttgct tttggtatca	2160
ccgttgtttt tgccggtgta tccgttgctg ccttcagca gttagtcggt attaacgccg	2220
tgctgtatta tgcaccgag atgttccaga atttaggttt tggagctgat acggcattat	2280
tgcagacat ctctatcgtt gtttgaact tcattctcac catgattgct tcccgtgttg	2340
ttgaccgtt cggccgtaaa cctctgctta tttgggtgct tctcggtatg gctgcaatga	2400
tggctgtttt aggtctgtgt ttctgggtca aagtcggtgg tgttttgct tttggtctgt	2460
tgcttcttta tattgcagtc tttggtatgt catggggccc tgtctgctgg gttgttctgt	2520

cagaaatgtt cccgagttcc atcaagggcg cagctatgcc tatcgctgtt accggacaat	2580
ggtttagctaa tatcttgggt aacttcctgt ttaaggttgc cgatggttct ccagcattga	2640
atcagacttt caaccacggt ttctcctatc tcgttttcgc agcattaagt atcttaggtg	2700
gcttgattgt tgctcgcttc gtgccggaac ccaaaggtcg gagcctggat gaaatcgagg	2760
agatgtggcg ctcccagaag tagttaaact tgctttggct gaatcctttt gtctttttta	2820
gataagtctt aaccaattat actttttgtt tacaacgatg gtataaagcg ggcggactta	2880
ttttacctgt tgggtagcct tctgatttca gaaaggaatt attatggaaa ttgttgcgat	2940
tgacatcggt ggaacgcgat cgcgtttctc tattgcggaa gtaagcaatg gtcgggttct	3000
ttctcttgga gaagaaacaa cttttaaaac ggcagaacat gctagcttgc agttagcttg	3060
ggaacgtttc ggtgaaaaac tgggtcgtcc tctgccacgt gccgcagcta ttgcatgggc	3120
tggcccggtt catggtgaag ttttaaaact taccaataac ccttgggtat taagaccagc	3180
tactctgaat gaaaagctgg acatcgatac gcatgttctg atcaatgact tcggcgcggt	3240
tgccacgcg gttgcgcata tggattcttc ttatctggat catatttttg gtcctgatga	3300
agcgttctct agcgtgggtg ttatcactat tcttgggtccg ggaacgggct tgggtgttgc	3360
ccatctgttg cggactgaag gccgttattt cgtcatcgaa actgaaggcg gtcatatcga	3420
ctttgctccg cttgacagac ttgaagacaa aattctggca cgtttacgtg aacgtttccg	3480
ccgcgtttct atcgaacgca ttatttctgg cccgggtctt ggtaatatct acgaagcact	3540
ggctgccatt gaaggcgttc cgttcagctt gctggatgat attaaattat ggcagatggc	3600
tttgggaagt aaagacaacc ttgctgaagc cgttttggat cgcttctgct tgagccttgg	3660
cgctatcgct ggigatcttg ctttggcaca ggttcgaacc agtgttgta ttggcggttg	3720
tgctggctct cgtatcgctt cccatttggc agaactcgtt ttccgtcagc gctttgttcc	3780
aaaaggacgc tttgaacgcg tcatgtccaa gattccggtt aagttgatta cttatccgca	3840
gcctggactg ttgggtgcgc agctgcctat gccacaaat attctgaagt tgaataatat	3900
tttttaatat tatgaactga atttaagagg ctgccttccg ataaaatcgg gaggtggcct	3960
tttttatatt ttttactaaa aaatgaagac aaaaaagtct taagtaagaa taatattatt	4020
attaactttt gatataattt gtattagtgg atccgccctc ccgctggaaa ttgaagccat	4080
cgcagtacgt agtgcgtaaa gcctcgtgag cgggacggtc gtaaggtcgt tccgtccac	4140
ttcactgaac ggcaatccga ggggtgtggat atgattagtg cattcgatat tttcaaatt	4200
gggattggtc cctccagttc gcataccgtg gggccaatga atgccggaac aagttttatt	4260

gatcggctgg aaagtagcgg cttattaacc gcgacgagcc atattgtggt cgatctgtac 4320

gggtcgttgt cactgacggg caaaggccat gccacggatg tcgccatcat catgggactg 4380

gcaggaacaa gtccgcagga tggtgtcatt gatgagatcc ctgcatttat agagttagta 4440

acgcgcagcg ggcggtgcc agtggcatct ggtgcgcata ttgttgattt tctgttagca 4500

aagaacatta tcttccatcc cgaaatgttg cctcgccatg agaacggaat gcggatcact 4560

gcctggaagg gacaggaaga gctattaagt aaaac 4595

<210> 40

<211> 1069

<212> DNA

<213> Escherichia coli

<220><221> gene

<222> (1)..(1069)

<

223> DNA sequence of the galP region

<400> 40

acttttgctg tgaacatttc ccgtgggaaa aaaccgacaa agcgcagctg ctgcgcgatg 60

ctgccggtct gaagtaatct ttcttcacct gcgttcaaag gccagcctcg cgctggcctt 120

tttcttttgg ataggcgctc acgcccgcac cggcaaaaaa accgcccgca caataacatc 180

attcttctct atcacgtttc accgcagatt atcatcacia ctgaaaccga ttacaccaac 240

cacaacagac aaagatttgt aatattttca tattattatt cggttttcac agttgttaca 300

tttcttttca gtaaagtctt aattgcagat aacagcggtt aatctatgat gatataactc 360

aattattttc atgcacttaa atcataacta agataaatgt tagtgtaagc gattacactg 420

atgtgatttg cttcacatct ttttacgtcg tactcaccta tcttaattca caataaaaaa 480

taaccatatt ggagggcgatc atgcctgacg ctaaaaaaca ggggcgggtca aacaaggcaa 540

tgacgtttga aataggcgct cagattaat ctccccaagc ttcctcccat cgcggaggaa 600

gccacctctt gcagtcactt tttcttcgct ctatcctctg ccgctatgaa aacatcccgt 660

ctccctatcg ccatccaaca ggccgttatg cgtcgcctgc gggaaaaact cgcccaggcc 720

aacctgaagc tagggcgtaa ctaccggag ccaaaactct cttacacca gcgcggaacc 780

tccgccggaa cggcctggct ggaaagctat gaaattcgcc tcaatcccgt tttgctgttg 840

gaaaacagtg aagcttttat tgaagaagtg gtaccgcacg aactggcaca tttgctggtg 900

tggaaacatt tcggccgcgt agcggccatc ggcaaagagt ggaagtggat gatggaaaac 960

gtgctgggtg ttcccgcccc tcgtacgcat cagttcgaac tgcaatccgt gcgtcgcaac 1020

acccttccct accgctgcaa gtgccaggag catcagctta ccgtacgcc 1069

<210> 41

<211> 6100

<212> DNA

<213> Escherichia coli

<220><221> gene

<222> (1)..(6100)

<

223> MYR352 adhE::P15-catAX , PR-aroY, P26-quiC

<400> 41

ttgattttca taggttaagc aaatcatcac cgactgact atactctcgt attcgagcag 60

atgatttact aaaaaagttt aacattatca ggagagcatt agcttgctat tgacgacagc 120

tatggttcac tgtccaccaa ccaaaactgt gctcagtacc gccaatattt ctcccttgag 180

gggtacaaag aggtgtccct agaagagatc cacgctgtgt aaaaatttta caaaaaggta 240

ttgactttcc ctacagggtg tgtaataatt taattacagg cgggggcaac cccgcctgtt 300

ctgcagagga ggaatatagc catggaagtg aaaatcttca acaccagga tgttcaggat 360

tttctgcgtg ttgcaagcgg tctggaacaa gagggtggtg atccgcgtgt taaacaaatt 420

attcatcgtg tttcgagcga cctgtataaa gcaattgaag atctgaatat caccagcgac 480

gaatatgtgg caggcgttgc atatctgaat cagctgggtg caaatcaaga agcaggtctg 540

ctgagtccgg gtctgggttt tgatcattat ctggatatgc gtatggatgc agaagatgca 600

gcactgggta ttgaaaatgc aacaccgcgt accattgaag gtccgctgta tgttgcgggt 660

gcaccggaag gcgttggtta tgcacgatg gatgatggtg gcatccgaa tggatcatacc 720

ctgattctgc atggcaccat ttttgatgca gatggtaaac cgctgccgaa tgcaaaagtt 780

gaaatttggc atgcaaacac caaaggcttt tatagccatt ttgatccgac cggatgaacag 840

caggccttta atatgcgtcg tagcattatt accgatgaga atggtcagta tcgtgttcgt 900

accattctgc ctgccggtta tggttgtcct ccggaaggte cgaccagca actgctgaac 960

caactgggtc gtcatggtaa tcgtccggca catattcatt atttgttag cgcagatggt 1020

caccgtaaac tgaccacca gattaatgtt gccggtgac cgtataccta tgatgatttt 1080

gcatatgcca cccgtgaagg tctggttggt gatgcagttg aacataccga tccggaagca 1140

attaaagcca atgatgtgga aggtcctttt gccgaaatgg tgtttgatct gaaactgacc 1200

cgctcgtgtg atgggtgtga taatcagggt gtggatcgtc cgcgtctggc agtttaatac 1260

accaaagtgg ttcaaaatta tcaggcgagt gatcatgac actggcctgt ttttatttca 1320

gggaagggtg gagacaatta cgtggataat cagatcatcc aagaaaccgt ggataaaatt	1380
ctgagcgttc tgccgaatca ggcaggtcag ctggcacgtc tgggtcgtct gatgcaattt	1440
gcatgcgac cgaccattac cgttattggc aaatataacc atggtaaaag ccgtctgctg	1500
aatgaactga ttggcacca tatctttagc gttgcagata aacgtgaaac cattcagctg	1560
gccgaacata aacaggatca ggttcgttgg ctggatgcac ctggtctgga tgccgatgtt	1620
gcagcagttg atgategtca tgcatttgaa gcagtttga cccaggcaga tattcgtctg	1680
tttgttcata gcgttcgtga aggtgaactg gatgcaaccg aacaccatct gctgcaacag	1740
ctgattgaag atgccgatca tagccgtcgt cagaccattc tggttctgac ccagattgat	1800
cagattccgg atcagacat cctgacacag attaaaacca gcattgcaca gcaggttccg	1860
aaactggata ttggggcagt tagcgcaacc cgtcatcgtc agggcattga aaacggtaaa	1920
accctgctga tcgaaaaaag cggttattgt gactgcgcc ataccctgga acaggcactg	1980
gcacaggtgc cgagcgcacg tacctatgaa aaaaatcgtc tgctgtcaga tctgcacat	2040
cagctgaaac aactgctgct ggatcagaaa catgttctgc aacaactgca acagacacag	2100
caacagcagc tgcatgattt tgataccggt ctgattaaca ttctggacaa aattcgtgtt	2160
gatctggaac cgattgtgaa tattgatggt caggatcaag cactgaatcc ggatagcttt	2220
gcaacatgt ttaaaaacac cgcagcaaaa cagcagcgtg ccaaagttca gattgcatat	2280
agccgtgcat gcattgaaat caacagccat ctgattcgcc atgggtgtgt tggctgcct	2340
gcggaacagc agaccacat taaaagcatt gataccgtga ttgttgccgt gtttggtatc	2400
agcgttaaat ttctgatca gctgcgtgcc ctgtttata ccgataccga acgtcagcgt	2460
ctgcaacgtg aatttcgttt ctattttgaa aaaagtgccg gtcgcatgat tctggcagca	2520
aaaattgaac agaccatgcg tcagcagggc tgtattcaga atgccatgat ggcactgcaa	2580
caaatggaaa ggcagcata aaaacacgga cgccgcaaac ggctccgaa ttctttggtc	2640
gaccgttaaa tctatcaccg caagggataa atatctaaca ccgtgcgtgt tgactatfff	2700
acctctggcg gtgataatgg ttgcatgtac taatctagat aaggaatata gccatgaccg	2760
caccgattca ggatctgct gatgcaattg cctgctgca acagcatgat aatcagtatc	2820
tggaaaccga tcatccggtt gatccgaatg cagaactggc aggcgtttat cgtcatattg	2880
gtgccgttg caccgttaaa cgtecgacc gtattgttcc ggcaatgatg ttttaataaca	2940
ttaaaggtta tccgcacagc cgtattctgg ttggtatgca tgcaagccgt cagcgtgcag	3000
cactgctgct gggttgtgaa gcaagtcagc tggcactgga agtttggtaaa gcagttaaaa	3060
aaccggttgc accggtggtt gttccggcaa gcagcgacc gtgtcaagag cagatttttc	3120
tggcagatga tccggatttt gatctcgta cctgctgcc tgcacatacc aataccccga	3180

ttgatgcagg tccgtttttt tgtctgggtc tggccctggc aagcgatccg gtggatgcaa	3240
gcctgaccga tgttaccatt catcgtctgt gtgttcaggg tcgtgatgaa ctgagcatgt	3300
tcctggcagc aggtcgccat attgaagttt ttcgtcagaa agcagaagca gcaggtaaac	3360
cgctgccgat taccattaat atgggtcttg acccagcaat ctatatggc gcatgttttg	3420
aagcaccgac caccctgtt ggttataatg aactgggtgt tgccggtgca ctgcgtcagc	3480
gtccggttga actggttcag ggtgttagcg ttcgggaaaa agcaattgca cgtgccgaaa	3540
ttgttatga aggtgaactg ctgcctgggt ttcgtgttcg tgaagtcag cataccaatt	3600
caggtcatgc aatgccggaa ttccgggtt attgtggtgg tgcaaatccg agcctgccgg	3660
ttattaaagt taaagccgtt accatgcgca ataacgcaat tctgcaaacc ctggttggtc	3720
cggtgaaga acataccacc ctggcaggtc tgccgaccga agcaagcatt tggaatgcag	3780
ttgaagcagc aattccgggt tttctgcaaa atgtttatgc ccataccgca ggcggtggtg	3840
aatttctggg tattctgcaa gtgaaaaaac gtcagcctgc cgatgaaggt cgtcagggtc	3900
aggcagccct gctggcgctg gcaacctata gcgaactgaa aaatatcatt ctggtggatg	3960
aggatgtgga catttttgat agtgatgata tttctgtggc aatgaccacc cgtatgcagg	4020
gtgatgttag cattaccacc attccgggtg ttcgcggtca tcagctggac ccgagccaga	4080
caccggaata ttcaccgagc attcgtggta atggtattag ctgcaaaacc atctttgatt	4140
gtaccgttcc gtgggcactg aaaagccatt ttgaacgtgc accgtttgca gatgttgatc	4200
cgctccgtt tgcacctgaa tattttgcac gtctggaaaa aaatcagggc agcgcaaaat	4260
aagctaataa caggcctgct ggtaatcgca ggaattttta tttggatgga tccgcctacc	4320
tagcttccaa gaaagatata ctaacagcac aagagcggaa agatgttttg ttctacatcc	4380
agaacaacct ctgctaaaaa tctgaaaaa ttttgcaaaa agttgttgac tttatctaca	4440
aggtgtggta taataatctt aacaacagca ggacgctccc gggttgagga aaacctaag	4500
aaactgacca gcctgcgtgt tagcctgctg gcactgggtc tggttaccag cggttttgca	4560
gcagcagaaa cctataccgt tgatcgttat caggatgata gcgaaaaagg tagcctgcgt	4620
tgggcaattg aacagagcaa tgcaaatagc gcacaagaaa accagattct gattcaggca	4680
gttggtaaag caccgtatgt tatcaaagtt gataaacgcg tgcctccgat taaaagcagc	4740
gttaaaatca ttggcaccga gtgggataaa accggtgaat ttattgcaat tgatggcagc	4800
aactatatca aaggcgaagg tgaaaaagca gtccgggtg caaatccggg tcagtatggc	4860
accaatgttc gtacatgac cctgcctggt ctggttctgc aagatgttaa tgggtgtacc	4920

ctgaaaggtc tggatgttca tcgtttttgt attggtgttc tggtaatcg cagcagcaat 4980

aacctgattc agcataatcg tatcagcaac aattatggtg gtgccggtgt tatgattacc 5040

ggtgatgatg gtaaaggtaa tccgaccagc accaccacca ataataacaa agttctggat 5100

aacgtgttca tcgataatgg tgatggtctg gaactgacct gtggtgcagc attaatctg 5160

attgcaaata acctgtttac cagcacaaaa gccaatccgg aaccgagcca gggatttgaa 5220

attctgtggg gtaatgataa tgccgtggtg ggtaacaaat tcgaaaacta ttcagatggc 5280

ctgcaaatca attggggtaa acgtaactat atcgctata acgaactgac caataacagc 5340

ctgggtttca atctgacagg tgatggtaac attttcgaca gcaataaagt gcatggtaac 5400

cgatttggtg ttgccattcg tagtgaaaa gatgccaatg cacgtattac cctgacaaaa 5460

aatcagattt gggataacgg caaagatatc aaacgttgtg aagccggtgg tagctgtgtt 5520

ccgaatcagc gtctgggtgc aattgttttt ggtgttcgg cactggaaca tgaaggtttt 5580

gttggtagcc gtggcggtgg tgttgttatt gaaccggcaa aactgcaaaa aacctgcacc 5640

cagccgaacc agcagaattg taatgcaatt cctaatacagg gtattcagc accgaaactg 5700

acagttagca aaaaacagct gaccgttgaa gttaaaggca ccctaataca gcgttataat 5760

gtggaatttt ttggcaatcg taatgccagc agcagcgaag cagaacagta tctgggtagc 5820

attgttgttg ttaccgatca tcagggtctg gcaaaagcaa attgggctcc gaaagttagc 5880

atgccgagcg ttaccgcaaa tgtgacagat catctgggtg cgaccagcga actgagcagc 5940

gcagttaaaa tgcgttaaat gcatgcgcgc cgcgttcgcg cggcgctttt ttttggtact 6000

cagtagcgct gtctggcaac ataaacggcc cttcttgggc aatgccgatc agttaaggat 6060

tagttgaccg atccttaaac tgaggcacta taacggcttc 6100

<210> 42

<211> 594

<212> DNA

<213> Klebsiella pneumoniae

<220><221> gene

<222> (1)..(594)

<223> Nucleotide sequence of Klebsiella pneumoniae kpdB gene.

<400> 42

atgaaactga ttattgggat gacgggggcc accggggcac cgcttggggt ggcattgctg 60

caggcgctgc gcgatatgcc ggaggtggaa acccatctgg tgatgtcgaa atgggccaaa 120

accaccatcg agctggaaac gccctggacg gcgcgcgaag tggccgcgct ggcggacttt 180

tcccacagcc cggcagacca ggccgccacc atctcatccg gttcatttcg taccgacggc 240
atgatcgta ttcctgcag tatgaaaacg ctgcaggca ttcgcgcggg ttatgccgaa 300
gggctggtgg gccgcgcggc ggacgtggtg ctcaaagagg ggcgcaagct ggtgttggtc 360

ccgcgggaaa tgccgctcag cacgatccat ctggagaaca tgctggcgct gtcccgcgatg 420
ggcgtggcga tggccccgc gatgccagct tactacaacc acccgagac ggttgacgat 480
atcaccaatc atatcgtcac ccgggtgctg gatcagtttg gcctcgacta tcacaaagcg 540
cgccgctgga acggcttgcg caggcagaa caatttcac aggagatcga ataa 594

<210> 43

<211> 197

<212> PRT

<213> Klebsiella pneumoniae

<220><221> PEPTIDE

<222> (1)..(197)

<223> Amino acid sequence of KpdB protein of Klebsiella pneumoniae

<400

> 43

Met Lys Leu Ile Ile Gly Met Thr Gly Ala Thr Gly Ala Pro Leu Gly

1 5 10 15

Val Ala Leu Leu Gln Ala Leu Arg Asp Met Pro Glu Val Glu Thr His

20 25 30

Leu Val Met Ser Lys Trp Ala Lys Thr Thr Ile Glu Leu Glu Thr Pro

35 40 45

Trp Thr Ala Arg Glu Val Ala Ala Leu Ala Asp Phe Ser His Ser Pro

50 55 60

Ala Asp Gln Ala Ala Thr Ile Ser Ser Gly Ser Phe Arg Thr Asp Gly

65 70 75 80

Met Ile Val Ile Pro Cys Ser Met Lys Thr Leu Ala Gly Ile Arg Ala

85 90 95

Gly Tyr Ala Glu Gly Leu Val Gly Arg Ala Ala Asp Val Val Leu Lys

100 105 110

Glu Gly Arg Lys Leu Val Leu Val Pro Arg Glu Met Pro Leu Ser Thr

115 120 125

Ile His Leu Glu Asn Met Leu Ala Leu Ser Arg Met Gly Val Ala Met

130 135 140

Val Pro Pro Met Pro Ala Tyr Tyr Asn His Pro Glu Thr Val Asp Asp

145 150 155 160

Ile Thr Asn His Ile Val Thr Arg Val Leu Asp Gln Phe Gly Leu Asp

165 170 175

Tyr His Lys Ala Arg Arg Trp Asn Gly Leu Arg Thr Ala Glu Gln Phe

180 185 190

Ala Gln Glu Ile Glu

195

<210> 44

<211> 570

<212> DNA

<213> Escherichia coli

<220><221> gene

<222> (1)..(570)

<223> Nucleotide sequence of Escherichia coli ubiX gene

<400> 44

atgaaacgac tcattgtagg catcagcggg gccagcggcg cgatttatgg cgtgcgctta 60

ttacaggttc tgcgcgatgt cgcagatata gaaacgcata tggatgatgag ccaggcggcg 120

cgccagacct tatccctcga aacgggtttt tccctgcgcg aagtcagagc attagctgat 180

gtcacgcacg atgcgcgcga tattgccgcc agcatctctt ccggttcttt ccagacgctg 240

gggatggtta ttttacctg ttcaatcaaa accctttccg gcattgtcca tagctacacc 300

gatgggtttac tgacccgtgc ggcagatgtg gtgctgaaag agcgtcgccc gttggtgctc 360

tgcggtcggtg aaacaccatt gcacttaggc catctgcgtt taatgactca ggcagcagaa 420

atcgggtcggg tgattatgcc tcccgttccg gcgttttata atcgcccaca gtcccttgat 480

gatgtgataa atcagacggt taatcgtgtt cttgaccagt ttgcgataac ctttcctgaa 540

gatctctttg cccgtggca gggcgcataa 570

<210> 45

<211> 189

<212> PRT

<213> Escherichia coli

<220><221> PEPTIDE

<222> (1)..(189)

<400> 45

Met Lys Arg Leu Ile Val Gly Ile Ser Gly Ala Ser Gly Ala Ile Tyr

1	5	10	15
Gly Val Arg Leu Leu Gln Val Leu Arg Asp Val Ala Asp Ile Glu Thr			
	20	25	30
His Leu Val Met Ser Gln Ala Ala Arg Gln Thr Leu Ser Leu Glu Thr			
	35	40	45
Gly Phe Ser Leu Arg Glu Val Gln Ala Leu Ala Asp Val Thr His Asp			
	50	55	60
Ala Arg Asp Ile Ala Ala Ser Ile Ser Ser Gly Ser Phe Gln Thr Leu			
65	70	75	80

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

145	150	155	160
Asp Val Ile Asn Gln Thr Val Asn Arg Val Leu Asp Gln Phe Ala Ile			
	165	170	175
Thr Leu Pro Glu Asp Leu Phe Ala Arg Trp Gln Gly Ala			
	180	185	

<210> 46

<211> 594

<212> DNA

<213> Escherichia coli

<220><221> gene

<222> (1)..(594)

<223> Nucleotide sequence of Escherichia coli Wstrain elw gene

<400> 46

atgaaactga tcgtcgggat gacaggggct accggtgcgc ctcttgggtg ggcattactg 60

caagcgctgc gggagatgcc gaatgtcgag actcatctgg tgatgtcgaa gtgggcgaaa 120

accaccattg aactggaaac gccttacagc gctcgcgatg ttgctgcct cgcagacttc 180

agccataacc cggcggatca ggcggcgatc atctcatccg gttcttttcg taccgacggc 240

atgatcgta ttccgtgcag tatgaaaacg ctgcgcgta tccgcgctgg ttacgccgat 300

ggcctgtag ggcgcgcgcg ggacgtcgtg ctcaaagaag gccgcaaact ggtgctggtg 360

ccgcgtgaaa tgccgcttag caccatccat ctcgaaaata tgctcgact ttcacgcatg 420

ggcgtggcga tggtgccgcc gatgcctgcc ttttataacc atcccgaac ggtagatgac 480

attgtccacc atgtggtagc ccgcgtgctg gatcaatttg gcctcgaaca tccccacgcc 540

aggcgtggc aaggattgcc gcaggcccgg aattttctc aggagaatga ataa 594

<210> 47

<211> 197

<212> PRT

<213> Escherichia coli

<220><221> PEPTIDE

<222> (1)..(197)

<223> Amino acid sequence of Elw protein of Escherichia coli W strain

<400> 47

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

1 5 10 15

Val Ala Leu Leu Gln Ala Leu Arg Glu Met Pro Asn Val Glu Thr His

20 25 30

Leu Val Met Ser Lys Trp Ala Lys Thr Thr Ile Glu Leu Glu Thr Pro

35 40 45

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

50 55 60

Ala Asp Gln Ala Ala Ile Ile Ser Ser Gly Ser Phe Arg Thr Asp Gly

65 70 75 80

Met Ile Val Ile Pro Cys Ser Met Lys Thr Leu Ala Gly Ile Arg Ala

85

90

95

Gly Tyr Ala Asp Gly Leu Val Gly Arg Ala Ala Asp Val Val Leu Lys

100

105

110

Glu Gly Arg Lys Leu Val Leu Val Pro Arg Glu Met Pro Leu Ser Thr

115

120

125

Ile His Leu Glu Asn Met Leu Ala Leu Ser Arg Met Gly Val Ala Met

130

135

140

Val Pro Pro Met Pro Ala Phe Tyr Asn His Pro Glu Thr Val Asp Asp

145

150

155

160

Ile Val His His Val Val Ala Arg Val Leu Asp Gln Phe Gly Leu Glu

165

170

175

His Pro His Ala Arg Arg Trp Gln Gly Leu Pro Gln Ala Arg Asn Phe

180

185

190

Ser Gln Glu Asn Glu

195

<210> 48

<211> 579

<212> DNA

<213> Klebsiella oxytoca

<220><221> gene

<222> (1)..(579)

<223> Nucleotide sequence of Klebsiella oxytoca kox gene

<400> 48

atgacggcac gcatcatcat tggatcagc ggcgcacccg ggtttcagta cggcgtaag 60

gcgctggagc tactgcgcc gcatccgtt gaagtcacc tggtcgtctc taaaggcgcg 120

gaaaaaacct gcgagctgga gacggatcac cgcttgacg aggtgatggc gctggccgac 180

gtggtgcatc ccatcgcaa tcttggggcg gctatctcca gcggttcgtt taaaacgtg 240

ggaatgttga tcgcgccgtg ttcaatgcgt tcttaggcg ccatcgcca ctgcctgacc 300

gacaacctgc tcacccgcg tcgggacgtg gtgctgaaag agcgtcgccg cctggtgctg 360

ctggcccgga aaacaccgt gaaccttggc catatccga atatggccgc cgtaaccgaa 420

atgggcggaa ttatctttcc gccgtcccg gcactttacc agcgtccga aacggcggac 480

gacatcggtta cccatagcgt caatcgcgcg ctgatatgtg ttgacctgca ggtcaacaac 540

atcccccgct ggggtgaagg cgagctacgt ttttaattaa 579

<210> 49

<211> 192

<212> PRT

<213> Klebsiella oxytoca

<220><221> PEPTIDE

<222> (1)..(192)

<223> Amino acid sequence of Kox protein of Klebsiella oxytoca.

<400> 49

Met Thr Ala Arg Ile Ile Ile Gly Ile Ser Gly Ala Ser Gly Phe Gln

1 5 10 15

Tyr Gly Val Lys Ala Leu Glu Leu Leu Arg Pro His Pro Val Glu Val

20 25 30

His Leu Val Val Ser Lys Gly Ala Glu Lys Thr Cys Glu Leu Glu Thr

35 40 45

Asp His Arg Leu Asp Glu Val Met Ala Leu Ala Asp Val Val His Pro

50 55 60

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

65 70 75 80

Gly Met Leu Ile Ala Pro Cys Ser Met Arg Ser Leu Gly Ala Ile Ala

85 90 95

His Cys Leu Thr Asp Asn Leu Leu Thr Arg Ala Ala Asp Val Val Leu

100 105 110

Lys Glu Arg Arg Arg Leu Val Leu Leu Ala Arg Glu Thr Pro Leu Asn

115 120 125

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

130 135 140

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

145 150 155 160

Asp Ile Val Thr His Ser Val Asn Arg Ala Leu Asp Leu Phe Asp Leu

165

170

175

Gln Val Asn Asn Ile Pro Arg Trp Gly Glu Gly Glu Leu Arg Phe Asn

180

185

190

<210> 50

<211> 564

<212> DNA

<213> Lactobacillus plantarum

<220><221> gene

<222> (1)..(564)

<223> Nucleotide sequence of Lactobacillus plantarum lpl gene

<400> 50

atgaaacgaa ttgttgtggg aatcacggga gcgtccggtg cgatttacgc ggtcgactta 60

ttagaaaagt tacatcagcg gccagatggt gaagttcatc tggtaatgag tgcgtgggct 120

aaaaaaaaact tggagttaga gactgattac tcgctcgcgc agctgacggc gctcgcggat 180

gctacttatac gggctaataa ccaaggcgca gcgattgcca gcggttcgtt tttgaatgac 240

ggaatgggtc ttgtccagc tagtatgaag acggtagcag ggattgcgta cggttcggt 300

gataatttaa tatcgcgggc tgctgatgtc acgattaaag aacaacgtaa acttgtgatt 360

gttcacgtg aaacaccgtt aagcgtgatt catttagaaa atctaacgaa gttggcaaaa 420

ctcggtgccc aaattattcc accgattccc gcgttttata atcatccgca atccattcag 480

gatctgggtc atcatcaaac catgaaaatt ttagatgcgt ttcattattca taatgaaact 540

gatcgccgtt gggaggggga ttaa 564

<210> 51

<211> 187

<212> PRT

<213> Lactobacillus plantarum

<220><221> PEPTIDE

<222> (1)..(187)

<223> Amino acid sequence of Lpl protein of Lactobacillus plantarum

<400> 51

Met Lys Arg Ile Val Val Gly Ile Thr Gly Ala Ser Gly Thr Ile Tyr

1

5

10

15

Ala Val Asp Leu Leu Glu Lys Leu His Gln Arg Pro Asp Val Glu Val

20 25 30
His Leu Val Met Ser Ala Trp Ala Lys Lys Asn Leu Glu Leu Glu Thr

35 40 45
Asp Tyr Ser Leu Ala Gln Leu Thr Ala Leu Ala Asp Ala Thr Tyr Arg

50 55 60
Ala Asn Asp Gln Gly Ala Ala Ile Ala Ser Gly Ser Phe Leu Asn Asp

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

85 90 95
Tyr Gly Phe Gly Asp Asn Leu Ile Ser Arg Ala Ala Asp Val Thr Ile

100 105 110

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

115 120 125

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

130 135 140

Ile Ile Pro Pro Ile Pro Ala Phe Tyr Asn His Pro Gln Ser Ile Gln

145 150 155 160

Asp Leu Val Asn His Gln Thr Met Lys Ile Leu Asp Ala Phe His Ile

165 170 175

His Asn Glu Thr Asp Arg Arg Trp Glu Gly Asp

180 185

<210> 52

<211> 200

<212> DNA

<213> Escherichia coli

<220><221> promoter

<222> (1)..(200)

<223> Nucleotide sequence of Pgi promoter

<400> 52

agcggggcgg ttgtcaacga tggggtcatg cggatttttc atccactcct ggcggtcagt 60

agttcagcta ataatgctt cactgcgcta agggtttaca ctcaacatta cgctaacggc 120

actaaaacca tcacattttt ctgtgactgg cgctacaatc ttccaaagtc acaatttca 180

tgacagaggag gaatatagcc

200

<210> 53

<211> 3537

<212> DNA

<213> Saccharomyces cerevisiae

<220><221> gene

<222> (1)..(3537)

<223> Nucleotide sequence of Saccharomyces cerevisiae pyc gene

<400> 53

atgtcgcaaa gaaaattcgc cggcttgaga gataacttca atctcttggg tgaaaagaac 60

aaaatatattgg tggctaataag aggagaaatt ccaatcagaa tttttcgtac cgctcatgaa 120

ctgtctatgc agacggtagc tatatattct catgaagatc gtctttcaac gcacaaacaa 180

aaggctgacg aagcatatgt cataggtgaa gtaggccaat atacccccgt cggcgcttat 240

ttggccattg acgaaatcat ttccattgcc caaaaacacc aggtagattt catccatcca 300

ggttatgggt tcttgtctga aaattcggaa ttgcccagaa aagtagtgaa ggccggtatc 360

acttggattg gccctccagc tgaagttatt gactccgtgg gtgataaggt ctcagctaga 420

aacctggcag caaaagctaa tgtgcccacc gtctctggta caccaggtcc tatagaaact 480

gtagaggaag cacttgactt cgtcaatgaa tacggctacc cggatgatcat taaggccgcc 540

tttggtggtg gtggttagagg tatgagagtc gttagagaag gtgacgacgt ggcagatgcc 600

tttcaacgtg ctacctccga agccccgtact gccttcggta atggtacatg ctttgtggaa 660

agattcttgg acaagccaaa gcatattgaa gttcaattgt tggccgataa ccacggaaac 720

gtggttcacg ttttcgaaag agactgttcc gtgcagagaa gacacaaaaa ggttgtcgaa 780

gtggccccag caaagacttt accccgtgaa gtccgtgacg ccattttgac agatgcagtt 840

aaattggcca aagagtgtgg ctacagaaat gcgggtactg ctgaattctt ggttgataac 900

caaaatagac actatttcat tgaaattaat ccaagaatcc aagtgaaca taccatcaca 960

gaagaaatta ccggtataga tatttgtggcg gtcagatcc aaattgcggc aggtgcctct 1020

ctaccccagc tgggcctatt ccaggacaaa attacgactc gtggctttgc cattcagtgc 1080

cgtattacca cggaagaccc tgctaagaac ttccaaccag ataccgtag aatagaagtg 1140

taccgttctg caggtggtaa tgggtttaga ctggatgggt gtaacgccta tgcaggaaca 1200

ataatctcac ctcatcaga ctcaatgctg gtcaaatgct catgtccgg ttccacctac 1260

gaaatcggtc gtagaaaaat gattcgtgca ttaatcgagt tcagaattag aggtgtcaag 1320

accaacattc ccttcctatt gactcttttg accaatccag tatttattga gggtacatac	1380
tggacgactt ttattgacga caccaccaca ctgttccaaa tggtttcac acaaacaga	1440
gccccaaaac ttttacatta cctcgccgac gtggcagtc atggttcac tatcaagggt	1500
caaatgggt tgccaaaatt aaaatcaaat ccaagtgtcc cccatttgca cgatgctcag	1560
ggcaatgtca tcaacgttac aaagtctgca ccaccatccg gatggaggca agtgctacta	1620
gaaaaggggc cagctgaatt tgccagacaa gttagacagt tcaatggtac tttattgatg	1680
gacaccacct ggagagacgc tcatcaatct ctacttgca caagagtcag aaccacgat	1740
ttggctacaa tcgtccaac aaccgcacat gcccttgag gtcgtttcgc cttagaatgt	1800
tgggtgggt ccacattcga tgttcaatg agatttttgc atgaggatcc atgggaacgt	1860
ttgagaaaat taagatctct ggtgccta attcattcc aaatgttatt gcgtggtgcc	1920
aatggtgtgg cttattcttc attgcctgac aatgctattg accatttcgt caagcaagcc	1980
aaggataatg gtgttgatat atttagagtc tttgatgcct taaatgactt ggaacaattg	2040
aaggtcgggt tagatgctgt gaagaaggca ggtggtgttg tagaagccac tgtttgttc	2100
tctggggata tgcttcagcc aggcaagaaa tacaatttgg attactactt ggaaattgct	2160
gaaaaaattg tccaaatggg cactcatatc ctgggtatca aagatatggc aggtaccatg	2220
aagccagcag ctgccaact actgattgga tctttgaggg ctaagtacc tgatctcca	2280
atacatgttc aactcacga ttctgcaggt actgctgttg catcaatgac tgcgtgtgct	2340
ctggcgggcg ccgatgtcgt tgatgttgcc atcaactcaa tgtctggttt aacttcacaa	2400
ccatcaatca atgctctgtt ggcttcatta gaaggtaata ttgacactgg tattaacgtt	2460
gagcatgtcc gtgaactaga tgcatattgg gcagagatga gattgttata ctcttgtttc	2520
gaggctgact tgaagggccc agatccagaa gtttatcaac atgaaatccc aggtggtcaa	2580
ttgacaaact tgttgttca agccaacaa ttgggtcttg gagaacaatg ggccgaaaca	2640
aaaagagctt acagagaagc caattattta ttgggtgata ttgtcaaagt taccccaact	2700
tcgaaggctg ttggtgatct ggcacaattt atggtctcca ataaattaac ttccgatgat	2760
gtgagacgcc tggctaattc tttggatttc cctgactctg ttatggattt cttcgaaggc	2820
ttaatcgccc aaccataagg tgggttcca gaaccattta gatcagacgt ttaaggaac	2880
aagagaagaa agttgacttg tcgtccaggc ctggaactag agccatttga tctcgaaaaa	2940
attagagaag acttgacgaa tagatttggg gatgttgatg agtgcgacgt tgcttcttat	3000
aacatgtacc caagagtta tgaagacttc caaaagatga gagaaacgta tgggtgattta	3060
tctgtatttg caacaagaag ctttttgtct ccactagaga ctgacgaaga aattgaagtt	3120

gtaatcgaac aaggtaaaac gctaattatc aagctacagg ctgtgggtga ttgaacaaa 3180

aagaccggtg aaagagaagt ttactttgat ttgaatggtg aaatgagaaa aattcgtgtt 3240

gtgacagat cacaaaaagt ggaaactgtt actaatcca aagcagacat gcatgatcca 3300

ttacacattg gtgcaccaat ggcagggtgc attgttgaag ttaaagttca taaaggatca 3360

ctaataaaga agggccaacc ttagccgta ttaagcgcca tgaaaatgga aatgattata 3420

tctttccat ccgatggaca agttaagaa gtgtttgtct ctgatggtga aatgtggac 3480

tctttctgatt tattagtct attagaagac caagttctg ttgaaactaa ggcata 3537

<210> 54

<211> 1178

<212> PRT

<213> Escherichia coli

<220><221> PEPTIDE

<222> (1)..(1178)

<223> Amino acid sequence of Pyc protein of Saccharomyces cerevisiae

<400> 54

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

1 5 10 15

Gly Glu Lys Asn Lys Ile Leu Val Ala Asn Arg Gly Glu Ile Pro Ile

20 25 30

Arg Ile Phe Arg Thr Ala His Glu Leu Ser Met Gln Thr Val Ala Ile

35 40 45

Tyr Ser His Glu Asp Arg Leu Ser Thr His Lys Gln Lys Ala Asp Glu

50 55 60

Ala Tyr Val Ile Gly Glu Val Gly Gln Tyr Thr Pro Val Gly Ala Tyr

65 70 75 80

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

85 90 95

Phe Ile His Pro Gly Tyr Gly Phe Leu Ser Glu Asn Ser Glu Phe Ala

100 105 110

Asp Lys Val Val Lys Ala Gly Ile Thr Trp Ile Gly Pro Pro Ala Glu

115 120 125

Val Ile Asp Ser Val Gly Asp Lys Val Ser Ala Arg Asn Leu Ala Ala
130 135 140

Lys Ala Asn Val Pro Thr Val Pro Gly Thr Pro Gly Pro Ile Glu Thr
145 150 155 160

Val Glu Glu Ala Leu Asp Phe Val Asn Glu Tyr Gly Tyr Pro Val Ile
165 170 175

Ile Lys Ala Ala Phe Gly Gly Gly Gly Arg Gly Met Arg Val Val Arg
180 185 190

Glu Gly Asp Asp Val Ala Asp Ala Phe Gln Arg Ala Thr Ser Glu Ala
195 200 205

Arg Thr Ala Phe Gly Asn Gly Thr Cys Phe Val Glu Arg Phe Leu Asp
210 215 220

Lys Pro Lys His Ile Glu Val Gln Leu Leu Ala Asp Asn His Gly Asn
225 230 235 240

Val Val His Leu Phe Glu Arg Asp Cys Ser Val Gln Arg Arg His Gln
245 250 255

Lys Val Val Glu Val Ala Pro Ala Lys Thr Leu Pro Arg Glu Val Arg
260 265 270

Asp Ala Ile Leu Thr Asp Ala Val Lys Leu Ala Lys Glu Cys Gly Tyr
275 280 285

Arg Asn Ala Gly Thr Ala Glu Phe Leu Val Asp Asn Gln Asn Arg His
290 295 300

Tyr Phe Ile Glu Ile Asn Pro Arg Ile Gln Val Glu His Thr Ile Thr
305 310 315 320

Glu Glu Ile Thr Gly Ile Asp Ile Val Ala Ala Gln Ile Gln Ile Ala
325 330 335

Ala Gly Ala Ser Leu Pro Gln Leu Gly Leu Phe Gln Asp Lys Ile Thr
340 345 350

Thr Arg Gly Phe Ala Ile Gln Cys Arg Ile Thr Thr Glu Asp Pro Ala
355 360 365

Lys Asn Phe Gln Pro Asp Thr Gly Arg Ile Glu Val Tyr Arg Ser Ala
370 375 380

Gly Gly Asn Gly Val Arg Leu Asp Gly Gly Asn Ala Tyr Ala Gly Thr
 385 390 395 400
 Ile Ile Ser Pro His Tyr Asp Ser Met Leu Val Lys Cys Ser Cys Ser
 405 410 415
 Gly Ser Thr Tyr Glu Ile Val Arg Arg Lys Met Ile Arg Ala Leu Ile
 420 425 430
 Glu Phe Arg Ile Arg Gly Val Lys Thr Asn Ile Pro Phe Leu Leu Thr
 435 440 445
 Leu Leu Thr Asn Pro Val Phe Ile Glu Gly Thr Tyr Trp Thr Thr Phe
 450 455 460
 Ile Asp Asp Thr Pro Gln Leu Phe Gln Met Val Ser Ser Gln Asn Arg
 465 470 475 480
 Ala Gln Lys Leu Leu His Tyr Leu Ala Asp Val Ala Val Asn Gly Ser
 485 490 495
 Ser Ile Lys Gly Gln Ile Gly Leu Pro Lys Leu Lys Ser Asn Pro Ser
 500 505 510
 Val Pro His Leu His Asp Ala Gln Gly Asn Val Ile Asn Val Thr Lys
 515 520 525
 Ser Ala Pro Pro Ser Gly Trp Arg Gln Val Leu Leu Glu Lys Gly Pro
 530 535 540
 Ala Glu Phe Ala Arg Gln Val Arg Gln Phe Asn Gly Thr Leu Leu Met
 545 550 555 560
 Asp Thr Thr Trp Arg Asp Ala His Gln Ser Leu Leu Ala Thr Arg Val
 565 570 575
 Arg Thr His Asp Leu Ala Thr Ile Ala Pro Thr Thr Ala His Ala Leu
 580 585 590
 Ala Gly Arg Phe Ala Leu Glu Cys Trp Gly Gly Ala Thr Phe Asp Val
 595 600 605
 Ala Met Arg Phe Leu His Glu Asp Pro Trp Glu Arg Leu Arg Lys Leu
 610 615 620
 Arg Ser Leu Val Pro Asn Ile Pro Phe Gln Met Leu Leu Arg Gly Ala

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

 690 695 700
 Leu Gln Pro Gly Lys Lys Tyr Asn Leu Asp Tyr Tyr Leu Glu Ile Ala
 705 710 715 720
 Glu Lys Ile Val Gln Met Gly Thr His Ile Leu Gly Ile Lys Asp Met
 725 730 735
 Ala Gly Thr Met Lys Pro Ala Ala Ala Lys Leu Leu Ile Gly Ser Leu
 740 745 750
 Arg Ala Lys Tyr Pro Asp Leu Pro Ile His Val His Thr His Asp Ser
 755 760 765

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

 835 840 845
 Pro Glu Val Tyr Gln His Glu Ile Pro Gly Gly Gln Leu Thr Asn Leu
 850 855 860
 Leu Phe Gln Ala Gln Gln Leu Gly Leu Gly Glu Gln Trp Ala Glu Thr
 865 870 875 880
 Lys Arg Ala Tyr Arg Glu Ala Asn Tyr Leu Leu Gly Asp Ile Val Lys

885 890 895
 Val Thr Pro Thr Ser Lys Val Val Gly Asp Leu Ala Gln Phe Met Val
 900 905 910

 Ser Asn Lys Leu Thr Ser Asp Asp Val Arg Arg Leu Ala Asn Ser Leu
 915 920 925
 Asp Phe Pro Asp Ser Val Met Asp Phe Phe Glu Gly Leu Ile Gly Gln
 930 935 940
 Pro Tyr Gly Gly Phe Pro Glu Pro Phe Arg Ser Asp Val Leu Arg Asn
 945 950 955 960
 Lys Arg Arg Lys Leu Thr Cys Arg Pro Gly Leu Glu Leu Glu Pro Phe
 965 970 975
 Asp Leu Glu Lys Ile Arg Glu Asp Leu Gln Asn Arg Phe Gly Asp Val

 980 985 990
 Asp Glu Cys Asp Val Ala Ser Tyr Asn Met Tyr Pro Arg Val Tyr Glu
 995 1000 1005
 Asp Phe Gln Lys Met Arg Glu Thr Tyr Gly Asp Leu Ser Val Leu Pro
 1010 1015 1020
 Thr Arg Ser Phe Leu Ser Pro Leu Glu Thr Asp Glu Glu Ile Glu Val
 1025 1030 1035 1040
 Val Ile Glu Gln Gly Lys Thr Leu Ile Ile Lys Leu Gln Ala Val Gly
 1045 1050 1055

 Asp Leu Asn Lys Lys Thr Gly Glu Arg Glu Val Tyr Phe Asp Leu Asn
 1060 1065 1070
 Gly Glu Met Arg Lys Ile Arg Val Ala Asp Arg Ser Gln Lys Val Glu
 1075 1080 1085
 Thr Val Thr Lys Ser Lys Ala Asp Met His Asp Pro Leu His Ile Gly
 1090 1095 1100
 Ala Pro Met Ala Gly Val Ile Val Glu Val Lys Val His Lys Gly Ser
 1105 1110 1115 1120
 Leu Ile Lys Lys Gly Gln Pro Val Ala Val Leu Ser Ala Met Lys Met

 1125 1130 1135

Glu Met Ile Ile Ser Ser Pro Ser Asp Gly Gln Val Lys Glu Val Phe
 1140 1145 1150
 Val Ser Asp Gly Glu Asn Val Asp Ser Ser Asp Leu Leu Val Leu Leu
 1155 1160 1165
 Glu Asp Gln Val Pro Val Glu Thr Lys Ala

1170 1175
 <210> 55
 <211> 6273
 <212> DNA
 <213> Escherichia coli
 <220><221> misc_feature
 <222> (1)..(6273)

<223> DNA sequence of the plasmid pCAT350

<400> 55
 tgaatgacct ttaatagatt atattactaa ttaattgggg accctagagg tccccctttt 60
 tattttaaaa attttttcac aaaacggttt acaagcatat gttggccgat tcattaatgc 120
 agctggcacg acaggtttcc cgactggaaa gcgggcagtg agcgcaacgc aattaatgtg 180
 agtttagctca ctcataggc accccaggct ttacacttta tgcttccggc tcgtatgttg 240
 tgtggaattg tgagcggata acaatttcac acaggaaaca gctatgacca tgattacgcc 300
 aagcttgcat gcctgcaggt cgactctaga ggatccccgg gtaccgagct cgaattcact 360
 ggccgtcggt ttacaacgtc gtgactggga aaacctggc gttacccaac ttaatgcct 420
 tgcagacat ccccccttcg ccagctggcg taatagcgaa gagggccgca ccgatcgccc 480
 ttcccaacag ttgcgcagcc tgaatggcga atggcgcttg atgcggtatt ttctccttac 540
 gcatctgtgc ggtatttcac accgcatatg gtgcactctc agtacaatct gctctgatgc 600
 cgcatagtta agccagcccc gacacccgcc aacaccgct gacgaattcg ttgacagtaa 660
 gacgggtaag cctgttgatg ataccgctgc ctactgggt gcattagcca gtctgaatga 720
 cctgtcacgg gataatccga agtggtcaga ctggaaaatc agagggcagg aactgctgaa 780
 cagcaaaaag tcagatagca ccacatagca gaccgcctat aaaacgcctt gagaagcccc 840
 tgacgggctt ttcttgtatt atgggtagtt tccttgcatg aatccataaa aggcgcctgt 900
 agtgccattt accccattc actgccagag ccgtgagcgc agcgaactga atgtcacgaa 960
 aaagacagcg actcaggtgc ctgatggctg gagacaaaag gaatattcag cgatttggcc 1020
 gagcttgcca ggggtgtact taagccttta gggttttaag gtctgttttg tagaggagca 1080

aacagcgttt gcgacatcct tttgtaatac tgcggaactg actaaagtag tgagttatac	1140
acagggctgg gatctattct ttttatcttt ttttattctt tctttattct ataaattata	1200
accacttgaa tataaacaaa aaaaacacac aaaggtctag cggaatttac agagggtcta	1260
gcagaattta caagttttcc agcaaaggtc tagcagaatt tacagatacc cacaactcaa	1320
aggaaaagga ctagtaatta tcattgacta gcccatctca attggtatag tgattaaaat	1380
cacctagacc aattgagatg tatgtctgaa ttagttgttt tcaaagcaaa tgaactagcg	1440
attagtcgct atgacttaac ggagcatgaa accaagctaa ttttatgctg tgtggcacta	1500
ctcaacccca cgattgaaaa ccctacaagg aaagaacgga cggatcgtt cacttataac	1560
caatacgctc agatgatgaa catcagtagg gaaaatgctt atggtgtatt agctaaagca	1620
accagagagc tgatgacgag aactgtggaa atcaggaatc ctttggttaa aggctttgag	1680
attttccagt ggacaaacta tgccaagttc tcaagcgaaa aattagaatt agtttttagt	1740
gaagagatat tgccttatct tttccagtta aaaaaattca taaaataaa tctggaacat	1800
gttaagtctt ttgaaaacaa atactctatg aggatttatg agtggttatt aaaagaacta	1860
acacaaaaga aaactcacaa ggcaaatata gagattagcc ttgatgaatt taagttcatg	1920
ttaatgcttg aaaataacta ccatgagttt aaaaggctta accaatgggt ttgaaacca	1980
ataagtaaag atttaaacac ttacagcaat atgaaattgg tggttgataa gcgaggccgc	2040
ccgactgata cgttgatttt ccaagttgaa ctagatagac aaatggatct cgtaaccgaa	2100
cttgagaaca accagataaa aatgaatggt gacaaaatac caacaacat tacatcagat	2160
tcctacctac gtaacggact aagaaaaaca ctacacgatg ctttaactgc aaaaattcag	2220
ctcaccagtt ttgaggcaaa atttttgagt gacatgcaaa gtaagcatga tctcaatggt	2280
tcgtttctcat ggctcacgca aaaacaacga accacactag agaacatact ggctaaatac	2340
ggaaggatct gaggttctta tggcaaacac ggacgccgca aacggcgtcc gaatttcttg	2400
gtcgaccgtt aaatctatca ccgcaaggga taaatatcta acaccgtgcg tgttgactat	2460
tttacctctg gcggtgataa tggttgcatg tactaatcta gataaggaat atagccatgg	2520
aagtgaaaat cttaacacc caggatgttc aggattttct gcgtgttgca agcggctctgg	2580
aacaagaggg tggtaatccg cgtgttaaac aaattattca tcgtgttctg agcgacctgt	2640
ataaagcaat tgaagatctg aatatcacca gcgacgaata ttgggcaggc gttgcatatc	2700
tgaatcagct ggggtgcaat caagaagcag gtctgctgag tccgggtctg ggttttgatc	2760
attatctgga tatgcgtatg gatgcagaag atgcagcact gggtattgaa aatgcaacac	2820

cgcgtagcat tgaaggtccg ctgtatgttg cgggtgcacc ggaaagcgtt ggttatgcac	2880
gcatggatga tggtagcgat ccgaatggtc ataccctgat tctgcatggc accatttttg	2940
atgcagatgg taaaccgctg ccgaatgcaa aagttgaaat ttggcatgca aacaccaaag	3000
gcttttatag ccatTTTgat ccgaccggtg aacagcaggc ctttaatatg cgtcgtagca	3060
ttattaccga tgagaatggt cagtatcgtg ttcgtaccat tctgcctgcc ggttatggtt	3120
gtcctccgga aggtccgacc cagcaactgc tgaaccaact gggtcgtcat ggtaatcgtc	3180
cggcacatat tcattatTTT gttagcgcag atggtcaccg taaactgacc acccagatta	3240
atgttgccgg tgatccgtat acctatgatg atTTTgcata tgccaccctg gaaggtctgg	3300
ttgttgatgc agttgaacat accgatccgg aagcaattaa agccaatgat gtggaaggtc	3360
cttttgccga aatgggtgtt gatctgaaac tgaccctctt ggttgatggt gttgataatc	3420
aggtttgtgga tcgtccgctt ctggcagttt aatacaccaa aatggttcaa aattatcagg	3480
cagtgatca tgatcactgg cctgttttta tttcaggga gggaggagac aattacgtgg	3540
ataatcagat catccaagaa accgtggata aaattctgag cgttctgccg aatcaggcag	3600
gtcagctggc acgtctggtg cgtctgatgc aatttgcatg cgatccgacc attaccgtta	3660
ttggcaaata taaccatggt aaaagccgtc tgctgaatga actgattggc accgatatct	3720
ttagcgttgc agataaacgt gaaaccattc agctggccga acataaacag gatcaggttc	3780
gttggctgga tgcacctggt ctggatgccg atgttgagc agttgatgat cgtcatgcat	3840
ttgaagcagt ttggaccgag gcagatattc gtctgtttgt tcatagcgtt cgtgaagggtg	3900
aactggatgc aaccgaacac catctgctgc aacagctgat tgaagatgcc gatcatagcc	3960
gtcgtcagac cattctggtt ctgaccaga ttgatcagat tccggatcag accatcctga	4020
cacagattaa aaccagcatt gcacagcagg ttccgaaact ggatatttgg gcagtttagc	4080
caaccgtca tcgtcagggc attgaaaacg gtaaaacct gctgatcgaa aaaagcggta	4140
ttggtgcact gcgccatacc ctggaacagg cactggcaca ggtgccgagc gcacgtacct	4200
atgaaaaaaaa tcgtctgctg tcagatctgc accatcagct gaaacaactg ctgctggatc	4260
agaaacatgt tctgcaacaa ctgcaacaga cacagcaaca gcagctgcat gatTTTgata	4320
ccggtctgat taacattctg gacaaaattc gtgttgatct ggaaccgatt gtgaatattg	4380
atggtcagga tcaagcactg aatccggata gctttgcaac catgtttaaa aacaccgcag	4440
caaaacagca gcgtgccaaa gttcagattg catatagccg tgcatgcatt gaaatcaaca	4500
gccatctgat tcgcatggt gttgttggtc tgctgcgga acagcagacc accattaaaa	4560

gcattgatac cgigattgtt gccgtgtttg gtatcagcgt taaatttcgt gatcagctgc 4620
 gtgccctgtt ttataccgat accgaacgtc agcgtctgca acgtgaattt cgtttctatt 4680
 ttgaaaaaag tgccggtcgc atgattctgg cagcaaaaat tgaacagacc atgcgtcagc 4740
 agggctgtat tcagaatgcc atgatggcac tgcaacaaat ggaaagcgca gcataagtct 4800
 gacaggtgcc ggatttcata tccggcactt actttcctta actcttcgcc ttaacgcaaa 4860
 atctcacact gatgatcctg aatttcctcg gctgaagcac ggttaagcgt cagtagattt 4920
 cgttgtgtcg ccagcaatac aaatgatgga cataagcctg ttcggttcgt aagctgtaat 4980

gcaagtagcg tatgcgtcca cgcaactggt ccagaacctt gaccgaacgc agcgggtgta 5040
 acggcgcagt ggcggttttc atggcttgtt atgactgttt ttttgggta cagtctatgc 5100
 ctccggcatc caagcagcaa gcgcgttacg ccgtgggtcg atgtttgatg ttatggagca 5160
 gcaacgatgt tacgcagcag ggcatcgcct ctaaaacaaa gttaaacatc atgagggag 5220
 cggatgatcg cgaagtatcg actcaactat cagaggtagt tggcgtcatc gagcgccatc 5280
 tcgaaccgac gttgctggcc gtacatttgi acggctccgc agtggatggc ggcctgaagc 5340
 cacacagtga tattgatttg ctggttacgg tgaccgtaag gcttgatgaa acaacgcggc 5400

gagctttgat caacgacctt ttggaaactt cggcttcccc tggagagagc gagattctcc 5460
 gcgctgtaga agtcaccatt gttgtgcacg acgacatcat tccgtggcgt tatccagcta 5520
 agcgcgaact gcaatttga gaattggcagc gcaatgacat tcttgcaggt atcttcgagc 5580
 cagccacgat cgacattgat ctggctatct tcttgacaaa agcaagagaa catagcgttg 5640
 ccttggtagg tccagcggcg gaggaactct ttgatccggt tcctgaacag gatctatttg 5700
 aggcgctaaa tgaaacctta acgctatgga actcgccgcc cgactgggct ggcatgagc 5760
 gaaatgtagt gcttacgttg tcccgcatit ggtacagcgc agtaaccggc aaaatcgcg 5820

cgaaggatgt cgtgccgac tgggcaatgg agcgcctgcc ggcccagtat cagcccgtca 5880
 tacttgaagc tagacaggct tatcttggac aagaagaaga tcgcttggcc tcgcgcgag 5940
 atcagttgga agaatttgtc cactacgtga aaggcgagat caccaaggta gtcggcaaat 6000
 aatgtctaac aattcgttca agccgacgcc gcttcgcggc gcggcttaac tcaagcgtta 6060
 gatgcactaa gcacataatt gctcacagcc aaactatcag gtcaagtctg cttttattat 6120
 ttttaagcgt gcataataag ccctacacaa attgggagat atatcatgaa aggctggctt 6180
 tttcttgta tcgcaatagt tggcgaagta atcgcaacat ccgcatataa atctagcgag 6240

ggctttacta agctgatccg gtggatgacc ttt 6273

<210> 56

<211> 3943

<212> DNA

<213> Escherichia coli

<220><221> misc_feature

<222> (1)..(3943)

<223> DNA sequence of the plasmid pCP165

<400> 56

tgaatgacct ttaatagatt atattactaa ttaattgggg accctagagg tccccTTTT	60
tattttaaaa attttttcac aaaacggttt acaagcatat gttggccgat tcattaatgc	120
agctggcacg acaggtttcc cgactggaaa gcgggcagtg agcgcaacgc aattaatgtg	180
agttagctca ctcattaggc accccaggct ttacacttta tgcttccggc tcgtatgttg	240
tgtggaattg tgagcggata acaatttcac acaggaaaca gctatgacca tgattacgcc	300
aagcttgcat gcctgcaggt cgactctaga ggatccccgg gtaccgagct cgaattcact	360
ggccgtcgtt ttacaacgtc gtgactggga aaacctggc gttacccaac ttaatcgctt	420
tgcagacat cccccTTTcg ccagctggcg taatagcgaa gagggccgca ccgatcgccc	480
ttccaacag ttgcgcagcc tgaatggcga atggcgctg atgcggtatt ttctccttac	540
gcatctgtgc ggtatttcac accgcatatg gtgcactctc agtacaatct gctctgatgc	600
cgcatagtta agccagcccc gacacccgcc aacacccgct gacgaattcg ttgacagtaa	660
gacgggtaag cctgttgatg ataccgctgc cttactgggt gcattagcca gtctgaatga	720
cctgtcacgg gataatccga agtggtcaga ctggaaaatc agagggcagg aactgctgaa	780
cagcaaaaag tcagatagca ccacatagca gacccgccat aaaacgccct gagaagcccc	840
tgacgggctt ttcttgtatt atgggtagt ttcttgcag aatccataaa aggcgcctgt	900
agtgccattt accccattc actgccagag ccgtgagcgc agcgaactga atgtcacgaa	960
aaagacagcg actcaggtgc ctgatggctg gagacaaaag gaatattcag cgatttgccc	1020
gagcttgcca ggggtgtact taagccttta gggttttaag gtctgttttg tagaggagca	1080
aacacggtt gcgacatcct tttgtaatac tgcggaactg actaaagtag tgagttatac	1140
acagggtcgg gatctattct ttttatcttt ttttattctt tctttattct ataaattata	1200
accacttgaa tataaacaaa aaaaacacac aaaggtctag cggaatttac agagggtcta	1260
gcagaattta caagttttcc agcaaaggtc tagcagaatt tacagatacc cacaactcaa	1320
aggaaaagga ctagtaatta tcattgacta gcccattctc atttggtatag tgattaaaat	1380
cacctagacc aattgagatg tatgtctgaa ttagttgttt tcaaagcaaa tgaactagcg	1440
attagtcgct atgacttaac ggagcatgaa accaagctaa ttttatgctg tgtggcacta	1500

ctcaacccca cgattgaaaa ccctacaagg aaagaacgga cggatcgtt cacttataac	1560
caatacgtct agatgatgaa catcagtagg gaaaatgctt atgggtgtatt agctaaagca	1620
accagagagc tgatgacgag aactgtggaa atcaggaatc ctttggttaa aggctttgag	1680
attttccagt ggacaaacta tgccaagtgc tcaagcgaag aattagaatt agtttttagt	1740
gaagagatat tgccttatct tttccagtta aaaaaattca taaaataata tctggaacat	1800
gttaagtctt ttgaaaacaa atactctatg aggatttatg agtggttatt aaaagaacta	1860
acacaaaaga aaactcaca ggcaaatata gagattagcc ttgatgaatt taagtcatg	1920
ttaatgcttg aaaataacta ccatgagttt aaaaggctta accaatgggt ttgaaacca	1980
ataagtaaag atttaaacac ttacagcaat atgaaattgg tggttgataa gcgaggccgc	2040
ccgactgata cggtgatctt ccaagttgaa ctagatagac aaatggatct cgtaaccgaa	2100
cttgagaaca accagataaa aatgaatggg gacaaaatac caacaacat tacatcagat	2160
tcctacctac gtaacggact aagaaaaaca ctacacgatg ctttaactgc aaaaattcag	2220
ctcaccagtt ttgaggcaaa atttttgagt gacatgcaaa gtaagcatga tctcaatgg	2280
tcgttctcat ggctcacgca aaaacaacga accacactag agaacatact ggctaaatac	2340
ggaaggatct gaggttctta tggcaaacac ggacgccgca aacggcgctc gaatttcttg	2400
gtcgaccgtt aaatctatca ccgcaaggga taaatatcta acaccgtgcg tgttgactat	2460
tttacctctg gcggtgataa tgggtgcatg tactaatcta gataaggaat atagccttag	2520
atttgactga aatcgtagc taaaaagcgt acaataaagg ctccacgaaa gtggggcctt	2580
ttttagcgcg agagcctttt ttgtcagcta tctatatgga cataagcctg ttcggttcgt	2640
aagctgtaat gcaagtagcg tatgcgctca cgcaactggg ccagaacctt gaccgaacgc	2700
agcgggtgga acggcgcatg ggcggttttc atggcttggt atgactgttt ttttggggta	2760
cagtctatgc ctcgggcatc caagcagcaa gcgcgttacg ccgtgggtcg atgtttgatg	2820
ttatggagca gcaacgatgt tacgcagcag ggcatcgcc ctaaaacaaa gttaaacatc	2880
atgaggggaag cggtagatcg cgaagtatcg actcaactat cagaggtagt tggcgtcatc	2940
gagcgccatc tcgaaccgac gttgctggcc gtacatttgt acggctccgc agtggtggc	3000
ggcctgaagc cacacagtga tattgatttg ctggttacgg tgaccgtaag gcttgatgaa	3060
acaacgcggc gagctttgat caacgacctt ttggaaactt cggcttcccc tggagagagc	3120
gagattctcc gcgtctaga agtcaccatt gttgtgcacg acgacatcat tccgtggcgt	3180
tatccagcta agcgcgaact gcaatttggg gaatggcagc gcaatgacat tcttgcaggt	3240
atcttcgagc cagccacgat cgacattgat ctggctatct tgctgacaaa agcaagagaa	3300
catagcgttg ccttggtagg tcacgcggcg gaggaactct ttgatccggt tcctgaacag	3360

gatctatattg aggcgcctaaa tgaacacctta acgctatgga actcgccgcc cgactgggct 3420
ggcgtatgagc gaaatgtagt gcttacgttg tcccgcattt ggtacagcgc agtaaccggc 3480
aaaatcgcg cgaaggatgt cgctgccgac tgggcaatgg agcgccctgcc ggcccagtat 3540
cagcccgtca tacttgaagc tagacaggct tatcttggac aagaagaaga tcgcttggcc 3600

tcgcgcgcag atcagttgga agaatttgtc cactacgtga aaggcgagat caccaaggta 3660
gtcggcaaat aatgtctaac aattcggttca agccgacgcc gcttcgcggc gcggcttaac 3720
tcaagcgta gatgcactaa gcacataatt gctcacagcc aaactatcag gtcaagtctg 3780
cttttattat ttttaagcgt gcataataag ccctacacaa attgggagat atatcatgaa 3840
aggctggctt tttcttgta tcgcaatagt tggcgaagta atcgcaacat ccgcattaaa 3900
atctagcgag ggctttacta agctgatccg gtggatgacc ttt 3943

<210> 57

<211> 10863

<212> DNA

<213> Escherichia coli

<220><221> misc_feature

<222> (1)..(10863)

<223> DNA sequence of the plasmid pCP140

<400> 57

tggacataag cctgttcggt tcgtaagctg taatgcaagt agcgtatgcg ctcacgcaac 60
tgggtccagaa ccttgaccga acgcagcggg ggtaacggcg cagtggcggt tttcatggct 120
tgttatgact gtttttttgg ggtacagtct atgcctcggg catccaagca gcaagcgcgt 180
tacgccgtgg gtcgatgttt gatgttatgg agcagcaacg atgttacgca gcagggcagt 240
cgccctaaaa caaagttaaa catcatgagg gaagcgggtga tcgccgaagt atcgactcaa 300

ctatcagagg tagttggcgt catcgagcgc catctcgaac cgacgttgct ggccgtacat 360
ttgtacggct ccgcagtgga tggcggcctg aagccacaca gtgatattga tttgctggtt 420
acggtgaccg taaggcttga tgaacaacg cggcgagcct tgatcaacga ctttttgaa 480
acttcggctt cccctggaga gagcgagatt ctccgcgtg tagaagtcac cattgttgtg 540
cacgacgaca tcattccgtg gcgttatcca gctaagcgcg aactgcaatt tggagaatgg 600
cagcgcaatg acattcttgc aggtatcttc gagccagcca cgatcgacat tgatctggct 660
atcttctgta caaaagcaag agaacatagc gttgccttgg taggtccagc ggcggaggaa 720

ctctttgatc cggttcctga acaggatcta tttgaggcgc taaatgaaac cttaacgcta 780

tggaactcgc cgcccactg ggctggcgat gagcgaaatg tagtgcttac gttgtccgc	840
atttggtaca gcgcagtaac cggcaaaatc gcgccgaagg atgtcgctgc cgactgggca	900
atggagcgcc tgccggccca gtatcagccc gtcatacttg aagctagaca ggcttatctt	960
ggacaagaag aagatcgctt ggccctcgcg gcagatcagt tggaagaatt tgtccactac	1020
gtgaaaggcg agatcaccaa ggtagtcggc aaataatgtc taacaattcg ttcaagccga	1080
cgccgcttcg cgccgcggt taactcaagc gttagatgca ctaagcacat aattgctcac	1140
agccaaacta tcagggtcaag tctgctttta ttatttttaa gcgtgcataa taagccctac	1200
acaaattggg agatataatc tgaaaggctg gctttttctt gttatcgcaa tagttggcga	1260
agtaatcgca acatccgcat taaaatctag cgagggtctt actaagctga tccggtggat	1320
gaccttttga atgaccttta atagattata ttactaatta attggggacc ctgaggtcc	1380
ccttttttat tttaaaaatt tttcacaaa acggtttaca agcatagctt ggccgattca	1440
ttaatgcagc tggcacgaca ggtttcccga ctggaaagcg ggacgtgagc gcaacgcaat	1500
taatgtgagt tagctcactc attaggcacc ccaggcttta cactttatgc ttccggctcg	1560
tatgttgtgt ggaattgtga gcggataaca atttcacaca ggaaacagct atgaccatga	1620
ttacgccatc gaccgttaaa tctatcaccg caaggataa atatctaaca ccgtgcgtgt	1680
tgactatttt acctctggcg gtgataatgg ttgcatgtac taatctagat aaggaatata	1740
gccatggaag tgaaaatctt caacacccag gatgttcagg atttctcgcg tgttgcaagc	1800
ggtctggaac aagagggtgg taatccgctg gttaacaaaa ttattcatcg tgttctgagc	1860
gacctgtata aagcaattga agatctgaat atcaccagcg acgaatatg ggcaggcgtt	1920
gcataatctga atcagctggg tgcaaatcaa gaagcaggtc tgctgagtc gggtctgggt	1980
tttgatcatt atctggatat gcgtatggat gcagaagatg cagcactggg tattgaaaat	2040
gcaacaccgc gtaccattga aggtccgctg tatgttcggg gtgcaccgga aagcgttggt	2100
tatgcacgca tggatgatgg tagcgatccg aatggtcata cctgattct gcatggcacc	2160
atttttgatg cagatggtaa accgtgccg aatgcaaaag ttgaaatttg gcatgcaaac	2220
accaaaggct tttatagcca ttttgatccg accggtgaac agcaggcctt taatatgcgt	2280
cgtagcatta ttaccgatga gaatggtcag tatcgtgttc gtaccattct gcctgccggt	2340
tatggtgtgc ctccggaagg tccgaccag caactgtga accaactggg tcgtcatggt	2400
aatcgtccgg cacatatcca ttattttgtt agcgcagatg gtcaccgtaa actgaccacc	2460
cagattaatg ttgccgtga tccgtatacc tatgatgatt ttgcatatgc caccgtgaa	2520
ggtctggttg ttgatgcagt tgaacatacc gatccggaag caattaaagc caatgatgtg	2580
gaaggctcctt ttgccgaaat ggtgtttgat ctgaaactga cccgtctggt tgatggtgtt	2640

gataatcagg ttgtggatcg tccgcgtctg gcagtttaat acaccaaaat ggttcaaaat	2700
tatcaggcga gtgatcatga tcactggcct gtttttattt caggaaggg tggagacaat	2760
tacgtggata atcagatcat ccaagaaacc gtggataaaa ttctgagcgt tctgccgaat	2820
caggcaggtc agctggcacg tctgggtcgt ctgatgcaat ttgcatgcga tccgaccatt	2880
accgttattg gcaaatataa ccatggtaaa agccgtctgc tgaatgaact gattggcacc	2940
gatatcttta gcgttcgaga taaacgtgaa accattcagc tggccgaaca taaacaggat	3000
caggttcgtt ggctggatgc acctggctcg gatgccgatg ttgcagcagt tgatgatcgt	3060
catgcatttg aagcagtttg gaccaggca gatattcgtc tgtttgttca tagcgttcgt	3120
gaaggtgaac tggatgcaac cgaacacat ctgctgcaac agctgattga agatgccgat	3180
catagccgtc gtcagacat tctggttctg acccagattg atcagattcc ggatcagacc	3240
atcctgacac agattaaaac cagcattgca cagcaggttc cgaaactgga tatitgggca	3300
gttagcgcaa cccgtcatcg tcagggcatt gaaaacggtg aaacctgct gatcgaaaaa	3360
agcggatttg gtgcactcg ccataccctg gaacaggcac tggcacaggt gccgagcgca	3420
cgtacctatg aaaaaaatcg tctgctgtca gatctgcacc atcagctgaa acaactgctg	3480
ctggatcaga aacatgttct gcaacaactg caacagacac agcaacagca gctgcatgat	3540
tttgataccg gtcgattaa cattctggac aaaattcgtg ttgatctgga accgattgtg	3600
aatattgatg gtcaggatca agcactgaat ccggatagct ttgcaacat gtttaaaaac	3660
accgcagcaa aacagcagcg tgccaaagt t cagattgcat atagccgtgc atgcattgaa	3720
atcaacagcc atctgattcg ccatgggtgtt gttggctcgc ctgcggaaca gcagaccacc	3780
attaaaagca ttgataccgt gattgttgcc gtgtttggta tcagcgttaa atttcgtgat	3840
cagctgcgtg cctgtttta taccgatacc gaacgtcagc gctgcaacg tgaatttcgt	3900
ttctattttg aaaaaagtgc cggtcgcatg attctggcag caaaaattga acagaccatg	3960
cgtcagcagg gctgtattca gaatgccatg atggcactgc acaaatgga aagcgagca	4020
taaaaacacg gacgccgcaa acggcgtcg aatttcttgg tcgactctag aggatccccg	4080
ggtaccgagc tcgaattcac tggccgtcgt tttaaacgt cgtgactggg aaaacctgg	4140
cgttaccaa cttaatcgcc ttgcagcaca tcccccttc gccagctggc gtaatagcga	4200
agaggcccg accgatcgcc ctteccaaca gtgcgcagc ctgaatggcg aatggcgct	4260
gatgcggtat ttctcctta cgcactcttg cggatattca caccgcatat ggtgcactct	4320
cagtacaatc tgctctgatg ccgcatagtt aagccagccc cgacaccgc caacaccgc	4380
tgacgaattc gttgacagta agacgggtaa gcctgttgat gataccgctg cttactggg	4440

tgcattagcc agtctgaatg acctgtcacg ggataatccg aagtggtcag actggaaaat	4500
cagagggcag gaactgctga acagcaaaaa gtcagatagc accacatagc agacccgcca	4560
taaaacgccc tgagaagccc gtgacgggct tttcttgat tatgggtagt ttccttgcat	4620
gaatccataa aaggcgctg tagtgccatt tacccccatt cactgccaga gccgtgagcg	4680
cagcgaactg aatgtcacga aaaagacagc gactcagtg cctgatggtc ggagacaaaa	4740
ggaatattca gcgatttgcc cgagcttgcg aggggtgtac ttaagccttt agggttttaa	4800
ggtctgtttt gtagaggagc aaacagcgtt tgcgacatcc ttttgaata ctgcggaact	4860
gactaaagta gtgagttata cacagggctg ggatctattc tttttatctt tttttattct	4920
ttctttattc tataaattat aaccacttga atataaaca aaaaaacaca caaaggtcta	4980
gcggaattta cagagggctc agcagaattt acaagttttc cagcaaaggt ctagcagaat	5040
ttacagatac ccacaactca aaggaaaagg actagtaatt atcattgact agcccatctc	5100
aattggtata gtgattaaaa tcacctagac caattgagat gtatgtctga attagtgtt	5160
ttcaaagcaa atgaactagc gattagtgc tatgacttaa cggagcatga aaccaagcta	5220
attttatgct gtgtggcact actcaacccc acgattgaaa acctacaag gaaagaacgg	5280
acggtatcgt tcacttataa ccaatacgct cagatgatga acatcagtag ggaaaatgct	5340
tatgggtgat tagctaaagc aaccagagag ctgatgacga gaactgtgga aatcaggaat	5400
cctttggtta aaggctttga gattttccag tggacaaact atgccaagtt ctcaagcgaa	5460
aaattagaat tagtttttag tgaagagata ttgccttacc ttttcagtt aaaaaattc	5520
ataaaatata atctggaaca tggttaagtct ttgaaaaca aatactctat gaggatttat	5580
gagtggttat taaaagaact aacacaaaag aaaactcaca aggcaaatat agagattagc	5640
cttgatgaat ttaagttcat gttaatgctt gaaaataact accatgagtt taaaaggctt	5700
aaccaatggg ttttgaaacc aataagtaaa gatttaaaaa cttacagcaa tatgaaattg	5760
gtggttgata agcgaggccg cccgactgat acgttgattt tccaagtga actagataga	5820
caaatggatc tcgtaacgga acttgagaac aaccagataa aatgaatgg tgacaaaata	5880
ccaacaacca ttacatcaga ttccctacct cgtaacggac taagaaaaac actacacgat	5940
gctttaactg caaaaattca gctcaccagt tttaggcaa aatttttgag tgacatgcaa	6000
agtaagcatg atctcaatgg ttctgtctca tggctcacgc aaaaacaacg aaccacacta	6060
gagaacatac tggctaataa cggaaggatc tgaggttctt atggcccggc gtagcccaaa	6120
acgcgctgtc gtcaagtcgt taagggcgtg ccttcatca tccgatctgg agtcaaatg	6180

tcctcacgta aagagcttgc caatgctatt cgtgcgctga gcatggacgc agtacagaaa	6240
gccaaatccg gtcacccggg tgcccctatg ggtatggctg acattgccga agtcctgtgg	6300
cgtgatttcc tgaaacacaa cccgcagaat ccgctcctggg ctgaccgtga ccgcttcgtg	6360
ctgtccaacg gccacggctc catgctgac tacagcctgc tgcacctcac cggttacgat	6420
ctgccgatgg aagaactgaa aaacttccgt cagctgcact ctaaaactcc gggccacccg	6480
gaagtaggtt ataccgctgg tgtggaaacc accaccggtc cgctgggtca gggatttggc	6540
aacgcagtcg gtatggcgat tgcagaaaaa acgctggcgg cgcagtttaa ccgtccaggt	6600
cacgacattg tcgaccacta cacctacgcc ttcattggcg acggctgcat gatggaaggc	6660
atctcccacg aagtttgcct tctggcgggt acgctgaagc tgggtaaact gattgcgttc	6720
tacgatgaca acggtatctc aatcgatggc cacgttgaag gctggttcac tgacgacacc	6780
gcaatgcgtt tcgaagctta cggctggcac gtatttcgac acatcgacgg tcatgacgcg	6840
gcatccatca aacgcgcagt agaagaagcg cgcgcagtga ctgacaaaacc gtccctgctg	6900
atgtgcaaaa ccatcatcgg ttctcggttc ccgaacaaag ccggtaccca cgactccac	6960
ggtgcgcgcg tgggcgacgc tgaaattgcc ctgaccgcg aacagctggg ctggaaatac	7020
gcgccgttcg aaatcccgtc tgaaatctat gctcagtggg atgcgaaaga agcaggccag	7080
gcgaaagaat ctgcatggaa tgagaagttt gcggcttacg cgaaagctta tccgcaggaa	7140
gcggctgaat ttacccgccg tatgaaaggc gaaatgccgt ctgacttcga cgccaaagcg	7200
aaagagttaa tcgctaaact gcaggtctat ccggcgaaaa tcgccagccg taaagcgtcg	7260
cagaatgcta tcgaagcgtt cggcccgtg ttgcctgaat tcctcgccgg ctctgctgac	7320
ctggcaccgt ctaacctgac cctgtggtct ggttctaaag caatcaacga agatgctgca	7380
ggttaactaca tccactacgg tgttcgcgag ttcggtatga ccgcgattgc taacggtatc	7440
tccttcacg gtggtttcct gccgtacacc tccaccttcc tgatgttcgt ggaatacgea	7500
cgtaacgccg tacgtatggc tgcgtgatg aaacagcgtc aggtgatggt ttacaccac	7560
gactccatcg gtctggcgca agatggcccg actcaccagc cggttgagca ggtcgtttct	7620
ctgcgcgtga cccgaacat gtctacatgg cgtccgtgtg accaggttga atccgcggtc	7680
gcgtggaaat acggcgttga gcgtcaggac ggcccactg cgcttatect ctcccgtcag	7740
aaactggcgc agcaggaacg aactgaagag caactggcaa acatcgcgcg cgggtggttat	7800
gtgctgaaag actgcgcggc tcagccggaa ctgattttca tcgtaccgg ttcagaagtt	7860
gaactggctg ttgctgccta cgaaaaactg actgccgaag gcgtgaaagc gcgcgtggtg	7920
tccatgccgt ctaccgacgc atttgacaag caggatgctg cttaccgtga atccgtactg	7980
ccgaaagcgg ttactgcacg cgttgctgta gaagcgggta ttgctgacta ctggtacaag	8040

tatgttggcc tgaacgggtc tatcgtcggc atgaccacct tcggtgaatc tgctccggca	8100
gagctgctgt ttgaagagtt cggcttcact gttgataacg ttgttgcgaa agcaaaagaa	8160
ctgctgtaat tagcatttcg ggtaaaaagg tcgcttcggc gacctttttt attaccttga	8220
tatgtccgtt tgcggacaag caatagataa agcgtgttgt agatcacaaa tatttatatg	8280
caataaatat caattatgta atatgcatca cgatatgcgt attgacattt gttgttataa	8340
ctataactca atgttatata agaaattaac tcgaggtat tgacgacagc tatggttcac	8400
tgtccaccaa ccaaaactgt gctcagtacc gccaatatct ctccttgag gggtacaaag	8460
aggtgtccct agaagagatc cagctgtgt aaaaatttta caaaaaggta ttgactttcc	8520
ctacagggtg tgtaataatt taattacagg cgggggcaac cccgcctgtt ctagaggagg	8580
aggaatcgcc atggagagga ttgtcgttac tctcggggaa cgtagtacc caattacat	8640
cgcactcgtt ttgtttaatg aaccagcttc attcttaccg ctgaaatcgg gcgagcaggt	8700
catgttggc accaacgaaa ccctggctcc tctgtatctc gataaggctc gcggcgtact	8760
tgaacaggcg ggtgttaacg tcgatagcgt taccctccct gacggcgagc agtataaaag	8820
cctggctgta ctcgataccg tctttacggc gttgtttaca aagccgcatg gtcgcgatac	8880
tacgttggtg gcgcttgccg gcggcgtagt gggcgatctg accggcttcg cggcgcgag	8940
ttatcagcgc ggtgttcgtt tcattcaagt cccgacgacg ttactgtcgc aggtcgattc	9000
ctccgttggc ggcaaaactg cggtaacca tccctcggc aaaaacatga ttggcgcgtt	9060
ctaccagcct gcttcagttg tggtagatct cgaactgtctg aaaacgcttc ccccgctga	9120
gttagcgtcg gggctggcag aagtcacaa atacggcatt attcttgacg gtgcgttttt	9180
caactggctg gaagagaatc tggatgcgtt gttgcgtctg gacggtccgg caatggcgta	9240
ctgtattcgc cgttgttgtg aactgaaggc agaagttgtc gccccgacg agcgcgaaac	9300
cgggttacgt gctttactga atctgggaca cacctttggt catgccattg aagctgaaat	9360
ggggtatggc aattggttac atggtgaagc ggtcgtcgc ggtatggtga tggcgcgcg	9420
gagctcgga cgtctcgggc agtttagttc tgccgaaacg cagcgtatta taacctgct	9480
cacgcgggct gggttacgg tcaatgggcc gcgcgaaatg tccgcgcagg cgtatttacc	9540
gcataatgctg cgtgacaaga aagtccttgc gggagagatg cgcttaattc ttccgttggc	9600
aattggttaag agtgaagttc gcagcggcgt ttgcacgag cttgtttta acgcatgtc	9660
cgattgtcaa tcagcgtaat catcgttcat gcctgatgcc gctatgtagg ccgataagg	9720
cgttcacgcc gcatccggca accgatgcct gatgcgacgc ggtcgcgtct tatcaggcct	9780
acaggtcgat gccgatatgt acatcgtatt cggcaattaa tacatagcac tcgaggccta	9840
cctagcttcc aagaaagata tcctaacagc acaagagcgg aaagatgttt tgttctacat	9900

ccagaacaac ctctgctaaa attcctgaaa aatittgcaa aaagtgttg actttatcta 9960

caaggtgtgg tataataatc ttaacaacag caggacgctc tagaatgaaa accgtaactg 10020

taaaagatct cgtcatttgt acgggcgcac ctaaaatcat cgtctcgctg atggcgaaag 10080

atatcgccag cgtgaaatcc gaagctctcg cctatcgtga agcggacttt gatattctgg 10140

aatggcgtgt ggaccactat gccgacctct ccaatgtgga gtctgtcatg gcggcagcaa 10200

aaattctccg tgagaccatg ccagaaaaac cgtctgtgtt taccttccgc agtgccaaag 10260

aaggcggcga gcaggcgatt tccaccgagg cttatatgtc actcaatcgt gcagccatcg 10320

acagcggcct ggttgatatg atcgatctgg agttatttac cggatgatgat cagggttaaag 10380

aaaccgtcgc ctacgccac gcgcatgatg tgaaagtagt catgtccaac catgacttcc 10440

ataaaacgcc ggaagccgaa gaaatcattg cccgtctgcg caaaatgcaa tccttcgacg 10500

ccgatattcc taagattgcg ctgatgccgc aaagtaccag cgatgtgctg acgttgcttg 10560

ccgcgacctt ggagatgcag gagcagtatg ccgatcgtcc aattatcacg atgtcgatgg 10620

caaaaactgg cgtaatttct cgtctggctg gtgaagtatt tggctcggcg gcaacttttg 10680

gtgcggtaaa aaaagcgtct gcgccagggc aaatctcggt aaatgatttg cgcacggtat 10740

taactatttt acaccaggca taagcaataa tatttcggcg ggaacaccct ccccgccgaa 10800

ctaaaaaata tattcaatcg tatttaataa aaatatttcg tgagtctctg tgcgctaatt 10860

ctc 10863

<210> 58

<211> 5500

<212> DNA

<213> Escherichia coli

<220><221> gene

<222> (1)..(5500)

<223> Nucleotide sequence of ppc::PR-pyc

<400> 58

cgaaccgacg tactacaac cggtagcgcg acataaaggt catatctcta acgcatccg 60

tattcagggc cagtcggggc actccagcga tccagcacgc ggagttaac ctatcgaaact 120

aatgcacgac gccatcgggc atattttgca attgcgcgat aacctgaaag aacgttatca 180

ctacgaagcg tttaccgtgc catacctac gctcaacctc gggcatattc acggtggcga 240

cgctttctaac cgtatttgcg cttgtctgtga gttgcatatg gatattcgtc cgctgccttg 300

catgacactc aatgaactta atggtttgct caacgatgca ttggctccgg tgagcgaacg 360

ctggccgggt cgctgacgg tcgacgagct gcatccgccg atccctggct atgaatgccc	420
accgaatcat caactgggtg aagtgggtga gaaattgtct ggagcaaaaa ccgaagtgg	480
gaactactgt accgaagcgc cgtttattca aacgttatgc ccgacgtgg tgttggggcc	540
tggtcaatt aatcaggctc atcaacctga tgaatatctg gaaacacggt ttatcaagcc	600
cacccgcgaa ctgataaccc aggttaattca ccatttttgc tggcattaaa acgtaggccg	660
gataaggcgc tcgcgccga tccggcactg ttgccaaact ccagtgccgc aataatgtcg	720
gatgcgatac ttgcgcatct tatccgacct acacctttgg tgttacttgg ggcgattttt	780
taacatttcc ataagttacg cttattttaa gcgtcgtgaa tttaatgacg taaattcctg	840
ctatttattc gticgtttaa tctatcaccg caagggataa atatctaaca ccgtgcgtgt	900
tgactatttt acctctggcg gtgataatgg ttgcatgtac taatctagat aaggaatata	960
gccatgtcgc aaagaaaatt cgccggcttg agagataact tcaatctctt gggtgaaaag	1020
aacaaaatat tggtaggctaa tagaggagaa attccaatca gaatttttcg taccgtcat	1080
gaactgtcta tgcagacggt agctatatat tctcatgaag atcgtcttcc aacgcacaaa	1140
caaaaggctg acgaagcata cgtcataggt gaagtaggcc aatatacccc cgtcggcgct	1200
tatttggcca ttgacgaaat catttccatt gcccaaaaac accaggtaga tttcatccat	1260
ccaggttatg ggttcttgtc tgaaaattcg gaatttgccg acaaagtagt gaaggccggt	1320
atcacttgga ttggccctcc agctgaagtt attgactccg tgggtgataa ggtctcagct	1380
agaaacctgg cagcaaaagc taatgtgccc accgttctcg gtacaccagg tcctatagaa	1440
actgtagagg aagcacttga ctctgtcaat gaatacggct acccggtgat cattaaggcc	1500
gcctttgggt gtggtggtag aggtatgaga gtctgttagag aaggtgacga cgtggcagat	1560
gcctttcaac gtgtacctc cgaagcccggt actgccttcg gtaatggtac ctgctttgtg	1620
gaaagattct tggacaagcc aaagcatatt gaagttcaat tgttggccga taaccacgga	1680
aacgtggttc atcttttcga aagagactgt tccgtgcaga gaagacacca aaaggttgtc	1740
gaagtggccc cagcaaaagc tttaacccgt gaagtccgtg acgccatttt gacagatgca	1800
gttaaattgg ccaaagagtg tggctacaga aatgcgggta ctgctgaatt ctgggtgat	1860
aacaaaaata gacactattt cattgaaatt aatccaagaa tccaagtgga acataccatc	1920
acagaagaaa ttaccggtat agatatgtg gcggtcaga tccaaattgc ggcaggtgcc	1980
tctctacccc agctgggcct attccaggac aaaattacga ctctggctt tgccattcag	2040
tgccgtatta ccacggaaga cctgctaag aacttccaac cagataccgg tagaatagaa	2100
gtgtaccgtt ctgcaggtgg taatgggtgt agactggatg gtggtaacgc ctatgcagga	2160
acaataatct cactcatta cgactcaatg ctggtcaaat gctcatgctc cggttcacc	2220

tacgaaatcg ttcgtagaaa aatgattcgt gcattaatcg agttcagaat tagaggtgtc	2280
aagaccaaca ttcccttcct attgactctt ttgaccaatc cagtatttat tgagggtaca	2340
tactggacga cttttattga cgacacccca caactgttcc aaatggtttc atcacaaaac	2400
agagcccaaa aacttttaca ttacctgcc gacgtggcag tcaatggttc atctatcaag	2460
ggtcaaatgg gcttgccaaa attaaaatca aatccaagtg tccccattt gcacgatgct	2520
cagggcaatg tcatcaacgt taaaaagtct gcaccacat ccggatggag gcaagtgcta	2580
ctagaaaagg ggccagctga atttgccaga caagttagac agttcaatgg tactttattg	2640
atggacacca cctggagaga cgctcatcaa tctctacttg caacaagagt cagaaccac	2700
gatttggcta caatcgctcc aacaaccgca catgcccttg caggtcgttt cgccttagaa	2760
tgttggggtg gtgccacatt cgatgttgca atgagatttt tgcatgagga tccatgggaa	2820
cgtttgagaa aattaagatc tctggtgcct aatattccat tccaaatgtt attgcgtggt	2880
gccaatgggtg tggcttattc ttcatgtcct gacaatgcta ttgaccattt cgtcaagcaa	2940
gccaaaggata atgggtgtga tataattaga gtctttgatg ccttaaatga cttggaacaa	3000
ttgaaggctg gtgtagatgc tgtgaagaag gcaggtgggtg ttgtagaagc cactgtttgt	3060
ttctctgggg atatgcttca gccaggcaag aaatacaatt tggattacta cttggaatt	3120
gctgaaaaaa ttgtccaaat gggcactcat atcctgggta tcaaagatat ggcaggtacc	3180
atgaagccag cagctgccaa actactgatt ggatctttga gggctaagta ccctgatctc	3240
ccaatacatg ttacacatca cgattctgca ggtactctg ttgcatcaat gactgcgtgt	3300
gctctggcgg gcgccgatgt cgttgatgtt gccatcaact caatgtctgg tttaacttca	3360
caaccatcaa tcaatgctct gttggcttca ttagaaggta atattgacac tggatattaac	3420
gttgagcatg tccgtgaact agatgcata tgggcagaga tgagattgtt atactctgt	3480
ttcgaggctg acttgaaggg ccagatcca gaagtttate aacatgaaat cccaggtggt	3540
caattgacaa acttgttgtt tcaagcccaa caattgggtc ttggagaaca atgggccgaa	3600
acaaaaagag cttacagaga agccaattat ttattgggtg atattgtcaa agttaccca	3660
acttcgaagg tcgttgggtg tctggcaciaa tttatggtct ccaataaatt aacttccgat	3720
gatgtgagac gcctggctaa ttctttgat ttcctgact ctgttatgga tttcttcgaa	3780
ggcttaatcg gccaaacata tgggtgggtc ccagaacat ttagatcaga cgttttaagg	3840
aacaagagaa gaaagttagc ttgtcgtcca ggcctggaac tagagccatt tgatctcgaa	3900
aaaattagag aagacttgca gaatagattt ggtgatgttg atgagtcga cgttgcttct	3960

tataacatgt acccaagagt ttatgaagac ttccaaaaga tgagagaaac gtatggtgat 4020
 ttatctgtat tgccaacaag aagctttttg tctccactag agactgacga agaaattgaa 4080
 gttgtaatcg aacaaggtaa aacgctaatt atcaagctac aggctgtggg tgatttgaac 4140
 aaaaagaccg gtgaaagaga agtttacttt gatttgaatg gtgaaatgag aaaaattcgt 4200
 gttgctgaca gatcacaaaa agtggaact gttactaaat ccaaagcaga catgcatgat 4260
 ccattacaca ttggtgcacc aatggcaggt gtcattgttg aagttaaagt tcataaagga 4320
 tcactaataa agaagggcca acctgtagcc gtattaagcg ccatgaaaat ggaaatgatt 4380

atatcttctc catccgatgg acaagttaaa gaagtgtttg tctctgatgg tgaaaatgtg 4440
 gactcttctg atttattagt tctattagaa gaccaagtgc ctggtgaaac taaggcatga 4500
 tcttctctct ctgcaaaccc tcgtgctttt gcgcgagggt tttctgaaat acttctgttc 4560
 taacaccctc gttttcaata tatttctgtc tgcattttat tcaaattctg aatatacctt 4620
 cagatatact taaggaattg tcgttacatt cggcgatatt tttcaagac aggttcttac 4680
 tatgcattcc acagaagtc aggctaaacc tctttttagc tggaaagccc tgggttgggc 4740
 actgctctac ttttggtttt tctctactct gctacaggcc attatttaca tcagtggtta 4800

tagtggcact aacggcattc gcgactcgct gttattcagt tcgctgtggt tgatcccgtt 4860
 attctctttt ccgaagcgga ttaaaattat tgccgcagta atcggcgtgg tgctatgggc 4920
 ggcctctctg gcggcgctgt gctactacgt catctacggt caggagtctc cgcagagcgt 4980
 tctgtttgtg atgttcgaaa ccaacaccaa cgaagccagc gagtatttaa gccagtattt 5040
 cagcctgaaa attgtgctta tcgcgtggc ctatacggcg gtggcagttc tgctgtggac 5100
 acgcctgcgc ccggtctata ttccaaagcc gtggcggtat gttgtctctt ttgccctgct 5160
 ttatggcttg attctgcatc cgatcgccat gaatacgttt atcaaaaaca agccgtttga 5220

gaaaacgttg gataacctgg cctcgcgtat ggagcctgcc gcaccgtggc aattcctgac 5280
 cggctattat cagtatcgtc agcaactaaa ctgcctaaca aagttactga atgaaaataa 5340
 tgccctgccg cactggcta atttcaaaga tgaatcgggt aacgaaccgc gcactttagt 5400
 gctggtgatt ggcgagtcga cccagcgagg acgcatgagt ctgtacggtt atccgcgtga 5460
 aaccacgccg gagctggatg cgctgcataa aaccgatccg 5500

<210> 59

<211> 200

<212> DNA

<213> Escherichia coli

<220><221> promoter

<222> (1)..(200)

<223> Nucleotide sequence of acpP promoter

<400> 59

cacaaaatgc tcatgttgcg cgcagtctgc gtggttatga gtaataatta gtgcaaatg	60
atttgcgtta ttggggggta aggcctcaaa ataacgtaaa atcgttgtaa gacctgccgg	120
gatttagttg caaatttttc aacattttat acactacgaa aaccatcgcg aaagcgagtt	180
ttgataggaa atttaagagt	200

<210> 60

<211> 200

<212> DNA

<213> Escherichia coli

<220><221> promoter

<222> (1)..(200)

<223> Nuclotide sequence of rplU promoter

<400> 60

ttatggttag gaatataggg tgattgtact gaaaaaatgg cacagataaa cgttaccgta	60
caagttgtgt ttttttctt cgtgtattga ctgtagcact tgtcaaaggc gtgcgttttg	120
cgtaatatcc gcgccctatt gtgaatattt atagcgact ctgaatcatt gaaaagggtg	180
gcgcggaagc ggagttttat	200

<210> 61

<211> 60

<212> DNA

<213> Artificial Sequence

<220><223> Primer

<220><221> primer_bind

<222> (1)..(60)

<223> Sequence of DNA primer MS604

<400> 61

aacgccgtat aatgggcgca gattaagagg ctacagtggg cttacatggc gatagctaga	60
	60

<210> 62

<211> 60

<212> DNA
 <213> Artificial Sequence
 <220><223> Primer
 <220><221> primer_bind
 <222> (1)..(60)
 <223> Sequence of DNA Primer MS605
 <400> 62
 tgtcggatcg ataataggg caaacaac gcgcattccg gaaaacgatt ccgaagccca 60
 60

<210> 63
 <211> 60
 <212> DNA
 <213> Artificial Sequence
 <220><223> Primer
 <220><221> primer_bind
 <222> (1)..(60)
 <223> Sequence of DNA Primer MS608
 <400> 63
 aaagtctgcc tgcaagtctg acagggaac ttttggg cttacatggc gatagctaga 60
 60

<210> 64
 <211> 60
 <212> DNA
 <213> Artificial Sequence
 <220><223> Primer
 <220><221> primer_bind
 <222> (1)..(60)
 <223> Sequence of DNA Primer MS609
 <400> 64
 ttgcaaat gccctgaac agggcaacag cggagtccg gaaaacgatt ccgaagccca 60
 60

<210> 65
 <211> 25

<212> DNA
 <213> Artificial Sequence
 <220><223> Primer
 <220><221> primer_bind
 <222> (1)..(25)
 <223> Sequence of DNA Primer MS461
 <400> 65

ggctatatattc cttatctaga ttagt

25

<210> 66
 <211> 25
 <212> DNA
 <213> Artificial Sequence

<220><223> Primer
 <220><221> primer_bind
 <222> (1)..(25)
 <223> Sequence of DNA Primer MS346
 <400> 66

gtctgacagg tgccggattt catat

25

<210> 67
 <211> 45
 <212> DNA
 <213> Artificial Sequence

<220><223> Primer
 <220><221> primer_bind
 <222> (1)..(45)
 <223> Sequence of DNA Primer RP712
 <400> 67

tctagataag gaatatagcc atgaccgcac cgattcagga tctgc

45

<210> 68
 <211> 45

<

212> DNA
 <213> Artificial Sequence
 <220><223> Primer

<220><221> primer_bind
 <222> (1)..(45)
 <223> Sequence of DNA Primer RP714
 <400> 68
 aaatccggca cctgtcagac ttattttgcg ctaccctggg ttttt 45
 <210> 69
 <211> 60
 <212> DNA
 <213> Artificial Sequence
 <220><223> Primer
 <220><221> primer_bind
 <222> (1)..(60)
 <223> Sequence of DNA Primer RP731
 <400> 69
 catgtactaa tctagataag gaatatagcc atgaaactga ttattgggat gacggggggcc 60
 60
 <210> 70
 <211> 60
 <212> DNA
 <213> Artificial Sequence
 <220><223> Primer
 <220><221> primer_bind
 <222> (1)..(60)
 <223> Sequence of DNA Primer RP732
 <400> 70
 gccggatatg aaatccggca cctgtcagac ttattcgatc tcctgtgcaa attgttctgc 60
 60
 <210> 71
 <211> 45
 <212> DNA
 <213> Artificial Sequence
 <220>
 ><223> Primer
 <220><221> primer_bind

<222> (1)..(45)

<223> Sequenc of DNA Primer MS669

<400> 71

tctagataag gaatatagcc atgaaacgac tcattgtagg catca 45

<210> 72

<211> 45

<212> DNA

<213> Artificial Sequence

<220><223> Primer

<220><221> primer_bind

<222> (1)..(45)

<223> Sequence of DNA Primer MS666

<400> 72

accgaacagg cttatgtcca gatagcagg atagcggttg aatcg 45

<210> 73

<211> 24

<212> DNA

<213> Artificial Sequence

<220><223> Primer

<220><221> primer_bind

<222> (1)..(24)

<223> Sequence of DNA Primer RP607

<400> 73

tggacataag cctgttcggt tcgt 24

<210> 74

<211> 25

<212> DNA

<213> Artificial Sequence

<220><223> Primer

<220><221> primer_bind

<222> (1)..(25)

<223> Sequence of DNA Primer RP621

<400> 74

ttagatttga ctgaaatcgt acagt

25

<210> 75

<211> 45

<212> DNA

<213> Artificial Sequence

<220><223> Primer

<220><221> primer_bind

<222> (1)..(45)

<223> Sequence of DNA Primer MS676

<400> 75

tctagataag gaatatagcc atgaaactga tcgtcgggat gacag

45

<210> 76

<211> 45

<212> DNA

<213> Artificial Sequence

<220><223> Primer

<220><221> primer_bind

<222> (1)..(45)

<223> Sequence of DNA Primer MS680

<400> 76

acgatttcag tcaaatctaa ttattcattc tctgagaaa aattc

45

<210> 77

<211> 45

<212> DNA

<213> Artificial Sequence

<220><223> Primer

<220><221> primer_bind

<222> (1)..(45)

<223> Sequence of DNA Primer MS686

<400> 77

tctagataag gaatatagcc atgacggcac gcatcatcat tggta

45

<210> 78

<211> 45

<212> DNA
 <213> Artificial Sequence
 <220><223> Primer
 <220><221> primer_bind
 <222> (1)..(45)
 <223> Sequence of DNA Primer MS684
 <400> 78
 acgatttcag tcaaattctaa ttaattaaaa cgtagctcgc cttca 45

<210> 79
 <211> 45
 <212> DNA
 <213> Artificial Sequence
 <220><223> Primer
 <220><221> primer_bind
 <222> (1)..(45)
 <223> Sequence of DNA Primer MS692
 <400> 79
 tctagataag gaatatagcc atgaaacgaa ttgttggtggg aatca 45

<210> 80
 <211> 45
 <212> DNA
 <213> Artificial Sequence
 <220><223> Primer
 <220><221> primer_bind
 <222> (1)..(45)
 <223> Sequence of DNA Primer MS691
 <400> 80
 acgatttcag tcaaattctaa ttaatccccc tcccaacggc gatca 45

<210> 81
 <211> 45
 <212> DNA
 <213> Artificial Sequence
 <220><223> Primer

<220><221> primer_bind
 <222> (1)..(45)
 <223> Sequence of DNA Primer RP677
 <400> 81
 acgatttcag tcaaattctaa ttaatccccc tccaacggc gatca 45
 <210> 82
 <211> 25
 <212> DNA
 <213> Artificial Sequence
 <220><223> Primer
 <220><221> primer_bind
 <222> (1)..(25)
 <223> Sequence of DNA Primer RP671
 <400> 82
 ttaatcgcct tgcagcacat ccccc 25

 <210> 83
 <211> 24
 <212> DNA
 <213> Artificial Sequence
 <220><223> Primer
 <220><221> primer_bind
 <222> (1)..(24)
 <223> Sequenc of DNA primer RP664
 <400> 83
 acgaaccgaa caggcttatg tcca 24
 <210> 84
 <211> 45
 <212> DNA
 <213> Artificial Sequence
 <220><223> Primer
 <220><221> primer_bind
 <222> (1)..(45)
 <223> Sequence of DNA Primer RP702
 <400> 84

gccgtcggtt tacaacgtcg gatccgccta cctagcttcc aagaa 45

<210> 85
 <211> 45
 <212> DNA
 <213> Artificial Sequence
 <220><223> Primer
 <220><221> primer_bind
 <222> (1)..(45)
 <223> Sequence of DNA Primer RP783
 <400> 85

cctacaatga gtcgtttcat taggttttcc tcaaccggg agcgt 45

<210> 86
 <211> 45
 <212> DNA
 <213> Artificial Sequence
 <220><223> Primer
 <220><221> primer_bind
 <222> (1)..(45)
 <223> Sequence of DNA primer RP781
 <400> 86

cccgggttga ggaaaaccta atgaaacgac tcattgtagg catca 45

<210> 87
 <211> 45
 <212> DNA
 <213> Artificial Sequence
 <220><223> Primer
 <220><221> primer_bind
 <222> (1)..(45)
 <223> Sequence of DNA Primer 780
 <400> 87

atgtgctgca aggcgattaa gatagcaggt atagcggttg aatcg 45

<210> 88
 <211> 45

<212> DNA
 <213> Artificial Sequence
 <220><223> Primer
 <220><221> primer_bind
 <222> (1)..(45)
 <223> Sequence of DNA Primer RP700
 <400> 88
 gccgtcgttt tacaacgtcg agcggggcgg ttgtcaacga tgggg 45

<210> 89
 <211> 45
 <212> DNA
 <213> Artificial Sequence
 <220><223> Primer
 <220><221> misc_feature
 <222> (1)..(45)
 <223> Sequence of DNA Primer RP784
 <400> 89
 cctacaatga gtcgtttcat ggctatatattc ctctctgca tgaga 45

<210> 90
 <211> 45
 <212> DNA
 <213> Artificial Sequence
 <220><223> Primer
 <220><221> primer_bind
 <222> (1)..(45)
 <223> Sequenc of DNA Primer RP779
 <400> 90
 tgcagaggag gaatatagcc atgaaacgac tcattgtagg catca 45

<210> 91
 <211> 45
 <212> DNA
 <213> Artificial Sequence
 <220><223> Primer

<220><221> primer_bind
 <222> (1)..(45)
 <223> Sequence of DNA Primer MS1383
 <400> 91
 gatggggtgt ctgggtaat atgtcgcaaa gaaaattcgc cggct 45
 <210> 92
 <211> 45
 <212> DNA
 <213> Artificial Sequence
 <220><223> Primer
 <220><221> primer_bind
 <222> (1)..(45)
 <223> Sequence of DNA Primer MS1384
 <400> 92
 gggtttgag aagaggaaga tcatgcctta gtttaacag gaact 45

 <210> 93
 <211> 45
 <212> DNA
 <213> Artificial Sequence
 <220><223> Primer
 <220><221> primer_bind
 <222> (1)..(45)
 <223> Sequence of DNA Primer MS1429
 <400> 93
 attcctgcta tttattcggtt cgtaaactct atcaccgcaa gggat 45
 <210> 94
 <211> 45
 <212> DNA
 <213> Artificial Sequence
 <220><223> Primer
 <220><221> primer_bind
 <222> (1)..(45)
 <223> Sequence of DNA Primer MS1430
 <400> 94

gccaattttc ttgcgacat ggctatatc cttatctaga ttagt

45