



(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2018-0047228
(43) 공개일자 2018년05월10일

(51) 국제특허분류(Int. Cl.)
C12N 15/63 (2006.01) *C12P 7/16* (2006.01)
(52) CPC특허분류
C12N 15/63 (2013.01)
C12P 7/16 (2013.01)
(21) 출원번호 10-2016-0143113
(22) 출원일자 2016년10월31일
심사청구일자 2016년10월31일

(71) 출원인
고려대학교 산학협력단
서울특별시 성북구 안암로 145, 고려대학교 (안암동5가)
(72) 발명자
오민규
경기도 성남시 분당구 정자일로 248, 602동 1601호 (정자동, 파크뷰)
정휘민
서울특별시 성북구 아리랑로 89, 109동 1902호 (돈암동, 일신건영휴먼빌아파트)
(74) 대리인
이처영, 장제환

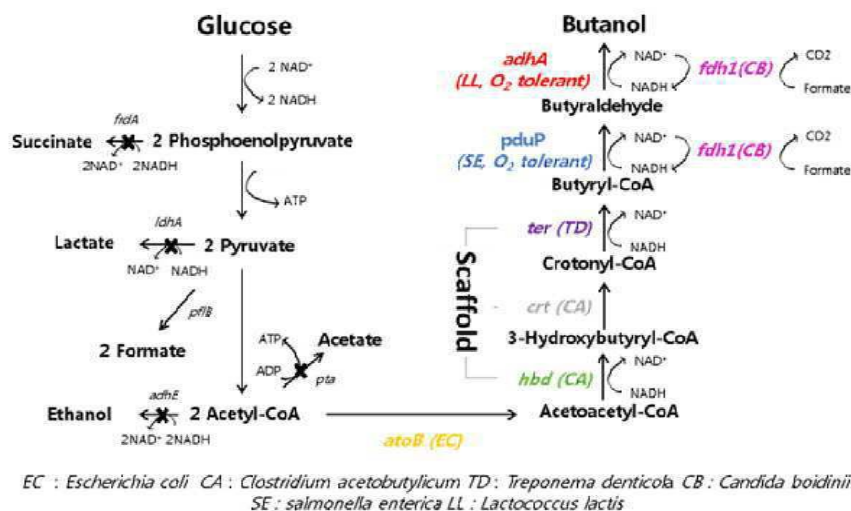
전체 청구항 수 : 총 11 항

(54) 발명의 명칭 호기조건에서 부탄올을 생성하는 변이 미생물 및 이를 이용한 부탄올의 제조방법

(57) 요약

본 발명은 혐기조건에서 글루코오스로부터 부탄올 생성능을 가지는 변이 미생물에서 알코올 디하이드로게네이즈를 코딩하는 유전자와 프로피온알데하이드 디하이드로게네이즈를 코딩하는 유전자가 호기성 미생물 유래 유전자로 치환되어 있고, 시트레이트 신타아제 유전자의 발현이 하향조절되어 있는 변이 미생물 및 이를 이용한 부탄올의 제조방법에 관한 것으로, 발명에 따르면, 저가의 배양매지를 이용할 수 있는 호기발효를 통하여 부탄올을 고수율로 제조할 수 있어, 원료비용 절감효과를 얻을 수 있으며, 종래 필수적이었던 혐기조건에 구애받지 않아 부탄올 제조공정의 다양성을 꾀할 수 있다.

대표도 - 도2



(52) CPC특허분류

C12Y 101/01001 (2013.01)
C12Y 101/01157 (2013.01)
C12Y 103/99002 (2013.01)
C12Y 203/01016 (2013.01)
C12Y 203/03001 (2013.01)
C12Y 402/01017 (2013.01)
Y02E 50/10 (2013.01)

(72) 발명자

허민지

서울특별시 동대문구 무학로42길 31 (제기동)

엄재용

인천광역시 남동구 백범로124번길 126, 413동 509
 호 (만수동, 만수주공4단지아파트)

이상우

경기도 남양주시 도농로 34, 508동 804호 (도농동,
 부영아파트1단지)

이 발명을 지원한 국가연구개발사업

과제고유번호 20143030091040

부처명 산업통상부

연구관리전문기관 한국에너지기술평가원

연구사업명 에너지기술개발사업

연구과제명 케톤-알코올 발효공정과 화학축매공정을 융합한 고탄소 바이오연료 생산 기술

기 여 율 1/2

주관기관 고려대학교 산학협력단

연구기간 2014.12.01 ~ 2017.11.30

이 발명을 지원한 국가연구개발사업

과제고유번호 20133030000300

부처명 산업통상부

연구관리전문기관 한국에너지기술평가원

연구사업명 에너지기술개발사업

연구과제명 전기활성 미생물을 이용한 저급탄소원의 바이오연료 전환기술

기 여 율 1/2

주관기관 고려대학교 산학협력단

연구기간 2013.12.01 ~ 2016.09.30

명세서

청구범위

청구항 1

글루코오스로부터 아세틸-CoA를 생성하는 경로를 가지는 미생물에서 다음 유전자가 도입되어 있는 호기조건에서 부탄올 생성능을 가지는 변이 미생물;

- (a) 호기성 균주 유래 알코올 디하이드로게네이즈를 코딩하는 유전자;
- (b) 호기성 균주 유래의 프로피온알데하이드 디하이드로게네이즈를 코딩하는 유전자;
- (c) 부틸-CoA 디하이드로게네이즈를 코딩하는 유전자;
- (d) 크로토네이즈(crotonase)를 코딩하는 유전자;
- (e) 하이드록시부티릴-CoA 디하이드로게네이즈를 코딩하는 유전자; 및
- (f) 티올레이즈(thiolase)를 코딩하는 유전자.

청구항 2

글루코오스로부터 아세틸-CoA를 생성하는 경로를 가지는 미생물에서 다음 유전자가 도입되어 있고, 시트레이트 신타아제(citrate synthase)를 코딩하는 유전자의 발현이 하향조절되어 있는 호기조건에서 부탄올 생성능을 가지는 변이 미생물;

- (a) 호기성 균주 유래 알코올 디하이드로게네이즈를 코딩하는 유전자;
- (b) 호기성 균주 유래의 프로피온알데하이드 디하이드로게네이즈를 코딩하는 유전자;
- (c) 부틸-CoA 디하이드로게네이즈를 코딩하는 유전자;
- (d) 크로토네이즈(crotonase)를 코딩하는 유전자;
- (e) 하이드록시부티릴-CoA 디하이드로게네이즈를 코딩하는 유전자; 및
- (f) 티올레이즈(thiolase)를 코딩하는 유전자.

청구항 3

제1항 또는 제2항에 있어서, 숙신산, 락테이트, 에탄올 및 아세테이트로 구성된 군에서 선택되는 부산물의 생합성에 관여하는 효소를 코딩하는 유전자가 결실되어 있는 것을 특징으로 하는 변이미생물.

청구항 4

제1항 또는 제2항에 있어서, 푸마레이트 리덕테이즈, 락테이트 디하이드로 게네이즈, 포스페이트 아세틸 트랜스퍼레이즈 및 아세트알데하이드/알코올 디하이드로게네이즈로 구성된 군에서 선택되는 효소를 코딩하는 유전자가 결실되어 있는 것을 특징으로 하는 변이 미생물.

청구항 5

제1항 또는 제2항에 있어서, 상기 알코올 디하이드로게네이즈를 코딩하는 유전자는 살모넬라 유래인 것을 특징으로 하는 변이 미생물.

청구항 6

제1항 또는 제2항에 있어서, 상기 프로피온알데하이드 디하이드로게네이즈를 코딩하는 유전자는 락토코커스 유래인 것을 특징으로 하는 변이 미생물.

청구항 7

제1항 또는 제2항에 있어서, 포메이트 디하이드로게네이즈(formate dehydrogenase)를 코딩하는 유전자가 추가로 도입되어 있는 변이 미생물.

청구항 8

제2항에 있어서, 시트레이트 신타아제(citrate synthase)를 코딩하는 유전자 발현의 하향조절은 CRISPR/Cas 시스템에 의해 수행되는 것을 특징으로 하는 변이 미생물.

청구항 9

제2항에 있어서, 시트레이트 신타아제(citrate synthase)를 코딩하는 유전자 발현의 하향조절은 상기 시트레이트 신타아제(citrate synthase)를 코딩하는 유전자의 5'-UTR 부분이 서열번호 1~4 중 어느 하나로 표시되는 올리고뉴클레오타이드로 치환된 것을 특징으로 하는 변이 미생물.

청구항 10

다음 단계를 포함하는 부탄올의 제조방법;

- (a) 제1항 내지 제9항 중 어느 한 항의 변이 미생물을 배양하여 부탄올을 생성시키는 단계; 및
- (b) 생성된 부탄올을 수득하는 단계.

청구항 11

제10항에 있어서, 상기 배양은 호기 또는 미세호기 조건에서 수행되는 것을 특징으로 하는 방법.

발명의 설명

기술 분야

[0001]

본 발명은 호기조건에서 부탄올 생성능을 가지는 변이 미생물에 관한 것으로, 더욱 자세하게는 혐기조건에서 글루코오스로부터 부탄올 생성능을 가지는 변이 미생물에서 알코올 디하이드로게네이즈를 코딩하는 유전자와 프로피온알데하이드 디하이드로게네이즈를 코딩하는 유전자가 호기성 미생물 유래 유전자로 치환되어 있고, 시트레이트 신타아제 유전자의 발현이 하향조절되어 있는 변이 미생물 및 이를 이용한 부탄올의 제조방법에 관한 것이다.

배경 기술

[0002]

n-부탄올은 가솔린과 비슷한 에너지밀도를 가지는 알코올로서 가솔린과 혼합하여 바이오연료로 활용될 수 있는 물질이다. n-부탄올은 자연적으로 Clostridium 종에서 아세트-부탄올-에탄올 발효를 통해 생산되는 것으로 잘 알려져 있으며, 많은 연구가 선행되어 왔다. 하지만 Clostridium 종은 대사공학적 접근이 제한적이며, 혐기조건에서만 배양할 수 있는 특징을 지닌다. 따라서 E. coli를 포함한 다양한 종의 미생물에서 n-부탄올 생산을 위한 연구가 진행되어 왔다. 예를 들어 E. coli로부터 n-부탄올을 생산하기 위해서 Coenzyme A 기반의 대사회로나

reverse beta oxidative 대사회로, ACS 기반의 대사회로등이 개발되어왔다(Lan, E. I., and Liao, J. C., *Bioresour. Technol.* 135:339, 2013).

- [0003] 대표적으로 E. coli에서 아세트산, 젖산, 에탄올, 숙신산등의 부산물 대사회로를 제거하여 30g/L 이상의 부탄올을 생산한 연구 성과가 발표되었으며, 또한 세포내의 redox 정도를 조절하는 것이 부탄올 생산에 도움이 된다는 연구 결과가 존재한다(Lim, JH *et al.*, *Metab. Eng.* 20:56, 2013; Sani, M.*et al.*, *Biotechnol. Biofuels* 9:69, 2016).
- [0004] 대사공학적 도구로 특정 유전자의 발현을 높이기 위해 플라스미드를 사용하는 방법이 주로 사용되어왔다. 하지만 플라스미드 기반의 유전자 과발현 방법은 수반되는 비용과 불안정성 문제로 인해 산업적으로 활용되기 힘든 측면이 있다. 따라서 플라스미드가 아닌 미생물의 지놈을 직접 편집하는 기술이 필요로 된다. 전통적으로 미생물의 지놈 상에 목적 유전자 삽입하는 낙인(Knock in) 혹은 불필요한 유전자를 삭제하는 낙 아웃(Knock out) 방법이 사용되어왔다. 또한 최근 유전자 조작 도구의 발전으로 여러 유전자들의 발현을 효율적으로 조절하는 것이 가능해 졌으며, 유전자 발현의 정도를 예측하고 구현하는 기술이 주목받고 있다. 이 방법을 통해서 유전자의 발현을 On-Off가 아닌 적정 수준으로 조절할 수 있기에 대사회로 최적화에 효과적으로 활용될 수 있다. 대장균에서 small RNA와 Hfq 단백질을 활용하여 유전자의 발현량을 조절하는 기술이나 dCas9을 활용하여 유전자를 낙다운(Knock down) 하려는 연구가 시도되어왔다. small RNA나 dCas9을 활용한 유전자 낙다운 방법은 hfq 및 dCas9과 같은 부가적인 단백질의 발현이 필수적이기 때문에 세포 성장에 불리한 면이 존재하며, 정확도가 높지 않아 세포 내 여러 부위를 표적하여 예상치 못한 문제점을 발생시킬 수 있다. 따라서 5' UTR 부근의 유전적 염기서열을 임의로 조작 및 치환하여 외부 단백질의 발현 없이도 여러 유전자의 발현을 조절하는 기술이 유리하다. 5' UTR 부근을 조작하는 방법은 CRISPR/Cas9을 활용한 지놈 편집방법을 통해 구현 가능하다.
- [0005] CRISPR/Cas 시스템은 박테리아나 고세균류(archaea)가 바이러스의 공격 또는 다른 미생물로부터 전해지는 접합 플라스미드(conjugated plasmid)로부터 스스로를 방어하기 위한 면역기작이다. 세포에 유입된 외래 유전자는 endonuclease인 Cas protein에 의해 30 bp 내외로 잘리게 되고, 이것은 박테리아 게놈 내에 여러 개의 짧은 염기서열이 반복되는 구간인 CRISPR (Clustered Regularly Interspaced Short Palindromic Repeats)에 sequence로 보존하게 된다. 이렇게 CRISPR 에 삽입된 DNA를 spacer 라 명명하며, 재차 같은 염기서열을 지니는 DNA 혹은 RNA가 침입했을 때 그것을 인식하는 역할을 한다. CRISPR에 해당되는 DNA 서열은 우선 전사기작을 통해 precursor-crRNA(pre-crRNA)로 발현되고, 이는 다시 cas nuclease에 의해 mature-crRNA (crRNA)로 변형된 후, Cas protein과 결합하여 인식된 유전자 서열을 절단하게 된다. CRISPR/Cas9 시스템은 유전자 편집 가위로 실험자가 원하는 유전자 서열을 목표화 할 수 있다는 장점을 가지고 있다(Bikard, D *et al.*, *Nucleic Acids Res.* 41, 7429, 2013; Cleto, S *et al.*, *ACS Synth. Biol.* 5, 375, 2016).
- [0006] 위 시스템을 활용하여 목적하는 염기서열에 이중가닥 절단을 만들고 세포가 가지는 DNA 복구 기작을 활용하여 지놈 편집이 가능하다. 세포는 비상동 말단접합 이라는 기작을 활용하여 손상된 DNA를 주형 가닥 없이 복구가 가능하며 그 결과 염기서열의 삽입이나 삭제가 일어나게 된다. 또한 상동재조합이라는 기작을 통해 손상된 DNA 수복이 가능하며 그 과정에서 주형 DNA와 같은 염기서열이 손상된 부위에 삽입된다. 이와 같이, CRISPR/Cas9과 세포가 가지는 DNA 복구기작을 활용하여 유전자의 낙인, 낙 아웃, 치환 등 다소 긴 유전서열을 편집할 수 있을 뿐만아니라 점 돌연변이(oaint mutation)와 같은 짧은 유전자 서열 또한 편집 가능하다.
- [0007] 한편, 지금까지 연구된 부탄올 생산균주는 Coenzyme A 대사회로를 통한 혐기조건에서 생산되어 왔다. 미생물의 혐기발효를 통해서 세 포 내 높은 수준의 NADH 농도 기대할 수 있어 대사회로에 요구되는 다량의 NADH를 충족시킬 수 있는 장점이 있다. 또한 대사회로에 사용되는 효소들 (Clostridium 유래의 alcohol dehydrogenase)이 산소에 민감한 특징을 지니기 때문에 혐기조건 발효가 진행되어 왔다. 하지만 혐기발효에서는 TCA 회로 및 산화적 인산화가 작동되지 않기 때문에 세 포 내 낮은 에너지 수준이 문제가 될 수 있으며, 대사회로에 관련된 여러 단백질을 발현 시 미생물 성장에 저해를 가져오게 된다. 이러한 이유로 대부분의 부탄올 생산균주는 효모추출물이나 트립톤과 같은 단백질 수화물이 포함된 고 영양배지 조건에서 진행되어 왔다. 이와 같이 복합배지를 사용한 바이오 부탄올 생산 공정은 원료로 사용되는 배지의 비용의 증가를 가져올 뿐만 아니라 배양 후 불순물 정제를 하는 단계에 있어서 경제적으로 불리한 요소로 작용되어 왔다.
- [0008] 이에, 본 발명자들은 호기발효를 통하여 저가의 배지에서 부탄올을 제조하는 방법을 개발하고자 예의 노력한 결과, 혐기조건에서 글루코오스로부터 부탄올 생성능을 가지는 변이 미생물에서 알코올 디하이드로게네이즈를 코딩하는 유전자와 프로피온알데하이드 디하이드로게네이즈를 코딩하는 유전자를 호기성 미생물 유래 유전자로 치환하고, 시트레이트 신타아제 유전자의 발현이 하향조절시키는 경우, 호기조건에서도 높은 수율로 부탄올을 생

성시킬 수 있다는 것을 확인하고 본 발명을 완성하게 되었다.

발명의 내용

해결하려는 과제

- [0009] 본 발명의 목적은 호기조건에서 부탄올 생성능을 가지는 변이 미생물을 제공하는데 있다.
- [0010] 본 발명이 다른 목적은 상기 변이 미생물을 이용한 부탄올의 제조방법을 제공하는데 있다.

과제의 해결 수단

- [0011] 상기 목적을 달성하기 위하여, 본 발명은 글루코오스로부터 아세틸-CoA를 생성하는 경로를 가지는 미생물에서 다음 유전자가 도입되어 있는 호기조건에서 부탄올 생성능을 가지는 변이 미생물을 제공한다:
- [0012] (a) 호기성 균주 유래 알코올 디하이드로게네이즈를 코딩하는 유전자;
- [0013] (b) 호기성 균주 유래의 프로피온알데하이드 디하이드로게네이즈를 코딩하는 유전자;
- [0014] (c) 부틸-CoA 디하이드로게네이즈를 코딩하는 유전자;
- [0015] (d) 크로토네이즈(crotonase)를 코딩하는 유전자;
- [0016] (e) 하이드록시부틸-CoA 디하이드로게네이즈를 코딩하는 유전자; 및
- [0017] (f) 티올레이즈(thiolase)를 코딩하는 유전자.
- [0018] 본 발명은 또한, 글루코오스로부터 아세틸-CoA를 생성하는 경로를 가지는 미생물에서 다음 유전자가 도입되어 있고, 시트레이트 신타아제(citrate synthase)를 코딩하는 유전자의 발현이 하향조절되어 있는 호기조건에서 부탄올 생성능을 가지는 변이 미생물을 제공한다:
- [0019] (a) 호기성 균주 유래 알코올 디하이드로게네이즈를 코딩하는 유전자;
- [0020] (b) 호기성 균주 유래의 프로피온알데하이드 디하이드로게네이즈를 코딩하는 유전자;
- [0021] (c) 부틸-CoA 디하이드로게네이즈를 코딩하는 유전자;
- [0022] (d) 크로토네이즈(crotonase)를 코딩하는 유전자;
- [0023] (e) 하이드록시부틸-CoA 디하이드로게네이즈를 코딩하는 유전자; 및
- [0024] (f) 티올레이즈(thiolase)를 코딩하는 유전자.
- [0025] 본 발명은 또한, (a) 상기 변이 미생물을 배양하여 부탄올을 생성시키는 단계; 및 (b) 생성된 부탄올을 수득하는 단계를 포함하는 부탄올의 제조방법을 제공한다.

발명의 효과

- [0026] 본 발명에 따르면, 저가의 배양매지를 이용할 수 있는 호기발효를 통하여 부탄올을 고 수율로 제조할 수 있어, 원료비용 절감효과를 얻을 수 있으며, 종래 필수적이었던 혐기조건에 구애받지 않아 부탄올 제조공정의 다양성을 꾀할 수 있다.

도면의 간단한 설명

- [0027] 도 1은 EMJ40 균주의 부탄올 대사회로를 나타낸 것이다.
- 도 2는 EMJ50 균주의 부탄올 대사회를 나타낸 것이다.
- 도 3의 a는 호기와 혐기조건에서 EMJ40의 성장과 부탄올 생산성을 확인한 결과를 나타낸 것이고, b는 EMJ40 균주와 EMJ50의 부탄올 생산과 수율 비교한 결과를 나타낸 것이다.
- 도 4는 인공의 염기서열로 치환된 citrate synthase의 5'UTR 과 예상되는 발현정도를 나타낸 것이다.
- 도 5의 a는 EMJ50~EMJ55의 성장곡선을 비교한 결과를 나타낸 것이고, b는 EMJ50~EMJ55의 Citrate synthase의 활성정도를 비교한 결과를 나타낸 것이고, c는 EMJ50~EMJ55의 부탄올 생산성과 수율을 비교한 결과를 나타낸 것이다.

다.

발명을 실시하기 위한 구체적인 내용

- [0028] 종래에는 혐기발효를 통해 부탄올을 생산하였으며, 혐기조건에서의 에너지 부족 현상을 극복하기 위해 고영양 배지를 사용하여 배양을 진행했다. 본 발명에서는 최소 영양배지에서 부탄올을 생산하였으며, 호기조건에서 실험을 진행하여 부족한 에너지를 산화적 인산화를 통해 충족시켰다. 또한 기존의 부탄올 대사회로에 관여하는 효소들 중 호기성 조건에서 활성이 낮아지는 *Clostridium* 유래의 알코올 디하이드로게네이즈(adhE2)를 *Salmonella* 유래의 CoA 의존성 프로피온알데하이드 디하이드로게네이즈(CoA dependent propionaldehyde dehydrogenase, pduP)와 *Lactococcus* 유래의 알코올 디하이드로게네이즈(adhA)로 대체해 주어 호기성 조건에서도 활성을 지니는 대사회로를 구축하였다. 더불어 호기성 조건에서 부탄올의 전구체인 acetyl-CoA가 TCA 회로로 과량 유입되는 문제를 해결하기 위해 TCA 회로에 관여하는 첫 번째 효소인 citrate synthase(gltA)의 발현을 조절하였다. CRISPR/Cas9 시스템을 활용한 지놈 편집기술을 통해 대장균 지놈상에 존재하는 gltA의 5' untranslated region (UTR) 유전적 염기서열의 변형하였다. 다양한 인공의 UTR 서열을 합성하여 gltA의 5'UTR 부분을 조작하였고, 생산 균주의 성장과 목표물질인 부탄올의 생산이 최적으로 발생하는 균주를 개발하였다.
- [0029] 따라서, 본 발명은 일 관점에서, 글루코오스로부터 아세틸-CoA를 생성하는 경로를 가지는 미생물에서 다음 유전자가 도입되어 있는 호기조건에서 부탄올 생성능을 가지는 변이 미생물에 관한 것이다:
- [0030] (a) 호기성 균주 유래 알코올 디하이드로게네이즈를 코딩하는 유전자(adhA);
- [0031] (b) 호기성 균주 유래의 프로피온알데하이드 디하이드로게네이즈를 코딩하는 유전자(pduP);
- [0032] (c) 부틸-CoA 디하이드로게네이즈를 코딩하는 유전자(ter);
- [0033] (d) 크로토네이즈(crotonase)를 코딩하는 유전자(crt);
- [0034] (e) 하이드록시부틸-CoA 디하이드로게네이즈를 코딩하는 유전자(hbd); 및
- [0035] (f) 티올레이즈(thiolase)를 코딩하는 유전자(atoB).
- [0036] 본 발명에서는 고 영양배지가 아닌 조건에서도 세포 내 높은 수준의 에너지 레벨을 유지하고자 호기 조건에서의 부탄올 생산을 시도했다. 호기조건에서는 해당과정과 TCA 회로로부터 생성된 NADH가 ATP로 전환되는 산화적 인산화가 진행되기 때문에 세포 내 높은 ATP를 형성할 수 있다. 하지만 이를 구현하기 위해서는 앞서 언급한 바와 같이 호기조건에서 발생할 수 있는 효소의 비활성화 문제를 해결해야 한다. 이를 위해 산소에 민감한 alcohol dehydrogenase를 대체 할 수 있는 효소를 도입하였다. 또한 호기조건에서 부탄올의 전구체인 acetyl-CoA가 TCA 회로로 과량 유입되어 부탄올 생산성이 저해될 수 있기 때문에 세포의 성장과 부탄올 생산성이 모두 최적이 되는 TCA 회로 작동이 필수적이다. TCA 회로의 첫 번째 스텝에 관여하는 효소의 발현량을 지놈 편집도구를 활용하여 정교하게 조절함으로써 최적의 TCA 회로 속도를 결정할 수 있다.
- [0037] 본 발명의 변이미생물은 숙신산, 락테이트, 에탄올 및 아세테이트로 구성된 군에서 선택되는 부산물의 생합성에 관여하는 효소를 코딩하는 유전자가 결실되어 있는 것을 특징으로 할 수 있으며, 바람직하게는 푸마레이트 리덕테이즈, 락테이트 디하이드로 게네이즈, 포스페이트 아세틸 트랜스퍼레이즈 및 아세트알데하이드/알코올 디하이드로게네이즈로 구성된 군에서 선택되는 효소를 코딩하는 유전자가 결실되어 있는 것을 특징으로 할 수 있다.
- [0038] 본 발명에 있어서, 상기 알코올 디하이드로게네이즈를 코딩하는 유전자는 살모넬라 유래인 것을 특징으로 할 수 있으며, 상기 프로피온알데하이드 디하이드로게네이즈를 코딩하는 유전자는 락토코커스 유래인 것을 특징으로 할 수 있다.
- [0039] 본 발명에 있어서, NADH 공급을 위하여, 폴메이트 디하이드로게네이즈(formate dehydrogenase)를 코딩하는 유전자(fdh)가 추가로 도입되어 있는 것을 특징으로 할 수 있다.
- [0040] 본 발명의 일 양태에서, EMJ40 균주는 포도당으로부터 부탄올을 생산을 생산하기 위한 대사회로가 도입된 대장균으로서, 부산물인 숙신산, 젖산, 에탄올, 아세트산의 생산 유전자가 제거된 돌연변이 균주에 atoB, hbd, crt, ter adhE2 및 NADH 공급을 위한 fdh가 도입된 균주이고(도 1), EMJ50 균주는 부탄올 대사회로의 마지막 두 단계를 *Clostridium* 유래의 adhE2가 아닌 *Salmonella* 유래의 pduP, *Lactococcus* 유래의 adhA로 교체한 균주이다. *Clostridium* 유래의 adhE2는 산소에 민감성을 지니는 효소이기 때문에 호기 발효를 위해 pduP와 adhA로 교체하였다 (도 2).

- [0041] EMJ40을 혐기조건과 호기조건에서 배양한 결과를 보면, 혐기조건에서의 부탄올 생산성이 호기조건에서보다 3.6배 높은 것을 확인할 수 있다. 하지만 호기조건에서의 균주생장이 혐기조건에 비해서 46배 더 높은 것을 확인할 수 있었다 (도3a). 산소에 대한 내성을 가지는 대사회로를 도입한 EMJ50 균주를 호기조건에서 배양했을 시 호기조건에서 배양한 EMJ40에 비해 59배 많은 양의 부탄올이 생산된 것을 확인할 수 있었다. 이로서 EMJ50에 도입한 대사회로가 산소에 대한 내성을 가지는 것을 추론할 수 있다. 하지만 부탄올의 생산성이 증대 되었음에도 불구하고 여전히 호기조건에서의 부탄올 수율은 혐기조건의 수율에 비해 여전히 낮다는 것을 확인할 수 있었다(도 3b). 이는 호기조건에서 TCA 회로가 활성화 된다는 생물학적 이론에 미루어 볼 때 부탄올의 전구체인 acetyl-CoA가 부탄올 대사회로 보다 TCA회로로 다량 유입된다는 사실을 유추해 볼 수 있다.
- [0042] 따라서 TCA 회로의 첫 번째 단계인 citrate synthase (gltA)의 발현량을 조절하여 TCA 회로로 유입되는 acetyl-CoA 양을 줄이면서 부탄올 대사회로의 대사체 흐름을 증가시키고자 하였다.
- [0043] 다른 관점에서, 본 발명은 본 발명은 또한, 글루코오스로부터 아세틸-CoA를 생성하는 경로를 가지는 미생물에서 다음 유전자가 도입되어 있고, 시트레이트 신타아제(citrate synthase)를 코딩하는 유전자의 발현이 하향조절되어 있는 호기조건에서 부탄올 생성능을 가지는 변이 미생물에 관한 것이다:
- [0044] (a) 호기성 균주 유래 알코올 디하이드로게네이즈를 코딩하는 유전자;
- [0045] (b) 호기성 균주 유래의 프로피온알데하이드 디하이드로게네이즈를 코딩하는 유전자;
- [0046] (c) 부틸-CoA 디하이드로게네이즈를 코딩하는 유전자;
- [0047] (d) 크로토네이즈(crotonase)를 코딩하는 유전자;
- [0048] (e) 하이드록시부틸-CoA 디하이드로게네이즈를 코딩하는 유전자; 및
- [0049] (f) 티올레이즈(thiolase)를 코딩하는 유전자.
- [0050] 본 발명에 있어서, 상기 시트레이트 신타아제(citrate synthase)를 코딩하는 유전자의 발현의 하향조절은 CRISPR/Cas9 시스템에 의하여 수행되는 것을 특징으로 할 수 있다. 상기 시트레이트 신타아제(citrate synthase)를 코딩하는 유전자의 5'-UTR 부분이 서열번호 1~4 중 어느 하나로 표시되는 올리고뉴클레오타이드로 치환된 것을 특징으로 할 수 있다.
- [0051] 박테리아 내의 단백질 발현을 조절하는 방법 중 하나로 해당 유전자의 5'UTR 염기서열을 조작하는 방법이 알려져 있다. gltA의 5'UTR의 발현세기를 “UTR designer” 라는 툴을 활용해 예측하였을 때 그 값이 273860이라는 것을 확인했고, 이보다 낮은 세기의 발현을 위한 인공의 UTR 염기서열을 제작하여 균주내로 도입하고자 하였다. CRISPR/Cas9 시스템을 활용한 게놈편집 기술을 통해 EMJ50 균주의 gltA 5'UTR 염기서열에 4가지 인공의 UTR 염기서열을 치환 시켜 주었고, 발현양에 따라 EMJ51,52,53,54라고 명명했다. 또한 gltA가 제거된 돌연변이를 제작하여 EMJ55라고 명명했다 (도 4)
- [0052] 다른 관점에서, 본 발명은 (a) 상기 변이 미생물을 배양하여 부탄올을 생성시키는 단계; 및 (b) 생성된 부탄올을 수득하는 단계를 포함하는 부탄올의 제조방법에 관한 것이다.
- [0053] 본 발명의 일 양태에서 EMJ50-EMJ55 까지 6개의 균주를 호기조건에서 배양하고 균주들의 생장곡선, citrate synthase 발현 양, 부탄올 생산성 및 수율을 측정하였을 때, EMJ50의 citrate synthase 발현양을 기준으로 했을 때 EMJ51-EMJ55에서 각각 74%, 54%, 45%, 33%, 17%의 citrate synthase 활성도가 관찰되었다 (도 5b). EMJ50-EMJ55 생장패선에 대해서 EMJ50-EMJ53까지는 배양 시 생장이 저해되는 것을 발견하지 못했지만 EMJ54의 경우 EMJ50에 비해 약 33% 감소된 생장을 확인 할 수 있었으며, EMJ55에서는 생장이 89% 감소한 것을 확인하였다(도 5a). EMJ50-55의 부탄올 생산성에 대해서는 아치형 그래프를 얻을 수 있었으며, EMJ52에서 가장 높은 부탄올 생산성을 확인했다. 또한 부탄올 수율에 대해서는 citrate synthase가 낮 다운 됨에 따라서 증가하는 현상을 관찰할 수 있었다 (도 5c).
- [0054] 본 발명에서, 용어 “벡터 (vector)” 는 적합한 숙주 내에서 DNA를 발현시킬 수 있는 적합한 조절 서열에 작동 가능하게 연결된 DNA 서열을 함유하는 DNA 제조물을 의미한다. 벡터는 플라스미드, 파지 입자, 또는 간단하게 잠재적 게놈 삽입물일 수 있다. 적당한 숙주로 형질전환되면, 벡터는 숙주 게놈과 무관하게 복제하고 기능할 수 있거나, 또는 일부 경우에 게놈 그 자체에 통합될 수 있다. 플라스미드가 현재 벡터의 가장 통상적으로 사용되는 형태이므로, 본 발명의 명세서에서 “플라스미드 (plasmid)” 및 “벡터 (vector)” 는 때로 상호 교환적으로 사용된다. 그러나, 본 발명은 당업계에 알려진 또는 알려지게 되는 바와 동등한 기능을 갖는 벡터의 다른

형태를 포함한다. 포유동물 세포 배양물 발현을 위한 전형적인 발현 벡터는 예를 들면 pRK5 (EP 307,247호), pSV16B (WO 91/08291호) 및 pVL1392 (Pharmingen)을 기초로 한다.

[0055] “발현 조절 서열 (expression control sequence)”이라는 표현은 특정한 숙주 생물에서 작동가능하게 연결된 코딩 서열의 발현에 필수적인 DNA 서열을 의미한다. 그러한 조절 서열은 전사를 실시하기 위한 프로모터, 그러한 전사를 조절하기 위한 임의의 오퍼레이터 서열, 적합한 mRNA 리보솜 결합 부위를 코딩하는 서열 및 전사 및 해독의 종결을 조절하는 서열을 포함한다. 예를 들면, 원핵생물에 적합한 조절 서열은 프로모터, 임의로 오퍼레이터 서열 및 리보솜 결합 부위를 포함한다. 진핵세포는 프로모터, 폴리아데닐화 시그널 및 인헨서가 이에 포함된다. 플라스미드에서 유전자의 발현 양에 가장 영향을 미치는 인자는 프로모터이다. 고 발현용의 프로모터로서 SR α 프로모터와 사이토메가로바이러스 (cytomegalovirus) 유래 프로모터 등이 바람직하게 사용된다.

[0056] 본 발명의 DNA 서열을 발현시키기 위하여, 매우 다양한 발현 조절 서열중 어느 것이라도 벡터에 사용될 수 있다. 유용한 발현 조절서열의 예에는, 예를 들어, SV40 또는 아데노바이러스의 초기 및 후기 프로모터들, lac 시스템, trp 시스템, TAC 또는 TRC 시스템, T3 및 T7 프로모터들, 파지 람다의 주요 오퍼레이터 및 프로모터 영역, fd 코드 단백질의 조절 영역, 3-포스포글리세레이트 키나제 또는 다른 글리콜분해 효소에 대한 프로모터, 상기 포스포타제의 프로모터들, 예를 들어 Pho5, 효모 알파-교배 시스템의 프로모터 및 원핵세포 또는 진핵 세포 또는 이들의 바이러스의 유전자의 발현을 조절하는 것으로 알려진 구성과 유도의 기타 다른 서열 및 이들의 여러 조합이 포함된다. T7 RNA 폴리메라아제 프로모터 Φ 10은 이. 콜라이에서 단백질 NSP를 발현시키는데 유용하게 사용될 수 있다.

[0057] 핵산은 다른 핵산 서열과 기능적 관계로 배치될 때 “작동가능하게 연결 (operably linked)” 된다. 이것은 적절한 분자 (예를 들면, 전사 활성화 단백질)은 조절 서열(들)에 결합될 때 유전자 발현을 가능하게 하는 방식으로 연결된 유전자 및 조절 서열(들)일 수 있다. 예를 들면, 전서열(pre-sequence) 또는 분비 리더 (leader)에 대한 DNA는 폴리펩타이드의 분비에 참여하는 전단백질로서 발현되는 경우 폴리펩타이드에 대한 DNA에 작동가능하게 연결되고; 프로모터 또는 인헨서는 서열의 전사에 영향을 끼치는 경우 코딩서열에 작동가능하게 연결되거나; 또는 리보솜 결합 부위는 서열의 전사에 영향을 끼치는 경우 코딩 서열에 작동가능하게 연결되거나; 또는 리보솜 결합 부위는 번역을 용이하게 하도록 배치되는 경우 코딩 서열에 작동가능하게 연결된다. 일반적으로, “작동가능하게 연결된”은 연결된 DNA 서열이 접촉하고, 또한 분비 리더의 경우 접촉하고 리딩 프레임 내에 존재하는것을 의미한다. 그러나, 인헨서 (enhancer)는 접촉할 필요가 없다. 이들 서열의 연결은 편리한 제한 효소 부위에서 라이게이션(연결)에 의해 수행된다. 그러한 부위가 존재하지 않는 경우, 통상의 방법에 따른 합성 올리고뉴클레오타이드 어댑터 (oligonucleotide adaptor) 또는 링커(linker)를 사용한다.

[0058] 본원 명세서에 사용된 용어 “발현 벡터”는 통상 이종의 DNA의 단편이 삽입된 재조합 캐리어 (recombinant carrier)로서 일반적으로 이중 가닥의 DNA의 단편을 의미한다. 여기서, 이중 DNA는 숙주 세포에서 천연적으로 발견되지 않는 DNA인 이형 DNA를 의미한다. 발현 벡터는 일단 숙주 세포내에 있으면 숙주 염색체 DNA와 무관하게 복제할 수 있으며 벡터의 수 개의 카피 및 그의 삽입된 (이종) DNA가 생성될 수 있다.

[0059] 당업계에 주지된 바와 같이, 숙주세포에서 형질감염 유전자의 발현 수준을 높이기 위해서는, 해당 유전자가, 선택된 발현 숙주 내에서 기능을 발휘하는 전사 및 해독 발현 조절 서열에 작동가능하도록 연결되어야만 한다. 바람직하게는 발현 조절서열 및 해당 유전자는 세균 선택 마커 및 복제 개시점 (replication origin)을 같이 포함하고 있는 하나의 발현 벡터 내에 포함되게 된다. 발현 숙주가 진핵세포인 경우에는, 발현 벡터는 진핵 발현 숙주 내에서 유용한 발현 마커를 더 포함하여야만 한다.

[0060] 본 발명에서 도입되는 효소의 유전자를 발현시키기 위해 매우 다양한 발현 숙주/벡터 조합이 이용될 수 있다. 진핵 숙주에 적합한 발현 벡터에는, 예를 들어 SV40, 소 유두종바이러스, 아데노바이러스, 아데노-연관 바이러스(Adeno-associated virus), 시토메갈로바이러스 및 레트로바이러스로부터 유래된 발현 조절 서열을 포함한다. 세균 숙주에 사용할 수 있는 발현 벡터에는 pBluescript, pGEX2T, pUC벡터, col E1, pCR1, pBR322, pMB9 및 이들의 유도체와 같이 E. coli에서 얻는 것을 예시할 수 있는 세균성 플라스미드, RP4와 같이 보다 넓은 숙주 범위를 갖는 플라스미드, λ gt10과 λ gt11, NM989와 같은 매우 다양한 파지 람다(phage lambda) 유도체로 예시될 수 있는 파지 DNA, 및 M13과 필라멘트성 단일가닥의 DNA 파지와 같은 기타 다른 DNA 파지가 포함된다. 효모 세포에 유용한 발현 벡터는 2 μ 플라스미드 및 그의 유도체이다. 곤충 세포에 유용한 벡터는 pVL 941이다.

[0061] 상술한 발현 벡터에 의해 형질전환 또는 형질감염된 숙주 세포는 본 발명의 또 다른 측면을 구성한다. 본원 명세서에 사용된 용어 “형질전환”은 DNA를 숙주로 도입하여 DNA가 염색체의 인자로서 또는 염색체 통합완성에 의해 복제가능하게 되는 것을 의미한다. 본원 명세서에 사용된 용어 “형질감염”은 임의의 코딩 서열이 실제

로 발현되든 아니든 발현 벡터가 숙주 세포에 의해 수용되는 것을 의미한다.

[0062] 발명의 숙주 세포는 원핵 또는 진핵생물 세포일 수 있다. 또한, DNA의 도입효율이 높고, 도입된 DNA의 발현효율이 높은 숙주가 통상 사용된다. 이. 콜라이, 슈도모나스, 바실러스, 스트렙토마이세스, 진균, 효모와 같은 주지의 진핵 및 원핵 숙주들, 스포도프테라 프루기페르다(SF9)와 같은 곤충 세포, CHO 및 생쥐 세포같은 동물 세포, COS 1, COS 7, BSC 1, BSC 40 및 BMT 10과 같은 아프리카 그린 원숭이 세포, 및 조직배양된 인간 세포는 사용될 수 있는 숙주 세포의 예이다. 본 발명의 NSP 단백질을 코딩하는 cDNA를 클로닝할 때에는 동물세포를 숙주로 하는 것이 바람직하다. 본 발명에서는 어류 기원의 CHSE-214, FHM, RTG-2 및 EPC를 예시하였으나 물론 이에 제한되는 것은 아니다. COS 세포를 이용하는 경우에는 COS 세포에서 SV40 라지 T안티겐(large T antigen)이 발현하고 있으므로 SV40의 복제개시점을 갖는 플라스미드는 세포중에서 다수 카피(copy)의 에피솜(episome)으로 존재하도록 되고 통상보다 고 발현이 기대될 수 있다. 도입된 DNA 서열은 숙주 세포와 동일한 종으로부터 얻을 수 있거나, 숙주 세포와 다른 종의 것일 수 있거나, 또는 그것은 어떠한 이중 또는 상동성 DNA를 포함하는 하이브리드 DNA 서열일 수 있다.

[0063] 물론 모든 벡터와 발현 조절 서열이 본 발명의 DNA 서열을 발현하는데 모두 동등하게 기능을 발휘하지는 않는다는 것을 이해하여야만 한다. 마찬가지로 모든 숙주가 동일한 발현 시스템에 대해 동일하게 기능을 발휘하지는 않는다. 그러나, 당업자라면 과도한 실험적 부담없이 본 발명의 범위를 벗어나지 않는 채로 여러 벡터, 발현 조절 서열 및 숙주 중에서 적절한 선택을 할 수 있다. 예를 들어, 벡터를 선택함에 있어서는 숙주를 고려하여야 하는데, 이는 벡터가 그 안에서 복제되어야만 하기 때문이다. 벡터의 복제 수, 복제 수를 조절할 수 있는 능력 및 당해 벡터에 의해 코딩되는 다른 단백질, 예를 들어 항생제 마커의 발현도 또한 고려되어야만 한다. 발현 조절 서열을 선정함에 있어서도, 여러 가지 인자들을 고려하여야만 한다. 예를 들어, 서열의 상대적 강도, 조절가능성 및 본 발명의 DNA 서열과의 상용성 등, 특히 가능성있는 이차 구조와 관련하여 고려하여야 한다. 단 세포 숙주는 선정된 벡터, 본 발명의 DNA 서열에 의해 코딩되는 산물의 독성, 분비 특성, 단백질을 정확하게 폴딩시킬 수 있는 능력, 배양 및 발효 조건들, 본 발명 DNA 서열에 의해 코딩되는 산물을 숙주로부터 정제하는 것의 용이성 등의 인자를 고려하여 선정되어야만 한다. 이들 변수의 범위내에서, 당업자는 본 발명의 DNA 서열을 발효 또는 대규모 동물 배양에서 발현시킬 수 있는 각종 벡터/발현 조절 서열/숙주 조합을 선정할 수 있다. 발현 클로닝에 의해 NSP 단백질의 cDNA를 클로닝 하려고 할 때의 스크리닝법으로서 바인딩법(binding법), 페닝법(panning법), 필름에멀션법(film emulsion 법)등이 적용될 수 있다.

[0064] 이하, 실시예를 통하여 본 발명을 더욱 상세히 설명하고자 한다. 이들 실시예는 오로지 본 발명을 예시하기 위한 것으로, 본 발명의 범위가 이들 실시예에 의해 제한되는 것으로 해석되지 않는 것은 당업계에서 통상의 지식을 가진 자에게 있어서 자명할 것이다.

[0065] 실시예 1: 글루코오스로부터 부탄올 생성능을 가지는 균주의 제작

[0066] 글루코오스로부터 부탄올을 생성할 때에 부산물인 숙신산, 젖산 에탄올, 아세트산의 생산을 억제하기 위하여, 대장균 MG1655(DE3)에서 frd 유전자, ldhA 유전자, pta 유전자 및 adhE 유전자가 결실된 대장균 DSM101 균주(Baek, J.M *et al.*, *Biotechnol. Bioeng.*110:2790-2794, 2013)에서 atoB 유전자, hbd 유전자 crt 유전자, ter 유전자, adhE2 유전자 및 NADH 공급을 위한 fdh 유전자를 도입하여, EMJ 40균주를 제작하였다.

[0067] 표 1에는 본 발명에서 제작된 균주와 플라스미드의 특징을 기재하였다.

[0068] EMJ40 균주는 도 1에 나타난 바와 같이, 글루코오스로부터 부탄올을 생성하는 생합성 경로가 도입된 균주이다.

[0069] 상기 EMJ40 균주에서 도입된 외래 유전자는 atoB 유전자는 대장균 유래 유전자를 도입하였고, adhE2, crt 및 hbd는 클로스트리디움 아세토부틸리쿰(*Clostridium acetobutylicum*) 유래 유전자를 도입하였으며, ter은 트레포네마 덴티콜라(*Treponema denticola*) 유래 유전자를 도입하였고, fdh1은 캔디다 보이디니(*Candida boidinii*) 유래 유전자를 도입하였다.

[0070] atoB(서열번호 39), adhE2 및 fdh1(서열번호 38)는 각각의 균주의 게놈을 해당 프라이머(서열번호 6~11)를 이용하여 PCR로 증폭한 단편을 pACYCDuet(Novagen, USA)에 도입하여, pACYC-AEF를 제작한 후 도입하였다.

[0071] 대사경로의 효율을 높이기 위한 스캐폴드 단백질(GBD1SH31PDZ2)을 코딩하는 유전자는 pJD758에 삽입하여 도입하였다(Dueber, J. E. *et al.*, *Nat. Biotechnol.*27:753, 2009).

[0072] ter(서열번호 44), crt(서열번호 43) 및 hbd(서열번호 42)는 pCDF-HCT의 형태로 도입하였다(Baek, J.M *et al.*, *Biotechnol. Bioeng.*110:2790-2794, 2013).

표 1

명칭	특정	입수처
Strains		
DSM01	MG1655(DE3) Δ frdA::FRT Δ pta::FRT Δ ldhA::FRT Δ adhE::FRT	Baek <i>et al.</i> (2013)
EMJ30	DSM01 Δ gltA	본 발명
EMJ31	DSM01, but GltA activity adjusted to 75%	본 발명
EMJ32	DSM01, but GltA activity adjusted to 55%	본 발명
EMJ33	DSM01, but GltA activity adjusted to 46%	본 발명
EMJ34	DSM01, but GltA activity adjusted to 34%	본 발명
EMJ40	DSM01/pJD758, pCDF-HCT, pACYC-AEF	본 발명
EMJ50	DSM01/pJD758, pCDF-HCT, pACYC-APAF	본 발명
EMJ51	EMJ31/pJD758, pCDF-HCT, pACYC-APAF	본 발명
EMJ52	EMJ32/pJD758, pCDF-HCT, pACYC-APAF	본 발명
EMJ53	EMJ33/pJD758, pCDF-HCT, pACYC-APAF	본 발명
EMJ54	EMJ34/pJD758, pCDF-HCT, pACYC-APAF	본 발명
EMJ55	EMJ30/pJD758, pCDF-HCT, pACYC-APAF	본 발명
Plasmids		
pZA31MCS	Expression vector, Cm^R , p15A ori	EXPRESSYS
pZS21MCS	Expression vector, Km^R , pSC101 ori	EXPRESSYS
pACYCDuet	Expression vector, Cm^R , p15A ori	Novagen
pKD46	Red recombinase expression vector, Amp^R	Datsenko and Wanner <i>et al.</i> (2000)
pZA-Cas9	pZA31MCS, but $P_{LtetO-1}::cas9-P_{sacB}::sacB$	본 발명
pZS-CRISPR	pZS21MCS, but TracerRNA-crRNA- $P_{sacB}::sacB$	본 발명
pZS-CRISPRgltA	pZS21MCS, but TracerRNA-crRNA(<i>gltA</i>)- $P_{sacB}::sacB$	본 발명
pJD758	ColE1 ori, Amp^R , $P_{tet}::scaffold$ protein	Dueber <i>et al.</i> (2009)
pCDF-HCT	CDF ori, Sm^R , $P_{lac}::hbdGBD_L-crtSH3_L-terPDZ_L$	Beak <i>et al.</i> (2013)
pACYC-AEF	pACYCDuet, but $P_{lac}::atoB-adhE2-fdhI$	본 발명
pACYC-APAF	pACYCDuet, but $P_{lac}::atoB-pduP-adhA-fdhI$	본 발명

[0073]

[0074]

[0075]

[0076]

[0077]

[0078]

[0079]

[0080]

[0081]

[0082]

[0083]

[0084]

[0085]

[0086]

[0087]

atoB-F tatagtcgacaaggagatataATGAAAAATTGTGTCATCGTC(서열번호 6)

atoB-R tatagcgccgcTTAATTCACCGTTCAATCA(서열번호 7)

adhE2-F ggagatatacatatggcaATGAAAGTTACAAATCAAAAAGAAC(서열번호 8)

adhE2-R atatctccttTTAAAATGATTTTATATAGATATCCTTAAG(서열번호 9)

fdh1-F aatcatttttaaaggagatataATGAAGATCGTTTTAGTCTTATATG(서열번호 10)

fdh1-R cggttttctttaccagacTTATTTCTATCGTGTTTACCG(서열번호 11)

EMJ50 균주는 호기성 조건에서 부탄올 생성을 증가시키기 위하여, 부탄올 대사회로의 마지막 두 단계를 클로스트리디움 유래의 adhE2가 아닌 살모넬라 엔테리카(Salmonella enterica)유래의 pduP, lactococcus 유래의 adhA로 교체한 균주이다. 클로스트리디움 아세토부틸리쿰 유래의 adhE2는 산소에 민감성을 지니는 효소이기 때문에 호기 발효를 위해 pduP(서열번호 41)와 adhA(서열번호 40)로 교체하였다 (도 2).

pduP, adhA, atoB 및 fdh1 유전자는 각각의 유래 균주의 게놈에서 해당 프라이머(서열번호 6~15)를 이용하여 PCR로 증폭하였으며, pACUCduet 벡터(Novagen, USA)에 도입시켜, pACYC-APAF(서열번호 35)를 제작하였으며, DSM101 균주에 상기 pACYC-APAF, pJD758(서열번호 37) 및 pCDF-HCT(서열번호 36)을 도입하여 EMJ 50을 제작하였다.

pduP-F tataggatccaaggagatataATGAATACTTCTGAACTCGAAACC(서열번호 12)

pduP-R tatagagctcTTAGCGAATAGAAAAGCCGTTG(서열번호 13)

adhA-F tataagaaggagatatacaATGAAAGCAGCAGTAGTAAGAC(서열번호 14)

adhA-R gatcttcattatctccttTTATTTAGTAAAATCAATGACCATTG(서열번호 15)

실시예 2: EMJ40 및 EMJ50 균주를 이용한 부탄올의 생산

EJM 40의 부탄올 생성능을 확인하기 위하여, 5 mL M9G 배지에서 하룻밤 전배양한 후, 250mL 플라스크에서 50mL M9G 배지로 배양하였다. 플라스크는 호기성 또는 혐기조건에서 37℃, 250 rpm 에서 0.05 mM IPTG 및 108 nM

anhydrotetracycline (aTc) 을 첨가하고 6시간 배양하였다

[0088] M9G 배지의 조성은 다음과 같다:

[0089] 12.8 g/L Na_2HPO_4 , 3 g/L KH_2PO_4 , 0.5 g/L NaCl, 1 g/L NH_4Cl , 1 mM MgSO_4 , 0.1 mM CaCl_2 , 10 mg/L thiamine, 25 g/L glucose, 0.5 g/L yeast extract, and 1000 X trace elements($27 \text{ g/L FeCl}_3 \cdot 6\text{H}_2\text{O}$, $2 \text{ g/L ZnCl}_2 \cdot 4\text{H}_2\text{O}$, $2 \text{ g/L CaCl}_2 \cdot 2\text{H}_2\text{O}$, $2 \text{ g/L Na}_2\text{MoO}_4 \cdot 2\text{H}_2\text{O}$, $1.9 \text{ g/L CuSO}_4 \cdot 5\text{H}_2\text{O}$, and $0.5 \text{ g/L H}_3\text{BO}_3$) 50 $\mu\text{g/mL}$ ampicillin, 50 $\mu\text{g/mL}$ chloramphenicol 및 100 $\mu\text{g/mL}$ spectinomycin.

[0090] 도 3의 a에 나타난 바와 같이, EMJ40 균주를 혐기조건과 호기조건에서 배양한 결과를 보면, 혐기조건에서의 부탄올 생산성이 호기조건에서보다 3.6배 높았다.

[0091] 하지만 호기조건에서의 균주생장이 혐기조건에 비해서 46배 더 높은 것을 확인할 수 있었다.

[0092] 산소에 대한 내성을 가지는 대사회로를 도입한 EMJ50 균주를 호기조건에서 배양 했을 시 호기조건에서 배양한 EMJ40에 비해 59배 많은 양의 부탄올이 생산된 것을 확인할 수 있다(도 3 b).

[0093] 이로서 EMJ50에 도입한 대사회로가 산소에 대한 내성을 가지는 것을 추론할 수 있다. 하지만 부탄올의 생산성이 증대되었음에도 불구하고 여전히 호기조건에서의 부탄올 수율은 혐기조건의 수율에 비해 여전히 낮다는 것을 확인할 수 있다 (도3b).

[0094] **실시예 3: CRISPR/Cas9 시스템을 활용 이용한 TCA 회로가 조절된 변이균주의 제작**

[0095] 호기조건에서 TCA 회로가 활성화 된다는 생물학적 이론에 미루어 볼 때 부탄올의 전구체인 acetyl-CoA가 부탄올 대사회로보다 TCA회로로 다량 유입된다는 사실을 유추해 볼 수 있다.

[0096] 따라서 TCA 회로의 첫 번째 단계인 citrate synthase (gltA)의 발현량을 조절하여 TCA 회로로 유입되는 acetyl-CoA 양을 줄여 부탄올 대사회로로의 대사체 흐름을 증가시키고자 하였다.

[0097] 박테리아 내의 단백질 발현을 조절하는 방법 중 하나로 해당 유전자의 5'UTR 염기서열을 조작하는 방법이 알려져 있다. gltA의 5'UTR의 발현세기를 "UTR designer" 라는 툴을 활용해 예측하였을 때 그 값이 273860이라는 것을 확인했고, 이보다 낮은 세기의 발현을 위한 인공의 UTR 염기서열을 제작하여 균주내로 도입하였다.

[0098] pCas9 plasmid (Addgene, Cambridge, MA,USA)의 cas9, crRNA 및 tracrRNA 부위를 프라이머들을 이용하여, PCR 로 증폭시켰으며, sacB는 *Bacillus subtilis* 게놈 DNA를 이용하여 증폭하였다. cas9 및 sacB의 PCR 산물을 Gibson assembly (New England Biolabs, USA)를 이용하여, pZA31MCS(EXPRESSYS)에 삽입하고, pZA-Cas9 로 명명 하였다(표 1).

[0099] 이와 유사하게, crRNA, tracrRNA 및 sacB를 pZS21MCS에 삽입하고, pZS-CRISPR로 명명하였다(표 1). gltA 타겟팅-crRNA sequence에 gltA를 삽입하기 위하여 gltA 5' -UTR sequence (gltA crRNA-S 및 gltA crRNA-A)의 양 가닥을 합성하였다(Bioneer Inc., 한국) 합성된 DNA를 pZS-CRISPR에 어닐링시키고, 라이게이션 한 후, BsaI로 절단하여, crRNA targeting gltA-5' -UTR을 포함하는 pZS-CRISPRgltA를 제작하였다.

[0100] Cas9 뉴클레아제에 의한 분해 후의 상동성 재조합을 위한 레스큐 DNA는

[0101] 22개의 오버랩핑 서열을 가지는 올리고머 DNA를 합성하였으며, 올리고머 쌍은 다음과 같다: gltA Rescue F 및 0.6R, gltA Rescue F 및 0.3R, gltA

[0102] Rescue F2 및 0.5R, gltA Rescue F2 및 0.4R.

[0103] 합성된 올리고머는 96℃에서 변성 후, 천천히 냉각시켜 어닐링시키고, Klenow fragment (Takara Bio,Japan)를 이용하여, 신장시켜 DNA를 합성하였다.

[0104] 숙주세포로는 DSM01균주를 사용하였으며, 먼저, 일렉트로포레이션으로 pKD46(Red recombinase expression vector, Datsenko, KA. and Wanner, BL. *Proc. Natl. Acad. Sci. U S A.* 97: 6640, 2000)과 pZA-Cas9를 차례로 DSM01에 도입시키고, 상기 균주에 gltA Rescue DNA와 pZS-CRISPRgltA를 함께 형질전환시킨 후, 카나마이신(50 $\mu\text{g/mL}$) 및 클로람페니콜(50 $\mu\text{g/mL}$)을 포함하는 LB 고체 배지에서 스크리닝하였다.

[0105] CRISPR/Cas9을 위한 플라스미드들을 포함하는 스크리닝된 변이균주는 40 g/L의 슈크로오스를 포함하는 LB 배지에 도말하여 37℃에서 하룻밤 배양하였으며, 두 가지 항생제에 대하여 내성이 있는 콜로니를 선별하였다. 선별

된 균주의 게놈이 변이는 결과는 *gltA*-5' -F 및 *gltA*-5' -R 프라이머(서열번호 32 및 32)를 이용하여 PCR을 수행한 후, PCR 산물을 CosmoGenetech (Seoul, Korea)에서 시퀀싱하여 확인하였다.

- [0106] *cas9* F aaaagtgcacATGGATAAGAAATACTCAATAGGCT(서열번호 16)
- [0107] *cas9* R aaaactgcagTCAGTCACCTCCTAGCTGAC(서열번호 17)
- [0108] *tracrRNA* F ataaaagcttTTACGAAATCATCCTGTGGAG(서열번호 18)
- [0109] *tracrRNA* R taatggatccTTTGCCTCCTAAAATAAAAAGTT(서열번호 19)
- [0110] *crRNA* F aattggtaccAGTATATTTTAGATGAAGATTATTTCTTA(서열번호 20)
- [0111] *crRNA* R attaaagcttATCACACTACTCTTCTTTGCCTA(서열번호 21)
- [0112] *gltA* *crRNA* S aaacAGGTTGATGTGCGAAGGCAAATTAAGTTCg(서열번호 22)
- [0113] *gltA* *crRNA* A aaaacGAACTTAAATTTGCCTTCGCACATCAACCT(서열번호 23)
- [0114] *sacB* F ataagcagcatgcctgtTACCTGCCGTTCACTATTATTAG(서열번호24)
- [0115] *sacB* R cacatagacagcctgaATCGGCATTTTCTTTTGC(서열번호25)
- [0116] *gltA* Rescue Uni F
- [0117] ccaaataacaaacgggtaaagccaggttgatgtgcgaaggcaaatttaagttcccgagctcttacgtgtaggttaaaggagcat(서열번호 26)
- [0118] *gltA* Rescue 0.6 R
- [0119] tcaacagctgtgtccccgttgagggtgagttttgcttttgatcagccatctctgatgctccttttaacctacagcg(서열번호 27)
- [0120] *gltA* Rescue 0.3 R
- [0121] tcaacagctgtgtccccgttgagggtgagttttgcttttgatcagccatcaacgatgctccttttaacctacagcg(서열번호 28)
- [0122] *gltA* Rescue Uni F2
- [0123] ccaaataacaaacgggtaaagccaggttgatgtgcgaaggcaaatttaagttcccgagctcttacgagggtggtgtaaaggagcat(서열번호 29)
- [0124] *gltA* Rescue 0.5 R
- [0125] tcaacagctgtgtccccgttgagggtgagttttgcttttgatcagccatggccgatgctcctttacaccagccgcg(서열번호 30)
- [0126] *gltA* Rescue 0.4 R
- [0127] tcaacagctgtgtccccgttgagggtgagttttgcttttgatcagccatctcagatgctcctttacaccagccgcg(서열번호 31)
- [0128] *gltA*-5' -F AAAGTTGTTACAAACATTACCAGGAA(서열번호 32)
- [0129] *gltA*-5' -R TTCACCATTCAGCAGGATGTA(서열번호 33)
- [0130] *gltA*-5' -seq TACCCAGGTTTCCCTCTT(서열번호 34)
- [0131] CRISPR/Cas9 시스템을 활용한 게놈편집 기술을 통해 EMJ50 균주의 *gltA* 5'UTR 염기서열(서열번호 5)에 4가지 인공의 UTR 염기서열(서열번호 1~4)을 치환시켜 주었고, 발현양에 따라 각각 EMJ51, EMJ52, EMJ53 및 EMJ54라고 명명했다. 또한 *gltA*가 제거된 돌연변이를 제작하여 EMJ55라고 명명했다 (도 4)
- [0132] **실시예 4: 게놈편집 균주의 부탄올 생성능 확인**
- [0133] 실시예 3에서 제작된 EMJ50-55의 6개의 균주를 실시예 2와 동일한 호기조건에서 배양하고 균주들의 생장곡선, citrate synthase 발현 양, 부탄올 생산성 및 수율을 측정하였다.
- [0134] 먼저 EMJ50의 citrate synthase 발현양을 기준으로 했을 때 EMJ51, EMJ52, EMJ53, EMJ54 및 EMJ55에서 각각 74%, 54%, 45%, 33% 및 17%의 citrate synthase 활성도가 관찰되었다 (도 5b). EMJ50~EMJ55의 생장 패션을 확인한 결과, EMJ50~EMJ53까지는 배양 시 생장이 저해되는 것을 발견하지 못했지만 EMJ54의 경우 EMJ50에 비해 약 33% 생장이 감소하는 것을 확인할 수 있었으며, EMJ55에서는 생장이 89% 감소한 것을 확인하였다(도 5a). EMJ50~EMJ55의 부탄올 생산성에 대해서는 아치형 그래프를 얻을 수 있었으며, EMJ52에서 가장 높은 부탄올 생산성을 확인하였다. 또한 부탄올 수율에 대해서는 citrate synthase가 낙 다운 됨에 따라서 증가하는 현상을 관찰

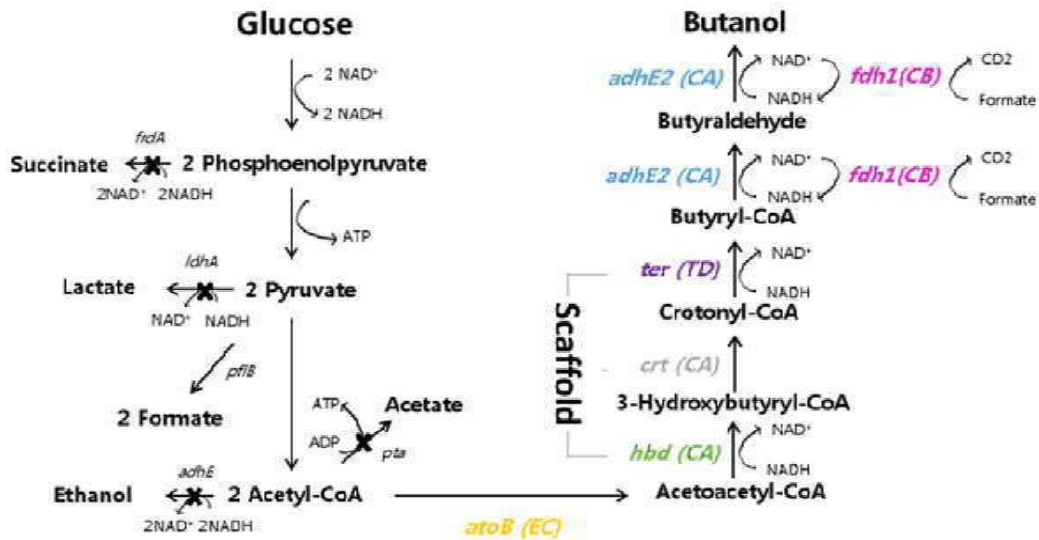
할 수 있었다 (도 5c).

[0135]

이상으로 본 발명 내용의 특정한 부분을 상세히 기술하였는바, 당업계의 통상의 지식을 가진 자에게 있어서, 이러한 구체적 기술은 단지 바람직한 실시양태일 뿐이며, 이에 의해 본 발명의 범위가 제한되는 것이 아닌 점은 명백할 것이다. 따라서 본 발명의 실질적인 범위는 첨부된 청구항들과 그것들의 등가물에 의하여 정의된다고 할 것이다.

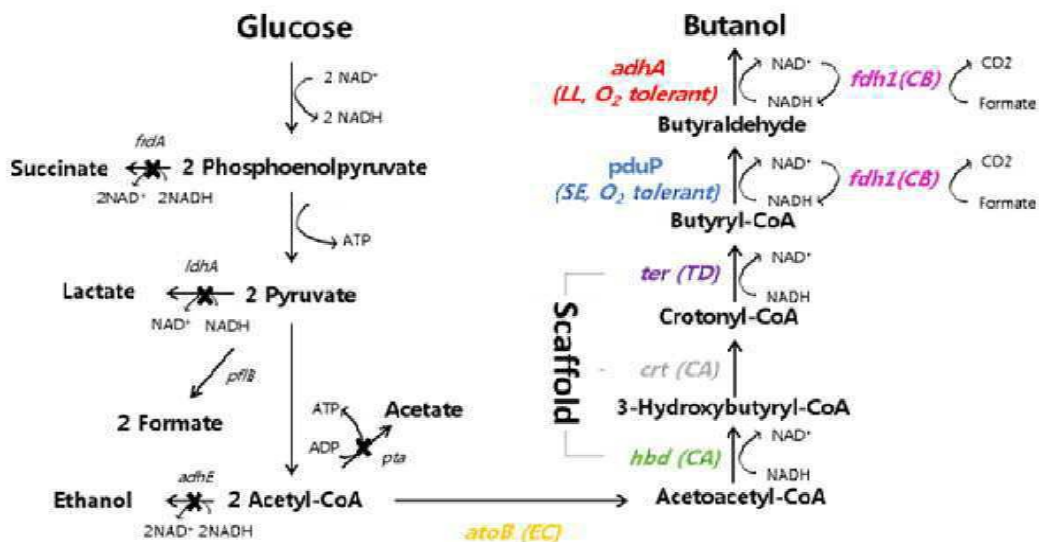
도면

도면1



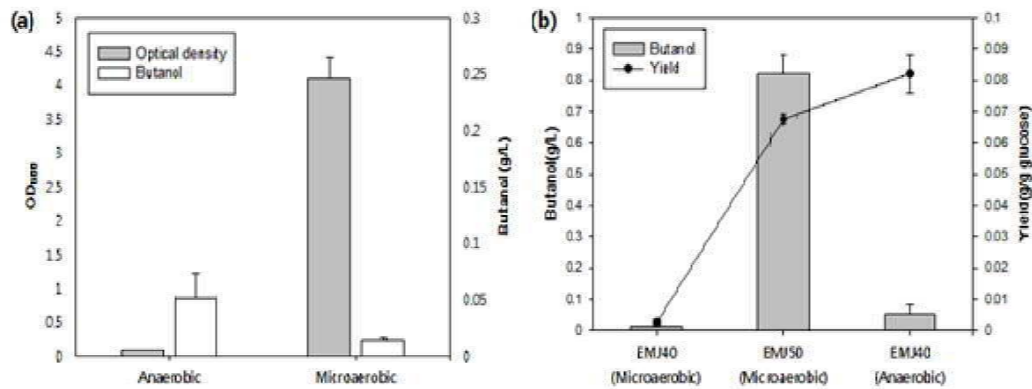
EC : *Escherichia coli* CA : *Clostridium acetobutylicum* TD : *Treponema denticola* CB : *Candida boidinii*

도면2

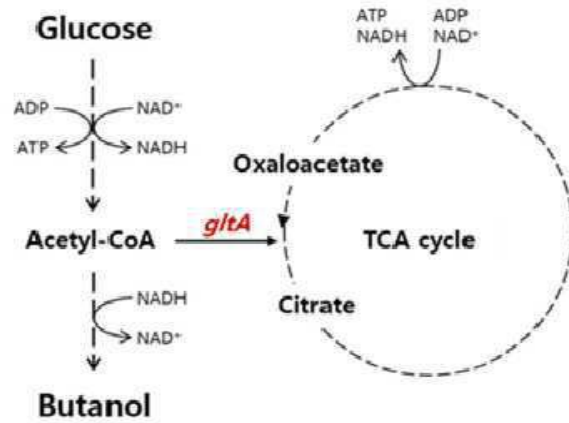


EC : *Escherichia coli* CA : *Clostridium acetobutylicum* TD : *Treponema denticola* CB : *Candida boidinii*
SE : *salmonella enterica* LL : *Lactococcus lactis*

도면3

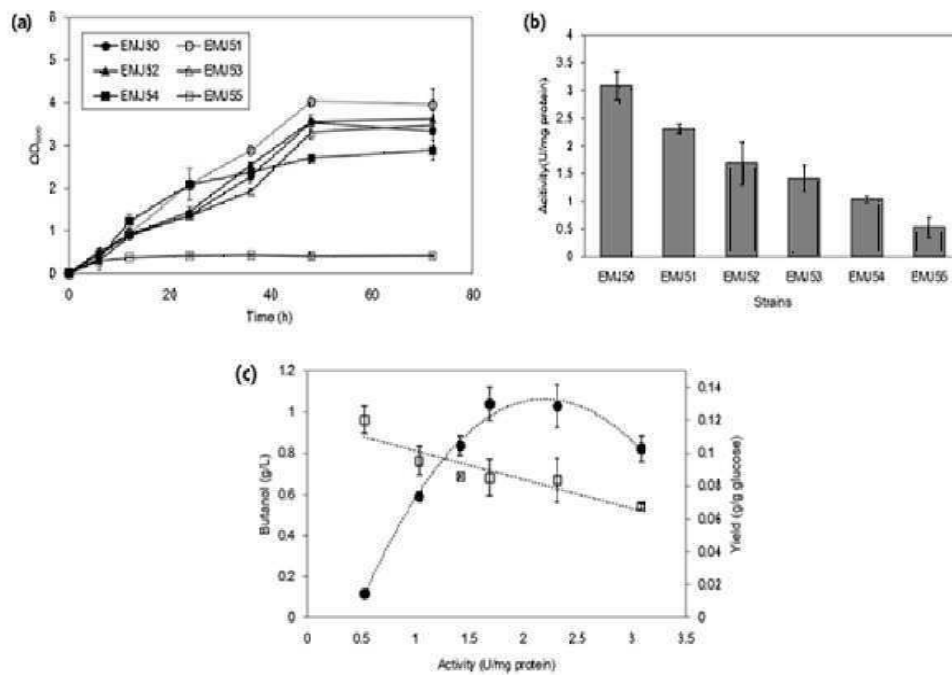


도면4



Strain	Edited <i>gltA</i> 5'-UTR sequence	Expected expression level
EMJ50	ATTAAAGTTC CGG CAGTCTTACGCAATAAGGCGCT TAAGGAGAC CTTAA ATG	273860.05
	↓ (modified by CRISPR/Cas9)	↓
EMJ51	ATTAAAGTTC CGG CAGTCTTACGCTGTAGGTTAA AAAGGAGCAT CAGAG ATG	144913.04
EMJ52	ATTAAAGTTC CGG CAGTCTTACGCTGTAGGTTAA AAAGGAGCAT CGTTG ATG	85182.74
EMJ53	ATTAAAGTTC CGG CAGTCTTACGCGGCTGGTGT AAAGGAGCAT CTGAG ATG	27064.8
EMJ54	ATTAAAGTTC CGG CAGTCTTACGCGGCTGGTGT AAAGGAGCAT CGGCC ATG	2809.68

도면5



서열 목록

- <110> Korea University Research & Business Foundation
- <120> Variant Microorganism Producing Butanol in Aerobic Condition and Method for Preparing Butanol Using the Same
- <130> P16-B294
- <160> 44
- <170> Kopatent In 2.0
- <210> 1
- <211> 48
- <212> DNA
- <213> Artificial Sequence
- <220><223> 5'UTR of EMJ51
- <400> 1
- athtaagttc cggcagtcctt acgctgtagg ttaaaaggag catcagag 48
- <210> 2
- <211> 48
- <212> DNA
- <213> Artificial Sequence
- <220><223> 5'UTR of EMJ52

<400> 2
 atttaagttc cggcagtcctt acgctgtagg ttaaaaggag catcgttg 48

<210> 3
 <211> 48
 <212> DNA
 <213> Artificial Sequence
 <220><223> 5`UTR of EMJ53

<400> 3
 atttaagttc cggcagtcctt acgcggctgg tgtaaaggag catctgag 48

<210> 4
 <211> 48
 <212> DNA
 <213> Artificial Sequence
 <220><223> 5`UTR of EMJ54

<400> 4
 atttaagttc cggcagtcctt acgcggctgg tgtaaaggag catcggcc 48

<210> 5
 <211> 47
 <212> DNA
 <213> Artificial Sequence
 <220><223> 5`UTR of EMJ50

<400> 5
 atttaagttc cggcagtcctt acgcaataag gcgctaagga gacctta 47

<210> 6
 <211> 42
 <212> DNA
 <213> Artificial Sequence
 <220><223> primer

<400> 6
 tatagtcgac aaggagatat aatgaaaaat tgtgtcatcg tc 42

<210> 7
 <211> 32
 <212> DNA
 <213> Artificial Sequence

<220><223> primer
 <400> 7
 tatagcggcc gcttaattca accgttcaat ca 32

<210> 8
 <211> 43
 <212> DNA
 <213> Artificial Sequence
 <220><223> primer
 <400> 8
 ggagatatac atatggcaat gaaagttaca aatcaaaaag aac 43
 <210> 9
 <211> 40
 <212> DNA
 <213> Artificial Sequence
 <220><223> primer
 <400> 9
 atatctcctt ttaaaatgat tttatataga tatccttaag 40
 <210> 10
 <211> 47
 <212> DNA
 <213> Artificial Sequence
 <220><223> primer
 <400> 10
 aatcatttta aaaggagata taatgaagat cgttttagtc ttatatg 47
 <210> 11
 <211> 39
 <212> DNA
 <213> Artificial Sequence
 <220><223> primer
 <400> 11
 cggtttcttt accagactta tttcttatcg tgtttaccg 39
 <210> 12
 <211> 45

<212> DNA
 <213> Artificial Sequence
 <220><223> primer
 <400> 12
 tataggatcc aaggagatat aatgaatact tctgaactcg aaacc 45
 <210> 13
 <211> 32
 <212> DNA
 <213> Artificial Sequence
 <220><223> primer
 <400> 13
 tatagagctc ttagcgaata gaaaagccgt tg 32
 <210> 14
 <211> 41
 <212> DNA
 <213> Artificial Sequence
 <220><223> primer
 <400> 14
 tataagaagg agatatacaa tgaaagcagc agtagtaaga c 41
 <210> 15
 <211> 46
 <212> DNA
 <213> Artificial Sequence
 <220><223> primer
 <400> 15
 gatcttcatt atatctcctt ttatttagta aaatcaatga ccattc 46
 <210> 16
 <211> 35
 <212> DNA
 <213> Artificial Sequence
 <220><223> primer
 <400> 16
 aaaagtcgac atggataaga aatactcaat aggct 35

<210> 17
 <211> 30
 <212> DNA
 <213> Artificial Sequence
 <220><223> primer
 <400> 17
 aaaactgcag tcagtcacct cctagctgac 30
 <210> 18
 <211> 31
 <212> DNA
 <213> Artificial Sequence
 <220><223> primer
 <400> 18
 ataaaagctt ttacgaaatc atcctgtgga g 31
 <210> 19
 <211> 34
 <212> DNA
 <213> Artificial Sequence
 <220><223> primer
 <400> 19
 taatggatcc ttttgcctcc taaaataaaa agtt 34

 <210> 20
 <211> 39
 <212> DNA
 <213> Artificial Sequence
 <220><223> primer
 <400> 20
 aattggtacc agtatatattt agatgaagat tatttctta 39
 <210> 21
 <211> 34
 <212> DNA
 <213> Artificial Sequence
 <220><223> primer
 <400> 21

attaaagctt atcacactac tcttcttttg ccta	34
<210> 22	
<211> 35	
<212> DNA	
<213> Artificial Sequence	
<220><223> primer	
<400> 22	
aaacaggttg atgtgcgaag gcaaatttaa gttcg	35
<210> 23	
<211> 35	
<212> DNA	
<213> Artificial Sequence	
<220><223> primer	
<400> 23	
aaaacgaact taaatttgcc ttcgcacatc aacct	35
<210> 24	
<211> 42	
<212> DNA	
<213> Artificial Sequence	
<220><223> primer	
<400> 24	
ataagcagca tcgctgtta cctgccgttc actattattt ag	42
<210> 25	
<211> 34	
<212> DNA	
<213> Artificial Sequence	
<220><223> primer	
<400> 25	
cacatagaca gcctgaatcg gcattttctt ttgc	34
<210> 26	
<211> 86	
<212> DNA	
<213> Artificial Sequence	

<220><223> gltA Rescue oligomer

<400> 26

ccaaataaca aacgggtaaa gccaggttga tgtgcgaagg caaatTTaag ttcccgcagt 60

cttacgctgt aggttaaaag gagcat 86

<210> 27

<211> 77

<212> DNA

<213> Artificial Sequence

<220><223> gltA Rescue oligomer

<400> 27

tcaacagctg tgtccccgtt gagggtagt tttgctttg tatcagccat ctctgatgct 60

ccttttaacc tacagcg 77

<210> 28

<211> 77

<212> DNA

<213> Artificial Sequence

<220><223> gltA Rescue oligomer

<400> 28

tcaacagctg tgtccccgtt gagggtagt tttgctttg tatcagccat caacgatgct 60

ccttttaacc tacagcg 77

<210> 29

<211> 86

<212> DNA

<213> Artificial Sequence

<220><223> gltA Rescue oligomer

<400> 29

ccaaataaca aacgggtaaa gccaggttga tgtgcgaagg caaatTTaag ttcccgcagt 60

cttacgcggc tgggtgtaaag gagcat 86

<210> 30

<211> 77

<212> DNA

<213> Artificial Sequence

<220><223> gltA Rescue oligomer

<400> 30
tcaacagctg tgtccccgtt gagggtagt tttgctttg tatcagccat ggccgatgct 60
cctttacacc agccgcg 77
<210> 31
<211> 77
<212> DNA
<213> Artificial Sequence
<220><223> gltA Rescue oligomer
<400> 31
tcaacagctg tgtccccgtt gagggtagt tttgctttg tatcagccat ctcagatgct 60
cctttacacc agccgcg 77
<210> 32
<211> 26
<212> DNA
<213> Artificial Sequence
<220><223> primer
<400> 32
aaagttgtta caaacattac caggaa 26
<210> 33
<211> 21
<212> DNA
<213> Artificial Sequence
<220><223> primer
<400> 33
ttcaccattc agcagatgt a 21
<210> 34
<211> 20
<212> DNA
<213> Artificial Sequence
<220><223> primer
<400> 34
taccaggtt ttccctctt 20
<210> 35

<211> 8677

<212> DNA

<213> Artificial Sequence

<220><223> pACYC APAF

<400> 35

cgccagggtg gtttttcttt tcaccagtga gacgggcaac agctgattgc ctttcaccgc	60
ctggccctga gagagtigca gcaagcggtc cacgctggtt tgccccagca ggcgaaaatc	120
ctgtttgatg gtggttaacg gcgggatata acatgagctg tcttcggtat cgtcgtatcc	180
cactaccgag atgtccgcac caacgcgcag cccggactcg gtaatggcgc gcattgcgcc	240
cagcgccatc tgatcgttgg caaccagcat cgcagtggga acgatgccct cattcagcat	300
ttgcatggtt tgttgaaaac cggacatggc actccagtcg ccttcccgtt ccgctatcgg	360
ctgaatttga ttgcgagtga gatatttatg ccagccagcc agacgcagac gcgccgagac	420
agaacttaat gggcccgccta acagcgcgat ttgctggtga cccaatgcga ccagatgctc	480
cacgcccagt cgcgtaccgt cttcatggga gaaaaataata ctgttgatgg gtgtctggtc	540
agagacatca agaaataacg ccggaacatt agtgcaggca gcttcacag caatggcatc	600
ctggatcatc agcggatagt taatgatcag cccactgacg cgttgcgcga gaagattgtg	660
caccgccgtt ttacagcgtt cgacgccgtt tcgttctacc atcgacacca ccacgtggc	720
accagttga tcggcgcgag atttaatcgc cgcgacaatt tgcgacggcg cgtgcagggc	780
cagactggag gtggcaacgc caatcagcaa cgactgtttg cccgccagtt gttgtgccac	840
gcggttggga atgtaattca gctccgcat cgccgcttcc actttttccc gcgttttcgc	900
agaaacgtgg ctggcctggt tcaccacgcg ggaaacggtc tgataagaga caccggcata	960
ctctgcgaca tcgtataacg ttactggttt cacattcacc acctgaatt gactctcttc	1020
cgggcgctat catgccatc cgcgaaaggt ttgcgccat tcgatggtgt ccgggatctc	1080
gacgtcttcc cttatgcgac tcctgcatta ggaaattaat acgactcact atagggaat	1140
tgtgagcgga taacaattcc cctgtagaaa taattttgtt taactttaat aaggagatat	1200
accatgggca gcagccatca ccatcatcac cacagccagg atccaaggag atataatgaa	1260
gatcgtttta gtcttatatg atgctggtaa gcacgctgct gatgaagaaa aatttatatg	1320
ttgtactgaa aataaattag gtattgctaa ttgggttaaaa gatcaaggtc atgaactaat	1380
tactacttct gataaagaag gtgaacaag tgaattggat aaacatatcc cagatgctga	1440
tattatcatc accactcctt tccatcctgc ttatatcact aaggaaagac ttgacaaggc	1500
taagaactta aaatcagtcg ttgtcgctgg tgttggttct gatcacattg atttagatta	1560

tattaatcaa acaggtgaaga aaatctcagt cctggaagtt acaggttcta atgttgtctc	1620
tggttgctgaa cacgttgica tgaccatgct tgccttggtt agaaatttcg ttccagcaca	1680
tgaacaaatt attaaccacg attgggaggt tgcgtctatc gctaaggatg cttacgatat	1740
cgaaggtaaa actatcgcta ccattgggtgc tggtagaatt gggtacagag tcttggaag	1800
attactccca tttaatccaa aagaattatt atactacgat tatcaagctt taccaaaaga	1860
agctgaagaa aaagttggtg ctagaagagt tgaaaatatt gaagaattag ttgctcaagc	1920
tgatatcggt acagttaatg ctccattaca cgcaggtaca aaaggtttaa ttaataagga	1980
attattatct aaatttaaaa aagggtgcttg gttagtcaat accgcaagag gtgctatttg	2040
tggttgctgaa gatgttgtag cagctttaga atctgggtcaa ttaagaggtt acggtggtga	2100
tggttggttc ccacaaccag ctccaaagga tcacccatgg agagatatga gaaataaata	2160
tggtgctggt aatgccatga ctctcacta ctctggtact actttagacg ctcaaaacag	2220
atagctgtaa ggtactaaaa atattttgga atcattcttt accggtaaat ttgattacag	2280
accacaagat attatcttat taaatggtga atacgttact aaagcttacg gtaaacacga	2340
taagaaataa gagctcggcg cgcctgcagg tcgacaagga gatataatga aaaattgtgt	2400
catcgtcagt gcggtacgta ctgctatcgg tagttttaac gggtcactcg cttccaccag	2460
cgccatcgac ctgggggcga cagtaattaa agccgccatt gaacgtgcaa aaatcgattc	2520
acaacacgtt gatgaagtga ttatgggtaa cgtgttacaa gccgggctgg ggcaaaatcc	2580
ggcgcgtcag gactgttaa aaagcgggct ggcagaaacg gtgtgcggat tcacgtgcaa	2640
taaagtatgt ggttcgggtc ttaaaagtgt ggcgcttgcc gccagggcca ttcaggcagg	2700
tcaggcgcag agcattgtgg cggggggtat ggaaaatatg agtttagccc cctacttact	2760
cgatgcaaaa gcacgctctg gttatcgtct tggagacgga caggtttatg acgtaatcct	2820
gcgcgatggc ctgatgtcgg ccacccatgg ttatcatatg gggattacg ccgaaaacgt	2880
ggctaaagag tacggaatta cccgtgaaat gcaggatgaa ctggcgctac attcacagcg	2940
taaagcgga gccgcaatg agtcgggtgc ttttacagcc gaaatcgtcc cggtaaatgt	3000
tgtcactcga aagaaaacct tcgtcttcag tcaagacgaa ttcccgaag cgaattcaac	3060
ggctgaagcg ttaggtgcat tgcgcccggc cttcgataaa gcaggaacag tcaccgtg	3120
gaacgcgtct ggtattaacg acggtgctgc cgtcttggtg attatggaag aatctgcggc	3180
gctggcagca ggccttacc ccctggctcg cattaaaagt tatgccagcg gtggcgtgcc	3240
ccccgcattg atgggtatgg ggccagtacc tgccacgcaa aaagcgttac aactggcggg	3300

gctgcaactg gcggatatgt atctcattga ggctaataaa gcatttgctg cacagttcct	3360
tgccgttggg aaaaacctgg gctttgattc tgagaaagt aatgtcaacg gcggggccat	3420
cgcgctcggg catcctatcg gtgccagtgg tgcctgtatt ctggtcacac tattacatgc	3480
catgcaggca cgcgataaaa cgctggggct ggcaaacctg tgcattggcg gcggtcaggg	3540
aattgcgatg gtgattgaac ggttgaatta agcggccgca taatgcttaa gtcgaacaga	3600
aagtaatcgt attgtacacg gccgcataat cgaattaat acgactcact atagggaat	3660
tgtgagcgga taacaattcc ccatcttagt atattagtta agtataagaa ggagatatac	3720
aatgaaagca gcagtagtaa gacacaatcc agatggttat gcggaccttg ttgaaaagga	3780
acttcgagca atcaaaccta atgaagcttt gcttgacatg gagtattgtg gagtctgtca	3840
taccgatttg cacgttcgag caggtgatta tggcaacaaa gcaggactg tctttgtca	3900
tgaaggaatt ggaattgtca aagaaattgg aactgatga agctcgctc aagttgtga	3960
tcgggtttca gtggcttgg tctttgaagg atgtgtcac tgtgaatact gtgtatctgg	4020
taatgaaact tttgtcgag aagttaaaaa tgcaggatat tcagttgatg gcggaatggc	4080
tgaagaagca attgttgtg ccgattatgc tgtcaaagtt cctgacggac ttgaccaat	4140
tgaagctagc tcaattactt gtgctggagt aacaacttac aaagcaatca aagtatcagg	4200
agtaaaacct ggtgattggc aagtaatttt tgggtctgga ggacttgga atttagcaat	4260
tcaatatgct aaaaatgttt ttggagcaaa agtaattgct gttgatatta atcaagataa	4320
attaaattta gtaaaaaaa ttggagctga tgtgattatc aattctgtg atgtaaatcc	4380
agttgatgaa attaaaaaa taactggcgg cttaggagca caaagtcaa tagtttgtgc	4440
tgttgcaagg attgcttttg aacaagcgg tgcctctttg aaacctatgg gcaaaatggt	4500
tgctgtggca ctcccaata ctgagatgac ttatcagtt ccaacagttg tttttgacgg	4560
agtgagggtt gcaggttcac ttgtcggaac aagacttgac ttggcagaag cttttcaatt	4620
tggagcagag ggtaaggtaa aaccaattgt tgcgacacgc aaactggaag aaatcaatga	4680
tattattgat gaaatgaagg caggaaaaat tgaaggccga atggtcattg attttactaa	4740
ataaaaggag atataatgaa tacttctgaa ctcgaaacat taatcagaac cattcttagc	4800
gagcagttaa ccacgccggc acaaacgccg gtccagcctc aaggcaaagg gatttttcag	4860
tccgttagcg aggccatcga tggcggccac caagcgtttt tacgttatca gcagtgtccg	4920
ctaaaaactc gtagtgctat aataagcgcg atgagacaag agctgacgcc tctgctggcg	4980
cccctggcgg aagaaagtgc caatgaaacg gggatgggca acaaagaaga taagtttctc	5040
aaaaacaagg ctgcattgga caatacaccg ggggtagaag atctcaccac aacagctctt	5100
accggcgacg gcggtatggt tctgtttgag tactcaccgt ttggagtta aggttcggtc	5160

gcccctagca ccaacccac tgaaactata atcaataaca gtatctctat gctagcagcg	5220
ggcaatagta tttatttttc ccctcatccg ggagctaaaa aagtttctct gaagctgatt	5280
agcttgattg aagaaatagc cttcaggtgc tgtggaatcc gcaatttagt ggtgaccgtg	5340
gcggaaccaa ccttcgaggc taccagcag atgatggccc atccacgaat cgcagtactt	5400
gccattacag gcggcccggg cattgttgca atgggaatga aaagcggtaa gaaagtata	5460
ggagctggtg cgggtaatcc gccatgtatc gttgacgaga cggcggacct ggtgaaagcg	5520
gcggaagaca tcatcaacgg ggcttcattc gattacaatc tgccctgcat tgccgaaaag	5580
agccttattg ttgtggagag tgtacgaga cggctggtgc agcaaatgca aacatttggt	5640
gcactgttgt tatcacctgc agataccgat aaattacgcg ctgtctgcct gcctgaagga	5700
caggcaaata aaaaacttgt cggcaaaagt ccatcggaac tgctggaagc cgcaggaatt	5760
gctgtccctg caaaagcgcc gcgtctttta attgctctgg ttaacgctga tgatccgtgg	5820
gtcacaagcg aacagttgat gcctatgctg ccagtggtaa aagtttcaga tttcgattcc	5880
gcactggcgc tggccttgaa ggttgaagag gggtctgcat atactgctat tatgcactcg	5940
cagaacgtgt cagcctgaa tctcgggca cgcactttac aaacatctat tttcgtaaaa	6000
aatggtccct catatgccgg gattggtgtt ggcggcgaag gctttaccac cttcactatt	6060
gcaacgcaa caggtgaagg gacaacatct gcgcgtactt ttgctcgttc ccggcgtgt	6120
gtactgacga atggtttttc tattcgtaa gtaccctga gtctggtaaa gaaaccgctg	6180
ctgcgaaatt tgaacgccag cacatggact cgtctactag cgcagcttaa ttaacctagg	6240
ctgctgccac cgtgagcaa taactagcat aacccttgg ggctctaaa cgggtcttga	6300
ggggtttttt gctgaaacct caggcatttg agaagcacac ggtcacactg cttccgtag	6360
tcaataaacc ggtaaacgag caatagacat aagcggctat ttaacgacct tgccctgaac	6420
cgacgaccgg gtcgaatttg ctttcgaatt tctgccattc atccgcttat tatcattat	6480
tcaggcgtag caccaggcgt ttaagggcac caataactgc cttaaaaaaa ttacgccccg	6540
ccctgccact catcgagta ctgttgtaat tcattaagca ttctgccgac atggaagcca	6600
tcacagacgg catgatgaac ctgaatgcc agcggcatca gcacctgtc gccttgcgta	6660
taatatattgc ccatagtga aacggggcg aagaagtgt ccatattggc cacgtttaaa	6720
tcaaaactgg tgaaactcac ccagggattg gctgagacga aaaacatatt ctcaataaac	6780
cccttaggga aataggccag gttttaccg taacacgcca catcttgca atatatgtgt	6840
agaaactgcc ggaatcgtc gtggtattca ctccagagcg atgaaaacgt ttcagtttgc	6900
tcatgaaaa cggtgtaaca aggggtaaca ctatccata tcaccagtc accgtctttc	6960
attgccatac ggaactccg atgagcattc atcaggcggg caagaatgtg aataaaggcc	7020

ggataaaact tggcttatt tttctttacg gtctttaaaa aggccgtaat atccagctga 7080

 acggtctggt tataggtaca ttgagcaact gactgaaatg cctcaaaatg tttctttacga 7140
 tgccattggg atatatcaac ggtggtatat ccagtgattt tttctccat tttagcttcc 7200
 ttagctcctg aaaatctcga taactcaaaa aatagcccg gtagtgatct tatttcatta 7260
 tggtgaaagt tggaaacctt tacgtgccga tcaacgtctc atttctgcca aaagtggcc 7320
 cagggttcc cggtatcaac agggacacca ggatttattt attctgcga gtgacttcc 7380
 gtcacaggta tttattcggc gcaaagtgcg tcgggtgatg ctgccaactt actgatttag 7440
 tgtatgatgg tgtttttgag gtgctccagt ggcttctgtt tctatcagct gtccctcctg 7500

 ttcagctact gacggggtgg tgcgtaacgg caaaagcacc gccggacatc agcgctagcg 7560
 gagtgtatac tggttacta tgttggcact gatgagggtg tcagtgaagt gcttcatgtg 7620
 gcaggagaaa aaaggctgca ccggtgcgtc agcagaatat gtgatacagg atatatccg 7680
 ctctctcgt cactgactcg ctacgtcgg tcgttcgact gcggcgagcg gaaatggctt 7740
 acgaacgggg cgagatttc ctggaagatg ccaggaagat acttaacagg gaagtgagag 7800
 ggccgcgga aagccgtttt tccatagct ccgccccct gacaagcatc acgaaatctg 7860
 acgtcaaat cagtggtggc gaaacccgac aggactataa agataccagg cgtttccct 7920

 ggcggtccc tcgtgcgtc tctgttctt gcctttcggg ttaccggtgt cattccgctg 7980
 ttatggccgc gtttgtctca ttccacgct gacactcagt tccgggtagg cagttcgctc 8040
 caagctggac tgtatgcacg aacccccgt tcagtcgac cgctgcgctt tatccggtaa 8100
 ctatcgtctt gagtccaacc cggaagaca tgcaaaagca ccaactggcag cagccactgg 8160
 taattgattt agaggagtta gtcttgaagt catgcgccg ttaaggctaa actgaaagga 8220
 caagttttgg tgactgcgt cctccaagcc agttacctg gttcaaagag ttggtagctc 8280
 agagaacctt cgaaaaaccg ccctgcaagg cggtttttc gtttccagag caagagatta 8340

 cgcgcagacc aaaacgatct caagaagatc atcttattaa tcagataaaa ttttctaga 8400
 tttcagtga atttatctt tcaaatgtag cacctgaagt cagccccata cgatataagt 8460
 tgtaattctc atgttagtca tgccccgcgc ccaccggaag gagctgactg ggttgaaggc 8520
 tctcaagggc atcggtcgag atcccggtgc ctaatgagt agctaactta cattaattgc 8580
 gtgctgctca ctgcccgtt tccagtcggg aaacctgtc tgccagctgc attaatgaat 8640
 cggccaacgc gcggggagag gcggtttgcg tattggg 8677

 <210> 36
 <211> 6590

<212> DNA

<213> Artificial Sequence

<220><223> pCDF HCT

<400> 36

taatacgact cactataggg gaattgtgag cggataacaa ttccctctta gaaataattt	60
tgtttaactt taagaaggag atataccatg gaactaaaca atgtcatcct tgaagaggaa	120
ggtaaagttg ctgtagttac cattaacaga cctaaagcat taaatgcgtt aaatagtgat	180
acactaaaag aaatggatta tgttataggt gaaattgaaa atgatagcga agtacttgca	240
gtaattttta ctggagcagg agaaaaatca tttgtagcag gagcagatat ttctgagatg	300
aaggaaatga ataccattga aggtagaaaa ttccgggatac ttggaaataa agtgtttaga	360

agattagaac ttcttgaaaa gcctgtaata gcagctgtta atggttttgc tttaggaggc	420
ggatgcgaaa tagctatgtc ttgtgataata agaatagtct caagcaacgc aagatttggt	480
caaccagaag taggtctcgg aataacacct ggttttggtg gtacacaaag actttcaaga	540
ttagttggaa tgggcatggc aaagcagctt atatttactg cacaaaatat aaaggcagat	600
gaagcattaa gaatcggact tgtaataaag gtagtagaac ctagtgaatt aatgaataga	660
gcaaaagaaa ttgcaacaaa aattgtgagc aatgctccag tagctgttaa gttaaagaaa	720
caggctatta atagaggaat gcagtgatgat attgatactg ctttagcatt tgaatcagaa	780

gcatttggag aatgcttttc aacagaggat caaaaggatg caatgacagc ttcatagag	840
aaaagaaaaa ttgaaggttt caaaaataga taggaggtaa ggatccgaat tcaaggagct	900
gtttaatgaa aaaggtatgt gttatagggt cagggtactat ggggttcagga attgctcagg	960
catttgcagc taaaggattt gaagtatgat taagagatat taaagatgaa tttgttgata	1020
gaggattaga ttttatcaat aaaaatcttt ctaaattagt taaaaaagga aagatagaag	1080
aagctactaa agttgaaatc ttaactagaa ttcccggaac agttgacctt aatatggcag	1140
ctgattgcga tttagttata gaagcagctg ttgaaagaat ggatattaaa aagcagattt	1200

ttgctgactt agacaatata tgcaagccag aaacaattct tgcatcaaat acatcatcac	1260
tttcaataac agaagtggca tcagcaacta aaagacctga taaggttata ggtatgcatt	1320
tctttaatcc agctcctgtt atgaagcttg tagaggtaat aagaggaata gctacatcac	1380
aagaaacttt tgatgcagtt aaagagacat ctatagcaat agggaaagat cctgtagaag	1440
tagcagaagc accaggattt gttgtaaata gaatattaat accaatgatt aatgaagcag	1500
ttggtatatt agcagaagga atagcttcag tagaagacat agataaagct atgaaacttg	1560

gagctaataccaatggga ccattagaat taggtgattt tataggtctt gatatatgtc	1620
ttgctataat ggatgtttta tactcagaaa ctggagattc taagtataga ccacatacat	1680
tacttaagaa gtatgtaaga gcaggatggc ttggaagaaa atcaggaaaa ggtttctacg	1740
attattcaaa ataagagctc ggcgcgccctg caggctcgaca agcttgcggc cgcataatgc	1800
ttaagtcgaa cagaaagtaa tcgtattgta cacggccgca taatcgaaat taatacgact	1860
cactataggg gaattgtgag cggataacaa ttccccatct tagtatatta gttaagtata	1920
agaaggagat atacatatga ttgtaaaacc aatggtagg aacaatattt gtctaaacgc	1980
tcacccgcaa ggatgcaaaa aaggcgttga ggatcaaata gactacacaa aaaagagaat	2040
taccgctgag gtaaaagccg gagcaaaagc ccctaaaaac gtgctggttc tcggctgctc	2100
gaacggttac ggacttgcaa gccggataac ggcagcattc ggctatgggg ccgccactat	2160
cggcggtttcc ttgaaaaag ccggaagcga aacaaagtac ggcacaccgc gctggtacaa	2220
caacctggcc ttgacgagg ctgccaaaag ggaaggcctt tattccgtaa ctatagacgg	2280
agacgccttt tccgatgaaa tcaaggcaca agtaatcgaa gaagccaaaa agaaaggaat	2340
taaattcgat cttatagttt acagcttggc aagccctgta agaaccgatc ccgacacagg	2400
cataatgcac aagtccgtct taaaaccctt cggtaaaaca ttacaggca agacagtcga	2460
tccttttacg ggagaactaa aagaaatctc cgccgaacct gcaaacgatg aagaagccgc	2520
tgcaaccgtt aaggttatgg gaggagaaga ctgggaacgc tggataaagc agctttcaaa	2580
agaaggtctt ttagaagaag gctgcattac cctagcctat tcctatatcg gccctgaggc	2640
cactcaagcc ctctaccgaa agggcaccat aggaaaggca aaagaacacc ttgaagcaac	2700
tgcccaccgc ctaaacaaag aaaaccctc aatacgggcc ttcttttcgg tgaacaaggg	2760
cttggttaaca agggcaagtg cggtaatccc cgtaatccc ctataacctg ctctctgtt	2820
taaggttatg aaagaaaaag gaaaccacga gggctgtatc gagcagatta cccgccttta	2880
tgccgaaaga ctctaccgta aagacggcac catccccgtc gatgaagaaa acagaatccg	2940
tatcgacgac tgggagcttg aagaagacgt tcaaaaggcg gtttcggctt taatggaaaa	3000
agtaaccggc gaaaatgccg aaagcctaac cgaccttgca ggctaccgcc acgacttttt	3060
agcctcaaac ggctttgatg tagaaggcat caactacgaa gccgaggtag aaaggttcga	3120
caggatttaa agatctcaat tggatatcgg ccggccacgc gatcgctgac gtcggtacct	3180
tcgagctcgg taaagaaacc gctgctgcga aatttgaacg ccagcacatg gactcgtcta	3240
ctagcgcagc ttaattaacc taggtcgtg ccaccgtga gcaataacta gcataacccc	3300

ttggggcctc taaacgggtc ttgaggggtt ttttgctgaa acctcaggca ttgagaagc	3360
acacggctcac actgcttccg gtagtcaata aaccggtaaa ccagcaatag acataagcgg	3420
ctatttaacg accctgccct gaaccgacga cgggtcctc gtggccggat ctgcccggcc	3480
ctcggttga acgaattgtt agacattatt tgccgactac cttggtgac tcgcctttca	3540
cgtagtggac aaattcttcc aactgatctg cgcgcgaggc caagcatct tcttctgtc	3600
caagataagc ctgtctagct tcaagtatga cgggtgata ctgggccggc aggcgtcca	3660
ttgccagtc ggcagcgaca tcttcggcg cgattttgcc ggtiactgcg ctgtacaaa	3720
tgccggacaa cgtaagcact acatttcgct catcgccagc ccagtcgggc ggcgagttcc	3780
atagcgttaa ggtttcattt agcgctcaa atagatcctg ttcaggaacc ggatcaaaga	3840
gttctctcgc cgctggacct accaaggcaa cgctatgtt tcttgctttt gtcagcaaga	3900
tagccagatc aatgtcgatc gtggctggct cgaagatacc tgcaagaatg tcattgcgt	3960
gccattctcc aaattgcagt tcgcgttag ctggataacg ccacggaatg atgtcgtcgt	4020
gcacaacaat ggtgacttct acagcgcgga gaatctcgt ctctccaggg gaagccgaag	4080
tttccaaaag gtcgttgatc aaagctcgcc gcgttgtttc atcaagcctt acggtcacgc	4140
taaccagcaa atcaatatca ctgtgtggct tcaggccgcc atccactgcg gagccgtaca	4200
aatgtacggc cagcaacgac ggttcgagat ggcgctcgat gacccaact acctctgata	4260
gttgagtga tacttcggcg atcacgctt cctcactact cttcttttt caatattatt	4320
gaagcattta tcagggttat tgtctcatga gcggatacat atttgaatgt atttagaaaa	4380
ataaacaat agctagctca ctcggtcgct acgctccggg cgtgagactg cggcgggcgc	4440
tcgggacaca tacaaagtta cccacagatt ccgtggataa gcaggggact aacatgtgag	4500
gcaaaacagc agggccgcgc cgggtggcgtt tttccatagg ctccgccctc ctgccagagt	4560
tcacataaac agacgtttt ccggtgcac tgtgggagcc gtgaggtca accatgaatc	4620
tgacagtacg ggcgaaaccc gacaggactt aaagatcccc accgtttccg gcgggtcgt	4680
ccctcttgcg ctctctgtt ccgacctgc cgtttaccgg atacctgtt cgcctttctc	4740
ccttacggga agtgtggcg tttctcatag ctacacact ggtactcgcg ctcggtgtag	4800
gtcgttcgt ccaagctggg ctgtaagcaa gaactcccc ttcagccga ctgctgcgc	4860
ttatccggt actgttcaat tgagtccaac ccgaaaaagc acggtaaaac gccactggca	4920
gcagccattg gtaactggga gttcgagag gatttgttta gctaaacag cggttgctct	4980
tgaagtgtgc gccaaagtc ggctacactg gaaggacaga tttggttgc gtgctctgcg	5040
aaagccagt accacggtta agcagttccc caactgactt aacctcgat caaacacct	5100
ccccaggtgg tttttctgt tacagggcaa aagattacgc gcagaaaaa aggatctcaa	5160

gaagatcctt tgatcttttc tactgaaccg ctctagatctt cagtgcatt tatctcttca 5220
aatgtacac ctgaagtcag ccccatagca tataagttgt aattctcatg ttagtcatgc 5280
cccgcgccca ccggaaggag ctgactgggt tgaaggctct caagggcac ggtcgagatc 5340
ccggtgccta atgagtgagc taacttacat taattgcgtt gcgtcactg cccgctttcc 5400

agtccggaaa cctgtcgtgc cagctgcatt aatgaatcgg ccaacgcgcg gggagaggcg 5460
gtttgcgtat tggcgccag ggtggttttt cttttacca gtgagacggg caacagctga 5520
ttgcccttca ccgctggcc ctgagagagt tgcagcaagc ggtccacgt ggtttgcccc 5580
agcaggcgaa aatcctgttt gatggtggtt aacggcgga tataacatga gctgtcttcg 5640
gtatcgtct atccactac cgagatgtcc gcaccaacgc gcagcccgga ctcggtaatg 5700
gcgcgcattg cgccagcgc catctgatcg ttggcaacca gcatcgagt gggaacgatg 5760
ccctcattca gcatttgcatt ggtttgttga aaaccggaca tggcactcca gtcgccttcc 5820

cgttccgcta tcggctgaat ttgattgcga gtgagatatt tatgccagcc agccagacgc 5880
agacgcgcg agacagaact taatgggccc gctaacagcg cgatttgctg gtgaccaat 5940
gcgaccagat gtcaccgccc cagtcgcgta ccgtcttcat gggagaaaat aatactgttg 6000
atgggtgtct ggtcagagac atcaagaaat aacgccgga cattagtga ggcagcttcc 6060
acagcaatgg catcctggtc atccagcgga tagttaatga tcagccact gacgcgttgc 6120
gcgagaagat tgtgcaccgc cgctttacag gcttcgacgc cgcttcgttc taccatcgac 6180
accaccacgc tggcaccag ttgatcggcg cgagatttaa tcgcccgac aatttgcgac 6240

ggcgcgtgca gggccagact ggaggtggca acgccaatca gcaacgactg ttgcccgcg 6300
agttgttgtt ccacgcggtt gggaatgtaa ttcagctccg ccatcgccgc ttccactttt 6360
tcccgcgttt tcgcagaaac gtggctggcc tggttacca cgcgggaaac ggtctgataa 6420
gagacaccgg catactctgc gacatcgat aacgttactg gtttcacatt caccaccctg 6480
aattgactct ctccgggcg ctatcatgcc ataccgcga aggttttgcg ccattcgatg 6540
gtgtccggga tctcgacgt ctcccttatg cgactcctgc attaggaaat 6590

<210> 37
<211> 7316
<212> DNA

<213> Artificial Sequence
<220><223> pJD758
<400> 37
tcgcgcgttt cggtgatgac ggtgaaaacc tctgacacat gcagctccc gagacggtca 60

cagcttgtct gtaagcggat gccgggagca gacaagcccc tcagggcgcg tcagcgggtg	120
ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc	180
accacgcttt tcaattcaat tcatcathtt ttttttattc tttttttga tttcggtttc	240
tttgaaatth ttttgattcg gtaatctccg aacagaagga agaacgaagg aaggagcaca	300
gacttagatt ggtatatata cgcatatgta gtgttgaga aacatgaaat tgcccagtat	360
tcttaaccca actgcacaga acaaaaacct gcaggaaacg aagataaatc atgtcgaaag	420
ctacatataa ggaacgtgct gctactcatc ctagtctgtg tgctgccaaag ctatttaata	480
tcatgcacga aaagcaaaca aacttgtgtg cttcattgga tgttcgtacc accaaggaat	540
tactggagtt agttgaagca ttaggtccca aaatttgtht actaaaaaca catgtggata	600
tcttgactga tttttccatg gagggcacag ttaagccgct aaaggcatta tccgccaagt	660
acaatthttt actcttcgaa gacagaaaaat ttgctgacat tggtaatata gtcaaattgc	720
agtactctgc ggggtgtatac agaatagcag aatgggcaga cattacgaat gcacacgggtg	780
tggtgggccc aggtattgth agcggthtga agcaggcggc agaagaagta acaaaggaac	840
ctagaggcct tttgatgtha gcagaattgt catgcaaggg ctccttatct actggagaat	900
atactaagggt tactgttgac attgcgaaga gcgacaaaga ttttgthtct ggctthtattg	960
ctcaaagaga catgggtgga agagatgaag gttacgattg gttgattatg acaccgggtg	1020
tgggtthaga tgacaaggga gacgcattgg gtcaacagta tagaaccgtg gatgatgtgg	1080
tctctacagg atctgacatt attattgttg gaagaggact atttgcaaag ggaagggatg	1140
ctaaggtaga ggggtgaacgt tacagaaaag caggctggga agcatatttg agaagatgcg	1200
gccagcaaaa ctaaaaaact gtattataag taaatgcatg tataactaac tcacaaatta	1260
gagcttcaat ttaattatat cagttattac cctgcggtgt gaaataccgc acagatgcgt	1320
aaggagaaaa taccgcatca ggaaattgta aacgttaata ttttgthaaa attcgcgtta	1380
aatthttgtt aaatcagctc atthtttaac caataggccg aaatcggcaa aatcccttat	1440
aatcaaaaag aatagaccga gatagggttg agtgtgttgc cagthttgga caagagtcca	1500
ctattaaaga acgtggactc caacgtcaaa gggcgaaaaa ccgtctatca gggcgatggc	1560
ccactacgtg aaccatcacc ctaatcaagi tthttggggt cgaggtgccg taaagcacta	1620
aatcggaacc ctaaaggag ccccgattt agagcttgac ggggaaagcc ggcgaacgtg	1680
gcgagaaagg aagggaagaa agcgaaagga gcgggcgcta gggcgctggc aagtgtagcg	1740
gtcacgtgc gcgtaacac cacaccgcc gcgttaatg cgccgtaca gggcgctcc	1800
attgccatt caggctgcgc aactgttggg aaggcgatc ggtgcgggcc tcttcgtat	1860
tacgccagct ggcgaaaggg ggatgtgctg caaggcgatt aagttgggta acgccagggt	1920

tttcccagtc acgacgttgt aaaacgacgg ccagtgaatt gtaatacgac tcactatagg	1980
gcgaattgga gctccaccgc ggtggcggcc gctctagaac tagtggatcc actagttcta	2040
gagcgccga aagatgccga ttggggcg aatcctttat ttggcttca cctcatact	2100
attatcaggg ccagaaaaag gaagtgtttc cctccttctt gaattgatgt taccctcata	2160
aagcacgtgg cctcttatcg agaaagaaat taccgtcgct cgtgatttgt ttgcaaaaag	2220
aacaaaactg aaaaaacca gacacgctcg acttccctgtc ttctattga ttgcagcttc	2280
caatttcgtc acacaacaag gtcctagcga cggtcacag gttttgtaac aagcaatcga	2340
aggttctgga atggcgggaa agggtttagt accacatgct atgatgccca ctgtgatctc	2400
cagagcaaag ttctgtcgat cgtactgtta ctctctctct ttcaaacaga attgtccgaa	2460
tcgtgtgaca acaacagcct gttctcacac actcttttct tctaaccaag ggggtggttt	2520
agtttagtag aacctcgta aacttacatt tacatatata taaacttgca taaattggtc	2580
aatgcaagaa atacatatat ggctctttct aattcgtagt tttcaagtt cttagatgct	2640
ttcttttct cttttttaca gatcatcaag gaagtaatta tctacttttt acaacaata	2700
taaaacaatg tctttatctt caaagtgtc tgtccaagat ttggacttga aggacaagcg	2760
tgtcttcac agagttagt tcaacgtccc attgacgggt aagaagatca cttctaacca	2820
aagaattggt gctgctttgc caaccatcaa gtacgttttg gaacaccacc caagatacgt	2880
tgtcttggct tctcacttgg gtagacaaa cggtgaaaga aacgaaaaat actctttggc	2940
tccagtgtct aaggaattgc aatcattgtt gggttaaggat gtcaccttct tgaacgactg	3000
tgtcggcca gaagtgaag ccgctgtcaa ggcttctgcc ccaggttccg ttattttgtt	3060
ggaaaacttg cgttaccaca tcgaagaaga aggttccaga aaggtcgatg gtcaaaaggt	3120
caaggttcc aaggaagatg ttcaaaagtt cagacacgaa ttgagctctt tggctgatgt	3180
ttacatcaac gatgccttcg gtacctgtt cgttgagcga gttctcaaaa atgaacaaat	3240
gtcgacgggc acgatgagag aaagtgttt taactctttt ttaggagatt catcgggtat	3300
ttcttttgcg aaattagtct tcacggcaac taattttcgt caagattccg gagatgacgt	3360
tctggatgaa gatattaaac aaagggaaca gaaatacaat ggatatgcag aggctgaaaa	3420
caatccagga tccccgggg agctcatgga agacgccggt accgctcaca gagctcactc	3480
ttctatggtc ggtttcgaact tgccacaacg tctgcccgtt ttcttgttgg aaaaggaatt	3540
gaagtacttc ggtaaggctt tggagaacct aaccagacca ttcttggcca tcttaggtgg	3600
tgccaaggtt gctgacaaga ttcaattgat tgacaacttg ttggacaagg tcgactctat	3660

catcattggt gggtgatatg ctttcacatt caagaagggt ttggaaaaca ctgaaatcgg	3720
tgactccatc ttcgacaagg ctggtgctga aatcggtcca aagttgatgg aaaaggccaa	3780
ggccaagggt gtcgaagtcg tcttgccagt cgacttcac c attgctgatg ctttctctgc	3840
tgatgccaac accaagactg tcactgacaa ggaaggtatt ccagctggct ggcaagggtt	3900
ggacaatggt ccagaatcta gaaagtgtt tgctgctact gttgcaaagg ctaagaccat	3960
tgtctggaac ggccaccag gtgttttcga attcgaaaag ttcgctgctg gtactaaggc	4020
tttgtagac gaagttgca agagctctgc tgctggtaac accgtcatca ttggtggtgg	4080
tgacactgcc actgtcgcta agaagtacgg tgcactgac aagatctccc atgtctctac	4140
tggtggtggt gcttctttgg aattattgga aggttaaggaa ttgccagggtg ttgctttctt	4200
atccgaaaag aaataaattg aattgaattg aaatcgatag atcaattttt ttcttttctc	4260
tttcccatc ctttacgta aaataatagt ttattttatt ttttgaatat tttttatita	4320
tatactata tatagactat tatttatctt ttaatgatta ttaagatttt tattaaaaaa	4380
aaattcgctc ctcttttaat gcctttatgc agttttttt tccattcga tatttctatg	4440
ttcgggttca gcgtatttta agtttaataa ctgaaaaatt ctgcgttcgt taaagcttat	4500
cgataccgtc gacctcgagg gggggcccg cgtttttgtt ccttttagtg agggttaatt	4560
ccgagcttgg cgtaatcatg gtcatactg tttctgtgt gaaattgtta tccgctcaca	4620
attccacaca acataggagc cggaagcata aagtgtaaag cctgggggtgc ctaatgagtg	4680
aggtaactca cattaattgc gttgcgctca ctgcccgtt tccagtcggg aaacctgtcg	4740
tgccagctgc attaatgaat cggccaacgc gcggggagag gcggtttgcg tattgggcgc	4800
tcttcgctt cctcgctcac tgactcgctg cgctcggtcg ttcggctgcg gcgagcggta	4860
tcagctcact caaaggcgtt aatacggta tccacagaat caggggataa cgcaggaaag	4920
aacatgtgag caaaaggcca gcaaaaggcc aggaaccgta aaaaggccgc gttgctggcg	4980
ttttccata ggctcggccc ccctgacgag catcacaaaa atcgacgctc aagtcagagg	5040
tggcgaaacc cgacaggact ataaagatac caggcggtcc cccctggaag ctccctcgtg	5100
cgctctcctg ttccgacct gccgcttacc ggatacctgt ccgcctttct ccttcggga	5160
agcgtggcgc tttctcaatg ctacacgtgt aggtatctca gttcgggtga ggtcggtcgc	5220
tccaagctgg gctgtgtgca cgaaccccc gttcagcccg accgtgcgc ctatccggt	5280
aactatcgtc ttgagtccea cccgtaaga cagacttat cgccactggc agcagccact	5340
ggtaacagga ttagcagagc gaggtatgta ggcggtgcta cagagttctt gaagtgggtg	5400

cctaactacg gctacactag aaggacagta tttggtatct gcgctctgct gaagccagtt	5460
accttcggaa aaagagttgg tagctcttga tccggcaaac aaaccaccgc tggtagcgg	5520
ggtttttttg tttgcaagca gcagattacg cgcagaaaaa aaggatctca agaagatcct	5580
ttgatctttt ctacggggtc tgacgtcagc tggaaacgaaa actcacgtta agggattttg	5640
gtcatgagat tatcaaaaag gatcttcacc tagatccttt taaattaaaa atgaagtttt	5700
aaatcaatct aaagtatata tgagtaaact tggctcgaca gttaccaatg cttaatcagt	5760
gaggcaccta tctcagcgat ctgtctatit cgttcatcca tagttgcctg actgcccgtc	5820
gtgtagataa ctacgatacg ggagggtta ccatctggcc ccagtgtgc aatgataccg	5880
cgagaccac gctaccggc tccagattta tcagcaataa accagccagc cggaagggcc	5940
gagcgagaa gtggctctgc aactttatcc gcctccatcc agtctattaa ttgttgccgg	6000
gaagctagag taagtagtic gccagttaat agtttgcgca acgttggtgc cattgctaca	6060
ggcatcgtgg tgcacgctc gtcgtttggt atggcttcat tcagctccgg ttcccaacga	6120
tcaaggcgag ttacatgac ccccatgttg tgaaaaaag cggtagctc cttcggtcct	6180
ccgatcgttg tcagaagtaa gttggccgca gtgttatcac tcatggttat ggcagcactg	6240
cataattctc ttactgtcat gccatccgta agatgctttt ctgtgactgg tgagtactca	6300
accaagtcat tctgagaata gtgtatgcgg cgaccgagtt gctcttgccc ggcgtcaata	6360
cgggataata ccgcgccaca tagcagaact ttaaaagtgc tcatcattgg aaaacgttct	6420
tcggggcgaa aactctcaag gatcttaccg ctgttgagat ccagttcgat gtaaccact	6480
cgtgcacca actgatcttc agcatctttt actttcacca gcgtttctgg gtgagcaaaa	6540
acaggaaggc aaaatgccgc aaaaaaggga ataaggcgca cacggaaatg ttgaatactc	6600
atactcttc tttttcaata ttattgaagc atttatcagg gttattgtct catgagcgga	6660
tacatatitg aatgtattta gaaaaataaa caaatagggg ttccgcgcac atttccccga	6720
aaagtccac ctgggtcctt ttcatcacgt gctataaaaa taattataat ttaaattttt	6780
taataataat atataaatta aaaatagaaa gtaaaaaaag aaattaaaga aaaaatagtt	6840
tttgttttcc gaagatgtaa aagactctag ggggatcgcc aacaataact accttttate	6900
ttgtctttcc tgctctcagg tattaatgcc gaattgtttc atcttgtctg tgtagaagac	6960
cacacacgaa aatcctgtga ttttacattt tacttatcgt taatcgaatg tatatctatt	7020
taatctgctt ttcttgtcta ataaatatat atgtaaagta cgctttttgt tgaaattttt	7080
taaacctttg tttatttttt tttcttcatt ccgtaactct tctaccttct ttatttactt	7140
tctaaaaatc aaatacaaaa cataaaaaata aataaacaca gagtaaatc ccaaattatt	7200
ccatcattaa aagatacgag gcgcgtgtaa gttacaggca agcgatccgt cctaagaaac	7260

cattattatc atgacattaa cctataaaaa taggcgtatc acgaggccct ttcgtc 7316

<210> 38

<211> 1095

<212> DNA

<213> Artificial Sequence

<220><223> Fdh1 from *C. boidinii*

<400> 38

atgaagatcg ttttagtctt atatgatgct ggtaagcacg ctgctgatga agaaaaatta 60

tatgggttga ctgaaaataa attaggtatt gctaattggt taaaagatca aggtcatgaa 120

ctaattacta ctctgataa agaaggtgaa acaagtgaat tggataaaca tatcccagat 180

gctgatatta tcatcaccac tcttttccat cctgcttata tctaactagga aagacttgac 240

aaggctaaga acttaaaatc agtcgtttgc gctgggtgtg gttctgatca cattgattta 300

gattatatta atcaaacagg taagaaaatc tcagtcctgg aagttacagg ttctaagtgt 360

gtctctgttg ctgaacacgt tgtcatgacc atgcttgtct tggtagaaa ttctgttcca 420

gcacatgaac aaattattaa ccacgattgg gaggttgctg ctatcgctaa ggatgcttac 480

gatatcgaag gtaaaactat cgctaccatt ggtgctggta gaattggta cagagtcttg 540

gaaagattac tccatttaa tccaaaagaa ttattatact acgattatca agctttacca 600

aaagaagctg aagaaaaagt tgggtgctaga agagttgaaa atatgaaga attagtgtct 660

caagctgata tcgttacagt taatgctcca ttacacgcag gtacaaaagg tttaattaat 720

aaggaattat tatctaaatt taaaaaaggt gcttggtag tcaataccgc aagaggtgct 780

atttggtgtg ctgaagatgt tgcagcagct ttagaatctg gtcaattaag aggttacggt 840

ggtgatgttt ggttcccaca accagctcca aaggatcacc catggagaga tatgagaaat 900

aaatatgggtg ctggtaatgc catgactcct cactactctg gtactacttt agacgtcaa 960

acaagatacg ctgaaggtac taaaaatatt ttggaatcat tctttaccgg taaatttgat 1020

tacagaccac aagatattat cttattaaat ggtgaatacg ttactaaagc ttacggtaaa 1080

cacgataaga aataa 1095

<210> 39

<211> 1185

<212> DNA

<213> Artificial Sequence

<220><223> atoB from *E. coli*

<400> 39

atgaaaaatt gtgtcatcgt cagtgcggta cgtactgcta tcggtagttt taacggttca	60
ctcgtttcca ccagcgccat cgacctgggg gcgacagtaa ttaaagccgc cattgaacgt	120
gcaaaaaatcg attcacaaca cgttgatgaa gtgattatgg gtaacgtgtt acaagccggg	180
ctggggcaaa atccggcgcg tcaggcacig ttaaaaagcg ggctggcaga aacggtgtgc	240
ggattcacgg tcaataaagt atgtggttcg ggtcttaaaa gtgtggcgct tgccgccag	300
gccattcagg caggtcaggc gcagagcatt gtggcggggg gtatggaaaa tatgagtta	360
gccccctact tactcgatgc aaaagcacgc tctggttata gtcttgaga cggacaggtt	420
tatgacgtaa tcttcgcgca tggcctgatg tgcgccaccc atggttatca tatggggatt	480
accgccgaaa acgtggctaa agagtacgga attaccctg aaatgcagga tgaactggcg	540
ctacattcac agcgtaaagc ggcagccgca attgagtcg gtgcttttac agccgaaatc	600
gtcccggtaa atgttgtcac tcgaaagaaa accttcgtct tcagtcaaga cgaattcccg	660
aaagcgaatt caacggctga agcgtaggt gatttgcgc cggccttcga taaagcagga	720
acagtcaccg ctgggaacgc gtctggtatt aacgacggtg ctgccgtctt ggtgattatg	780
gaagaatctg cggcgctggc agcaggcctt accccctgg ctgcattaa aagttatgcc	840
agcgggtggcg tgccccccgc attgatgggt atggggccag tacctgccac gcaaaaagcg	900
ttacaactgg cggggctgca actggcggat attgatctca ttgaggctaa tgaagcattt	960
gtgacacagt tcttgccgt tgggaaaaac ctgggctttg attctgagaa agtgaatgtc	1020
aacggcgggg ccatcgctt cgggcacctt atcggtgcca gtggtgctcg tattctggtc	1080
acactattac atgcatgca ggcacgcgat aaaacgtgg ggctggcaac actgtgcatt	1140
ggcggcggtc agggaattgc gatggtgatt gaacggttga attaa	1185
<210> 40	
<211> 1023	
<212> DNA	
<213> Artificial Sequence	
<220><223> adhA from L. lactis	
<400> 40	
atgaaagcag cagtagtaag acacaatcca gatggttatg cggaccttgt tgaaggaa	60
cttcgagcaa tcaaacctaa tgaagctttg cttgacatgg agtattgtgg agtctgtcat	120
accgatttgc acgttcagc aggtgattat ggcaacaaag cagggactgt tcttggtcat	180
gaaggaattg gaattgtcaa agaaattgga actgatgtaa gctcgcttca agttggtgat	240
cgggtttcag tggcttggtt ctttgaagga tgtggtcact gtgaatactg tgtatctggt	300

aatgaaactt ttgtcgaga agttaaaaat gcaggatatt cagttgatgg cggaatggct	360
gaagaagcaa ttgttgttc cgattatgct gtcaaagttc ctgacggact tgaccaatt	420
gaagctagct caattacttg tgctggagta acaacttaca aagcaatcaa agtatcagga	480
gtaaacctg gtgattggca agtaatTTTT ggtgctggag gacttgaaa ttagcaatt	540
caatatgcta aaaaTttt ttgagcaaaa gtaattgctg ttgatattaa tcaagataaa	600
ttaaatttag ctaaaaaaat ttgagctgat gtgattatca attctggtga tgtaaatcca	660
gttgatgaaa ttaaaaaaat aactggcggc ttaggagcac aaagtgaat agtttgtgct	720
gttgcaagga ttgcttttga acaagcggtt gcttctttga aacctatggg caaaatggtt	780
gctgtggcac ttccaatac tgagatgact ttatcagttc caacagttgt ttttgacgga	840
giggaggttg caggttcaact tgctggaaca agacttgact tggcagaagc ttttcaattt	900
ggagcagagg gtaaggtaaa accaatgtt gcgacacga aactggaaga aatcaatgat	960
attattgatg aaatgaaggc aggaaaaaat gaaggccgaa tggtcattga ttttactaaa	1020
taa	1023
<210> 41	
<211> 1395	
<212> DNA	
<213> Artificial Sequence	
<220><223> pduP from S. enterica	
<400> 41	
atgaatactt ctgaactcga aacattaatc agaaccattc ttagcgagca gtaaccacg	60
ccggcacaaa cgccggtcca gcctcaaggc aaagggattt ttcagtcctg tagcgaggcc	120
atcgatgccg cccaccaagc gtttttacgt tatcagcagt gtcgctaaa aactcgtagt	180
gctataataa gcgcgatgag acaagagctg acgcctctgc tggcgcccct ggcggaagaa	240
agtgccaatg aaacggggat gggcaacaaa gaagataagt ttctcaaaaa caaggctgca	300
ttggacaata caccgggggt agaagatctc accacaacag ctcttaccgg cgacggcgg	360
atggttctgt ttgagtactc accgtttgga gttatagggt cggtcgcccc tagcaccaac	420
cccactgaaa ctataatcaa taacagtatc tctatgctag cagcgggcaa tagtatttat	480
ttttccctc atccgggagc taaaaaagt tctctgaagc tgattagctt gattgaagaa	540
atagccttca ggtgctgtgg aatccgcaat ttagtggtga ccgtggcgga accaaccttc	600
gaggctacc agcagatgat ggcccatcca cgaatcgag tacttgccat tacaggcggc	660
ccgggcattg ttgcaatggg aatgaaaagc ggtaagaaag tgataggagc tggcgcgggt	720

aatccgccat gtatcgttga cgagacggcg gacctggtga aagcggcgga agacatcatc	780
aacggggctt cattcgatta caatctgccc tgcattgccg aaaagagcct tattgttgtg	840
gagagtgtag cagaacggct ggtgcagcaa atgcaaacat ttggtgcact gttgttatca	900
cctgcagata ccgataaatt acgcgctgtc tgcttgcttg aaggacaggc aaataaaaaa	960
cttgtcggca aaagtccatc ggcaatgctg gaagccgcag gaattgctgt cctgcaaaa	1020
gcgccgctc ttttaattgc tctggttaac gctgatgac cgtgggtcac aagcgaacag	1080
ttgatgccta tgctgccagt ggtaaaagt ttcagatttcg attccgcact ggcgctggcc	1140
ttgaaggttg aagaggggct gcatcatact gctattatgc actcgagaa cgtgtcacgc	1200
ctgaatctcg cggcacgcac tttaaaaaca tctattttcg taaaaaatgg tccctcatat	1260
gccgggattg gtgttggcgg cgaaggcttt accaccttca ctattgcaac gccaacaggt	1320
gaagggacaa catctgcgcg tacttttgcg cgttcccggc gctgtgtact gacgaatggt	1380
ttttctattc gttaa	1395
<210> 42	
<211> 789	
<212> DNA	
<213> Artificial Sequence	
<220><223> hbd from C. acetobutyricum with GBD tag	
<400> 42	
aactaacaa tgtcatcctt gaaaaggaag gttaaagtgc ttagttacc attaacagac	60
ctaaagcatt aaatgcgtta aatagtata cactaaaaga aatggattat gttataggtg	120
aaattgaaaa tgatagcgaa gtacttgcag taattttaac tggagcagga gaaaaatcat	180
ttgtagcagg agcagatatt tctgagatga aggaaatgaa taccattgaa ggtagaaaat	240
tcgggatact tggaaataaa gtgttttagaa gattagaact tcttgaaaag cctgtaatag	300
cagctgttaa tggttttgct ttaggaggcg gatgcgaaat agctatgtct tgtgatataa	360
gaatagcttc aagcaacgca agatttggtc aaccagaagt aggtctcgga ataacacctg	420
gttttggtag tacacaaaga ctttcaagat tagttggaat gggcatggca aagcagctta	480
tatttactgc acaaaatata aaggcagatg aagcattaag aatcggaatt gtaaataagg	540
tagtagaacc tagtgaatta atgaatacag caaaagaaat tgcaaacaaa attgtgagca	600
atgctccagt agctgttaag ttaagcaaac aggctattaa tagaggaaatg cagtgtgata	660
ttgatactgc tttagcattt gaatcagaag catttggaga atgcttttca acagaggatc	720
aaaaggatgc aatgacagct ttcatagaga aaagaaaaat tgaaggcttc aaaaatagat	780

aggaggttaa	789
<210> 43	
<211> 862	
<212> DNA	
<213> Artificial Sequence	
<220><223> crt from C acetobutyricum with SH3 tag	
<400> 43	
aaggagctgt ttaatgaaaa aggtatgtgt tataggtgca ggtactatgg gttcaggaat	60
tgctcaggca ttgcagcta aaggatttga agtagtatta agagatatta aagatgaatt	120
tgttgataga ggattagatt ttatcaataa aaatctttct aaattagtta aaaaaggaaa	180
gatagaagaa gctactaaag ttgaaatctt aactagaatt tccggaacag ttgaccttaa	240
tatggcagct gattgcgatt tagttataga agcagctgtt gaaagaatgg atattaaaaa	300
gcagattttt gctgacttag acaatatatg caagccagaa acaattcttg catcaaatac	360
atcatcactt tcaataacag aagtggcatc agcaactaaa agacctgata aggttatagg	420
tatgcatttc ttaaatccag ctccgtttat gaagcttgta gaggtataaa gaggaatagc	480
tacatcacia gaaacttttg atgcagttaa agagacatct atagcaatag gaaaagatcc	540
tgtagaagta gcagaagcac caggatttgt tgtaaataga atattaatac caatgattaa	600
tgaagcagtt ggtatattag cagaaggaat agcttcagta gaagacatag ataaagctat	660
gaaacttggg gctaatacacc caatgggacc attagaatta ggtgatttta taggtcttga	720
tatatgtctt gctataatgg atgttttata ctcaaaaact ggagattcta agtatagacc	780
acatacatta cttagaagat atgtaagagc aggatggcctt ggaagaaaat caggaaaagg	840
tttctacgat tattcaaaat aa	862
<210> 44	
<211> 1191	
<212> DNA	
<213> Artificial Sequence	
<220><223> Ter from T. denticola with PDZ tag	
<400> 44	
attgtaaac caatggttag gaacaatatt tgtctaaacg ctcatccgca aggatgcaaa	60
aaaggcgttg aggatcaaat agagtacaca aaaaagagaa ttaccgctga ggtaaaagcc	120
ggagcaaaag ccctaataaa cgtgctgggt ctcggctgct cgaacggtta cggacttgca	180
agccggataa cggcagcatt cggctatggg gccgccacta tcggcgcttc ctttgaaaaa	240

gccggaagcg aaacaaagta cggcacaccc ggctggtaca acaacctggc ctttgacgag	300
gctgccaaaa gggaaggcct ttattccgta actatagacg gagacgcctt ttccgatgaa	360
atcaaggcac aagtaatcga agaagccaaa aagaaaggaa ttaaattcga ttttatagtt	420
tacagcttgg caagccctgt aagaaccgat cccgacacag gcataatgca caagtccgtc	480
ttaaaaccct tcggtaaaac atttacaggc aagacagtcg atccctttac gggagaacta	540
aaagaaatct ccgccgaacc tgcaaacgat gaagaagccg ctgcaaccgt taaggttatg	600
ggaggagaag actgggaacg ctggataaag cagctttcaa aagaaggtct tttagaagaa	660
ggctgcatta ccctagccta ttctatatc ggccctgagg ccactcaagc cctctaccga	720
aagggcacca taggaaaggc aaaagaacac cttgaagcaa ctgcccaccg cctaaacaaa	780
gaaaaccctg caatacgggc cttcgtttcg gtgaacaagg gcttggtaac aagggaagt	840
gcggtaatcc ccgtaattcc cctatactc gcttccttgt ttaaggttat gaaagaaaaa	900
ggaaaccacg agggctgtat cgagcagatt acccgctttt atgccgaaag actctaccgt	960
aaagacggca ccatccccgt cgatgaagaa aacagaatcc gtatcgacga ctgggagctt	1020
gaagaagacg ttcaaaaggc ggtttcggct ttaatggaaa aagtaaccgg cgaaaatgcc	1080
gaaagcctaa ccgaccttgc aggctaccgc cagcactttt tagcctcaa cggtttgat	1140
gtagaaggca tcaactacga agccgaggta gaaaggttcg acaggattta a	1191