

	(19) 대한민국특허청(KR) (12) 공개특허공보(A)	(11) 공개번호 10-2012-0099258 (43) 공개일자 2012년09월07일
(51) 국제특허분류(Int. Cl.) C12N 15/54 (2006.01) C12N 1/21 (2006.01)		(71) 출원인 미리안트 코퍼레이션 미국, 매사추세츠 02169-4801, 퀸시, 2 베타리마치 파크
(21) 출원번호 10-2012-7015612	(22) 출원일자(국제) 2010년11월17일	(72) 발명자 공, 웨이 미국, 매사추세츠 01801, 워번, 700 밀 스트리트 에이퍼티 # 12
심사청구일자 없음	(85) 번역문제출일자 2012년06월15일	돌, 수단슈 미국, 매사추세츠 01845, 노쓰 앤도버, 5 로얄 크레스트 드라이브 유닛 # 10 (뒷면에 계속)
(86) 국제출원번호 PCT/US2010/057119	(87) 국제공개번호 WO 2011/063055	(74) 대리인 허용록
국제공개일자 2011년05월26일	(30) 우선권주장 61/281,481 2009년11월18일 미국(US)	
전체 청구항 수 : 총 23 항		

(54) 발명의 명칭 **화합물들의 효과적인 생산을 위한 미생물 엔지니어링**

(57) 요약

본 발명은 유전적으로 엔지니어링된 및 대사적으로 진화된 미생물들로부터의 화합물들의 생산에 관한 것이다. 화학적 생산에서의 향상들이 달성되었고, 또한 그러한 향상들을 유도하는 특정 돌연변이들이 확인되었다. 구체적인 실시예들이 숙신산을 생산하도록 유전적으로 엔지니어링된 박테리아 스트레인의 대사적 진화 과정 동안 발생하였던 돌연변이들의 확인에 대하여 제공되었다. 본 발명은 또한 증가된 숙신산 생산을 위한 대사적 진화 과정 동안 선택되었던 돌연변이들의 산업적 적용 가능성을 평가하기 위한 방법을 제공한다. 본 발명은 추가적으로 대사적 진화 과정 동안 선택되었던 돌연변이들을 포함하고 또한 숙신산, 다른 유기산들 및 상업적 관심의 다른 화합물들의 향상된 생산에 대하여 기여하도록 엔지니어링된 미생물들을 제공한다.

(72) 발명자

그라바, 타미

미국, 메사추세츠 01867, 리딩, 3 아크스톤 씨아
이알 # 10

콜라드, 앤드류, 크리스토퍼

미국, 메사추세츠 02145, 섬머빌, 에이퍼티 # 5,
597 브로드웨이

페로, 재니스, 지.

미국, 메사추세츠 02420, 렉싱턴, 20 솔로몬 피어
스 로드

요킴, 알., 로저스

미국, 메사추세츠 02420, 렉싱턴, 4 오차드 레인

특허청구의 범위

청구항 1

증가된 수준의 포스포에놀 피루베이트 카르복시키나제 활성 및 피루베이트 키나제 활성을 코딩하는 유전자의 돌연변이를 포함하는 것을 특징으로 하는 비-자연적으로 발생하는 미생물.

청구항 2

제 1 항에 있어서,

상기 증가된 수준의 포스포에놀 피루베이트 카르복시키나제 활성은 *pck* 유전자의 자연적인 조절 서열들을, 포스포에놀 피루베이트 카르복시키나제 활성을 증가시키는 개조된 조절 서열들로 대체함으로써 야기되는 것을 특징으로 하는 비-자연적으로 발생하는 미생물.

청구항 3

제 1 항에 있어서,

피루베이트 키나제 활성을 코딩하는 유전자는 *pykA* 유전자 또는 그것의 상동체인 것을 특징으로 하는 비-자연적으로 발생하는 미생물.

청구항 4

제 1 항에 있어서,

피루베이트 키나제 활성을 코딩하는 유전자는 *pykF* 유전자 또는 그것의 상동체인 것을 특징으로 하는 비-자연적으로 발생하는 미생물.

청구항 5

제 1 항에 있어서,

PEP-의존성 포스포트랜스퍼라제 시스템의 기능성을 저하시키는 하나 이상의 유전자 변형을 더 포함하는 것을 특징으로 하는 비-자연적으로 발생하는 미생물.

청구항 6

제 1 항에 있어서,

- (a) PEP-의존성 포스포트랜스퍼라제 시스템의 기능성을 저하시키는 하나 이상의 유전자 변형; 및
- (b) 적어도 하나의 비-PTS 슈거 이송자의 활성을 증가시키는 하나 이상의 유전자 변형을 더 포함하는 것을 특징으로 하는 비-자연적으로 발생하는 미생물.

청구항 7

제 7 항에 있어서,

상기 비-PTS 슈거 이송자는 갈락토오스 퍼미아제인 것을 특징으로 하는 비-자연적으로 발생하는 미생물.

청구항 8

제 7 항에 있어서,

갈락토오스 퍼미아제의 활성의 증가는 갈락토오스 퍼미아제를 코딩하는 *galP* 유전자의 복제 수를 증가시키는 것에 의하여 달성되는 것을 특징으로 하는 비-자연적으로 발생하는 미생물.

청구항 9

제 7 항에 있어서,

갈락토오스 퍼미아제의 활성의 증가는, 하나 이상의 돌연변이에 의한 억제자의 음성 조절을 완화하는 것에 의하여 달성되는 것을 특징으로 하는 비-자연적으로 발생하는 미생물.

청구항 10

제 9 항에 있어서,

상기 갈락토오스 퍼미아제의 억제자는 *galS* 또는 *galR* 유전자에 의하여 인코딩되는 것을 특징으로 하는 비-자연적으로 발생하는 미생물.

청구항 11

제 1 항에 있어서,

발효적 경로에 관여되는 적어도 하나의 유전자들 내에서의 돌연변이를 더 포함하는 것을 특징으로 하는 비-자연적으로 발생하는 미생물.

청구항 12

제 1 항에 있어서,

트리카르복실산 회로의 작동과 관련된 적어도 하나의 유전자 내에서의 돌연변이를 더 포함하는 것을 특징으로 하는 비-자연적으로 발생하는 미생물.

청구항 13

제 1 항에 있어서,

RNA 폴리머라제의 서브유닛을 코딩하는 유전자 내에서의 돌연변이를 더 포함하는 것을 특징으로 하는 비-자연적으로 발생하는 미생물.

청구항 14

20 g/L 초과 유기산 또는 옥살로아세트산염(oxaloacetate)으로부터 유도되는 다른 화합물을 생산하고, 또한 RNA 폴리머라제의 서브유닛을 코딩하는 유전자 내에서의 돌연변이를 포함하는 것을 특징으로 하는 비-자연적으로 발생하는 미생물.

청구항 15

제 13 항 또는 제 14 항에 있어서,

RNA 폴리머라제의 서브유닛을 코딩하는 유전자는 *rpoA* 유전자인 것을 특징으로 하는 비-자연적으로 발생하는 미생물.

청구항 16

제 13 항 또는 제 14 항에 있어서,

RNA 폴리머라제의 서브유닛을 코딩하는 유전자는 *rpoC* 유전자인 것을 특징으로 하는 비-자연적으로 발생하는 미생물.

청구항 17

제 1 항에 있어서,

염색체 DNA 내에서의 결실을 더 포함하되, 상기 결실은 *ycdC*, *ycdD*, *ycdE*, 및 *ycdF* 유전자를 제거하는 것인 것을 특징으로 하는 비-자연적으로 발생하는 미생물.

청구항 18

제 1 항 또는 제 14 항에 있어서,

ftsI 유전자 내에서의 돌연변이를 더 포함하는 것을 특징으로 하는 비-자연적으로 발생하는 미생물.

청구항 19

제 1 항 또는 제 14 항에 있어서,

gldA 유전자 내에서의 돌연변이를 더 포함하는 것을 특징으로 하는 비-자연적으로 발생하는 미생물.

청구항 20

제 1 항 또는 제 14 항에 있어서,

dhaKLM 오페론 또는 PTS-의존성 디하이드록시아세톤 키나제의 서브유닛을 인코딩하는 유전자 내에서의 돌연변이를 더 포함하는 것을 특징으로 하는 비-자연적으로 발생하는 미생물.

청구항 21

20 g/L 초과와 유기산 또는 옥살로아세트산염으로부터 유도되는 다른 화합물을 생산하고, 또한 염색체 DNA 내에서의 결실을 포함하되, 상기 결실은 *ydcC*, *ydcD*, *ydcE*, 및 *ydcF* 유전자들 또는 그들의 상동체를 제거하는 것인 것을 특징으로 하는 비-자연적으로 발생하는 미생물.

청구항 22

- (a) 제 1 항 내지 제 14 항 중 어느 한 항에 따른 미생물을 배양하는 단계;
- (b) 탄소원을 제공하는 단계;
- (c) 상기 미생물이 상기 탄소원을 대사하도록 허용하는 단계; 및
- (d) 선택적으로, 숙신산을 분리하는 단계를 포함하는 것을 특징으로 하는 숙신산을 생산하는 방법.

청구항 23

- (a) 제1 부모 스트레인으로부터, 소정의 화합물과 다른 발효 생성물에 대한 생합성 경로의 단계를 촉진하는 효소를 인코딩하는 적어도 하나의 유전자를 삭제하는 단계로서, 제1 중간체 스트레인을 유도하고,
- (b) 상기 제1 스트레인 성장에 비해 향상된 성장을 보이는 제2 중간체 스트레인을 유도하기 위하여 상기 제1 중간체 스트레인에 대사적 진화를 수행하는 단계,
- (c) 상기 제1 부모 스트레인 및 제2 중간체 스트레인의 게놈 DNA 서열을 결정하는 단계,
- (d) 상기 대사적 진화의 과정 동안 선택되었던 돌연변이들을 식별하기 위하여 상기 게놈 DNA 서열들을 비교하는 단계,
- (e) 상기 돌연변이들 중 적어도 하나를 테스트하고, 또한 상기 돌연변이들의 어느 것이 상기 소정의 화합물의 생산 효율의 성장 속도를 향상시키는데 유리한 것인지를 식별하는 단계, 및
- (f) 제2 부모 스트레인 내에, 하나 이상의 상기 식별된 유리한 돌연변이들, 또는 상기 유리한 돌연변이들 중 하나와 기능적으로 유사한 하나 이상의 돌연변이들을 장착하는 단계를 포함하는 것을 특징으로 하는 발효에 의해 소정의 화합물을 생산하기 위한 목적의 미생물을 제조하는 방법.

명세서

기술분야

[0001] 본 출원은 2009년 11월 18일자로 출원된 미국 가출원 번호 61/281,481에 대한 우선권을 주장한다.

[0002] 본 발명은 수여 번호(Award Number) DE-EF0002878/001 하에 미국 에너지 부(US Department of Energy)로부터 수여된 계약 하에 미합중국 정부 지원과 함께 이루어졌다. 미합중국 정부는 본 발명에 대하여 일정한 권리를 갖는다.

배경기술

[0003] 지난 10년 동안, 미생물 게놈들(microbial genomes), 세포 내 생화학적 경로들, 대사 흐름 분석(metabolic flux analysis), 마이크로레이 분석에 관한 우리의 지식 및 실리콘 분석(silico analysis)에서의 폭발과 함께, 산업적 미생물학은 생체 촉매들(biocatalysts)을 사용하는 재생 가능한 공급 재료(renewable feedstock)로부터 화합물들을 제조하는 것으로 도전되어 왔다.

- [0004] 2004 미국 에너지 부 리포트, "바이오매스(biomass)로부터 최고의 가치가 추가된 화합물(Top value added chemicals from biomass)"로 명명된, 는 생체 촉매들을 사용하여 재생 가능한 공급 재료들로부터 생산될 수 있는 15개 빌딩 블록 화합물들(building block chemicals)을 확인하였다. 상기 15개 빌딩 블록들은 1,4-이중산들 (1,4-diacids)(숙신산의(succinic), 푸마르산의(fumaric) 및 말산의(malic)), 2,5-퓨란 디카르복실산(2,5-furan dicarboxylic acid), 3-하이드록시프로피온산(3-hydroxypropionic acid), 아스파르트산(aspartic acid), 글루칸산(glucaric acid), 글루탐산(glutamic acid), 이타콘산(itaconic acid), 레블린산(levulinic acid), 3-하이드록시부티로락톤(3-hydroxybutyrolactone), 글리세롤(glycerol), 소르비톨(sorbitol), 자일리톨(xylitol), 및 아라비니톨(arabinitol)이다. 미국 에너지 부에 의해 규정된 15개의 화합물들 중에서, 생체 촉매들을 사용하는 산업적 규모에서의 숙신산 생산이 상당히 발전되어 왔다(커즈락 및 웨스터-보츠(Kurzrock and Weuster-Botz), 2009; 리 외.(Lee *et al.*), 2008; 앤더슨 외.(Andersson, *et al.*), 2007; 루 외.(Lu *et al.*), 2009; 미국 특허 출원 공개 번호 2006/0073577).
- [0005] 소(cattle)의 반추위(rumen)에 존재하는 박테리아가 혐기성 성장 조건들 하에서의 숙신산을 생산하는 것으로 알려져 있다. *악티노바실러스 수시노진스(Actinobacillus succinogens)*, *애어로바이로스프릴룸 숙시니프로듀센스(Anaerobiosprillum succiniproducens)* 및 *만하이미아 숙시니프로듀센스(Mannheimia succiniproducens)*와 같은 많은 수의 반추위 박테리아가 분리되어 왔고 또한 숙신산 생산을 위한 생체 촉매로서 개발되어 왔다. 상업적인 가능성에 도달한 타이터(titers), 속도(rates) 및 수율(yields)로 숙신산의 생산을 가능하게 하는 대장균(*Escherichia coli*) 스트레인들(strains)이 미생물적 탄소 대사 및 미생물 유전학의 결합된 지식을 사용하여 또한 제조되어 왔다. 미국 특허 번호 7,223,567은 풍부한 성장 배지 내에서 숙신산을 생산하는 *이. 콜리(E. coli)* 스트레인 SBS550MG의 제조를 기재하고 있다. 미국 특허 번호 6,455,284는 숙신산 생산을 위한 생체 촉매로서 외인성 피루베이트 카르복실라제 유전자(exogenous pyruvate carboxylase gene)를 사용하는 *이. 콜리(E. coli)* 스트레인의 제조를 기재하고 있다. 피루베이트 카르복실라제는 야생형 *이. 콜리(E. coli)* 스트레인에는 결여되어 있다. *리조비움 엘티(Rhizobium elti)*와 같은 다른 미생물 기관들로부터 얻어진 피루베이트 카르복실라제 유전자가 숙신산 생산을 증진시키기 위하여 구조적 프로모터(constitutive promoter) 하에서 발현될 수 있다. PCT 특허 출원 번호들 WO/2008/115958 및 WO/2010/115067 은 최소한의 성장 배지 내에서 숙신산 생산을 위한 생체 촉매들인 *이. 콜리(E. coli)* 스트레인 KJ122의 제조를 기재하고 있다.
- [0006] 숙신산 및 다른 화합물들의 상업적으로 매력적인 생합성적 생산에 도달하기 위하여, 추가적인 유전적 조작들이 요구된다. 새로운 유전적 접근들을 사용하여 세포 내부에서 생화학적 경로들을 조작하는 방법에 의하여 미생물들(microorganisms)에 의한 숙신산염 및 다른 화합물들의 생산에 있어서의 전체적인 "효율성(Efficiency)"(이에 제한되는 것은 아니나. 타이터, 속도 및 수율을 포함하는 것으로 정의되는)을 향상시키고자 하는 요구가 여전히 존재한다. 화합물 생산이 세포 성장과 결합되어 있는 만큼, 숙신산 또는 다른 화합물 생산의 효율에 있어서의 향상은 바람직한 화합물을 생산하는 미생물적 세포들의 성장 속도를 향상시키는 방법에 의하여 또한 도달될 수 있다. *이. 콜리(E. coli)* 내에서의 숙신산 생산에 대한 이론적인 수율의 최대값은 소비되는 글루코오스의 1 그램에 대하여 1.12 그램의 숙신산의 질량 수율을 제공하는 글루코오스 1 몰 당 1.714 몰의 숙신산으로 계산된다(베무리, 외(Vemuri, *et al.*), 2002).
- [0007] 미국 특허 출원 공개 번호 2008/0009041은 야생형 *이. 콜리(E. coli)* 스트레인의 것보다 적어도 470kb 짧은 염색체 DNA를 포함하는 *이. 콜리(E. coli)* 스트레인을 기재하고 있다. 이 돌연변이체 *이. 콜리(E. coli)* 스트레인은 더 많은 세포 매스(cell mass)를 축적하고 또한 축적되는 트레오닌(threonine)의 양에 있어서 상당한 증가를 나타낸다.
- [0008] 미국 특허 출원 공개 번호 2009/0075333은 또한 부모의 스트레인과 비교하여 하나 이상의 동일한 또는 향상된 성장 속도, 변형 효율성(transformation efficiency), 단백질 발현, DNA 생산, DNA 수율 및/또는 DNA 퀄리티를 갖는 감소된 계놈 사이즈를 갖는 *이. 콜리(E. coli)* 스트레인을 기재하고 있다.
- [0009] 미국 특허 출원 공개 번호 2009/0221055는 유전자 파괴를 통하여 유도되는 다양한 효소들의 우수한 생산성을 갖는 새로운 *바실러스 서브틸리스(Bacillus subtilis)* 돌연변이체 스트레인을 기재하고 있다.
- [0010] 향상된 성장 속도에 도달하기 위하여 계놈 사이즈를 줄이는 이러한 접근법들은 주 발효조(fermentor) 내에서 성장하는 박테리아가 삭제될 수 있는 많은 비-본질적인 유전자들을 포함하고 있다는 사실을 바탕으로 한다. 자연적인 환경에서 살아가는 박테리아는 온도, 스트레스, 또는 식량 공급의 부족의 다른 환경적인 조건들에서 생존하기 위한 메커니즘을 제공하는 많은 조건-반응적인 유전자들을 가지고 있다. 이러한 유전자들을 복제하는 것은, 이것은 주 발효조 내에서는 불필요한 것이다, 그렇지 않을 경우 이러한 불필요한 유전자들의 결여에

서 보존될 수 있는 세포적인 에너지의 소비를 요구한다.

[0011] 게놈 사이즈를 감소시키는 것의 접근이 재조합 단백질들, 핵산 및 아미노산을 생산하는 스트레인들의 성장 능력을 향상시키기 위하여 유용할 수 있는 반면, 그것은 크레브 회로(Krebs cycle) (또한 트리카르복실산 회로(tricarboxylic acid cycle) 또는 TCA 회로로도 알려져 있는) 내에서의 중간체들(intermediates)인 숙신산, 푸마릭산, 및 말산과 같은 다른 화합물들을 생산하는 스트레인들의 성능을 향상시키기 위한 유용한 접근으로 보여지지는 않아 왔다. 이러한 및 다른 화합물들의 생산 효율은 또한 중앙의 대사적 경로들에서 다른 경로들을 통한 탄소 흐름의 속도에, 또한, 어떤 경우들에 있어서는, 혐기성 및 미세 호기성 성장 과정 동안에 선호되는 레독스 균형에 도달하는 것에 의존된다.

[0012] 미국 특허 출원 공개 2009/0075352, 및 리 외(Lee *et al.*)(2005)는 실리코 분석 내(in-silico analysis)를 바탕으로 하는 박테리아 스트레인을 향상시키기 위한 방법을 기재하고 있다. 이러한 접근에서, *만하이미아 속시니프로듀센스(Mannheimia succiniproducens)*의 게놈 서열, 상당한 양으로 숙신산을 생산하는, 은 야생형의 *이. 콜리(E. coli)* 스트레인을 숙신산을 생산하는 스트레인으로 변환하기 위한 목적을 위하여 *이. 콜리(E. coli)*에서 삭제되어야 할 최적의 유전자들을 식별하기 위하여 *이. 콜리(E. coli)*의 게놈 서열과 비교된다. 이러한 접근이 표면에서 매력적으로 보일지라도, 그것은 *엠. 속시니프로듀센스(M. succiniproducens)*로부터 유도되는 유전적 정보가 상업적으로 성장 가능한 숙신산염 생산 스트레인에 도달하기 위한 *이. 콜리(E. coli)*에 까지 추정될 수 있는가 보여지는 것으로 남아있다. 상기 언급된 미국 특허 출원로부터 추천되는 삭제들은 *ptsG*, *pykA*, 및 *pykF*의 조합이었다. 그러나, 그러한 스트레인으로부터 실질적으로 달성되는 숙신산염 타이터는 단지 8.16 mM (0.96 g/l)였고, 이것은 상업적으로 매력적인 생산을 위해 요구되는 수준 근처가 결코 아니며, 또한 상기 스트레인의 성장은 매우 열악하였다. 나아가, 상기 참조된 출원에서 추천되는 것과 유사한 삭제 돌연변이들(*pykA*, *pykF*, 및 *ptsHI*)의 조합은 보다 먼저 보고되어져 왔고, 또한 그 결과적인 스트레인은 글루코오스 상에서 매우 열악하게 성장하였다(폰세, 외.(Ponce *et al.*), 1995). 따라서, 리 외.(Lee *et al.*)(2005)에 의하여 기재된 돌연변이들의 조합은 상업적으로 성장할 수 있는 스트레인을 제조하는데 충분하지 않았고, 또한 상기 돌연변이들은 그들이 상업적으로 성장 가능한 스트레인을 위해 필요한지 또는 적합한지를 입증할 수 있는 스트레인 내용물(strain context) 중에서 테스트되지 않았다. 따라서, 본 기술 분야에서 숙련된 자에 있어서, 글루코오스 및 다른 값싼 탄소 공급원상에서 잘 성장하는 것 및 적어도 20 g/l 숙신산염(170 mM)을 생산하는 것을 요구되는, 상업적으로 성장 가능한 숙신산염 생산 스트레인에 대해 상기 개시들로부터의 명확한 경로가 존재하지 않는다.

[0013] 모든 참조 문헌들이 독자의 편의를 위하여 하기에 목록화되었다. 각 문헌은 그것의 전체로서 참조에 의하여 포함되었다.

선행기술문헌

특허문헌

- [0014] (특허문헌 0001) 미국 특허 번호 6,455,284
(특허문헌 0002) 미국 특허 번호 6,962,794
(특허문헌 0003) 미국 특허 번호 6,989,265
(특허문헌 0004) 미국 특허 번호 7,223,567
(특허문헌 0005) 미국 특허 번호 7,229,794
(특허문헌 0006) 미국 특허 번호 7,303,906
(특허문헌 0007) 미국 특허 번호 7,371,558
(특허문헌 0008) 미국 특허 번호 7,524,660
(특허문헌 0009) 미국 특허 번호 7,629,162
(특허문헌 0010) 미국 특허 출원 공개 번호 2004/0214294
(특허문헌 0011) 미국 특허 출원 공개 번호 2004/0146966
(특허문헌 0012) 미국 특허 출원 공개 번호 2005/0181488

(특허문헌 0013) 미국 특허 출원 공개 번호 2005/0176114
 (특허문헌 0014) 미국 특허 출원 공개 번호 2005/0221455
 (특허문헌 0015) 미국 특허 출원 공개 번호 2006/0073577
 (특허문헌 0016) 미국 특허 출원 공개 번호 2007/0111294
 (특허문헌 0017) 미국 특허 출원 공개 번호 2008/0009041
 (특허문헌 0018) 미국 특허 출원 공개 번호 2008/0176302
 (특허문헌 0019) 미국 특허 출원 공개 번호 2008/0293100
 (특허문헌 0020) 미국 특허 출원 공개 번호 2009/0047719
 (특허문헌 0021) 미국 특허 출원 공개 번호 2009/0075333
 (특허문헌 0022) 미국 특허 출원 공개 번호 2009/0075352
 (특허문헌 0023) 미국 특허 출원 공개 번호 2009/0221055
 (특허문헌 0024) 미국 특허 출원 공개 번호 2009/0325243
 (특허문헌 0025) 미국 특허 출원 공개 번호 2010/0143997
 (특허문헌 0026) 미국 특허 출원 공개 번호 2010/0261239
 (특허문헌 0027) 미국 특허 출원 공개 번호 2010/0248311
 (특허문헌 0028) 미국 특허 출원 공개 번호 2010/0279369
 (특허문헌 0029) 국제 특허 출원 공개 번호 WO2008/115958
 (특허문헌 0030) 국제 특허 출원 공개 번호 WO2010/115067
 (특허문헌 0031) 유럽 특허 출원 EP2,241,630

비특허문헌

[0015]

(비특허문헌 0001) 알트라스, N.E.(Altaras, N. E.), 카메론 D.C.(Cameron, D. C.) (1999) "에세리키아 콜리 내에서의 1,2-프로판디올 경로의 대사적 엔지니어링(Metabolic engineering of a 1,2-propanediol pathway in Escherichia coli.)" App Environ Microbiol 65 : 1180- 1185.

(비특허문헌 0002) 앤더슨, C(Andersson, C), 호지, D(Hodge, D.), 베글룬드, K.A.(Berglund, K.A.), 로바, U.(Rova, U.) (2007) "대사적으로 조작된 에세리키아 콜리를 사용하는 숙신산의 생산에 대한 다른 탄소원의 효과(Effect of different carbon sources on the production of succinic acid using metabolically engineered Escherichia coli.)" Biotechnol Prog 23: 381-388.

(비특허문헌 0003) 바바, T.(Baba, T.), 아라 T.(Ara, T.), 하세가와, M(Hasegawa, M.), 타카이, Y.(Takai, Y.), 오쿠무라, Y.(Okumura, Y.), 바바, M.(Baba, M.), 다체코, K.A.(Datseko, K.A.), 토미타, M.(Tomita, M.), 워너, B.L.(Wanner, B. L.), 모리, H.(Mori, H.) (2006) "프레임, 단일-유전자 노크아웃 돌연변이체 내에서의 에세리키아 콜리 K-12의 건조: 케이오 컬렉션(Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection)." Mol Syst Biol 2: 아티클 넘버: 2006.0008

(비특허문헌 0004) 바비츠크, P.(Babitzke, P.), 로메오, T.(Romeo, T.) (2007) "CsrB sRNA 과(family): RNA-결합 조절 단백질들의 서열화(sequestration of RNA-binding regulatory proteins)." Curr Opin Microbiol 10: 156-163.

(비특허문헌 0005) 코지, T.B.(Causey, T.B.), 샨무감, K.T.(Shanmugam, K.T.), 요마노, L.P.(Yomano, L.P.), 잉그람, L.O.(Ingram, L.O.) (2004) "글루코오스의 피루브산염으로의 효과적인 전환을 위한 에세리키아 콜리의 엔지니어링 (Engineering Escherichia coli for efficient conversion of glucose to pyruvate)." Proc Natl Acad Sci USA 101 :2235-2240.

- (비특허문헌 0006) 크로난, J.(Cronan, J.), 라포르테, D.(Laporte, D.) (1996) "트리카르복실산 회로 및 글리옥살산염 바이패스(Tricarboxylic acid cycle and glyoxylate bypass)" 에세리키아 콜리 및 살모넬라 내에서의(in *Escherichia Coli* and *Salmonella*). " 편집자들 나이트라트, F. 외.(editors Neidhardt, F., et al), ASM Press, 워싱턴 D.C, 미국(Washington, DC, USA).
- (비특허문헌 0007) 다첸코, K.A.(Datsenko, K. A.), 워너, B.L.(Wanner, B. L.) (2000) "PCR 생성물들을 사용하는 에세리키아 콜리 K-12 내에서의 염색체 유전자들의 단일-단계 비활성화(One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products)." *Proc Natl Acad Sci USA* 97: 6640-6645.
- (비특허문헌 0008) 프리취, P.S.(Fritsch, P. S.), 어바노우스키, M.L.(Urbanowski, M. L.), 스타우퍼, G.V.(Stauffer, G. V.) (2000) "metE 및 metH의 MetR-의존성 활성화에서의 RNA 폴리머라제의 역할: RNA 폴리머라제 내부에서의 C-터미널 도메인 및 방향성에서의 중요한 잔기들(Role of RNA polymerase subunits in MetR-dependent activation of metE and metH: Important residues in the C-terminal domain and orientation requirements within RNA polymerase)." *J Bacteriol* 182:5539-5550.
- (비특허문헌 0009) 홀크로프트, C.C.(Holcroft, C.C.), 이간, S. M.(Egan, S. M.) (2000) "회로적 AMP 수용체 단백질, RNA 폴리머라제 알파 서브유닛 C-터미널 도메인, 및 RhaR에 의한 rhaSR에서의 활성화의 상호 의존성(Interdependence of activation at rhaSR by cyclic AMP receptor protein, the RNA polymerase alpha subunit C-terminal domain, and RhaR)." *J Bacteriol* 182: 3529-3535.
- (비특허문헌 0010) 이케다, M.(Ikeda, M.), 오니쉬, J.(Ohnishi, J.), 하야시, M.(Hayashi, M.), 미즈하시, S.(Mitsuhashi, S.). (2006) "효과적인 L-라이신 생산을 위한 최소한으로 돌연변이화된 코리네박테리움 글루타미쿰 스트레인을 형성하는 게놈-바탕의 접근(A genome-based approach to create a minimally mutated *Corynebacterium glutamicum* strain for efficient L-lysine production)." *J Ind Microbiol Biotechnol* 33: 610-615
- (비특허문헌 0011) 잔타마, K.(Jantama, K.), 합트, M. J.(Haupt, M. J.), 스보로노스, S.A.(Svoronos, S. A.), 장, X.(Zhang, X.), 무어 J.C.(Moore, J. C), 산무감, K.T.(Shanmugam, K. T.), 잉그람, L. O.(Ingram, L. O.) (2008a) "숙신산염 및 말산염을 생산하는 에세리키아 콜리 C의 비재조합 스트레인들을 개발하기 위한 대사적 엔지니어링 및 대사적 진화의 조합(Combining metabolic engineering and metabolic evolutions to develop nonrecombinant strains of *Escherichia coli* C that produce succinate and malate)." *Biotechnol Bioeng* 99: 1140-1153.
- (비특허문헌 0012) 잔타마, K.(Jantama, K.), 장, X.(Zhang, X.), 무어 J.C.(Moore, J. C), 산무감, K.T.(Shanmugam, K. T.), 스보로노스, S.A.(Svoronos, S. A.), 잉그람, L.O.(Ingram, L. O.) (2008b) "에세리키아 콜리 C의 조작된 스트레인들 내에서 부산물을 제거하는 것 및 숙신산염 수율을 향상시키는 것(Eliminating side products and increasing succinate yields in engineered strains of *Escherichia coli* C)." *Biotechnol Bioeng* 101 : 881-893.
- (비특허문헌 0013) 커즈락, T.(Kurzrock, T.), 웨스터-보츠, D.(Weuster-Botz, D.) (2009) "발효 배지로부터의 숙신산의 회수(Recovery of succinic acid from fermentation broth)." *Biotechnol Lett* 32: 331-339.
- (비특허문헌 0014) 크나그, Y.(Knag, Y.), 더팜, T.(Durfeem T.), 그라스너, J.D.(Glasner, J.D.), 퀴, Y.(Qiu, Y.), 프리쉬, D.(Frisch, D.), 윈터버그, K. M.(Winterberg, K. M.), 블라트너, F.R.(Blattner, F. R.) (2004) "에세리키아 콜리 게놈의 시스템적인 돌연변이 유발(Systematic mutagenesis of the *Escherichia coli* genome)." *J Bacteriol* 186: 4921-4930.
- (비특허문헌 0015) 콜리스니첸코, V.(Kolisnychenko, V.), 플런케트 III, G.(Plunkett III, G.), 헤르리그니, C.D.(Herrigni, C. D.), 페허, T.(Feher, T.), 포스파이, J.(Posfai, J.), 블라트너, F. R.(Blattner, F. R.), 포스파이, G.(Posfai, G.) (2009) "감소된 에세리키아 콜리 게놈의 엔지니어링(Engineering a reduced *Escherichia coli* genome)." *Genome Res* 12: 640-647.
- (비특허문헌 0016) 리, S.J.(Lee, S.J.), 리 D-Y.,(Lee, D-Y.), 김, T. Y.(Kim, T. Y.), 김, B.H.(Kim, B.H., Lee J.), 리, S. Y.(Lee, S. Y.) (2005) "게놈 비교를 바탕으로 하는 및 실리코 유전자 노크아웃 시뮬레이션 내에서의, 향상된 숙신산 생산을 위한 에세리키아 콜리의 대사적 엔지니어링(Metabolic Engineering of *Escherichia coli* for enhanced production of succinic acid, based on genome comparison and in silico

gene knockout stimulation)." App Environ Microbiol 71 : 7880- 7887.

(비특허문헌 0017) 리, S.Y.(Lee, S. Y.), 리 D.Y.(Lee, D. Y.), 김, T.Y.(Kim, T. Y.) (2005) "스트레인 향상을 위한 시스템 바이오테크놀로지(Systems biotechnology for strain improvement)." Trends Biotech 23: 349-358.

(비특허문헌 0018) 리, S.Y.(Lee, S. Y.), 김, J.M.(Kim, J. M.), 리 J.W.(Lee, J. W.), 김, T.Y.(Kim, T. Y.), 양, Y.S.(Jang, Y.S.) (2008) "만하이미아 숙시니프로듀스스에 의한 숙신산 생산을 위한 집적된 바이오 프로세스에 대한 게놈 서열로부터(From genome sequence to integrated bioprocess for succinic acid production by Mannheimia succiniproducens)." App Microbiol Biotechnol 79: 11-22;

(비특허문헌 0019) 루, S.(Lu, S.), 아이트남, M.A.(Eiteman, M. A.), 알트만, E.(Altman, E.) (2009) "듀얼-상 에세리키아 콜리 발효 내에서의 숙신산염 생산에 영향을 미치는 pH 및 염기 카운터이온(pH and base counterion affect succinate production in dual-phase Escherichia coli fermentations)." J Ind Microbiol Biotechnol 36: 1101-1109.

(비특허문헌 0020) 마티네즈, A.(Martinez, A.), 그라바, T.B.(Grabar, T.B.), 산무감, K.T.(Shanmugam, K.T.), 요마노, L.P.(Yomano, L.P.), 요크, S.W.(York, S.W.), 잉그람, L.O.(Ingram, L.O.) (2007) "재조합 에세리키아 콜리에 의한 락테이트 및 에탄올 생산을 위한 낮은 염 배지(Low salt medium for lactate and ethanol production by recombinant Escherichia coli)." Biotechnol Lett 29:397-404.

(비특허문헌 0021) 밀라드, C.S.(Millard, C. S.), 차오, Y-P.(Chao, Y-P.), 리아오, J.C.(Liao, J. C.), 도넬리, M.I.(Donnelly, M. I.) (1996) "에세리키아 콜리 내에서의 포스포에놀 피루베이트 카복실라제의 과발현에 의한 숙신산의 향상된 생산(Enhanced production of succinic acid by overexpression of phosphoenolpyruvate carboxylase in Escherichia coli)" App Environ Microbiol 62: 1808-1810.

(비특허문헌 0022) 퍼네스티그, A.K.(Pernestig, A. K.), 게오게릴스, D.(Geogellis, D.), 로메오, T.(Romeo, T.), 스즈키, K.(Suzuki, K.), 토메니우스, FL.(Tomenius, FL), 노르막, S.(Normakr, S.), 멜레포스, O.(Melefors, O.) (2003) "에세리키아 콜리 BarA-UvrY 두 성분 시스템은 당분해 및 당신생 탄소원들 사이의 효과적인 변환을 위해 요구된다(The Escherichia coli BarA-UvrY two component system is needed for efficient switching between glycolytic and gluconeogenic carbon sources)." J Bacteriol 185:843-853.

(비특허문헌 0023) 폰스, E.(Ponce, E.), 플로레스, N.(Flores, N.), 마티네즈, A.(Martinez, A.), 발레, F.(Valle, F.), 볼리바, F.(Bolivar, F.) (1995) "에세리키아 콜리로부터의 두 개의 피루베이트 키나제 동위 효소들 구조적 유전자들의 클로닝: 피루브산염 생합성에서의 이들 효소들의 관련성 있는 역할들(Cloning of the two pyruvate kinase isoenzymes structural genes from Escherichia coli: the relative roles of these enzymes in pyruvate biosynthesis)." J Bacteriol 111: 5719-5722.

(비특허문헌 0024) 포스파이, G.(Posfai, G.), 플런켓 III, G.(Plunket III, G.), 페허, T.(Feher, T.), 프리쉬, D.(Frisch, D.), 케일, G.M.(Keil, G. M.), 우멘노퍼, K(Umenhoffer, K.), 콜리스니첸코, V.(Kolisnychenko, V.), 스탈, B.(Stahl, B.), 샤마, S.(Sharma, S.), 드 스루다, M.(de Srruda, M.), 버란드, V.(Burland, V.), 하쿰, S.W.(Harcum, S. W.), 블라트너, F. R.(Blattner, F. R.) (2006) "감소된-게놈 에세리키아 콜리의 신생 특성들(Emergent properties of reduced-genome Escherichia coli.)" Science 312: 1044-1046.

(비특허문헌 0025) 사이어, M.H.Jr.(Saier, M. H. Jr.), 램사이어, T.M.(Ramseier, T. M.)(1996) "엔테릭 박테리아의 분해 대사질 억제자/활성자 (Cra) 단백질(The catabolite repressor/activator (Cra) protein of enteric bacteria)." J Bacteriol 178: 3411-3417.

(비특허문헌 0026) 산체스 AM(Sanchez AM), 버네트 GN(Bennett GN), 산 KY(San KY) (2005) "숙신산염 수율 및 생산성을 향상시키기 위한 에세리키아 콜리 내에서의 혐기성 중심 대사적 경로의 새로운 경로 엔지니어링 디자인(Novel pathway engineering design of the anaerobic central metabolic pathway in Escherichia coli to increase succinate yield and productivity)." Metab Eng 7:229-239.

(비특허문헌 0027) 시볼트, C.(Siebold, C), 가르시아-알레스, L.F.(Garcia-Alles, L. F.), 에르니, B.(Erni, B.), 바우만, U.(Baumann, U.) (2003) "에세리키아 콜리 하이드록시아세톤 키나제의 서브유닛 K의 x-레이 구조 내에서 결합하는 공유 기질의 메카니즘(A mechanism of covalent substrate binding in the x-ray structure of subunit K of the Escherichia coli dihydroxyacetone kinase)." Proc Natl Aca Sci USA

100:8188-8192.

(비특허문헌 0028) 실하비, T.(Silhavy, T.), 버만, M.(Berman, M.), 엔퀴스트, L.(Enquist, L.) (1984) "유전자 융합으로의 실험(Experiments With Gene Fusions)." 콜드 스프링 하버 레버러토리 프레스, 콜드 스프링 하버, NY(Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY), pp 107-112

(비특허문헌 0029) 수베디, K.P.(Subedi, K. P.), 김, I.(Kim, I.), 김, J.(Kim, J.), 민 B.(Min, B.), 박, C.(Park, C.) (2007) "에세리키아 콜리 K12의 디하이드록시아세톤 및 메틸글리옥살 대사에서의 gldA의 역할 (Role of gldA in dihydroxyacetone and methylglyoxal metabolism of Escherichia coli K12)." FEMS Microbiol Lett 279: 180-187.

(비특허문헌 0030) 수즈키, K.(Suzuki, K.), 왕, X.(Wang, X.), 웨일바커, T.(Weilbacher, T.), 퍼네스티그, A.K.(Pernestig, A. K.), 멜레포스, O.(Melefors, O.), 조질리스, D.(Georgellis, D.), 바비츠크, P.(Babitzke, P.), 로메오, T.(Romeo, T.). (2002) "에세리키아 콜리의 CsrA/CsrB 및 BarA/UvrY 시스템의 조절적 회로(Regulatory circuitry of the CsrA/CsrB and BarA/UvrY systems of Escherichia coli)." J Bacteriol 184:5130-5140.

(비특허문헌 0031) 시바가미선다람, C.(Sivagamisundaram, C.) 우프, E.(Wooff, E.), 콜드함, N.G.(Coldham, N. G.), 스리타란, M.(Sritharan, M.), 휴인슨, R.G.(Hewinson, R. G.), 고돈, S.V.(Gordon, S. V.), 휠러, P.R.(Wheeler, P. R.) (2009) "마이코박테리움 튜버쿨로시스 복합체 내에서의 피루베이트 키나제의 비활성화의 전체적 효과들(Global effects of inactivation of the pyruvate kinase gene in the Mycobacterium tuberculosis complex)." J Bacteriol 191 :7545-7553.

(비특허문헌 0032) 트루니거, V.(Truniger, V.), 부스, W.(Boos, W.) (1994) "에세리키아 콜리의 글리세롤 디하이드로지나제의 구조적 유전자, gldA의 맵핑 및 클로닝(Mapping and cloning of gldA, the structural gene of the Escherichia coli glycerol dehydrogenase)." J Bacteriol 176: 1796-1800.

(비특허문헌 0033) 베무리, G.N.(Vemuri, G. N.), 아이트만, M.A.(Eiteman, M. A.), 알트만, E.(Altman, E.) (2002) "에세리키아 콜리의 대사적으로 조작된 스트레인들에 의한 숙신산 생산에 대한 성장 모드 및 피루베이트 카복실라제의 효과들(Effects of growth mode and pyruvate carboxylase on succinic acid production by metabolically engineered strains of Escherichia coli)." App Environ Microbiol 68: 1715-1727.

(비특허문헌 0034) 왕, Q.(Wang, Q.), 첸 X.(Chen, X.), 양, Y.(Yang, Y.), 자오, X.(Zhao, X.) (2006) 실리콘 도움의 대사적 분석에서의 게놈-스케일 및 숙신산염 생산을 향상시키기 위한 에세리키아 콜리의 유동 비교 (Genome-scale in silico aided metabolic analysis and flux comparisons of Escherichia coli to improve succinate production)." App Microbiol Biotechnol 2006 73: 887-894.

(비특허문헌 0035) 와이커트, M.J.(Weickert, M. J.), 아드히라, S.(Adhya, S.) (1993) "에세리키아 콜리 내에서의 Gal 억제자 및 동위 억제 유전자들의 전사에 대한 컨트롤(Control of transcription of Gal Repressor and isorepressor genes in Escherichia coli)." J Bacteriol 175: 251-258.

(비특허문헌 0036) 장, X.(Zhang, X.), 잔타마, K.(Jantama, K.), 무어, J.C.(Moore, J. C.), 자르보, L.R.(Jarboe, L. R.), 산무감, K. T.((Shanmugam, K. T.), 잉그램, L.O.(Ingram, L. O.) (2009a) "에세리키아 콜리 내에서의 숙신산염 생산에 대한 에너지-보존적 경로들의 대사적 진화(Metabolic evolution of energy-conserving pathways for succinate production in Escherichia coli)." Proc Natl Acad Sci USA 106: 20180-20185.

(비특허문헌 0037) 장, X.(Zhang, X.), 잔타마, K.(Jantama, K.), 산무감, K.T.(Shanmugam, K. T.), 잉그램, L.O.(Ingram, L. O.) (2009b) "미네랄 염 배지 내에서의 숙신산염 생산에 대한 에세리키아 콜리 리엔지니어링(Re-engineering Escherichia coli for succinate production in mineral salts medium)." App Environ Microbiol 75: 7807-7813.

발명의 내용

해결하려는 과제

화합물 합성을 위한 석유화학적 프로세스와 경쟁하기 위하여, 화합물 생산을 위한 더욱 효과적인 생체 촉매들

을 개발하기 위한 기술 분야에서의 요구가 존재한다. 제1 단계는 증진된 화합물 생산에 기여할 수 있는 숙주 염색체에 있어서의 유전적 변화들을 찾아내는 것이다. 이러한 환경의 관점에서, 본 발명의 목적은 상업적인 사용을 위해 선택되는 미생물 스트레인들에 의하여 화합물 생산을 향상시키는데 사용될 수 있는 새로운 돌연변이들을 찾아내는 것이다. 구체적인 실시예에서, 본 발명의 목적은 유기산들의 미생물적 생산의 효율성을 향상시키는 것이다. 더욱 구체적인 실시예에서, 상기 목적은 숙신산 생산의 효율성을 향상시키는 것이다. 본 발명의 또 다른 관점에서, 목적은 미생물로부터 대사적 진화와 같은 노동력-집약적인 방법들에 의존해야 하는 것 없이 일반적으로 숙신산과 같은 화합물들의 생산을 합리적으로 향상시키는 방법을 습득하는 것이다.

[0017] 리버스 엔지니어링(reverse engineering)의 타입이 높은 라이신 생산을 위해 돌연변이화되고 또한 스크린되어져 온 *코리넨박테리움 글루타미쿰(Corynebacterium glutamicum)* 스트레인들에 대하여 수행되어져 왔다. 상당히 개조된 라이신 생산 스트레인의 계놈은 야생형 스트레인의 그것과 비교되며, 또한 그 차이점들은 최소한으로 돌연변이화된 스트레인을 재 조작(re-engineer)하기 위하여 사용되었다(이케다 외.(Ikeda *et al.*), 2006). 그러나, 수백 개의 돌연변이들이 발견되었고, 또한 단지 이들 중 몇 개 만이 라이신 초과 생산에서 주요한 역할을 하는 것으로 보임에도 불구하고, 타당성을 위하여 수백 개의 돌연변이들의 모두를 테스트하는 것은 상당한 비실용적일 것이다. 표면 상으로, 여기 기재되는 본 발명은 이러한 종래 기술에 기재되어 있는 것과 유사한 것으로 보일 수 있으나, 주요 차이는 본 발명의 시작 스트레인(starting strain)은 돌연변이화되지 않았고, 단지 오히려 자발적으로 돌연변이하도록 허용되었다는 것이며, 또한 그것은 스크린되지 않았고, 단지 오히려 대사적 진화의 프로세스 내에서 선택되었다는 것이다. 이것은 본 발명을 구분하는 적어도 두 개의 주요한 차이점들을 도출한다: (1) 돌연변이들의 밀도가 상당히 낮으며, 거의 두 단위(two orders of magnitude) 정도로, 또한 (2) 본 발명에서 테스트되는 돌연변이들의 여덟 개 모두가 향상된 성장 또는 향상된 화합물 생산의 효율 중 어느 것에 대하여 관련되어 있다는 것이 증명된다. 따라서 본 발명의 방법들 및 물건들은 종래 기술의 그것들에 비하여 훨씬 향상된 것이다.

과제의 해결 수단

- [0018] 본 발명은 숙신산과 같은 화합물들을 생산하는 향상된 능력을 갖는 미생물들(microorganisms)을 제공한다.
- [0019] 본 발명은 또한 숙신산과 같은 화합물들을 생산하는 향상된 능력을 갖는 미생물들(microorganisms)을 제조하는 방법을 제공한다.
- [0020] 본 발명은 또한 숙신산과 같은 화합물들을 생산하는 향상된 능력을 갖는 미생물들(microorganisms)을 사용하여 숙신산을 제조하기 위한 방법을 제공한다.
- [0021] 본 발명의 일 실시예에서, 숙신산과 같은 화합물들을 생산하는 향상된 능력을 가지며 또한 돌연변이화된 *pykA* 유전자를 갖는 미생물(microorganism)이 제공된다.
- [0022] 본 발명의 또 다른 실시예에서, 숙신산과 같은 화합물들을 생산하는 향상된 능력을 가지며 또한 돌연변이화된 *pykF* 유전자를 갖는 미생물(microorganism)이 제공된다.
- [0023] 본 발명의 또 다른 실시예에서, 숙신산과 같은 화합물들을 생산하는 향상된 능력을 가지며 또한 돌연변이화된 *pykA* 및 *pykF* 유전자들을 갖는 미생물(microorganism)이 제공된다.
- [0024] 본 발명의 일 실시예에서, 숙신산과 같은 화합물들을 생산하는 향상된 능력을 가지며 또한 비활성화되거나 또한 돌연변이화된 *galS* 유전자들을 갖는 미생물(microorganism)이 제공된다.
- [0025] 본 발명의 일 실시예에서, 숙신산과 같은 화합물들을 생산하는 향상된 능력을 가지며 또한 비활성화되거나 또한 돌연변이화된 *galR* 유전자들을 갖는 미생물(microorganism)이 제공된다.
- [0026] 본 발명의 또 다른 실시예에서, 숙신산과 같은 화합물들을 생산하는 향상된 능력을 가지며 또한 *galP* 유전자의 다중 복제들(multiple copies)을 갖는 미생물(microorganism)이 제공된다.
- [0027] 본 발명의 또 다른 실시예에서, 숙신산과 같은 화합물들을 생산하는 향상된 능력, *galP* 유전자의 다중 복제들(multiple copies) 및 그 세포 내부로의 글루코오스 이송에 대한 돌연변이화된 포스포트랜스퍼라제 시스템(phosphotransferase system)(PTS)을 갖는 미생물(microorganism)이 제공된다.
- [0028] 본 발명의 또 다른 실시예에서, 숙신산과 같은 화합물들을 생산하는 향상된 능력 및 염색체 DNA의 48kb 영역 내에서의 유전자들 *ycdC* 내지 *ycdF*를 포함하는 다중 유전자들을 포함하는 계놈의 큰 영역의 결실(deletion)을 갖는 미생물(microorganism)이 제공된다.

- [0029] 본 발명의 일 실시예에서, 숙신산과 같은 화합물들을 생산하는 향상된 능력 및 그 *rpoA* 유전자 내에서 미스센스 돌연변이(missense mutation)를 갖는 미생물(microorganism)이 제공된다.
- [0030] 본 발명의 또 다른 실시예에서, 숙신산과 같은 화합물들을 생산하는 향상된 능력 및 그 *rpoA* 유전자의 아미노산 포지션 322에서 돌연변이(mutation)를 갖는 미생물(microorganism)이 제공된다.
- [0031] 본 발명의 또 다른 실시예에서, 숙신산과 같은 화합물들을 생산하는 향상된 능력 및 그 *rpoA* 유전자의 아미노산 포지션 322에서 프로린을 류신으로의 돌연변이(proline to leucine mutation)를 갖는 미생물(microorganism)이 제공된다.
- [0032] 본 발명의 또 다른 실시예에서, 숙신산과 같은 화합물들을 생산하는 향상된 능력 및 그 *rpoC* 유전자 내에 돌연변이(mutation)를 갖는 미생물(microorganism)이 제공된다.
- [0033] 본 발명의 또 다른 실시예에서, 숙신산과 같은 화합물들을 생산하는 향상된 능력 및 *rpoC* 유전자의 F 영역 내에 돌연변이(mutation)를 갖는 미생물(microorganism)이 제공된다.
- [0034] 본 발명의 또 다른 실시예에서, 숙신산과 같은 화합물들을 생산하는 향상된 능력 및 *rpoC* 유전자의 F 영역 내 포지션 747에서의 아미노산 내에서 미스센스 돌연변이(missense mutation)를 갖는 미생물(microorganism)이 제공된다.
- [0035] 본 발명의 또 다른 실시예에서, 숙신산과 같은 화합물들을 생산하는 향상된 능력 및 *rpoC* 유전자의 F 영역 내 포지션 747에서 메티오닌을 이소류신으로의 돌연변이(methionine to isoleucine mutation)를 갖는 미생물(microorganism)이 제공된다.
- [0036] 본 발명의 일 실시예에서, 숙신산과 같은 화합물들을 생산하는 향상된 능력을 가지며 또한 그 *gldA* 유전자 내에 돌연변이(mutation)를 갖는 미생물(microorganism)이 제공된다.
- [0037] 본 발명의 일 실시예에서, 숙신산과 같은 화합물들을 생산하는 향상된 능력을 가지며 또한 그 *ftsI* 유전자 내에 돌연변이(mutation)를 갖는 미생물(microorganism)이 제공된다.
- [0038] 본 발명의 또 다른 실시예에서, 숙신산과 같은 화합물들을 생산하는 향상된 능력을 가지며 또한 그 *dhaM* 유전자 내에 돌연변이(mutation)를 갖는 미생물(microorganism)이 제공된다.

발명의 효과

- [0039] (1) 돌연변이들의 밀도가 상당히 낮으며, 거의 두 단위 정도로, 또한 (2) 본 발명에서 테스트되는 돌연변이들의 여덟 개 모두가 향상된 성장 또는 향상된 화학적 생산의 효율 중 어느 것에 대하여 관련되어 있다는 것이 증명된다. 따라서 본 발명의 방법들 및 물건들은 종래 기술의 그것들에 비하여 훨씬 향상된 것이다.

발명을 실시하기 위한 구체적인 내용

- [0040] 본 발명에 개시되는 것은 발효적 프로세스(fermentative process)를 통해 숙신산의 생산을 위하여 적합한 미생물들이다. 본 발명이 유전적으로 변형된 박테리아 스트레인들에 의하여 탄소 화합물로부터 상업적으로 상당한 양으로 숙신산의 생산을 위한 프로세스를 제공한다고 하더라도, 본 발명의 가르침은 다른 많은 수의 산업적 생산에 대하여 동일하게 적용될 수 있다.
- [0041] 본 발명의 기재의 목적을 위하여, 하기의 정의들이 사용될 것이다.
- [0042] 많은 수의 산업적으로 유용한 화합물들이 본 발명을 사용하여 제조될 수 있다. 그러한 화합물들의 예들은 이에 제한되는 것은 아니나, 에탄올(ethanol), 부탄올(butanols), 락트산염(lactate), 숙신산염(succinate), 푸마르산염(fumarate), 말산염(malate), 트레오닌(threonine), 메티오닌(methionine) 및 라이신(lysine)을 포함한다. 유기산들이 자유산들(free acids)으로서 및 염들(예로서, 이에 제한되는 것은 아니나, 소듐(sodium), 포타슘(potassium), 마그네슘(magnesium), 칼슘(calcium), 암모늄(ammonium), 클로라이드(chloride), 설페이트(sulfate), 카보네이트(carbonate), 바이카보네이트(bicarbonate) 등의 염들)로서 양자로 존재할 수 있기 때문에, 숙신산(succinic acid), 푸마르산(fumaric acid), 말산(malic acid), 아스파르트산(aspartic acid), 트레오닌(threonine), 메티오닌(methionine), 및 라이신(lysine)과 같은 화학명들은 그 자유산 및 그들의 어떤 염 양자를 포함하는 것으로 의미될 것이다. 이와 마찬가지로, 숙신산염(succinate), 푸마르산염(fumarate), 말산염(malate), 아스파르트산염(aspartate) 등과 같은 어떤 염은 자유산 역시 포함하는 것으로 의미될 것이다.

- [0043] 본 발명은 탄수화물 기질을 갖는 미네랄 염 배지 내에서의 혐기성 또는 미세 호기성 성장 조건들 하에서 숙신산 생산을 위한 높은 수율, 타이터 및 부피 생산성(volumetric productivity)을 나타내는 스트레인들을 얻기 위하여 특이적 유전자 변형들의 기술(technique of specific genetic modifications)과 대사적 진화(metabolic evolution)의 프로세스를 결합한다.
- [0044] 본 발명에서 사용되는 바와 같이, 용어 "타이터(titer)"는 발효 배지(fermentation broth) 내에서의 특정 화합물의 몰 농도를 의미한다. 따라서 본 발명에 따르는 숙신산의 생산을 위한 발효 프로세스에서, 100 mM의 숙신산 타이터는 측정의 시간에서의 발효 배지가 발효 배지의 리터(liter) 당 100 mMoles의 숙신산을 함유하였다는 것을 의미할 것이다.
- [0045] 본 발명에서 사용되는 바와 같이, 용어 "수율(yield)"은 발효 프로세스 동안에 소모되는 공급 재료(feedstock)의 몰 당 생산되는 특정 화합물의 몰을 의미한다. 따라서 공급 재료로서 글루코오스를 사용하는 숙신산의 생산을 위한 발효적 프로세스에서, 상기 용어 수율은 소모되는 글루코오스 몰 당 생산되는 숙신산의 몰 수(the number of moles)를 의미한다.
- [0046] 본 발명에서 사용되는 바와 같이, 용어 "부피 생산성(volumetric productivity)"은 단위 시간 당 단위 부피 당 생산되는 그램으로서의 특정 화합물의 양을 의미한다. 따라서 $0.9 \text{ g L}^{-1} \text{ h}^{-1}$ 의 숙신산의 부피 생산성 값은 0.9 그램의 숙신산이 성장의 1 시간 동안 발효 배지의 1 리터 내에 축적된다는 것을 의미할 것이다.
- [0047] 여기 기재된 바와 같은 용어들 "유전적으로 엔지니어링된(genetically engineered)" 또는 "유전적으로 변형된(genetically modified)"은 그 미생물들의 게놈 DNA를 조작하는 것을 통해 그 미생물들 내에서의 하나 이상의 효소들의 발현을 개조하는 것의 실행을 의미한다.
- [0048] 본 발명에서 사용되는 바와 같이, 용어 "유전자(gene)"는 그 유전자의 개방 리딩 프레임(reading frame) 및 상부 및 하부 조절 서열들(upstream and downstream regulatory sequences)을 포함한다. 상기 상부 조절 영역은 또한 그 유전자의 프로모터 영역을 의미한다. 상기 하부 조절 영역은 또한 그 터미네이터 영역을 의미한다.
- [0049] "대립 유전 형질 Allele)"은 특정 유전자의 DNA 서열의 둘 또는 초과 형태들 중 하나이다. 어떠한 돌연변이도 없는 유전자는 돌연변이를 포함하는 상응하는 유전자와 비교될 때 야생형 대립 형질(wild type allele)로서 참조된다.
- [0050] "상동체(Homolog)"는 공통의 조상의(ancestral) DNA 서열로부터 전해져 온 이차적 유전자(second gene)에 관련된 유전자이다. 이 용어, 상동체는 종분화(speciation)의 사건에 의하여 구분되는 유전자들 사이의 관계에 대하여 또는 유전적 듀플리케이션의 사건에 의하여 구분되는 유전자들 사이의 관계에 대하여 적용될 수 있다. "오쏘로그들(Orthologs)"은 종분화에 의해 공통의 조상 유전자로부터 진화된 다른 종(species) 내의 유전자들이다. 보통, 오쏘로그들은 진화의 과정에서 동일한 기능을 유지한다. 오쏘로그들에 대한 식별(identification)은 새롭게 서열화된 게놈들 내에서의 유전자 기능에 대한 신뢰할 만한 예측을 위하여 중요하다. 종분화는 그것이 발생하는 종으로부터의 새로운 방법으로 살아가는 것을 가능하게 하는 새로운 종의 기원이다. 이러한 프로세스의 일부로서 그것은 그 부모 종과의 유전적 교환에 대하여 어떤 장벽을 또한 획득하였다. "파라로그들(Paralogs)"은 게놈 내부에서의 듀플리케이션에 의해 관련되는 유전자들이다. 오쏘로그들은 진화의 과정에서 동일한 기능을 유지하는 반면, 파라로그들은 그들이 그 고유의 것에 관련되어 있기는 함에도 불구하고, 새로운 기능들을 진화시킨다.
- [0051] 구절 "기능적으로 유사한(functionally similar)"은 여기 기재된 방법들에 의하여 동일한 또는 다른 생물 내에서 발견되는 어떤 야생형 또는 돌연변이화된 DNA 서열, 유전자, 효소, 단백질과 동일하거나 또는 유사한 생물학적 기능을 갖는, 어떤 생물(organism)로부터의 어떤 야생형 또는 돌연변이화된 DNA 서열, 유전자, 효소, 단백질을 광범위하게 의미한다. 기능적으로 유사성은 서열 상동(homology)을 요구할 필요가 없다. 예로서, 본 발명에서, 우리는 *이. 콜리(E. coli.)* 내에서 총 피루베이트 키나제(pyruvate kinase) 활성을 감소시키는, 그러나 삭제하지는 않는 *pykA* 내에서의 돌연변이가 보다 효과적인 숙신산염 생산을 유도한다는 것을 발견하였다. 확장하여, 피루베이트 키나제를 감소시키는, 생물체의 또 다른 타입, 예로서 이스트 *사카로마이세스(yeast Saccharomyces)*, 에서의 돌연변이는 기능적으로 유사하며, 또한 *사카로마이세스(Saccharomyces)* 내부의 관련성 있는 유전자, 예로서, PYK1 또는 PYK2 유전자는 *이. 콜리(E. coli.)*의 *pykA* 유전자와 기능적으로 동일할 것이다.
- [0052] "개조된 활성(altered activity)"을 갖는 유전자 또는 단백질은 관련성 있는 야생형의 유전자 또는 단백질과

비교되었을 때에 측정 가능한 특성에서 측정 가능한 차이점을 생산하는 유전자 또는 단백질로서 광범위하게 정의된다. 상기 개조된 활성은 개조된 유전자 또는 단백질을 포함하는 스트레인의 숙신산염 생산에 대한 성장 속도 또는 효율을 증가시키거나 또는 감소시키는 일반적인 방법에서 그 자체로 명확할 수 있다. 다른 측정 가능한 특성들은, 그러나 이에 제한되는 것은 아닌, 효소 활성, 효소의 기질 특이성, 기질에 대한 친화도 또는 속도와 같은 효소의 동력학적 파라미터들, 효소의 안정도, 효소의 조절 특성들, 유전자 발현 수준, 다양한 조건들 하에서의 유전자 발현의 조절 등을 포함한다.

[0053] 본 발명에서 사용되는 바와 같이, 상기 용어 돌연변이는 개방 리딩 프레임, 상부 조절 영역 및 하부 조절 영역을 포함하는 유전자에 대하여 이루어지는 유전자 변형들을 의미한다. 상기 유전자 돌연변이들은 상방 조절(up regulation) 또는 하방 조절(down regulation) 또는 그 유전자의 개방 리딩 프레임의 전사의 완전한 저해 중 어느 것을 유도한다. 상기 유전자 돌연변이들은, 그 유전자의 전체적인 코딩하는 영역 또는 그 코딩하는 뉴클레오티드 서열 부분을 삭제함으로써, 또는 프레임 이동 돌연변이(frame shift mutation), 미스센스 돌연변이(missense), 또는 삽입(insertion)을 도입함으로써, 또는 중지 코돈(stop codon) 또는 그들의 조합들을 도입함으로써의 어느 것에 의하여 달성된다. 돌연변이들은 효소 반응 또는 세포 멤브레인을 관통하는 유기 분자들의 이송과 같은 생물학적 기능들에 직접적으로 관련되는 단백질들에 대하여 코딩하는 구조 유전자들(structural genes)에서 발생할 수 있다. 대체적으로, 돌연변이들은 그 생물학적 기능들에 직접적으로 관련된 단백질에 대하여 코딩하는 유전자들의 발현을 조절하는 단백질들에 대하여 코딩하는 조절 유전자들(regulatory genes)에서 발생할 수 있다. 다른 유전자들의 발현을 조절하는 단백질들은 조절 단백질들(regulatory proteins)로서 참조되며, 또한 이러한 조절 단백질들에 대하여 코딩하는 유전자들은 조절 유전자들(regulatory genes)로서 참조된다.

[0054] "돌연변이(Mutation)"는 또한 관련성 있는 야생형 생물의 그것에 대하여 상대적인 DNA 서열 내에서의 어떤 변화를 포함한다. 예로서, 스트레인 KJ122 내에서 발견되는 돌연변이는, 그 돌연변이된 영역의 DNA 서열이 부모의 야생형 스트레인, *이. 콜리(E. coli.) C*, 또한 ATCC 8739로도 알려져 있는, 의 그것과 비교될 때 발견될 수 있는 DNA 서열 내에서의 어떤 변화이다. 돌연변이는 어떤 수의 염기 짝들(base pairs)에 대한 추가적인 DNA의 삽입 또는 어떤 수의 염기 짝들(base pairs)에 대한 DNA의 결실(deletion)일 수 있다. 삽입 돌연변이의 특정한 타입은 유전자 듀플리케이트(gene duplicate)이다. 유전자는 자발적인 돌연변이적 사건에 의해 듀플리케이트될 수 있고, 이때 상기 유전자의 이차적 카피(secondary copy)는 고유 카피(original copy)에 인접하여 위치될 수 있거나, 또는 유전자는 유전적 조작에 의해 듀플리케이트 될 수 있고, 여기서 상기 유전자의 이차적 카피(secondary copy)는 상기 고유 카피(original copy)로부터 이격되어 있는 게놈 내의 사이트에 위치될 수 있다. 돌연변이는 예로서 아데닌으로부터 구아닌 염기로의 변화와 같이, 하나의 염기 타입으로부터 또 다른 염기 타입으로의 변화일 수 있다. 유전학의 표현(vernacular)에서, 돌연변이는 미스센스(missense)(코돈에 의하여 그를 위해 코딩된 아미노산을 변화시키는), 비인식의(nonsense)(코돈을 중지 코돈으로 변화시키는), 프레임 이동(frameshift)(세 개의 다중적인 것이 아닌 염기들의 수의 삽입 또는 결실인, 또한 상기 리딩 프레임을 변화시키고 및 돌연변이로부터 하부로 인코딩된 아미노산 서열을 개조하는, 또한 종종 상기 돌연변이로부터 하부로 중지 코돈을 도입하는), 또한 변환(inversion)(극성(polarity)에서 스위치된 그러나 삭제되지 않는 DNA 서열로부터 도출되는) 일 수 있다.

[0055] "제로 돌연변이(null mutation)"는 그 관련성 있는 유전자의 전체적인 개방 리딩 프레임의 결실의 그것과 실질적으로 동일한, 또는 그 관련성 있는 유전자의 모든 측정 가능한 활성을 제거하는, 표현형(phenotype)을 부여하는 돌연변이이다.

[0056] "돌연변이체(mutant)"는 그것의 게놈이 하나 이상의 돌연변이들을 포함하는 미생물이다.

[0057] 본 발명에서 사용되는 바와 같이, 용어 "외인성(exogenous)"은 세포의 외부로부터 유도되는 분자 또는 활성이 그 숙주 미생물체 내부로 도입되는 것을 의미하는 것으로 의도된다. 상기 미생물 세포 내부로 도입되는 외인성 핵산 분자의 경우에 있어서, 상기 도입된 핵산은 독립적인 플라스미드로서 존재할 수 있거나 숙주 염색체 DNA 내부로 집적될 수 있다(integrated). 단백질에 대하여 코딩하는 외인성 핵산은 프로모터 또는 터미네이터 서열들과 같은 그 고유의 조절 서열들과 함께 발현 가능한 형태로 미생물 세포 내부로 도입될 수 있다. 대체적으로, 상기 외인성 핵산 분자는 숙주 염색체 DNA 내부로 집적될 수 있고 또한 상기 숙주 조절 서열들의 조절 하에 있을 수 있다.

[0058] 용어 "내인성(endogenous)"은 그 숙주 세포 내부에 존재하는 분자들 또는 활성을 의미한다. 생합성적 활성에 대한 참조로서 사용될 때, 상기 용어 "외인성(exogenous)"은 그 숙주 기준 생물 내부로 도입되는 활성을 의미한다. 상기 소스(source)는 예로서 상기 숙주 미생물체 내부로의 도입에 따라 기준이 되는 활성을 발현하는

상동의(homologous) 또는 이종의(heterologous) 인코딩하는 핵산일 수 있다. 만약 단백질에 대하여 코딩하는 핵산이 미생물체의 동일한 종으로부터 얻어지는 것이라면, 그것은 상동성 DNA로서 참조된다. 만약 상기 핵산이 다른 미생물 종으로부터 유도되는 것이라면, 그것은 이종의 DNA로서 참조된다. 상기 DNA의 성질과 관계없이, 그것이 상동이거나 또는 이종의 것이거나, 숙주 세포 내부로 도입될 때에, 상기 DNA 및 도입된 DNA로부터 유도되는 활성은 외인성인 것으로 참조된다. 따라서, 본 발명의 인코딩하는 핵산의 외인성 발현은 이종의 및 상동의 양자의 인코딩하는 핵산 중 어느 것을 이용할 수 있다.

[0059] "미생물(Microorganism)"은 어떤 박테리아(bacterium), 이케온(archeon), 이스트(yeast), 사상균(filamentous fungus), 단세포 알가(unicellular alga), 또는 쌍편모충류(dinoflagellate)를 포함할 수 있다.

[0060] 본 발명에 대하여 적합한 재조합 미생물들은 많은 수의 박테리아 과들(bacterial families)으로부터, 바람직하게는 엔테로박테리아케 과(*Enterobacteriaceae* family)으로부터 유도된다. 상기 적합한 미생물들은 에세리키아(*Escherichia*), 에르위니아(*Erwinia*), 프로비덴시아(*Providencia*), 및 세라티아(*Serratia*) 속(genera)으로부터 선택된다. 상기 에세리키아(*Escherichia*) 속이 특히 바람직하다. 에세리키아(*Escherichia*) 속 내부에서의 에세리키아 콜리(*Escherichia coli*) 종이 특히 바람직하다. 이. 콜리 B(*E. coli* B), 이. 콜리 씨(*E. coli* C), 이. 콜리 더블유(*E. coli*) W 또는 기타와 같은 어떤 하나의 스트레인이 본 발명에 대하여 유용하다.

[0061] 상당한 양으로 유기산들을 생산하는 것을 가능하게 하는 이. 콜리(*E. coli*) 스트레인은 본 기술 분야에서 잘 알려져 있다. 예로서, 미국 특허 출원 공개 번호 2009/0148914는 화학적으로 순수한 아세트산염 및/또는 피루브산염의 생산에 대하여 생체 촉매로서의 이. 콜리(*E. coli*)의 스트레인들을 제공한다. 미국 특허 7,629,162는 락트산의 생산을 위하여 제조된 이. 콜리 K011(*E. coli* K011) 스트레인의 유도체들을 제공한다. 특허 협정 조약 번호들 WO 2008/115958 및 WO 2010/115067 하에 공개된 국제 특허 출원들은 pH-조절된 배지 발효 내에서 탄소의 소스로서 글루코오스를 포함하는 최소한의 미네랄 염 배지 내에서 숙신산염 및 말산염을 생산하도록 엔지니어링된 미생물을 제공한다.

[0062] 본 발명의 어떤 실시예들에서, 본 발명에 따라 변형될 수 있는 박테리아는 이에 제한되는 것은 아니나 아크로모박터 델마바(*Achromobacter delmarvae*), 아크로모박터 비스코서스(*Achromobacter viscosus*), 아크로모박터 락티쿰(*Achromobacter lacticum*), 악티노마두라 마두라(*Actinomadura madurae*), 악티노마이세스 바이오아케오크로모진스(*Actinomyces violaceochromogenes*), 에어로모나스 살모니시다(*Aeromonas salmonicida*), 아그로박테리움 투메파시언스(*Agrobacterium tumefaciens*), 아그로박테리움 라디오박터(*Agrobacterium radiobacter*), 알칼리진스 페칼리스(*Alcaligenes faecalis*), 아르트로박터 시트레우스(*Arthrobacter citreus*), 아르트로박터 투메스스(*Arthrobacter tumescens*), 아크로박터 파라피누스(*Arthrobacter paraffineus*), 아크로박터 하이드로카보글루타미커스(*Arthrobacter hydrocarboglutamicus*), 아크로박터 옥시단스(*Arthrobacter oxydans*), 아우레오박테리움 사퍼다(*Aureobacterium saperae*), 아조박터 인디커스(*Azotobacter indicus*), 바실러스 아밀로리퀴파시언스(*Bacillus amyloliquefaciens*), 바실러스 코아굴란스(*Bacillus coagulans*), 바실러스 서큘란스(*Bacillus circulans*), 바실러스 리체니포미스(*Bacillus licheniformis*), 바실러스 퍼밀러스(*Bacillus pumilus*), 바실러스 서브틸리스(*Bacillus subtilis*), 바실러스 티아미노리티커스(*Bacillus thiaminolyticus*), 브레비박테리움 암모니아진스(*Brevibacterium ammoniagenes*), 디베리카툼(*divaricatum*), 브레비박테리움 락토퍼멘툼(*Brevibacterium lactofermentum*), 브레비박테리움 플라븀(*Brevibacterium flavum*), 브레비박테리움 글로보썸(*Brevibacterium globosum*), 브레비박테리움 퍼스쿰(*Brevibacterium fuscum*), 브레비박테리움 케토글루타미쿰(*Brevibacterium ketoglutamicum*), 브레비박테리움 헬로콜룸(*Brevibacterium helvolum*), 브레비박테리움 퍼실룸(*Brevibacterium pusillum*), 브레비박테리움 테스타세움(*Brevibacterium testaceum*), 브레비박테리움 로세움(*Brevibacterium roseum*), 브레비박테리움 이마리오필리움(*Brevibacterium immariophilum*), 브레비박테리움 리넨스(*Brevibacterium linens*), 브레비박테리움 프로토파미아(*Brevibacterium protopharmiae*), 시트로박터 프룬디(*Citrobacter freundii*), 코린박테리움 아세토폴룸(*Corynebacterium acetophilum*), 코린박테리움 글루타미쿰(*Corynebacterium glutamicum*), 코린박테리움 칼루나(*Corynebacterium callunae*), 코린박테리움 아세토아시도필룸(*Corynebacterium acetoacidophilum*), 코린박테리움 아세토글루타미쿰(*Corynebacterium acetoglutamicum*), 엔테로박터 에어로진스(*Enterobacter aerogenes*), 에르위니아 아밀로보라(*Erwinia amylovora*), 에르위니아 카로토보라(*Erwinia carotovora*), 에르위니아 허비콜라(*Erwinia herbicola*), 에르위니아 크리산테미(*Erwinia chrysanthemi*), 에세리키아 프룬디(*Escherichia freundii*), 프라보박테리움 페레그리눔(*Flavobacterium peregrinum*), 플라보박테리움 퍼카툼(*Flavobacterium fucatum*), 플라보박테리움 오란티눔(*Flavobacterium aurantinum*), 플라보박테리움 레나눔(*Flavobacterium rhenanum*), 플라보박테리움 세와넨스(*Flavobacterium sewanense*), 플라보박테리움 브레베(*Flavobacterium breve*), 플라보박테리움 메닝고셉티쿰(*Flavobacterium meningosepticum*), 글루코노박터 옥시단스(*Gluconobacter oxydans*), 글루코노박터 아사이

(*Gluconobacter asaii*), *키타사토스포리아 파룰로사(Kitasatosporia parulosa)*, *클레브시엘라 옥시토카(Klebsiella oxytoca)*, *클레브시엘라 뉴모니아(Klebsiella pneumoniae)*, *마이크로박테리움 암모니아필룸(Microbacterium ammoniaphilum)*, *마이크로코커스* 종 CCM825(*Micrococcus* sp. CCM825), *모르가넬라 모르가니(Morganella morganii)*, *노카디아 오파카(Nocardia opaca)*, *노카디아 루고사(Nocardia rugosa)*, *플라노코커스 유시나투스(Planococcus eucinatus)*, *프로테우스 레트게리(Proteus rettgeri)*, *프로피오니박테리움 셰르마니(Propionibacterium shermanii)*, *슈도모나스 신크산타(Pseudomonas synxantha)*, *슈도모나스 아조토포만스(Pseudomonas azotoformans)*, *슈도모나스 플루오레스스(Pseudomonas fluorescens)*, *슈도모나스 오발리스(Pseudomonas ovalis)*, *슈도모나스 스투트제리(Pseudomonas stutzeri)*, *슈도모나스 아시도볼란스(Pseudomonas acidovolans)*, *슈도모나스 머시도렌스(Pseudomonas mucidolens)*, *슈도모나스 테스트스테로니(Pseudomonas testosteroni)*, *슈도모나스 에어루기노사(Pseudomonas aeruginosa)*, *로도코커스 에리트로폴리스(Rhodococcus erythropolis)*, *로도코커스 로도크루스(Rhodococcus rhodochrous)*, *로도코커스* 종 ATCC 15592(*Rhodococcus* sp. ATCC 15592), *로도코커스* 종 ATCC 19070(*Rhodococcus* sp. ATCC 19070), *세라티아 마세스스(Serratia marcescens)*, *스포로사키나 우레아(Sporosarcina ureae)*, *스타필로코커스 아우레우스(Staphylococcus aureus)*, *스트렙토마이세스 코엘리컬러(Streptomyces coelicolor)*, *스트렙토마이세스 플라벨루스(Streptomyces flavelus)*, *스트렙토마이세스 그리세올러스(Streptomyces griseolus)*, *스트렙토마이세스 리비단스(Streptomyces lividans)*, *스트렙토마이세스 올리바세우스(Streptomyces olivaceus)*, *스트렙토마이세스 타나시엔시스(Streptomyces tanashiensis)*, *스트렙토마이세스 버지니아(Streptomyces virginiae)*, *스트렙토마이세스 안티비오티커스(Streptomyces antibioticus)*, *스트렙토마이세스 카카오(Streptomyces cacaoi)*, *스트렙토마이세스 라벤둘라(Streptomyces lavendulae)*, *스트렙토마이세스 비리도크로모진스(Streptomyces viridochromogenes)*, *살모넬라 티피뮤리움(Salmonella typhimurium)*, *살모넬라 쇼트몰레리(Salmonella schottmulleri)*, *비브리오 메츠니코비(Vibrio metschnikovii)*, *비브리오 티로진스(Vibrio tyrogenes)*, *크산토모나스 시트리(Xanthomonas citri)* 및 기타의 것을 포함한다.

[0063] 본 발명은 발효 프로세스에 대한 기질로서 탄소원을 포함하는 미네랄 염 배지 내에서 발효적인 조건들 하에서 성장되는 숙신산에 대하여 인상적인 타이터, 높은 수율 및 상당한 부피 생산성을 나타내는 유전적으로 변형된 박테리아 스트레인들을 제공한다. 그 대상 발명의 미생물들은 헥소스(hexoses), 펜토스(pentoses), 디사카라이드(disaccharides) 및 글리세롤(glycerol)과 같은 다른 탄소 화합물들과 같은 다양한 슈거들을 사용하는 단일 단계 생산 프로세스에서 사용될 수 있다.

[0064] 어떤 내용들 중에서의 용어 "발효(fermentation)" 및 "발효하다(ferment)"는 생물체에 의한 혐기성 성장 또는 대사를 의미한다. 그러나, 본 개시에서, 우리는 용어들 "발효(fermentation)" 및 "발효하다(ferment)"가 호기성, 혐기성, 또는 미세 호기성, 또는 이들의 조합을 포함하는 어떤 조건 하에서의 미생물들의 상업적 또는 실험적 성장으로 광범히 하게 참조될 수 있는 것으로 사용할 것이다.

[0065] 많은 화합물들의 생합성 생산은 상기 생산 생물체의 성장이 산소 또는 공기가 결여되거나 또는 제한된 조건들 하에서 수행될 때 더욱 효과적으로 진행될 수 있다(예로서 더 높은 수율로서). 이것은 대부분 산소의 존재가 일반적으로 상대적으로 낮은 값 부산물인 이산화탄소로의 탄소원들의 대사의 결과를 가져오기 때문이다. 산소 또는 공기의 고의적인 공급의 결여에서의 발효가 통상적으로 "혐기성(anaerobic)"으로 불린다. 그러나, 엄격한 혐기성 조건들에 도달하는 것은 비용이 많이 들고 및 종종 달성하기가 어렵다. 나아가, 어떤 발효들을 위해, 엄격한 혐기성 조건들은 요구되지 않거나 또는 종종 최적의 것이 아니다. 산소 또는 공기가 엄격하게 결여되도록 이루어진 것이 아닌, 또는 산소 또는 공기가 고의적으로 낮은, 조절된 속도로 공급되는 발효에 대하여, 우리는 용어 "미세 호기성(microaerobic)"을 사용할 것이다.

[0066] 본 발명의 실행을 위한 적합한 미생물들은 호기성적으로(산소의 존재하에) 또는 혐기성적으로(산소의 완전한 결여에서) 또는 미세 호기성적으로(최소량의 산소 공급을 가지고) 성장될 수 있다. 본 발명의 바람직한 실시예에서, 숙신산의 생산을 위해 선택되는 미생물은 혐기성 조건에서 성장된다. 대체적으로 본 발명을 위한 적합한 미생물들은, 상기 미생물이, 상업적으로 상당한 양으로 숙신산의 생산에 도달하기 위하여 그것을 혐기성 성장 조건으로 이동시키기 전에 세포 성장의 어떤 수준에 도달할 때까지 호기성 성장 조건에서 최초로 성장되는, 이중-상 성장 레짐(double-phase growth regime) 내에서 성장될 수 있다. 본 발명의 미생물들에 의한 숙신산의 생산에 대한 이중-상 성장 과정에서, 숙신산의 생산 및 축적은 혐기성 발효적 성장 상 동안에 발생한다.

[0067] 본 발명에서, 유전자 돌연변이들의 고유의 및 유리한 조합들이 숙신산 생산으로의 탄소 흐름을 지시하기 위하여 사용되어 왔다. 이에 더하여, 상기 숙신산 생산은, 차례로 세포의 ATP 수준 및 레독스 균형과 결합되어 있

는 미생물 성장과 결합되어 있다.

[0068] 상기 용어 "레독스 균형(redox balance)"은 NAD^+ 에 대한 NADH의 적합한 비율을 유지하는 세포의 능력을 의미한다. 다른 의미에서, 상기 세포들은 NADH를 산화할 수 있어 그에 따라 혐기성 발효적 성장 동안에 탄수화물 기질들을 산화시키기 위한 충분한 NAD^+ 가 존재한다. 호기성 성장 과정에서, 상기 NAD^+ 풀(pool)은 NADH를 포함하는 산화성 인산화(phosphorylation)을 통하여 재생산된다. 그러나, 혐기성 성장 조건들 하에서 상기 NAD^+ 풀(pool)의 재생산은 NADH를 산화시킬 수 있는 세포 내부에서의 다양한 대사적 경로들을 통하여 탄소의 흐름을 조작하는 방법에 의하여만 단지 달성될 수 있다.

[0069] "탄소 분해 대사질 억제(Carbon catabolite repression)"는 하나 이상의 유전자들의 발현들이 그 상에서 생물체가 성장하는 탄소원에 의하여 부분적으로 조절되는 어떤 생물학적 시스템으로 광범위하게 참조된다. 예로서, *이. 콜리(E. coli)* 내에서, 야생형 생물체 내에서의 당신생(gluconeogenesis)으로 기능하는 유전자들은 그 배지 내의 글루코오스의 존재에 의해 억제된다. 특히, *pck* 유전자는 대부분의 야생형 *이. 콜리(E. coli)* 스트레인들 내에서 글루코오스에 의하여 하향 조절된다. 또 다른 예로서, 이스트 *사카로마이세스 세레비시아(yeast Saccharomyces cerevisiae)*의 야생형 스트레인들 내에서, 갈락토오스(galactose)와 같은 글루코오스 이외의 다른 탄소원들의 사용은 그 배지 내의 글루코오스 존재에 의하여 저해된다.

[0070] "트리카르복실산 회로(tricarboxylic acid cycle)" 또는 "TCA" 회로는 대부분의 미생물들 내에서 전체 또는 부분적으로 존재하는 생화학적 반응들의 세트이다. 많은 생물체에서, 예로서 유박테리아(eubacteria) 및 이스트(yeast), 상기 TCA 회로는 호기성 성장 과정에서 실질적(true) 회로로서 수행한다. 이러한 경우에서, 상기 TCA 회로는 두 개의 탄소 단위들(아세테이트와 같은)을 이산화탄소로 이화하는, 또한 글루탐산염(glutamate), succinate(숙신산염), 및 aspartate(아스파르트산염)과 같은 생합성에 대한 생화학적 중간체들을 생산하는, 두 개의 주요 기능들을 가진다. 혐기성 또는 미세 호기성 성장 과정에서, 상기 TCA 회로는 실질적인 회로로서 수행되지 않을 수 있고, 오히려 하나 이상의 선형적 경로들이 수행될 수 있다. 예로서, *이. 콜리(E. coli)* 내에서 낮은 산소의 조건들 하에서 상기 TCA 회로 효소들은, 말산염을 통하여 숙신산염을 유도하는 소위 환원성 경로, 및 시트레이트를 통하여 α -케토글루타르산염(α -ketoglutarate)을 유도하는 소위 산화성 경로를 포함하는, 두 개의 분지된 경로로서 작동한다. 이러한 경우에서, 상기 TCA 회로의 기능은 주로 생화학적 중간체들 글루탐산염(glutamate), 숙신산염(succinate), 및 아스파르트산염(aspartate)을 생산하는 것이다. 어떤 탄소는 글리옥실산 섀트(glyoxylate shunt)를 통하여 두 개의 브랜치들 사이에서 이송될 수 있다. 호기성, 혐기성, 및 미세 호기성 성장 과정에서, 생합성을 위하여 사용되는 생화학적 중간체의 각 분자에 대하여, 상기 TCA 회로 내에서의 분자는 대체되어야 하며, 그렇지 않을 경우 상기 회로 내에서의 화합물들은 고갈될 것이다. 이러한 대체 반응은 아나플루로틱 반응(anapleurotic reaction)이라 불린다.

[0071] 통상적인 아나플루로틱 반응들은, 피루브산염 또는 포스포에놀 피루브산염(phosphoenol pyruvate(PEP))과 같은 세 개 탄소의 화합물을 사용하고, 또한 이차적 기질로서 이산화탄소를 사용하여 그것을 옥살로아세트산염(oxaloacetate) 또는 말산염(malate)과 같은 네 개 탄소의 화합물로 전환한다.

[0072] 유박테리아(eubacteria) 또는 이스트(yeast)와 같은 많은 생물체들 내에서, TCA 회로 및 글리옥실산 섀트(glyoxylate shunt)의 효소들을 인코딩하는 유전자들의 발현은 글루코오스에 의하여 억제되고 및/또는 산소에 의하여 유도된다(크로난 및 라포르테(Cronan and Laporte), 1996). 이러한 조절의 컨트롤은 복잡하다. 문헌에 따르면, 숙신산염 및 다른 화합물들의 효율적인 생산은, 특히 낮은 산소의 조건들 하에서, 환원성 TCA 회로 효소들과 상기 글리옥실산 섀트의 그것들의 미세한 균형을 요구한다(버뮤리 외.(Vermuri et al.), 2002; 왕 외.(Wang et al.), 2006; 산체즈 외.(Sanchez et al.) 2005).

[0073] 그러나, 본 기술 분야의 상황은, 스트레인들을 최적의, 또는 최적에 근접하게, 숙신산염과 같은 화합물들의 생합성을 위한 TCA 회로 및 글리옥실산 섀트 경로들의 사용을 달성할 수 있도록 합리적으로 디자인할 만큼 충분히 발전되어 있지 않다. 본 발명의 목적들 중 하나는 TCA 회로의 화합물의 생산을 위하여 최적에 근접하게 상당히 진화된 미생물을 DNA 서열의 수준에서 절단하는 것으로, 이에 따라 미래에, 더욱 합리적인 스트레인 디자인 및 제조가 숙신산염 및 다른 화합물들의 생산을 위한 스트레인들을 엔지니어링하는데 있어서 사용될 수 있다. 본 발명에서 발견된 돌연변이들은, 여기 기재된 돌연변이들이, 그러한 화합물들을 생산하는 야생형 생물체, 유전자들 및 효소들이 복잡한 방법들로 네거티브하게 조절되는 글루코오스(또는 다른 슈거) 바탕의 배지 내 낮은 산소의 조건들 하에서 옥살로아세트산염으로 및 그로부터의 대사적 흐름을 증가시키는데 있어서 유용하기 때문에, TCA 회로 중간체들, 및 TCA 회로 중간체들의 유도체들로서, 이에 제한되는 것은 아니나, 숙신산염(succinate), 푸마르산염(fumarate), 말산염(malate), 글루탐산염(glutamate), 아스파르트산염

(aspartate), 트레오닌(threonine), 라이신(lysine), 메티오닌(methionine) 및 이소류신(isoleucine)을 포함하는, 특히 옥살로아세트산염(oxaloacetate)으로부터 유도되는 화합물들을 포함하는, TCA 회로와 관련된 화합물들의 생합성을 향상시키는데 있어 유용하다.

[0074] 박테리아에서 자주 발견되는 슈거들의 이송을 위한 하나의 메카니즘은 포스포트랜스퍼라제 시스템(phosphotransferase system (PTS))이다. 이것은 두 개-에너지-커플링 단백질들(two-energy-coupling proteins), 엔자임 I 및 HPr(Enzyme I and HPr), 및 몇 개의 슈거-특이적 엔자임 II 단백질들(sugar-specific Enzyme II proteins) 또는 단백질 복합체들(protein complexes)로 구성되며, 이는 전형적으로 세 개의 단백질 도메인들, EIIA, EIIB 및 EIIC로 이루어진다. EII 도메인들의 구조는 박테리아 사이에서 다르다. EII는 하나의 융합된 단백질 또는 다른 융합된 및 융합되지 않은 도메인들로 이루어진다. 멤브레인을 통한 특정 슈거의 위치이송(translocation)은 인테그랄 멤브레인 도메인(들) EIIC(integral membrane domain(s) EIIC) 및 종종 EIID에 의하여 촉진된다. 그러나, 함께 작용하여 상기 슈거 기질의 이송 및 인산화를 가져오며, 결과적으로 인산화된 탄수화물의 세포 간 풀(intracellular pool)을 유도하는 것은, 상기 세 개의 효소 도메인들 또는 단백질들의 복합체이다. 상기 PTS 를 위한 에너지 및 포스페이트의 소스는 포스포에놀 피루브산염(phosphoenol pyruvate (PEP))이다. 그러나 PEP 는 또한 말산염, 푸마르산염, 숙신산염, 아스파르트산염, 및 그러한 네 개-탄소 화합물들로부터 만들어지는 다른 화합물들과 같이 네 개-탄소 화합물들의 생합성을 위한 기질일 수 있고, 따라서 PTS 및 생합성 경로들 사이에서의 PEP에 대한 세포 내부에서의 경쟁이 존재할 수 있다. *이. 콜리(E. coli)* 내부에서, 글루코오스에 대한 PTS 시스템은 *ptsH*, *I*, *G* 및 *crr* 유전자들에 의해 코딩된다. 슈거들을 PTS를 포함하지 않는 미생물들로 이송시키기 위한 다른 메카니즘들이 존재한다. 우리는 기질로서 PEP를 직접적으로 사용하지 않는 이러한 다른 이송 메카니즘에 관여되는 단백질들을 "비-PTS(non-PTS)" 슈거 이송자(sugar transporters)라 부른다. 비-PTS 이송자들의 예들은 *이. 콜리(E. coli)*의 GalP 및 XylE와 같은 프로톤 심포터들(proton symporters), *사카로마이스 세레비시아(Saccharomyces cerevisiae)*의 HEX1,2,6, 및 7 및 *짐모모나스 모빌리스(Zymomonas mobilis)*의 Glf와 같은 촉진된 디퓨저들(facilitated diffusers), 및 *이. 콜리(E. coli)*의 XylFGH와 같은 ABC-타입 이송자들(ABC-type transporters)이다.

[0075] 박테리아의 슈거 이송자들의 가장 큰 분류군(group)은 ATP 결합 카세트(ABC) 이송자들로 알려져 있다. 상기 이름이 의미하는 바와 같이, ABC 이송자들은 박테리아 세포 내로 이송되는 슈거의 모든 분자에 대하여 APT 분자를 요구한다. XylFGH는 자일로스(xylose), 펜토스 슈거(pentose sugar)의 세포 내부로의 이송에 대한 ABC 이송자이다. AraFGH는 아라비노스(arabinose), 역시 또 다른 펜토스 슈거(pentose sugar)의 이송에 대한 이송자이다.

[0076] 비-PTS 박테리아 슈거 이송자들의 다른 타입은 주요 촉진자 상과(Major Facilitator Super family (MFS)) 하에 분류된다. 상기 MFS 슈거 이송자들 내에서, 이송자들의 두 개의 다른 카테고리들이 인식되어 있다. MFS 는 H^{+} -연결된 심포터들(symporters), Na^{+} -연결된 심포터들(symporters)-안티포터들(antiporters) 및 유니포터들 uniporters)을 포함한다. 상기 유니포터들은 슈거 이송을 위한 간단한 촉진제들이다. 상기 H^{+} -연결된 심포터들(symporters)은 세포 내부로 이송된 모든 슈거 분자에 대하여 프로톤을 요구한다. *이. 콜리(E. coli)* 내의 GalP 단백질은 갈락토오스, 헥소스 슈거의 세포 내로의 이송을 위한 심포터이다. GalP 단백질은 12 트랜스-멤브레인 루프들을 갖는 매우 잘 특성화된 심포터이다. GalP는 또한 세포 멤브레인을 통과하여 글루코오스를 이송시키는 능력을 가지고 있는 것으로 보고되어 있다. AraE는 세포 멤브레인을 통하여 아라비노스(arabinose)의 이송에 대한 프로톤-연결된 심포터이다. 마찬가지로 XylE 단백질은 자일로스(xylose)의 이송에 대한 프로톤-연결된 심포터이다.

[0077] 포스포트랜스퍼라제 시스템(phosphotransferase system (PTS))에 의하여 섭취되는 글루코오스의 제거는 미생물 세포 내로의 글루코오스 섭취에 대해 사용되는 에너지를 감소시키는데 도움을 줄 수 있다. PTS를 조작하는 것에 의해 보존되는 에너지는 유기산 생산의 효율을 향상시키기 위하여 보내어질 수 있다. 포스포트랜스퍼라제 시스템(phosphotransferase system) 유전자 *ptsH* 및 *ptsG*는 글루코오스 섭취에서 에너지를 보존하도록 및 이에 따라 그 미생물에 의한 숙신산 생산의 효율을 향상시키도록 조작될 수 있다. 따라서 미생물 대사적 경로들의 영역에서 입수 가능한 데이터를 탐색함으로써, 유전자들의 세트를 삭제할 수 있고, 따라서 대사적 경로들의 대부분을 차단하고 또한 탄소 흐름을 큰 효율로 숙신산 생산으로 보낼 수 있다. PTS-매개된 글루코오스 섭취의 저해로, 글루코오스 섭취를 위한 다른 시스템들이 활성화 될 수 있어 산업적으로 유용한 화합물들의 생산을 위한 세포 내에서의 글루코오스의 연속된 입수 가능성을 확실히 한다. 예로서, 글루코오스 퍼미아제(glucose perase), 글루코오스 유니포터(glucose uniporter)에 대한 *glf* 유전자는 PTS 매개된 글루코오스 섭취의 손실에 대하여 대체하는 것으로 보여져왔다. 마찬가지로 *galP* 및 *glk* 유전자들의 과 발현은 *이. 콜리(E. coli)*

coli)의 *pts*⁻ 스트레인 내에서의 글루코오스 섭취 및 인산화를 증진시키는 것으로 보고되어 있다. GalP는 갈락토오스, 헥사로스 슈거의 섭취에 대한 심포터이다. GalP는 *pts*⁻ 스트레인 내에서 글루코오스를 이송시키는 것으로 보고되어 있다. GalP 매개된 글루코오스 섭취의 중요성은 *pts*⁻ 돌연변이 내에서의 *galP* 유전자 비활성화가 치명적인 것으로 나타났다는 사실에 의하여 증명되었다(위 외.(Yi *et al.*), 2003). Glk는 그것이 당분해(glycolysis) 내로 들어갈 수 있기 이전에 글루코오스 분자의 인산화를 달성하는데 있어서 요구된다. *pts*⁻ 스트레인 내에서 GalP 단백질의 발현은 구조적 프로모터 하에서 외인성 유전자를 발현시킴으로써 또는 *galS* 및 *galR*과 같은 *galP* 유전자의 억제자에 대하여 코딩하는 유전자들 내에서의 돌연변이들을 통해 상기 *galP* 발현의 억제를 완화하는 방법에 의하여 중 어느 것으로 달성될 수 있다.

[0078] 미생물 세포 내에서의 대사적 경로들에 대한 우리의 현재의 이해를 사용하여, 상당한 양으로 숙신산을 생산하도록 디자인된 스트레인을 합리적으로 제조하는 것은 가능하다. 효율적인 숙신산을 생산하는 미생물의 제조를 위한 합리적인 디자인은, 숙신산이 혐기성 또는 미세-호기성 조건들 하에서 세포 내부에서 탄소의 흐름에 대한 다른 잠재적인 발효적 경로들이 유전적 조작을 통해 차단될 때에 축적될 수 있다는 가정을 바탕으로 한다. 세포 내부에서 탄소의 흐름에 대한 특정 경로들을 차단하기 위하여 요구되는 유전적 조작들은 차단되도록 소망되어지는 경로의 작동에 관여되는 효소들에 대하여 코딩하는 하나 이상의 유전자들 활성을 감소시키거나 또는 제거하는 것을 포함한다. 아세트산염, 푸마르산염, 에탄올 및 락트산염에 대한 탄소 흐름은 적합한 유전적 방법들을 통하여 차단될 수 있다. 예로서, *adhE*에 대한 유전자를 삭제함으로써, 혐기성 알콜 생산이 차단될 수 있고, 또한 아세트산염으로의 탄소 흐름은 *ack* 및/또는 *ptaA* 유전자들을 삭제함으로써 차단될 수 있다.

[0079] 이론적으로 세포 내에서 탄소 흐름에 대한 특정한 원하지 않는 경로들을 차단하고 또한 그 탄소 흐름을 숙신산 생산으로 보내는 것이 이론적으로 직접적인 것이라 하더라도, 실제로 있어서 합리적으로 디자인된 유전자 결실들은 바람직한 표현형(phenotype)의 결과를 가져오지 않는다. 잔타마 외(Jantama *et al.*)(2008a; 2008b)에 의하여 기재된 바와 같이, 합리적으로 디자인된 유전자 결실들은 처음에는 열악한 성장을 갖는 박테리아 스트레인들의 결과를 가져왔다. 열악한 성장을 갖는 유전자-삭제된 스트레인들은 후속적으로 성장 속도를 향상시키기 위한 대사적 진화 하에 놓여졌다. 따라서, 숙신산 생산을 위한 *이. 콜리(E. coli)* 스트레인을 제조하는데 있어서, *IdhA*, *adhE*, 및 *ack* 유전자들을 제거하는 결실들의 제1 라운드가 수행되었고, 대사적 진화의 제1 단계에 의해 후속되었다.

[0080] 대사적 진화의 제1 단계 후에, 유전자 결실들의 제2 라운드는 *focA* 및 *pflB* 유전자들을 제거하였다. 다음으로 대사적 진화의 제2 라운드가 수행되었다. 삭제의 제3 라운드에서, *mgsA* 유전자가 제거되었다. 다음으로 대사적 진화의 제3 단계가 수행되었다. 삭제의 제4 라운드에서, *poxB* 유전자가 제거되었다. 다음으로 대사적 진화의 제4 단계가 수행되었다.

[0081] "대사적 진화(metabolic evolution)"는 그것에 의하여 미생물(종종 유전적으로 조작된 미생물)의 배양균(culture)(보통, 그러나 필수적인 것은 아닌, 액상의 q배양균)이 회석 및 재-성장(소위 "이동(transfers)")의 반복된 사이클들 하에 놓여지는 프로세스이며, 이에 따라 많은 수의 이동 후에, 향상된 성장 특성들 및/또는 숙신산, 락트산, 에탄올, 또는 부탄올과 같은 성장에 대하여 결합되어 있는 화합물의 향상된 생산을 보유하는 스트레인을 얻는 것이 가능하다. 대사적 진화는 특히 미생물의 성장이 발효에 의하여 바람직한 화합물을 생산하는 것에 의존적이도록 되었을 때에 스트레인 향상을 위한 방법으로서 효과적이다. 예로서, 야생형 *이. 콜리(E. coli)*와 같은 이중 발효적 미생물(heterofermentative microorganism)은, 글루코오스와 같은 슈거 상에서 혐기성적으로 또는 미세 호기성적으로 성장될 때에, 화합물들, D-락트산염, L-락트산염, 포름산염, 숙신산염, 아세트산염, 이산화탄소, 수소, 및 에탄올의 혼합물을 생산한다. 이러한 화합물들에 대한 경로는 소위 "발효적 경로들(fermentative pathways)"로 명명된다. 다른 미생물들 내에서, 발효적 경로들은 부탄올, 다른 알콜들, 다른 유기산들, 에스테르들, 3-하이드록시 알칸산들(3-hydroxy alkanolic acids), 지방산들, 알칸들, 알켄들, 카로테노이드들(carotenoids), 아미노산들, 비타민들, 및 많은 이상의 것과 같은 큰 다양성의 화합물들을 유도할 수 있다. 야생형 *이. 콜리(E. coli)*에 의해 만들어지는 화합물들 중에서, 적어도 D-락트산염, 숙신산염 및 에탄올은 상업적으로 관심이 있으며, 또한 따라서 "바람직한 화합물(desired compound)"로서 간주될 수 있다. 원하지 않는 화합물들에 대한 경로들을 컨트롤하는 유전자들을 삭제하는 것에 의하여, 그 미생물은 상기 바람직한 화합물에 대한 하나 이상의 경로로만 남게지고, 또한 따라서 성장은 상기 바람직한 화합물에 의존적이거나 또는 그에 대하여 결합된다. 대체적으로, 그 숙주 미생물로부터의 모든 발효적 경로들은 삭제될 수 있고, 또한 그 숙주 미생물들 내에서는 자연적으로 생산되지 않는 바람직한 화합물을 유도하는 다른 미생물로부터의 새로운 외인성 경로가 유전적 엔지니어링에 의하여 도입될 수 있다. 대사적 진화는 다음으

로 성장 및 그에 따라 바람직한 화합물의 생산을 향상시키기 위하여 돌연변이화된 및 조작된 스트레인에 대하여 적용될 수 있다. 대사적 진화가 고의적인 돌연 변이 유발(mutagenesis)을 요구하지 않기 때문에, 진화된 스트레인 상의 돌연변이적 로드는 최소한이며, 또한 대부분의, 전체는 아니지만, 축적되는 돌연변이들은 바람직한 화합물의 성장 및/또는 생산에 대해 유리할 수 있다. 대사적 진화의 좋은 예는 스트레인 KJ122이며, 이는 바람직한 화합물로서 숙신산염을 생산하기 위하여 엔지니어링되고, 또한 그런다음 대사적 진화 하에 놓이게 된다(잔타마 외.(Jantama *et al.*), 2008b).

[0082] 본 명세서에서, 용어들 "화합물(chemical)" 및 "화합물(compound)"은 상호 교환적으로 사용될 것이며, 또한 양자는 통상적인 의미로 사용될 것이다.

[0083] 대사적 진화의 과정에서, 선택된 배양균은 신선한 최소 배지(minimal medium) 내로 연속적으로 반복적으로 이동되며 또한 희석되어 자발적인 돌연변이들이 발생하고 또한 집단으로 고정되는 클론들 사이에서 부수하는 경쟁(attendant competition)을 갖는 수 많은 성장의 세대들(generations)에 도달한다. 어떤 돌연변이들은 공급되는 탄소원(들)의 더욱 효율적인 사용, 그 배양균 내에서의 독성 화합물들을 내성하는 더 나은 능력, 및 바람직한 화합물, 여기서 이러한 예는 숙신산이었다, 생산에 있어서 더 높은 효율을 부여한다. 대사적 진화의 과정에서, 주의는 바람직한 표현형들을 가진 클론들을 선택하는 것에 모아지며, 그러나 이는 바람직하지 않은 부산물들의 낮은 또는 결여된 생산을 또한 포함한다.

[0084] 따라서, 일반적으로, 특정 화합물을 과생산하도록 디자인되고 또한 유전적으로 엔지니어링된 미생물은 적합한 성장 속도를 가지지 않을 수 있고, 또한 결과적으로 그 특정 화합물을 생산하기 위한 기대되는 또는 바람직한 효율을 보이지 않을 수 있다. 대사적 진화는 그 특정 화합물의 생산을 위한 향상된 속도와 함께 향상된 성장 및 생산 내성(product tolerance)을 갖는 스트레인을 위한 선택을 하는데에서 사용될 수 있다. 그러나, 대사적 진화가 자발적인 돌연변이들로부터의 결과이기 때문에, 향상된 성장, 생산 내성 및 화합물 생산의 효율을 가져오는 정확한 유전적 변화들은 분명하거나, 예측 가능하거나 또는 대사적 진화 방법에 의해 재생산 가능하지 않을 수 있다. 그 진화된 미생물의 게놈 DNA 서열의 결정은 축적되는 돌연변이들의 상세함을 제공할 수 있다. 그런다음, 상동의 야생형 대립 형질(allele)으로 진화된 스트레인 내부의 각 개별적인 돌연변이를 대체하는 것, 또는 대체적으로, 스트레인 성능의 비교 분석에 의해 후속되는, 나이브(naive)의, 비진화된 스트레인 내부로 돌연변이화된 대립 형질을 도입하는 것에 의해, 어떤 돌연변이들이 바람직한 표현형에 기여하는지를 나타낼 수 있다. 진화된 스트레인 내에 남아있는 어떤 돌연변이들은 바람직한 프로세스에 대해 불리한 것일 가능성이 있다. 이들은 진화 내 초기 단계에서는 유리한 것이었을 수 있거나, 또는 진화하기 위한 시간을 가지지 않았을 수 있다. 어떤 경우에는, 돌연변이들이 유리한 것인지 및 그들 돌연변이들의 정확한 특성을 아는 것은, (1) 미래의 유전적 조작을 위한 더욱 합리적인 디자인, 예를들어 긴 진화 프로세스에 대한 필요성 없이 숙신산 생산에 대한 새로운 스트레인 또는 미생물을 엔지니어링하는 것, 및 (2) 나아가, 차례로 더욱 합리적인 스트레인 향상들을 유도할 수 있는 스트레인 성능을 향상시키는 미묘한 또는 예측 가능하지 않은 돌연변이들에 대한 통찰력을 허용한다. 예로서, 억제 유전자 내에서 발견되는 프레임 이동 돌연변이가 진화된 스트레인 내에서 발견되는 경우라면, 그러면 유전적 엔지니어링이, 동일한 표현형을 부여하는, 복귀할 수 없는(non-reverable), 및 따라서 더욱 안정적인 대립 형질을 생성하는 그 억제 유전자의 완전한 삭제를 형성하는데 있어 사용될 수 있다.

[0085] 미생물이 특정한 바람직한 표현형을 획득하도록 강제하는 어떤 선택적인 압력을 사용하는 대사적 진화의 과정 동안에, 두 개의 가능한 변화들이 일어날 수 있다. 그 미생물은 간단히 그 자체로 선택적인 압력에 대하여 순응할 수 있고 또한 변화된 유전자형(genotype)을 보일 수 없다. 대체적으로, 그 미생물은 선택적인 압력 하에서 특정 유전적 변화들을 경험하고 더욱 영원한 변화된 표현형을 나타낼 수도 있다. 단지 순응(adaptation)이 존재할 때에 및 어떠한 유전적 변화가 존재하지 않을 때에 그 미생물은 상기 선택 압력이 일단 완화되면 그것의 본래의 표현형으로 복귀할 것이다. 이러한 미생물들은 "순응된(adapted)" 미생물들로 참조된다. "순응된" 미생물들은 변화된 표현형을 나타내기 위하여 또 다른 새로운 라운드(fresh round)의 선택 압력을 경험해야 한다. 다른 한편, 수반하는 유전적 변화가 존재할 때에, 그 변화된 표현형은 심지어 선택 압력이 존재하지 않을 때에도 계속해서 존재할 것이다. 특정한 유전적 변화들에 의해 수반되는 대사적 진화가 바람직하다. 대사적 진화의 과정 동안 안정적인 유전적 변화를 획득하는 미생물은 그 선택 압력을 갖는 신선한 배지로 그것을 이동시키기 이전에 어느 정도의 시간 동안 어떤 선택 압력 없이도 고유의 성장 배지 내에서 그 미생물을 성장시키는 방법에 의하여 용이하게 식별될 수 있다. 만약 이러한 미생물들이 우수한 성장 및 어떤 지연 기간(lag period) 없이 기대되는 표현형을 보일 수 있다면, 그 미생물은 대사적 진화의 과정 동안 그 표현형을 수반하는 변화된 유전자형을 획득하였다고 간주된다.

- [0086] 상당한 양으로 숙신산을 생산하는 미생물을 만들기 위하여, 당분해(glycolytic) 경로, 트리카르복실산 회로(또한 크레브(Krebs) 회로 또는 TCA 회로로 알려져 있는) 및 글리옥실산 선트(glyoxylate shunt), 를 포함하는, 많은 수의 미생물의 대사적 경로들에 관련되는 다양한 효소들이, 상기 구절들 내에서 참조에 의해 언급되고 또한 포함된 과학적 및 특허 문헌 내에 기재된 다양한 유전적 엔지니어링 기술들을 사용하여 조작될 수 있다. 다양한 미생물 대사적 경로들에 대한 상세는 르닝거에 의한 생화학의 이론(Principles of Biochemistry, by Lehninger) 및 루버트 스트라이어에 의한 미생물학(Biochemistry by Lubert Stryer) 과 같은 표준 생화학 교과서들에서 찾아질 수 있다. 세인트 루이스, MO, USA에 있는 시그마 화학 회사(Sigma Chemical Company in St. Louis, MO, USA)로부터의 지. 마이클(G. Michael)에 의한 생화학적 경로들 포스터는 또한 박테리아 세포를 갖는 다양한 생화학적 경로들에 대한 상세함을 제공한다.
- [0087] 호기성 성장 과정 동안, 미생물적 탄소 대사는 당분해, 트리카르복실산 회로 및 산화성 인산화를 포함한다. NADPH 및 NADH와 같은 환원된 효소 보조 인자들은 세포 성장을 위해 요구되는 ATP 생산에 의해 수반되는 산화성 인산화의 작동에 의해 재생산된다. 본 발명의 바람직한 실시예에서의 혐기성 성장 조건들 하에서, 환원된 보조 인자들 NADPH 및 NADH의 재생산은 숙신산 생산을 위한 트리카르복실산 회로의 그 탄소 흐름을 지시하는 것 및 NADP^+ 및 NAD^+ 의 재생산을 위한 모든 다른 발효적 경로들을 제거하는 것에 의해 달성된다.
- [0088] 바람직한 유기산의 타입에 의존적으로, 상기 대사적 경로들은 특히 미생물이 우리의 선택의 특정 유기산을 생산하도록 조작된다. 그 미생물들은 락트산, 아세트산, 말산, 피루브산, 포름산 및 숙신산을 포함하는 많은 수의 유기산들을 합성하는 것이 가능하다. 따라서 숙신산의 생산을 위한 생체 촉매의 개발에서, 아세트산, 락트산, 피루브산 및 푸마르산의 생산을 위한 경로들은 차단되고 또한 숙신산 생산으로의 탄소 흐름이 세포 내부의 탄소 대사에 포함되는 하나 이상의 효소들을 조작하는 것을 통해 촉진된다. 알려져 있는 유전적 엔지니어링 기술들을 사용하여 조작될 수 있는 미생물 발효적 경로 내에서 활성인 효소들의 목록은 이에 제한되는 것은 아닌, 이소시트레이트 신터타제(isocitrate synthetase (*aceA*)), 말레이트 신터타제(malate synthase (*aceB*)), 글리옥실산 선트 오페론(glyoxylate shunt operon (*aceBAK*)), 아세테이트 키나제-포스포트랜스아세틸라제(acetate kinase- phosphotransacetylase (*ackA-ptA*)); 아코니타제 하이드라타제 1 및 2(aconitase hydratase 1 and 2 (*acnA* and *acnB*)); 아세틸-CoA 신터타제(acetyl- CoA synthetase (*acs*)); 시트레이트 리아제(citrate lyase (*citDEF*)); 알콜 디하이드로지나제(alcohol dehydrogenase (*adhE*)); 시트레이트 신터타제(citrate synthase (*citZ*)); 푸마레이트 리듀스타제(fumarate reductase (*frd*)); 락테이트 디하이드로지나제(lactate dehydrogenases (*ldh*)); 말레이트 디하이드로지나제(malate dehydrogenase (*mdh*)); *aceBAK* 오페론 억제자 (*aceBAK* operon repressor (*icIR*)); 포스포에놀 피루베이트 카르복실라제(phosphoenol pyruvate carboxylase (*pepC*)); 피루베이트 포메이트 리아제(pyruvate formate lyase (*pfIB*)); 피루베이트 옥시다제(pyruvate oxidase (*poxB*)); 피루베이트 카르복시 키나제(pyruvate carboxy kinase (*pck*)); 및 피루베이트 카르복실라제(pyruvate carboxylase (*pvc*))를 포함한다.
- [0089] 탄소원들의 당분해는 포스포에놀 피루베이트(phosphoenol pyruvate(PEP))의 생산의 결과를 가져온다. PEP는 나아가 혼합된 산 경로(mixed acid pathway)에 의해 대사된다. 본 발명에서 사용되는 바와 같이, 상기 용어 "혼합된 산 경로(mixed acid pathway)"는 트리카르복실산 회로 및 혐기성 조건들에서 작동하는 다양한 발효적 경로들 양자를 통한 PEP로부터의 탄소의 흐름을 의미한다. 혐기성 조건들 하에서, 피루브산염의 대사를 위한 적어도 네 개의 다른 발효적 경로들이 인식될 수 있다. 피루브산염은 DNA를 사용하며 또한 이에 따라 탄소원의 계속적인 대사를 위해 요구되는 세포의 레독스 균형을 유지하기 위하여 NAD^+ 를 생산하여 락트산염으로 환원될 수 있다. 피루브산염으로부터 유도되는 아세틸-CoA은 또한 NADH의 NAD^+ 로의 산화에 의해 수반되는 에탄올을 생산하도록 환원될 수 있다. 피루브산염은 또한 포름산염 또는 아세트산염으로 전환될 수 있다.
- [0090] TCA 회로 내에서, 두 개의 다른 수단들이 인식된다. 이소시트르산염(isocitrate)을 통하여 옥살로아세트산염(oxaloacetate)로부터 숙신산으로의 탄소 흐름을 포함하는 산화성 수단으로서 참조되는 TCA 회로의 하나의 수단에서는, NADP^+ 이 NADPH의 결과적인 형성과 함께 이소시트르산염을 산화시키는데에 사용된다. 말산염 및 푸마르산염을 통하여 옥살로아세트산염으로부터 숙신산으로의 탄소 흐름을 포함하는 TCA 회로의 환원성 수단으로서 참조되는 TCA회로의 다른 수단에서는, NADH가 NAD^+ 를 생산하도록 산화되고 또한 이에 따라 세포가 레독스 균형을 유지하도록 돕는다.
- [0091] 본 발명의 일 실시예에서, 발효적 경로들을 통한 PEP로부터의 탄소 흐름이 그 발효적 경로들에 포함되는 효소들에 대하여 코딩하는 유전자들을 비활성화하는 수단에 의하여 방지된다. 이러한 발효적 경로들을 통한 탄

소 흐름을 차단하는데 적합한 유전자들은 *IdhA*, *pflB*, *adhE*, *pta*, *ackA*, 및 *poxB* 을 포함한다. 이러한 유전자들의 하나 이상의 제거는 상기 발효적 경로들을 통한 PEP로부터의 탄소 흐름을 감소시킬 것으로 기대된다. 본 발명의 또 다른 관점에서, 메틸글리옥살(methylglyoxal)의 락트산으로의 전환에 대하여 원인이 되는 메틸글리옥살 신타제(methylglyoxal synthase (*mgsA*))를 코딩하는 *mgsA* 유전자가 상기 발효적 경로들에 관여되는 여섯 개의 다른 유전자들의 비활성화와 함께 비활성화된다.

[0092] 본 발명의 또 다른 실시예에서, 상기 발효적 경로들에 관여되는 기능적인 상동체들(functional homologues)은 또한 하나 이상의 발효적 경로에 관여되는 것으로 잘 알려져 있는 유전자들을 비활성화시키는 것과 함께 비활성화된다. 아세테이트 키나제 활성을 갖는 프로피오네이트 키나제(propionate kinase)(*tdcD*)는 트레오닌의 분해를 위해서만 생산되는 *tdcD* 유전자에 의해 인코딩된다. 그러나, 10%(w/v) 글루코오스를 갖는 혐기성 성장 과정 동안에, *tdcD*의 발현은 기능적으로 *ackA*를 대체할 수 있다. 이에 더하여, 동일한 오페론 내에서 인접한 *tdcE* 유전자는 *pflB*와 유사하고 또한 피루베이트 포메이트-리아제 활성(pyruvate formate-lyase activity)을 갖는 α-케토부티레이트 포메이트 리아제(α-ketobutyrate formate lyase)를 인코딩한다. 본 발명의 일 관점에서, *tdcDE* 유전자들은 발효적 경로들로의 탄소의 유입을 방지하기 위하여 및 TCA 회로 내로의 탄소의 흐름을 확실하게 하기 위하여 비활성화된다.

[0093] 본 발명의 또 다른 실시예에서, 발효적 경로들을 통하는 탄소 흐름을 차단하는 것과 함께, TCA 회로 내부에서의 탄소의 흐름이 개조되어 이에 따라 숙신산의 생산으로 지시된 탄소의 흐름이 존재하게 된다. 본 발명의 일 관점에서 TCA 내부의 탄소의 흐름에 대한 조작은 하나 이상의 유전자들의 발현을 상향-조절하는 수단에 의하여 달성된다. 본 발명의 또 다른 관점에서, TCA 회로 내부에서 기능하는 하나 이상의 유전자들이 숙신산으로의 증가된 탄소 흐름을 촉진하기 위하여 비활성화될 수 있다.

[0094] 본 발명의 바람직한 관점에서, 말레이트 디하이드로지나제(malate dehydrogenase)에 대하여 인코딩하는 유전자 *mdh* 는 말산염의 푸마르산염 및 숙신산염으로의 전환을 향상시키도록 상향 조절될 수 있다. 말산염 및 푸마르산염을 통한 옥살로아세트산염으로부터 숙신산으로의 탄소의 흐름은 TCA 회로의 환원성 수단으로서 참조된다. 옥살로아세트산염으로부터 숙신산으로의 이러한 TCA 회로의 환원성 수단을 통한 탄소의 흐름은 생산되는 숙신산의 매 몰(every mole)에 대하여 두 몰(two moles)의 NADH를 소비할 것이며 또한 이에 따라 혐기성 조건 하에서 세포의 레독스 균형을 유지하는데 있어서 도움을 줄 것이다. 다른 한편, *mdh*의 상향-조절은 세포의 레독스 균형을 유지하기 위하여 요구되는 NAD^+ 를 재생산하는데 있어서 도움이 될 것이다. *mdh* 유전자 발현의 상향-조절은 *mdh* 유전자에 대한 자연의 프로모터를 어떤 다른 강한 프로모터 서열로 대체하는 수단에 의하거나 또는 대체적으로 *mdh* 유전자의 프로모터 영역을 *mdh* 유전자의 증가 전사(increase transcription)가 존재하도록 돌연변이시키는 수단에 의하여 달성될 수 있다. 대체적으로, *mdh* 유전자의 추가적인 카피들(copies)이 스트레인 내부에 부가될 수 있거나 또는 *mdh* 유전자의 발현을 조절하는 유전자들이 *mdh* 유전자의 발현(icsae expression)을 증가시키도록 조작될 수 있다. 본 발명의 바람직한 실시예에서, *mdh* 유전자 발현의 상향 조절은 유전적으로 그것의 프로모터 영역을 조작하는 수단에 의하여 달성될 수 있다.

[0095] TCA 회로의 조절 작동에서, 숙신산은 TCA 회로 수단의 산화성 작동을 통하여 생산된다. 시트레이트(citrate), 시스-아코니테이트(cis-aconitate), 이소시트레이트(isocitrate), α-케토글루타레이트(α-ketoglutarate), 및 숙시닐-CoA를 통한 옥살로아세트산염으로부터 숙신산으로의 탄소의 흐름은 TCA 회로의 산화성 수단으로서 참조된다. 숙신산은 또한 글리옥실레이트 바이패스(glyoxylate bypass)의 작동을 통하여도 생산될 수 있다. 글리옥실레이트 바이패스(glyoxylate bypass)의 작동의 과정에서, 이소시트레이트 리아제(isocitrate lyase)의 작용에 의하여, 숙신산염 및 글리옥살산염은 이소시트르산염(isocitrate)으로부터 생산된다. 이에 따라 TCA 회로의 산화성 수단의 작동으로부터 또는 글리옥실레이트 바이패스의 작동으로부터 생산된 숙신산염은, 숙시네이트 디하이드로지나제(succinate dehydrogenase (*sdh*))에 의하여 푸마르산 및 그런 다음 말산을 생산하기 위하여 작용될 수 있다. 따라서, 본 발명의 또 다른 실시예에서, 유전자 비활성화는 세포 간 숙신산 생산을 증가시키기 위하여 숙신산염의 탈수소화(dehydrogenation)를 방지하는 데에 사용될 수 있다.

[0096] 본 발명의 또 다른 관점에서, 글리옥실레이트 바이패스(glyoxylate bypass)를 통한 탄소의 흐름은 숙신산 생산에서의 증가를 달성하기 위하여 조작될 수 있다. 이소시트레이트 리아제 효소는 이소시트르산염(isocitrate)의 글리옥살산염(glyoxalate) 및 숙신산염(succinate)로의 절단을 촉매한다. 이소시트레이트 리아제는 *aceBAK* 오페론에 의하여 코딩된다. 이소시트레이트 리아제 활성은 *icIR* 유전자들에 의하여 억제된다. 다른 표현으로, *icIR* 유전자의 발현은 글리옥살레이트 선트의 작동을 방지한다. 본 발명의 일 관점에서, *icIR* 유전자는 발효적 대사에 관여되는 유전자들의 비활성화와 함께 비활성화된다.

- [0097] 본 발명의 또 다른 실시예에서, 발효적 경로들의 작동을 방지하고 유전자 조작을 통하여 숙신산 생산으로 TCA 회로 내부에서의 탄소의 흐름을 증가시키는 것과 함께, 숙신산 생산을 증가시키기 위하여 TCA 회로로부터 다른 대사적 경로들로의 외부로의 탄소 흐름 또한 유전적 수단들을 통하여 차단될 수 있다. 예로서, TCA 회로로부터 아미노산 대사로의 탄소의 흐름이 숙신산으로의 탄소 흐름을 향상시키기 위하여 차단될 수 있다. 아스파테이트 아미노트랜스퍼라제 유전자(aspartate aminotransferase gene (*aspC*))는 아스파르트산의 합성에서 글루탐산으로부터의 아미노 그룹을 옥살로아세트산으로 이동시키고 또한 이에 따라 TCA 회로로부터의 외부로의 탄소 흐름을 촉진한다. 본 발명의 일 관점에서, *aspC* 유전자의 비활성화는 TCA 회로의 산화성 또는 환원성 수단의 어느 것을 통하여 옥살로아세트산염으로부터 숙신산 생산으로 탄소 흐름을 향상시키기 위하여 TCA 회로로부터 탄소의 외부로의 흐름을 차단하는 것에 대하여 후속될 수 있다.
- [0098] TCA 회로로부터 탄소의 다른 외부로의 흐름은 말산염으로부터 일어난다. 말산의 효소(malic enzyme(*sfcA*))에 의한 말산염의 탈카르복실화(decarboxylation)는 피루브산염의 생산의 결과를 가져온다. 본 발명의 일 관점에서, *sfcA* 유전자에 대하여 코딩하는 유전자는 TCA 회로로부터 탄소의 외부로의 흐름을 줄이기 위하여 비활성화된다. 본 발명의 또 다른 관점에서, *aspC* 및 *sfcA* 유전자들 양자는 숙신산 축적을 증진시키도록 TCA 회로로부터 외부로의 탄소 흐름을 방지하기 위하여 비활성화된다.
- [0099] 본 발명의 또 다른 관점에서, TCA 회로로부터 외부로의 탄소 흐름은 시트르산의 옥살로아세트산염 및 아세트산염으로의 절단의 원인이 되는 시트레이트 리아제 유전자(citrate lyase gene (*citDEF*))를 비활성화시키는 것에 의해 방지된다. 발효적 경로들 및 그들의 기능적 유사물에 관여되는 유전자들을 비활성화시키는 것 이외에, 숙신산 수율과 결합되어 있는 성장은 미생물 세포들 내부의 카르복실화 효소들(carboxylating enzymes)의 유전적 조작에 의하여 더욱 향상될 수 있다. 유전적 및 효소 분석을 수행하는 것을 통한 대사적 진화 과정에서 발생했던 변화들을 특성화하는 과정에서, 세포 내에서의 카르복실화 효소들(carboxylating enzymes)은 또한 향상된 숙신산 수율에 달성하기 위한 유전적 조작에 대하여 또 다른 타겟이 될 수 있다.
- [0100] 당분해 중간체들(glycolytic intermediates) 포스포에놀 피루브산염(phosphoenol pyruvate (PEP)) 및 피루브산(pyruvic acid)은 TCA 회로의 탄소 흐름을 향상시키기 위하여 카르복실화될 수 있다. 정상적인 조건들 하에서, TCA 회로의 탄소 유입은, 시트르산을 생산하기 위하여, 피루브산염으로부터 유도된 아세틸-CoA를 TCA 회로에서의 중간체인 옥살로아세트산염과 결합시키는 시트레이트 신타제(citrate synthase)의 작용에 의하여 달성된다. 세포 내부에 존재하는 하나 이상의 카르복실화 효소들의 효율을 향상시키는 수단에 의하여 포스포에놀 피루브산염 및 피루브산염을 옥살로아세트산염으로 카르복실화하는 것이 가능하다. 이에 따라 카르복실화 반응으로부터 생산된 옥살로아세트산염은 숙신산을 생산하기 위하여 TCA 회로의 환원성 수단을 통하여 추가적으로 환원될 수 있다.
- [0101] 세포 내부에 존재하는 카르복실화 효소들을 조작하는 것은 혐기성 발효적 성장 과정에서의 숙신산 수율을 증가시키는 수단이다. 본 기술 분야에서는 외인성 소스로부터 피루베이트 카르복실라제(pyruvate carboxylase (*pyc*))를 도입하는 것에 의해 피루베이트를 옥살로아세트산으로 카르복시화하는 것이 가능하다는 것이 잘 알려져 있다. *이. 콜리(E. coli)*와 같은 유전적 조작에 적합한 미생물적 스트레인들은 *pyc* 유전자를 가지고 있지 않다. *리조피움 엘티(Rhizopium elti)* 및 *락토바실러스 락티(Lactobacillus lacti)*와 같은 다른 박테리아 종으로부터 유도된 *pyc* 유전자들이 숙신산 생산을 향상시키기 위하여 유전적으로 변형된 *이. 콜리(E. coli)* 스트레인 내부로 도입되어 왔다.
- [0102] 네 개의 다른 외인성 카르복실화 효소들이 *이. 콜리(E. coli)* 내에 알려져 있다. 두 개의 이들 효소들은 포스포에놀 피루브산염을 카르복실화하는데에 원인이 되고 또한 두 개의 다른 효소들은 피루베이트 키나제 효소의 작용으로부터 유도되는 포스포에놀 피루브산으로부터 유도된 피루브산염의 카르복실화의 원인이 된다. 효소 포스포에놀 피루베이트 카르복실라제(phosphoenol pyruvate carboxylase (*ppc*))는 포스포에놀 피루브산염을 숙신산염을 생산하는 TCA 회로의 환원성 수단으로 들어갈 수 있는 옥살로아세트산염으로 카르복실화한다. 두 번째 카르복실화 효소 포스포에놀 피루베이트 카르복실라제(phosphoenol pyruvate carboxylase (*ppk*))는 또한 옥살로아세트산염을 생산하기 위하여 피루브산염을 카르복실화하지만, 그러나 정상적으로는 그것이 글루코오스의 존재 하에서 발현되지 않을 때에 역 환원을 촉매한다. 두 개의 다른 카르복실화 효소들 즉 NADH-연결된 말산 효소(NADH-linked maleic enzyme (*maeB*)) 및 NADH-연결된 말산 효소(NADPH-linked maleic enzyme (*maeA* / *sfcA*))는 피루브산을 말산으로 카르복실화한다. 상기 *maeB* 및 *sfcA* 효소들은 피루베이트 키나제의 작용에 의하여 포스포에놀 피루브산염으로부터 유도되는 피루브산염을 카르복실화한다.
- [0103] 세포 내에 존재하는 네 개의 카르복실화 효소들의 어느 하나는 당분해 회로 중간체들로부터 TCA 회로의 탄소 흐름을 향상시키기 위하여 그것의 효소적 활성을 증가시키도록 유전적으로 조작될 수 있다. *이. 콜리(E. coli)*

coli) 내에 존재하는 네 개의 자연적인 카르복실화 효소들 중에서, PPC-축매화되는 반응이 강하게 선호된다. PEP 내에 포함된 에너지는 무기 인산염의 방출과 함께 이 반응에서 손실된다. 다른 세 개의 카르복실화 효소들, 즉, *pck*, *maeA* 및 *sfca(maeB)*는 이들 세 개의 카르복실화 효소들이 글루코오스에 의하여 억제되므로, 기질로서 글루코오스를 사용하는 발효적 성장 과정에서 기능할 것으로 기대되지 않는다. 이러한 세 개의 카르복실화 효소들은 세포들이 산화적으로 유기산들을 대사할 때에 당신생(*gluconeogenesis*) 과정에서 역 방향으로 기능할 것으로 생각된다.

[0104] 당신생적 PEP 카르복시키나제(*gluconeogenic PEP carboxykinase (pck)*)는 TCA 회로 내로 탄소의 흐름을 향상시키도록 유전적으로 조작될 수 있다. *pck*의 활성을 향상시키는데 있어서의 유리한 점은 이러한 효소가 포스포에놀 피루브산염을 옥살로아세트산염으로 카르복실화시키는 반면, 생산되는 옥살로아세트산염의 매 분자에 대하여 ATP 분자의 생산의 결과를 가져온다는 사실에 있다. ATP 수율에서의 증가는 세포들의 성장 속도를 증가시킬 수 있다.

[0105] 발효적 숙신산염 생산에 대한 자연적 당신생 *pck(native gluconeogenic pck)*의 유입은 *pck* 유전자의 전사에 긍정적으로 영향을 미치는 어떤 돌연변이에 의하여 달성될 수 있다. PCK 활성의 수준에서의 증가는 그 유전자의 발현을 증가시키는 것으로 알려져 있는 자연적 프로모터 또는 어떤 다른 프로모터 서열을 가지고 있는 다중복제 플라스미드(*multicopy plasmid*) 내에서의 *pck* 유전자의 발현에 의해 달성될 수 있다. 세포 내에서 *pck* 유전자의 발현을 증가시키기 위한 또 다른 방법은 *pck* 유전자의 추가적인 카피들(*copies*)을 집적하는 것이다. 본 발명의 또 다른 실시예에서, *pck* 유전자의 자연적인 프로모터가 활성의 수준을 증진시키는 것으로 알려져 있는 어떤 다른 프로모터 성분들에 의하여 대체될 수 있다. *pck* 유전자의 증가된 발현은 또한 유전자의 프로모터 영역 내에서의 돌연변이에 의하거나 또는 *pck* 유전자의 프로모터 영역과 상호 작용하는 것으로 알려져 있는 조절 성분들의 유전적 조작에 의하거나 중 어느 것에 의해 달성될 수 있다. *pck* 유전자의 조절 단백질에 대하여 코딩하는 유전자는 *pck* 유전자의 발현을 증가시키기 위한 어떤 방법으로 돌연변이화되거나 또는 삭제되거나 또는 과발현될 수 있다. 싱글 포인트 돌연변이(*single point mutation*) (*pck* 유전자의 ATG 시작 코돈에 관련된 포지션 64에서 G에서 A로의 전이)가 포스포에놀 피루베이트 카르복시키나제 효소 활성(*phosphoenol pyruvate carboxykinase enzyme activity*)에 있어서 상응하는 증가에 수반되는 *pck* 유전자의 전사를 증가시킬 수 있다. *pck* 유전자 발현에서의 유사한 증가는 *pck* 유전자의 발현을 조절하는 것으로 알려져 있는 단백질들에 대하여 코딩하는 유전자들을 유전적으로 조작하는 것에 의하여 또한 달성될 수 있다. 예로서, 크라(*Cra*) 단백질은 *이. 콜리(E. coli)* 내에서의 *pck* 유전자의 발현을 활성화하는 것으로 보여졌다(사이어 및 램사이어(*Saier and Ramseier*), 196). 유사하게 *csrA* 시스템(*csrA*, *csrB*, *csrC*, *csrD*, *uvrY* 또는 *barA*를 포함하는)은 *pck* 및 mRNA 안정화도를 개조하는 것에 의하여 글루코오스 대사에 관여되는 다른 유전자들의 수준을 조절하는 것으로 또한 보고되어져 왔다(바비츠키 및 로메오(*Babitzke and Romeo*), 2007; 퍼네스티그 외.(*Pernestig et al*), 2003; 스즈키 외.(*Suzuki K et al.*) 2002).

[0106] 본 발명에서, 유전적으로 조작된 및 바람직한 표현형을 보여주는 대사적으로 진화된 스트레인은, 예로서 숙신산 생산의 효율을 향상시켰고, 또한 유전적으로 조작된 및 대사적으로 진화된 스트레인의 제조에서 사용되는 야생형 스트레인은 게놈 DNA 서열 분석 하에 놓여진다.

[0107] 야생형 부모 스트레인 및 유전적으로 조작된 및 대사적으로 진화된 스트레인들의 전체적인 게놈 서열화는 DNA 서열화의 높은 생산량의 필드에서 경험히 쌓인 자에게 잘 알려져 있는 서열화 기술들에 의해 달성될 수 있다. 많은 수의 서열화 기술들이 비용-효율적인 방법으로, 예를 들어 일루미나, 주식회사 및 454, 주식회사(*Illumina, Inc., and 454, Inc.*)에 의해 제공되는 하드웨어, 소프트웨어, 및 방법들, 게놈 서열 정보를 제공할 수 있는 시장에서 입수 가능하다.

[0108] 야생형 및 진화된 박테리아 스트레인들로부터 얻어지는 서열 데이터가 적합한 소프트웨어를 사용하여 비교될 수 있으며 또한 두 스트레인들 사이에서의 유전적 변화들이 얻어질 수 있다. 두 개의 스트레인들 사이에서의 유전적 변화들의 모든 리스트로부터, 유전적 엔지니어링의 단계에서 의도적으로 도입된 유전적 변화들이 실증될 수 있고, 또한 대사적 진화의 프로세스 과정 동안 발생하였던 또한 고정되었던 추가적인 유전적 변화들이 확인될 수 있다.

[0109] 일단 대사적 진화 과정에서 발생되었던 돌연변이들이 확인되면, 바람직한 화합물의 생산을 향상시키는데 있어서의 이들 돌연변이들의 중요성은 역 유전적 분석법을 사용하여 결정된다. 예로서, 측정 가능한 타이틀들에서 숙신산을 생산하고 또한 그것의 염색체 DNA 내에서의 최소한의 수의 유전자 변화들을 포함하는 나이브(*naive*) 박테리아 스트레인이 사용될 수 있다. 다른 스트레인들 중에서, *pck* 및 *pstI*의 돌연변이화된 형태들 및 *pflB* 내에서의 결실을 포함하는 양자, *이 콜리(E. coli)* 스트레인 XZ722 및 동일한 공개된 스트레인 XZ721는 이리

한 목적을 위해 적합하다(장 외.(Zhang *et al.*), 2009). 스트레인 KJ122(또는 관심의 다른 어떤 스트레인)의 대사적 진화의 과정 동안 발생하였던 유전적 돌연변이는 본 기술 분야에 잘 알려져 있는 기술들을 사용하여 스트레인 XZ721 및 XZ722 내에 장착될 수 있다. 두 단계 유전자 대체 방법들이 본 기술 분야에 잘 알려져 있다(장 외.(Zhang *et al.*), 2009a; 장 외.(Zhang *et al.*), 2009b; 잔타마 외.(Jantama *et al.*), 2008b). 본 발명의 스트레인들을 제조하기 위하여 사용되는 "두 단계 유전자 대체 방법(two step gene replacement method)"에 대하여, 제1 단계는 60 g/l 수크로스에 대하여 민감성을 부여하는, *바실러스 서브틸리스*(*Bacillus subtilis*)로부터의 *sacB* 유전자와 함께, *cat*(30 mg/l 에서 클로람페니콜(chloramphenicol)에 대하여 저항성인) 또는 *kan*(50 mg/l 에서 카나마이신(kanamycin)에 대하여 저항성인)과 같은 약물 저항성 유전자를 포함하는 선택 가능한 및 카운터-선택 가능한 카세트를 장착시키는 것이다. *cat-sacB* 또는 *kan-sacB* 카세트가 템플레이트로서 각각 pLOI4151(SEQ ID NO. 33) 또는 pGW162 (SEQ ID NO. 34)를 사용하는 PCR에 의하여 얻어진다. PCR 프라이머들은, 서열을 기폭하는(priming) 약 18 내지 25 염기들에 더하여, 상동적 재조합을 촉진하기 위하여, 염색체 타겟 서열의 5' 및 3' 말단들에 대해 상동인 약 50 염기들을 포함한다. 선형의 PCR 생성물은 전기 천공법(electroporation)에 의해, 플라스미드 pKD46으로 이미 이동되었던 수여 스트레인(recipient strain) 내로 이동되며, 그것은 그 수여 스트레인의 염색체 내로의 바람직한 PCR 파편의 집적의 빈도를 증가시키기 위하여 파지 람다 Red 재조합 효소(phage lambda Red recombinase)를 발현한다. pKD46은 30°C에서 100 mg/l에서 암피실린(ampicillin) 저항성에 의하여 선택된다. 그것은 리플리케이션의 온도 민감성 기원을 가지며 또한 42°C에서의 성장에 의하여 그 숙주 스트레인으로부터 보존될 수 있다. pKD46은 다첸코 및 워너((Datsenko and Wanner (2000)))에 기재되어 있고, 또한 콜리 유전적 스톡 센터, 예일 대학, 뉴 헤븐, CT, USA(Coli Genetic Stock Center, Yale University, New Haven, CT, USA)로부터 입수 가능하다. 두 단계 유전자 대체 방법의 제2 단계에 대하여, 제2의 PCR 생성물이 템플레이트로서 예를들어 KJ122 염색체 DNA를 사용하여, 장착되는 것이 바람직한 대립 형질(allele)의 5' 및 3' 말단들에서 촉진하는 PCR 프라이머들의 제2 쌍을 사용하여 만들어진다. 이러한 두 개의 프라이머들의 쌍은, 또한 PCR 프라이머들의 제1 쌍과 유사한, 상동의 재조합을 촉진하기 위하여 염색체 타겟 서열의 5' 및 3' 말단들과 상동인 약 50 염기쌍들을 포함할 수 있다. 그러나, 만약 삽입된 대립 형질에 대한 염색체 타겟이 삽입될 대립 형질의 자연적 로커스(locus)라면, 제2 단계를 위한 상동의 서열들은 단지 증폭된 DNA 서열의 일부일 수 있다. 다른 표현으로, 염색체 내의 제2 사이트에 유전자의 제2 카피(copy)를 장착시키는데 있어서와 같이, 염색체 타겟에 대해 상동인 50 염기쌍 서열들은 염색체 타겟이 그 대립 형질의 자연적 로커스와 다른 사이트라면, 단지 PCR 프라이머들의 5' 말단들에 대하여 부가되는 것이 필요하다. 제2 단계에서 변화된 대립 형질의 장착은 선형의 제2 PCR 생성물을 상동의 재조합을 자극하기 위한 pKD46을 여전히 포함하는, 수여 스트레인(이것은 제1 단계에서 형성되었던 스트레인이다) 내로 전기 천공하는 것 및 60 g/l 수크로스를 포함하고 있는 플레이트들 상에서 선택하는 것에 의하여 달성될 수 있다. 이러한 두 단계 방법에서, 상동적 재조합이 바람직한 모든 조각들에 대하여, 세포들은 헬퍼 플라스미드 pKD46으로부터의 람다 Red 재조합 효소 유전자들의 발현을 유발하기 위하여 5 g/l 아라비노스를 포함하는 풍부 배지 내에서 성장된다. 자발적인 *sacB* 돌연변이들이 수크로스 선택의 과정에서 일어날 수 있기 때문에, 결과적인 클로니들은 클로람페니콜 또는 카나마이신 민감성에 대하여 또한 바람직한 대립 형질 변화를 획득한 클로니들을 식별하는 진단적 PCR에 의하여 스크린된다. 포인트 돌연변이들과 같은, 미묘한 돌연변이들에 대하여, 바람직한 대립 형질이 장착되었는지 분명히 하기 위하여 장착된 대립 형질의 DNA 서열을 획득하는 것이 필요하다. 이것은 장착된 대립 형질의 PCR 증폭된 카피를 획득하고 또한 그 PCR 생성물을 서열화하는 것에 의하여 달성될 수 있다. 상업적인 서비스로서 DNA 서열화를 수행하는 많은 기관들, 예로서 터프스 대학 코어 퍼실리티, 터프스 대학 메디칼 스쿨, 보스턴, MA, USA(Tufts University Core Facility, Tufts University Medical School, Boston, MA, USA), 이 존재한다.

[0110] 두 단계 유전자 대체에 대하여 사용되는 PCR 프라이머들의 예들은 표 2에 주어진다.

[0111] 각 돌연변이의 영향을 결정하는데 있어서의 대체적인 접근은, 상기 기재된 동일한 두 단계 유전자 대체 방법을 사용하여, 진화된 스트레인(예로서 KJ122) 내에서의 싱글 돌연변이 대립 형질을 부모 스트레인, *이. 콜리*(*E. coli*) C (ATCC 8739)로부터의 상동의 야생형 대립 형질로 대체하는 것이다. 스트레인 제조는 또한 일반화된 박테리오파지 형질 도입(transduction), 예로서 파지 P1 *vir*(실해비 외.(Silhavy *et al.*), 1984)를 사용하여 달성될 수 있다.

[0112] 기술들은 또한 *이. 콜리*(*E. coli*) 염색체 내에서의 특정한 작은 또는 큰 결실들(deletions)을 생성하는 것에 대하여 잘 알려져 있으며, 이들 기술들은 다른 박테리아, 아케아 이스트(archaea yeasts), 및 곰팡이(filamentous fungi)에 대하여 적용 가능하여야 한다. 예로서, 콜린스니첸코 외.(Kolisinychenko *et al.* (2002))는 염색체 DNA로부터의 특정 유전자 서열을 제거하는 것에 대하여 후속될 수 있는 기술을 기재하며,

또한 이는 참조로서 이에 의하여 포함된다.

- [0113] 야생형 또는 돌연변이 대립 형질을 포함하는 스트레인들의 동질 유전자형 쌍들이 제조된 후에, 상기 쌍의 멤버들은 적합한 작은 스케일 발효제들에서의 성장 및 화학적 생산에 대하여 테스트될 수 있다. 예로서, 혐기성 또는 미세-호기성 조건들 하에서의 pH 컨트롤에 의한 성장 및 숙신산 생산은 이미 기재된 바와 같이 테스트될 수 있다(장 외.(Zhang *et al.*), 2009a; 장 외.(Zhang *et al.*), 2009b; 잔타마 외.(Jantama *et al.*), 2008a; 잔타마 외.(Jantama *et al.*), 2008b). 작은 스케일 발효제들에서의 스트레인 성능을 테스트하기 위하여, 그것이 생산됨에 따라 숙신산을 중성화하는데 사용되는 알카라인 용액은 2.4 M 포타슘 카보네이트 플러스 1.2 M 포타슘 하이드록사이드일 수 있거나, 또는 대체적으로 그것은 스트레인 성능을 테스트 하기 위한 6 M 암모늄 하이드록사이드 플러스 3 M 암모늄 바이카보네이트일 수 있다.
- [0114] 한꺼번에 XZ722와 같은 숙신산염-생산 *이. 콜리(E. coli)* 스트레인 내로 하나의 돌연변이를 도입하는 것, 또는 한꺼번에(즉, 하나 이상의 돌연변이화된 유전자들의 야생형 대립 형질을 재-도입하는 것) 하나의 돌연변이를 KJ122와 같은 숙신산염-생산 *이. 콜리(E. coli)* 스트레인으로부터 제거하는 것에 의하여, 성장 또는 화학적 생산을 향상시키는데 기여하는 유전적 변화들을 결정할 수 있다. 이에 더하여, 돌연변이들의 조합들로 인하여 어떤 부가적인 또는 시너지적인 효과가 존재하는지를 결정하기 위하여 돌연변이의 하나 초과와 타입이 화학적-생산 스트레인 내로 도입되거나 또는 제거될 수 있다. 돌연변이들의 이러한 스크리닝을 통하여, 효과적인 성장 및 화학적 생산에 도달하기 위해 중요한 돌연변이들을 식별하는 것이 가능하다. 일단 화학적 생산을 향상시키기 위한 그러한 신규한 돌연변이들이 식별되면, 그러한 돌연변이들은 숙신산 또는 다른 화학적 생산을 위한 박테리아 스트레인을 합리적으로 디자인하고 및 그에 의해 바람직한 표현형을 갖는 스트레인을 선택하는데 있어서 대사적 진화의 프로세스에 대한 필요성을 제거하는데 유용한 유전자 변형들의 목록에 더해질 수 있다.
- [0115] 본 발명의 개시에서, 특정 실시예들은 *이. 콜리(E. coli)*의 스트레인에 의한 숙신산염의 생산에 관련되는 것으로 주어진다. 그러나, 본 발명에서 발견된 이론들은 광범위한 적용 가능성을 가지고, 따라서 특정 실시예들의 언급이 어떤 방법에서 상기 광범위한 의미들을 제한하는 데에 사용되어서는 안된다. 예로서, 특정 유전자에서의 특정 타입의 돌연변이가 *이. 콜리(E. coli)* 내에서의 숙신산염 생합성에 대하여 유용하다는 발견은 다른 박테리아 및 다른 미생물들에 대하여 과도한 실험 없이도 적용될 수 있다. 관심의 유전자 내에서의 돌연변이가 일단 확인되면, 유사 돌연변이, 또는 동일한 기능이나 목적을 달성하는 돌연변이가 광범위한 다양성의 미생물들 내로 도입될 수 있다.
- [0116] 예로서, 본 발명은 피루베이트 키나제를 인코딩하는 유전자의 발현에서의 감소가 *이. 콜리(E. coli)*의 매우 발전된 숙신산염 생산 스트레인의 내용 중에서 숙신산염 생산에 대하여 유용하다는 것을 개시한다. 현재 다른 미생물들 내에서 피루베이트 키나제를 인코딩하는 유전자를 식별하고 및 결실(deletion) 또는 다른 돌연변이에 의하여, 숙신산염 생산을 향상시키기 위하여, 새로운 미생물 내에서의 그 유전자의 발현을 감소시키는 것은 용이하다. *클레브시엘라(Klebsiella)*, *세라티아(Serratia)*, *시트로박터(Citrobacter)*, *에르위니아(Erwinia)*, 또는 *살모넬라(Salmonella)* 속(genus)과 같은, *이. 콜리(E. coli)*의 가까운 친척들에 대하여, 관련성 있는 유전자는 유전자 은행(GenBank) 또는 DNA이나 단백질 서열들에 대한 다른 공공의 데이터베이스를 통한 상동 관계 서치에 의하여 용이하게 발견될 수 있다. 무엇이 상동 관계이고, 또는 아니고를 정의하는데 사용될 수 있는 예리한 컷오프(cutoff)가 존재하지 않음에도 불구하고, 조회 서열의 "상동 관계(homology)"는 조회 서열에 진화적으로 관련된 DNA 서열이며, 또한 통상적으로 1) 유전자 은행을 통하여 입수 가능한 BLASTN 프로그램을 사용하는 결실들 및 삽입들, 또한 디폴트 파라미터들을 선택하는 것("예측되는(Expected)" 수를 10의 디폴트로부터 더 큰 수로 증가시키는 것과 마찬가지로, "기재(Decription)", "도시적 관찰(Graphical overview)", 및 "정렬(Alignment)" 셋팅들을 포맷 요구 형태(Format Request form)에서 1000 또는 초과로 셋팅하는 것은 더 많은 수의 상동체들(homologs)을 생성할 수 있다)을 허용하여, 그 안에서 길이에 있어서 50 또는 초과와 염기들인 서열, 또는 서열의 일부가 조회 서열과 50% 또는 초과와 동일성의 매치를 갖는 유전자 또는 DNA 서열로서, 또는 2) 유전자 은행을 통하여 입수 가능한 BLASTP 프로그램의 디폴트 파라미터들을 사용하는 결실들 및 삽입들을 허용하여, 길이에 있어서 25 또는 초과와 아미노산들이며, 또한 조회 서열과 25% 또는 초과와 동일성 또는 50% 또는 초과와 유사성을 갖는 단백질 서열, 또는 단백질 서열의 일부로서 발견된다. 상동 관계 서치를 수행하기 위하여 유전자 은행 데이터베이스 및 다양한 BLAST 프로그램을 사용할 때에, 조회의 가장 근접한 친척들의 계통들을 선택하지 않는 것(또는 선택 해제하는 것)이 종종 필요하다. 예로서, *에세리키아(Escherichia)* 속(genus)의 멤버들의 많은 수의 계통들은 상동체 히트 목록에서 *클레브시엘라(Klebsiella)*와 같은 더 먼 친척들에서 존재하는 *이. 콜리(E. coli)* 조회의 상동체들을 밀어낼 수 있으나, *에세리키아(Escherichia)*의 멤버들의 계통들을 선택 해제하는 것은 *클레브시엘라(Klebsiella)* "히트(hit)" 가

발견되는 것을 허용한다. 덜 관련된 미생물들에 대하여, 관련성 있는 유전자는 상동체가 아닐 수 있거나, 또는 DNA 또는 단백질 서열 상동체에 의해 식별 가능하지 않을 수 있으나, 그 유전자는 여전히 키 워드 서치에 의하여 발견될 수 있다. 예로서, 서치에서의 키워드들로서 "사카로마이스(Saccharomyces)" 및 "피루브베이트 키나제(pyruvate kinase)"를 사용하여, 유전자 은행을 서치하는 것에 의해 피루브베이트 키나제를 인코딩하는 *사카로마이스 세레비시아(Saccharomyces cerevisiae)* 유전자를 찾는 것은 용이하다. 그러한 서치가 행해졌고, 또한 CDC19(또한 PYK1으로서도 알려져 있는) 및 PYK2의 두 개 유전자들이 발견되었으며, 양자는 피루베이트 키나제에 대하여 코딩한다.

[0117] 간단하게, 본 발명은 대사적 진화의 노동-집약적인 프로세스를 통하지 않고 산업적으로 유용한 미생물에 대한 바람직한 표현형들을 수여하는 새로운 유전자 돌연변이들을 식별하기 위하여 감하는(subtractive) 또는 가하는(additive) 계층 분석을 제공한다. 이러한 방법에서 산업적 미생물들은 합리적 디자인을 사용하여 전체적으로 제조될 수 있다. 이러한 발견이 미생물들을 사용하는 유기산들의 생산, 및 더욱 특별하게 생체 효소를 사용하는 숙신산의 생산을 상세하게 기재함에도 불구하고, 산업적 미생물학의 분야에서 숙련된 자는 광범위한 다양성의 미생물 생체 효소들을 사용하여 여기 기재된 방법들을 다른 산업적 화합물들의 생산성을 향상시키기 위하여 적용할 수 있을 것이다.

[0118] 실험 섹션

[0119] 일반적 기술들

[0120] 스트레인, 배지 및 성장 조건들(Strains, media and growth conditions)

[0121] *이. 콜리* (*E. coli* C(ACCT 8739))의 새로운 유도체들이 성장-기반의 선택과 결합된 유전자 결실의 고유의 조합을 사용하여 숙신산염 생산을 위해 개발되었다.

[0122] 본 발명의 미생물은 미생물학 분야에서 잘 알려져 있는 많은 수의 다른 배양 배지에서 성장될 수 있다. 예로서, *이. 콜리* (*E. coli*)의 야생형 및 돌연변이체 스트레인들이 1%(w/v) 트립톤(tryptone), 0.5%(w/v) 이스트 추출물, 및 0.5%(w/v) NaCl을 포함하는 루리아-버타니(Luria-Bertani)(LB) 배지에서 성장되었다. 생체 효소로서 유전적으로 변형된 미생물을 포함하는 발효적 프로세스를 사용하는 유기산의 상업적 생산을 위하여, 탄소원으로 보충된 최소한의 미네랄 염 배지가 선호된다. LB 배지와 같은 풍부 배지에 반대되는 것으로서 최소 미네랄 염 배지의 사용은 상업적 스케일로 유기산들의 생산을 위한 경비를 감소시킨다.

[0123] 본 발명을 위해 적합한 최소한의 미네랄 배지들은 NBS 배지(코지 외.(Causey *et al.*), 2007) 및 AM1 배지(마티네즈(Martinez *et al.*), 2007)을 포함한다. NBS 배지는 1 mM 베타인(betaine), 25.72 mM KH_2PO_4 , 28.71 mM KH_2PO_4 , 26.50 mM $(\text{NH}_4)_2\text{HPO}_4$, 1 mM $\text{MgSO}_4 \cdot 7\text{H}_2\text{O}$, 0.1 mM $\text{CaCl}_2 \cdot 2\text{H}_2\text{O}$, 0.15 mM 티아민(Thiamine) HCl, 5.92 μM $\text{FeCl}_3 \cdot 6\text{H}_2\text{O}$, 0.84 μM $\text{CoCl}_2 \cdot 6\text{H}_2\text{O}$, 0.59 μM $\text{CuCl}_2 \cdot 2\text{H}_2\text{O}$, 1.47 μM ZnCl_2 , 0.83 μM $\text{Na}_2\text{MoO}_4 \cdot 2\text{H}_2\text{O}$, 및 0.81 μM H_3BO_3 을 포함한다. AM1 배지는 1 mM 베타인(betaine), 19.92 mM $(\text{NH}_4)_2\text{HPO}_4$, 7.56 mM $\text{NH}_4\text{H}_2\text{PO}_4$, 1.5 mM $\text{MgSO}_4 \cdot 7\text{H}_2\text{O}$, 1.0 mM 베타인(Betaine)-KCl, 8.88 μM $\text{FeCl}_3 \cdot 6\text{H}_2\text{O}$, 1.26 μM $\text{CoCl}_2 \cdot 6\text{H}_2\text{O}$, 0.88 μM $\text{CuCl}_2 \cdot 2\text{H}_2\text{O}$, 2.20 μM ZnCl_2 , 1.24 μM $\text{Na}_2\text{MoO}_4 \cdot 2\text{H}_2\text{O}$, 1.21 μM H_3BO_3 및 2.50 μM $\text{MnCl}_2 \cdot 4\text{H}_2\text{O}$ 을 포함한다. 미량 성분들은 1000X 스톡(stock)으로 준비되며 및 하기 성분들을 포함했다: 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 $\text{MnCl}_2 \cdot 4\text{H}_2\text{O}$.

[0124] 유기산의 미생물적 생산을 위한 미네랄 배지는 탄소원으로 보충된다. 본 발명에서 유용한 탄소원들은 이에 제한되는 것은 아니나 자일로스와 같은 펜토스 슈거들, 글루코오스, 프럭토스, 및 갈락토오스와 같은 헥소스 슈거들, 또는 수크로스 및 말토스와 같은 디사카라이드들을 포함한다. 상기 탄소원은 또한 글루코오스 및 자일로스의 조합과 같이 다른 슈거들의 조합을 제공하는 것에 의해 만족될 수도 있다. 탄소원은 또한 스타치(starch) 또는 리그노셀룰로오스(lignocellulose)의 가수분해로부터 유도될 수도 있다. 스타치(starch) 또는 리그노셀룰로오스(lignocellulose)와 같은 복합 탄수화물들의 가수분해는 본 기술 분야에서 잘 알려져 있는 열-화학적 전하 프로세스 또는 효소적 방법들 중 어느 것을 사용하여 달성될 수 있다. 미생물적 발효를 사용하는 유기산의 산업적인 생산을 위한 바람직한 탄소원은 농업적(agricultural) 또는 임업적(forestry) 폐기물로부터 유도되는 리그노셀룰로오스 가수분해물이다. 리그노셀룰로오스 가수분해물은 추가적으로 헥소스-풍부

한 및 펜토스-풍부한 분획을 생산하도록 추가적으로 분할될 수 있으며, 또한 그러한 분획들은 미생물적 발효 프로세스를 사용하는 유기산들의 상업적 생산을 위한 탄소의 소스로서 작용할 수 있다. 상기 리그노셀룰로오스 가수분해물은 추가적으로 어떤 농도들 이상에서 많은 수의 미생물들에 대하여 독성인 것으로 발견된 푸르푸랄(furfural)과 같은 어떤 화합물들을 제거하기 위하여 해독될 수 있다.

[0125] 스트레인 제조의 과정에서, 배양균들은 2%(w/v) 글루코오스 또는 5%(w/v) 아라비노스를 포함하는 루리아 배지(Luria broth)(10 gl^{-1} 디프코 트립톤(Difco tryptone), 5 gl^{-1} 디프코 이스트 추출물(Difco yeast extract) 및 5 gl^{-1} NaCl)내에서 30, 37, 또는 39°C에서 호기적으로 성장되었다. 항생 물질에 대한 내성(antibiotic resistance)을 인코딩하는 유전자들, 플라스미드들 또는 외부 유전자들 어느 것도 숙신산염 생산을 위해 개발된 최종 스트레인들 내에 존재하지 않는다. 그러나, 다른 스트레인들의 제조의 초기 단계들에서, 다양한 항생 물질에 대한 내성(antibiotic resistance) 마커들이 사용되었다. 앰피실린(ampicillin) (50 mg l^{-1}), 카나마이신(kanamycin) (50 mg l^{-1}), 또는 클로람페니콜(chloramphenicol) (40 mg l^{-1})이 항생 물질 선택 프로세스(antibiotic selection process)를 위해 요구되는 것으로서 부가되었다.

[0126] 시드(seed) 배양균들 및 발효제들은 37°C에서 100 rpm으로 글루코오스, 100 mM KHCO_3 및 1 mM 베타인 HCl을 포함하는 NBS 또는 AM1 미네랄 염 배지에서 성장되었다. 어떤 실시예들에서, 콘 스티프 리퀴(corn steep liquor)가 사용되었다. 그것은 콘 습식-제분 산업으로부터의 부산물이다. 이스트 추출물 및 펩톤(peptone)과 비교할 때, 그것은 비타민 및 미량 성분들의 저렴한 소스이다.

[0127] 발효적 숙신산염 생산에 대하여, 스트레인들은 다른 언급이 없다면 항생제 없이 37°C에서 10%(w/v) 글루코오스 및 100 mM 포타슘 바이카보네이트로 보충된 NBS 미네랄 염 배지(코지 외.(Causey *et al.*), 2004)에서 성장되었다. 발효를 위한 예비-접종물(Pre-inocula)이 신선한 클로니들을 250 ml 플라스크(100 ml NBS 배지, 2% 글루코오스) 내로 이동시키는 것에 의하여 성장되었다. 16 시간 후에(37°C, 120 rpm), 이러한 배양균은 0.033 g 세포 건조 무게(CDW) l^{-1} 의 접종물(inoculum)을 제공하기 위하여 300 ml NBS 배지(10% 글루코오스, 100 mM 포타슘 바이카보네이트)을 포함하는 작은 발효 용기 내로 희석되었다.

[0128] 성장 배지 내에서의 유기산들의 축적이 배지 내에서의 pH를 감소시키는 경향이 있기 때문에, 배양 배지(culture medium)에 대하여 요구되는 적합한 중성화시키는 제재들을 부가하는 것이 필요하다. 배양 용기의 pH는 pH 프로브를 사용하여 연속적으로 모니터링될 수 있고, 또한 적합한 염기가 그 성장 배지의 pH를 중성 pH 주위로 유지시키기 위하여 부가될 수 있다. 미생물 배양의 pH를 유지하기 위하여 적합한 염기들은 이에 제한되는 것은 아니나, NaOH, KOH, NH_4SO_4 , Na_2CO_3 , NaHCO_3 , 및 NH_4CO_3 를 포함한다. 이러한 목적을 위한 적합한 염기들은 단독으로 또는 조합하여 사용될 수 있다.

[0129] 어떤 실험들에서, 발효는 부가적인 CO_2 (1.2 M 포타슘 하이드록사이드 내에서의 2.4 M 포타슘 카보네이트)를 포함하는 염기를 부가함으로써 자동적으로 pH 7.0으로 자동적으로 유지되었다. 후속적으로, pH는 3M K_2CO_3 및 6N KOH의 1:1 혼합물을 부가하는 것에 의하여 유지될 수 있었다. 발효 용기들은 샘플 제거를 위한 벤트(vent)로서 기능하는 16 게이지 니들을 제외하고 밀봉되었다. 혐기 생활(anaerobiosis)이 CO_2 의 대기를 보장하도록 작용하는 부가된 바이카보네이트를 갖는 성장 과정에서 빠르게 도달되었다.

[0130] 세포 성장: 세포 매스(Cell mass)는 썬모 일렉트로닉 스펙트로닉 20 스펙트로포토미터(Thermo Electronic Spectronic 20 spectrophotometer)를 사용하여 550 nm(OD_{550})에서 광학적 밀도를 측정하는 것에 의하여 측정되었다.

[0131] 유기산 및 슈거 분석: 다양한 유기산들 및 슈거들의 농도가 HPLC에 의하여 측정되었다. 발효 배지 내에 존재하는 숙신산 및 다른 유기산들은 바이오라드 아미넥스 HPX-87H 컬럼을 구비한 에질런트 1200 HPLC 장치(Agilent 1200 HPLC apparatus with BioRad Aminex HPX-87H column) 상에서 분석되었다. 바이오라드 마이크로가드 카타이온(BioRad Microguard Cation) H^+ 가 보호 컬럼(guard column)으로서 사용되었다. HPLC 분석에 대한 표준들은 0.008 N 황산 내에서 준비되었다. HPLC 컬럼 온도는 50°C로 유지되었다. 0.008 N 농도에서의 황산은 0.6 m/분의 흐름 속도에서의 이동상으로서 사용되었다. 다양한 성분들의 정량화가 210 nm에서의 그들의 흡광도를 측정하는 것에 의하여 이루어졌다.

- [0132] 실시예
- [0133] 실시예 1
- [0134] KJ122 및 *이. 콜리(E. coli)* C 야생형(ATCC 8739)의 게놈 서열의 비교
- [0135] 일련의 유전적 조작들 및 많은 수 또는 라운드의 대사적 진화를 통하여 KJ122 스트레인이 *이. 콜리(E. coli)* C 스트레인으로 유도되었다. KJ122 스트레인의 제조에 대한 상세는 공개된 PCT 특허 출원 번호들로서 참조로서 여기에 포함된 WO/2008/15958 및 WO/2010/115067에서 제공된다. 잔타마 외.(Jantama *et al.*)(2008a 및 2008b)에 의한 두 개의 과학적 공개물들 및 장 외.(Zhang *et al.*)(2009a; 2009b)에 의한 두 개의 과학적 공개물이 또한 KJ122 스트레인의 제조를 기술하고 있다. 이들 과학적 공개물들 양자는 참조로서 여기 포함된다. 스트레인 KJ122 는 2008년 2월 20일 스트레인 지정 번호 B-50115로 USDA-ARS 배양균 컬렉션에 수탁되었다.
- [0136] KJ122의 게놈 서열은 터프스 대학 코어 퍼실리티(Tufts University Core Facility)에서 일루미나 서열화 시스템(Illumina sequencing system)을 사용하여 얻어졌고, 또한 상기 서열은 유전자 은행으로부터 입수 가능한 (수납 번호 CP000946) 주석 달린 야생형 *이. 콜리(E. coli)* C 서열과 비교되었다. 이러한 비교 게놈 분석으로부터, 이들 두 개의 스트레인들 사이에서의 많은 수의 유전적 차이점들이 검출되었다. 이러한 개시에서, 개방 리딩 프레임 내에서 발생하는 돌연변이들의 위치가, 염기 번호 1로서 개방 리딩 프레임(상기 참조된 유전자 은행 서열에서 주석 달린 바로서)의 개시 코돈의 제1 염기를 카운트하는, 많은 수의 뉴클레오타이드 염기를 바탕으로 하는 좌표들을 사용하여 주어질 것이다. KJ122 및 *이. 콜리(E. coli)* C 사이에서의 유전적 차이점의 목록으로부터, KJ122 제조 내의 다양한 단계들에서 집약적으로 도입되었던 결실들(deletions)이 입증되었고, 또한 대사적 진화의 프로세스 동안 발생하였던 유전적 변화들의 부가적인 목록이 확인되었다. KJ122의 게놈 내에 고정된 돌연변이들의 목록 내에 포함된 것은 1) 두 개의 피루베이트 키나제들 중 하나를 인코딩하는 *pykA* (SEQ ID NO.1) 내의 프레임 이동 돌연변이, 2) 슈거 운송자 유전자 *galP*와 같은 갈락토오스 유도성 유전자들의 억제자를 인코딩하는 *galS* 유전자(SEQ ID NO.2), 3) *이. 콜리(E. coli)* 게놈의 *ycdC* 내지 *ycdF*를 포함하는 많은 유전자들을 포함하는 *이. 콜리(E. coli)* 게놈의 영역에서의 48 킬로베이스 결실(SEQ ID NO. 3), 4) RNA 폴리머라제의 서브유닛을 인코딩하는 *rpoA* 유전자 (SEQ ID NO. 4)의 C-터미널 도메인 내에서의 포인트 미스센스 돌연변이(point missense mutation), 5) RNA 폴리머라제의 다른 서브유닛을 인코딩하는 *rpoC* 유전자 (SEQ ID NO. 5)의 F 영역 내에서의 포인트 미스센스 돌연변이(point missense mutation), 6) 글리세롤 디하이드로지나제를 인코딩하는 *gldA* 유전자 (SEQ ID NO. 6) 내의 프레임 이동 돌연변이, 7) PEP-의존성 디하이드로아세톤 키나제(PEP-dependent dihydroxyacetone kinase)의 서브유닛을 인코딩하는 *dhaM* 유전자 (SEQ ID NO. 7) 내의 프레임 이동 돌연변이, 및 8) 세포 벽 합성 및 세포 형태에 관여된 유전자를 인코딩하는 *ftsI* 유전자 (SEQ ID NO. 8) 내에서의 포인트 미스센스 돌연변이(point missense mutation)였다.
- [0137] 실시예 2
- [0138] KJ122 내에서 돌연변이화된 *pykA* 유전자의 클리어링
- [0139] 야생형 *이. 콜리(E. coli)* 및 많은 다른 박테리아가 글루코오스 포함 배지 내에서 성장될 때에, 세포 내부에서의 피루브산염 생성은 피루브산염의 분자의 결과적인 형성을 가져오는 글루코오스를 이송시키고 또한 인산화하기 위해 포스포에놀피루베이트(phosphoenolpyruvate (PEP)) 분자를 사용하는, 포스포트랜스퍼라제 시스템(phosphotransferase system (PTS))의 작용을 통해 일어난다. 피루브산염 형성은 또한 *pykA* 및/또는 *pykG* 유전자들에 의해 코딩되는 피루베이트 키나제 효소의 작용으로부터도 일어날 수 있다. PykA 및 PykF 단백질들은 동위 효소들(isoenzymes)이다. 글루코오스가 탄소의 소스일 때, 피루베이트 키나제들의 동위 효소들이 피루브산염 생합성에서 활발한 역할을 갖는다. 그러나, 트리플 돌연변이체, *pts*, *pykA* 및 *pykF*는, 피루브산염을 형성하는 세포의 능력이 결여되었거나 또는 많이 감소되었기 때문에 단독 탄소원으로서 글루코오스 상에서 성장할 수 없다(폰세 외.(Ponce *et al.*), 1995).
- [0140] KJ122가 PTS의 성분인(장 외.(Zhang *et al.*), 2009a 장 외.(Zhang *et al.*), 2009b), 단백질 *ptsI*에 대하여 코딩하는 *ptsI* 내에서의 돌연변이들을 포함하고 있기 때문에, 대사적 진화의 과정에서 *pykA* 유전자가 KJ122 내에서 또한 돌연변이화되었다는 것의 관찰은 기대되지 않았다. 이러한 돌연변이가 숙신산 생성에 대해 기여하는지를 결정하기 위하여, 새로운 스트레인 WG85a를 공급하는 야생형 *pykA* 대립 형질이 KJ122 내에 장착되었고, 또한 숙신산 생산에 대한 이러한 돌연변이의 효과의 관점에서 표현형적 변화들이 평가되었다. WG85a는 상기 기재된 두 단계 유전자 대체 방법을 사용하여 두 단계들에서 제조되었다. 제1 단계에서, WG84를

공급하기 위하여 *cat-sacB* 카세트가 스트레인 KJ122의 *pykA* 로커스에 장착되었다. 사용된 PCR 프라이머들은 SEQ ID NO.21 및 22였다(표 1 참조). 제2 단계에서, 스트레인 WG85a를 공급하기 위하여 *이. 콜리(E. coli)* C로부터의 야생형 *pykA* 유전자가 WG84 내로 장착되었다. 사용된 PCR 프라이머들은 SEQ ID NO.23 및 24였다(표 1 참조). KJ122에 대한 약 550 mM과 비교할 때 WG85a는 약 160 mM 숙신산염을 생산하였다. 분명하게도 KJ122 내에서의 *pykA* 돌연변이는 숙신산염 생산에 대하여 중요하다. KJ122 내에서의 *pykA* 돌연변이는 프레임 이동이기 때문에, *pykA*의 결실을 포함하는 스트레인은 유사하게 작용할 것이라는 것이 추론된다. 중간체 스트레인, WG84는 *pykA* 개방 리딩 프레임의 완전한 결실을 포함하며, 또한 예상되는 바와 같이, WG84는 540 mM의 숙신산염 타이터를 갖는 KJ122와 유사하게 작용하였고, 이는 KJ122 내에서의 프레임 이동 돌연변이가 제로 돌연변이(null mutation)와 유사하다는 것을 보여준다. 이러한 발견을 이룸과 함께, 발명자들은 다음으로 피루베이트 키나제 활성의 감소된 수준이 KJ122에 대하여 중요하며 또한 이러한 상태는 관련된 많은 접근법들 중 어떤 하나에 의해 달성될 수 있다는 것을 추론할 수 있었다. 예로서, KJ122 내에서의 *ptsI*^{*} 활성은 추가적으로 감소되거나 또는 제거될 수 있고, *PykF* 활성은 작아지거나 또는 제거될 수 있으며, 또는 *PykA*, *PykF*, 및/또는 *ptsI*의 감소된 활성의 조합이 현재 숙신산염 생성에 있어서 중요한 것으로 알려져 있는 총 피루베이트 키나제 활성에서의 감소를 달성하는데 있어서 사용될 수 있었다.

[0141] 제2 피루베이트 키나제 유전자, *pykF*에서의 결실이 *pykA*에서의 결실을 대체할 수 있는지를 결정하기 위하여, 스트레인 WG89를 공급하는 *pykF*에서의 결실이 스트레인 WG85a 내로 장착되었다. 이것은 두 단계 유전자 대체 방법을 사용하여 완성되었다. 제1 단계에 대하여, 스트레인 WG87을 공급하기 위하여 *kan-sacB* 카세트가 WG85a의 *pykF* 로커스에 장착되었다. 사용된 PCR 프라이머들은 SEQ ID NO.36 및 37이었다(표 1 참조). 제2 단계에 대하여, 새로운 스트레인 WG89를 공급하기 위하여 *pykF* 개방 리딩 프레임의 결실이 WG87 내로 장착되었다. 사용된 PCR 프라이머들은 SEQ ID NO.38 및 39였고(표 1 참조), 또한 템플레이트는 pGW191(SEQ ID NO. 35)이었다. 작은 스케일 발효제들에서, WG89는 더욱 열악하게 성장하였고, KJ122에 대한 510 mM과 비교하여, 단지 115 mM 숙신산염을 생산하였다. 따라서, *pykF*의 결실은 숙신산염 생산을 향상시키기 위하여 *pykA*의 결실 또는 프레임 이동에 대해 대체할 수 없다. WG89 및 WG85a가 다른 실험들에서 테스트되었다고 하더라도, WG89가 WG85a보다 더욱 열악하게 수행하였다는 것은 분명하고, 이는 KJ122에서의 성장 및 숙신산 생산에 대한 피루베이트 키나제의 최적의 수준이 존재하며, 그러나 *pykF*의 결실은 우수한 숙신산염 생산에 대하여 최적이지 않은 총 피루베이트 키나제 활성의 수준의 결과를 가져온다는 것을 보여준다. 따라서, 대사적 진화는 최적의 수준으로 근접한 피루베이트 키나제의 수준을 갖는 스트레인을 생산하였다. 나아가 미국 특허 출원 공개 2009/0075352 및 리 외(Lee *et al.*)(2005)에 제시된 바와 같이 피루베이트 키나제를 인코딩하는 모든 유전자들(*pykA*, *pykF* 및 하나 이상의 *pts* 유전자들)의 결실은, 총 피루베이트 키나제 활성이 그러한 스트레인 내에서 너무 낮을 수 있기 때문에, 숙신산염 생산을 향상시키지 않을 것이다. 새로운 스트레인, 예로서 *사카로마이스 세레비시아(Saccharomyces cerevisiae)* 내에서 새로운 숙신산염 생산 스트레인을 제조할 때에, 본 발명에 따르면, 원하지 않는 경로들을 차단한 후에, 총 피루베이트 키나제의 수준은 숙신산염 생산을 위한 우수한 효율을 제공하도록 조정될 것이다. 이는 예로서 *PYK2*, 비본질적인 피루베이트 키나제 유전자,를 삭제하는 것 및 그런 다음 *PYK1*의 앞부분에 길이를 변화시키는 프로모터들을 장착시키거나, 또는 *PYK1* 개방 리딩 프레임의 5' 말단 또는 3' 말단으로부터의 진행하는 결실(progressive deletion)을 만듦으로서 남아있는 *PYK1*의 발현 또는 활성 수준을 변화시키는 것, 및 향상된 숙신산염 생산에 대해 테스트하는 것에 의해 달성될 수 있다.

[0142] 실시예 3

[0143] XZ722 내에서의 *galS* 돌연변이화

[0144] *이. 콜리(E. coli)* C 및 KJ122 스트레인 사이에서의 비교 계놈 서열 분석은 *galS* 유전자 내에서의 프레임 이동 돌연변이를 나타내었다. 이 유전자는 갈락토오스 퍼미아제(galactose permease), GalP를 인코딩하는 *galS* 유전자의 발현을 억제하는 것으로 알려져 있다. GalP 단백질은, PTS의 성분인 PtsI 단백질에 대하여 코딩하는 *ptsI*가 돌연변이화되는 KJ122 스트레인에서의 경우와 같이, 글루코오스 이송에 대하여 완전히 기능적인 포스포트랜스퍼라제 시스템의 결여 하에서 적어도 부분적으로 원인이 되는 것으로 보고되어 있다(장 외.(Zhang *et al.*), 2009a). 따라서 KJ122 내에서의 *galS* 돌연변이가 글루코오스 섭취 또는 숙신산염 생산의 속도에 대한 어떤 다른 영향의 관점에서 어떤 기능적인 의미를 갖는다는 것이 가능하다. 이것은 새로운 스트레인 WG86a를 공급하기 위하여 *galS*의 KJ122 대립 형질을 *이. 콜리(E. coli)* 스트레인 XZ722($\Delta pf1B$, *ptsI*^{*}, *pck*^{*})으로 이동시키는 것 및 이러한 스트레인 내에서의 숙신산 생산의 속도에 대한 돌연변이의 영향을 평가하는 것에 의해

테스트된다. WG86a는 상기 기재된 두 단계 유전자 대체 방법을 사용하여 두 단계들에서 제조되었다. 제1 단계에서, 스트레인 WG83을 공급하기 위하여 *kan-sacB* 카세트가 스트레인 XZ722의 *galS* 로커스에 장착되었다. 사용된 PCR 프라이머들은 SEQ ID NO. 17 및 18이었다(표 1 참조). 제2 단계에서, WG86a를 공급하기 위하여 KJ122 유전자로부터의 *galS*의 프레임 이동된 대립 형질이 장착되었다. 사용된 PCR 프라이머들은 SEQ ID NO. 19 및 20이었다(표 1 참조). 작은 스케일 발효제들 내에서, XZ722로부터의 숙신산염 타이터는, WG86a의 그것이 210 mM이었던 것에 비해, 175 mM이었으며 이는 *galS* 돌연변이가 향상된 숙신산염 생산에 기여할 수 있다는 것을 제공한다. *galS* 돌연변이는 프레임 이동이며, 따라서 *galS*에서의 결실은 유사하게 작용할 것이고, 그러나 유전적으로 더욱 안정적인 것으로 추론될 수 있다. 중간체 스트레인 WG83은 *galS* 개방 리딩 프레임의 결실을 포함하고, 또한 그것은 WG86a와 유사하게 행동하였으며, 220 mM 숙신산염을 생산하였다. 이에 더하여, 발명자들은 *galP* 발현의 다른 억제자를 인코딩하는 것으로 알려져 있는, *galR* 유전자를 돌연변이하는 것이 또한 세포의 글루코오스 유입 능력을 증가시킴으로써 숙신산 생산을 증진시킬 것이라는 것을 추론할 수 있다. 나아가, 발명자들은 *galP* 유전자의 카피(copy) 수에 있어서의 증가가 동일한 효과를 가지며 또한 숙신산염 생산의 효율을 증가시키는 다른 접근법을 제공한다는 것을 추론할 수 있다. 조절 사이트를 가지거나 또는 가지지 않는, *galP* 유전자 및 프로모터와 터미네이터를 포함하는 인접하는 상부 및 하부 DNA는 PCR에 의하여 증폭될 수 있고 또한 다음으로 자연적인 *galP* 로커스로부터 이격되어 있는 하나 이상의 사이트들에서 *이. 콜리(E. coli)* 스트레인을 생산하는 숙신산염의 염색체에 장착될 수 있어, 하나 초과와 *galP* 유전자의 카피를 포함하는 스트레인을 제공한다. 결과적인 스트레인은 더 높은 수준의 GalP 단백질을 갖는 것에 의해 숙신산 생산에 대하여 향상될 것이다(실시예 11 참조).

[0145] 실시예 4

[0146] 염색체 DNA로부터의 48 kbp 영역 결실

[0147] *이. 콜리(E. coli)* C 및 KJ122 스트레인 사이에서의 비교 게놈 서열 분석은 대사적 진화의 프로세스 과정 동안 KJ122 내 게놈의 유전자들 *ycdC* 내지 *ycdF* 을 포함하는 48 킬로베이스(kilobase) 영역의 결실을 나타내었다. 이러한 결실은 *이. 콜리(E. coli)* ATCC 8739의 뉴클레오티드 좌표들 2,416,108 및 2,464,284(양자 모두 포함되는 뉴클레오티드들) 사이에서의 서열에 대응한다. 이러한 유전자 결실의 기능적인 중요성 및 숙신산 생산에 대한 그것의 영향력은 스트레인 WG110을 공급하는 KJ122 내에 48 킬로베이스 서열을 재장착하는 것에 의하여 측정된다. 이것은 유전자 대체 방법 및 P1 *vir* 형질 도입의 조합을 사용하는 두 단계들에서 완성된다. 제1 단계에 대하여, 스트레인 WG51을 공급하기 위하여 KJ122의 결실 중점들 사이에 *cat-sacB* 카세트가 장착되었다. 사용된 PCR 프라이머들은 SEQ ID NO. 31 및 32였다(표 1 참조). 제2 단계에서, 새로운 스트레인 WG110을 제공하기 위하여, 제공자로서 *이. 콜리(E. coli)* C와 함께 P1 *vir*을 사용하여, 수크로스 저항성 및 염색체 민감성에 대하여 WG51이 형질 도입되며, 이제 이것은 48 킬로베이스 서열이 KJ122 스트레인 배경 내에서의 그것의 자연적인 로커스에서 원상 복구되게 한다. *이. 콜리(E. coli)*의 스트레인들의 특정 영역들의 결실은 미국 특허 출원들 공개들 2009/0075333 및 2008/0009041에 기재되어 있는 방법들을 사용하여 달성될 수 있다. KJ122의 특정 48 킬로베이스 결실은 (1) WG51로부터 상기 수여 스트레인(recipient strain) 내로 형질 도입하여, 클로람페니콜(chloramphenicol) 저항성에 대하여 선택하고, 및 (2) KJ122로부터 단계 1 내에서 제조된 스트레인으로 형질 도입하여, 수크로스 저항성에 대하여 선택하며 또한 클로람페니콜(chloramphenicol) 민감성에 대하여 스크리닝하는 것에 의한 두 단계들로 다른 수여 *이. 콜리(E. coli)* 스트레인들 내에 장착될 수 있다.

[0148] 분명하게도, 48 킬로베이스 결실 하에 포함되는 유전자들은 최소한의 글루코오스 배지 내에서의 성장을 위하여 또는 숙신산 생산을 위하여 필수적이지 않다. 48 kb 결실의 유리한 점은 결과적인 염색체가 약간 짧아, 이것이 성장의 속도에서의 약간의 증가를 제공할 것이라는 점이다. 나아가, 결실 하의 적어도 4 유전자들이 존재하고, 이들은 단독으로 삭제될 때에, 최소한의 글루코오스 배지 내에서의 성장에 대해 더 높은 밀도의 결과를 가져온다. 각각 유전자들 *ycdI*, *ycdL*, *ycdO*, 및 *ycdV* 내에서의 결실들을 포함하는, 케이오 결절 컬렉션(Keio deletion collection)의 스트레인들 JW5226, JW1427, JW5229, 및 JW1438은, 모두 최소한의 글루코오스 배지 내에서, 48 시간에서 상당히 더 높은 OD600까지로 성장한다(바바, 티. 외(Baba, T. et al)(2006) 내의 보충적 표 3 참조).

[0149] 실시예 5

- [0150] KJ122 내에서의 돌연변이화된 *rpoA* 유전자의 큐어링
- [0151] RNA 폴리머라제의 코어 엔자임은 화학양론 $\alpha_2\beta\beta'\omega$ 내의 네 개의 다른 폴리펩티드 서브 유닛들: 알파(α), 베타(β), 베타'(β') 및 오메가(ω)를 포함한다. RNA 폴리머라제의 알파 서브 유닛은 *rpoA* 유전자에 의해 코딩된다. 이러한 서브 유닛은 상기 효소의 조합을 위하여 요구되며 또한 그것은 어떤 조절 단백질들과 상호 작용한다. KJ122 및 *이. 콜리(E. coli)* C 스트레인들 사이에서의 비교 게놈 서열 분석은 상기 *rpoA* 유전자 내에서의 미스센스 포인트 돌연변이(missense point mutation)를 나타내었다. 더욱 구체적으로, 이러한 돌연변이는 RNA 폴리머라제 알파 서브 유닛 C-터미널 도메인(α -CTD) 내에서의 아미노산 포지션 322에 위치된다. 야생형 *이. 콜리(E. coli)* C 스트레인 내에서의 프롤린 잔기는 상기 KJ122 스트레인 내 류신 잔기로 변환되었다.
- [0152] 프롤린 322을 알라닌으로 변화시키는 다른 돌연변이는 종래 *metE* 유전자의 발현에 대하여 부정적인 효과를 가지는 것으로 기재되었다(프리트쉬 외.(Fritsh *et al.*), 2000). 다른 보고에서, RpoA 단백질 내의 이 위치에서 프롤린에서 알라닌으로의 돌연변이는 *이. 콜리(E. coli)* 내의 *rhaS* 프로모터의 사이클릭인 AMP 수용자 단백질(CRP) 활성화를 감소시키는 것으로 발견되었다. KJ122의 RpoA 단백질 내에서 프롤린에서 류신으로의 포인트 돌연변이의 영향은 새로운 스트레인 RY859A1을 공급하기 위해 *rpoA* 돌연변이의 상기 야생형 대립 형질을 KJ122 내로 이동시키는 것 및 숙신산 생산에 대한 이 돌연변이의 영향을 결정하는 것에 의해 평가될 수 있었다. RY859A1은 P1 *vir* 형질 도입을 사용하는 두 단계들로 제조되었다. 일차적으로, 연결된 *aroE::kan* 대립 형질이 제공자로서 케이오 결실 컬렉션(Keio deletion collection)(콜리 유전적 스톡 센터, 예일 대학, 뉴 헤븐, CT, USA(Coli Genetic Stock Center, Yale University, New Haven, CT, USA)로부터 입수 가능한)으로부터의 JW3242를 사용하여 KJ122 내로 형질 도입되었고, 또한 카나마이신 저항성에 대하여 선택하였다. 제2 단계에 대하여, 야생형 *이. 콜리(E. coli)* C가 제공자로서 사용되었고, 선택은 성장 및 최소한의 글루코오스 배지 상에서의 성장에 대한 것이었다. RY859A1의 *rpoA* 유전자 서열이 결정되었고, 또한 야생형인 것으로 보여졌다. 작은 스케인 발효제들 내 48 시간에서, RY859A1은 370 mM의 숙신산염을 생산하였고, 이는 385 mM을 생산하였던 KJ122 보다 약간 적은 것이었다. 발효는 듀플리케이트에서 수행되었고, 또한 재생 가능하였다.
- [0153] 실시예 6
- [0154] KJ122 내에서 돌연변이화된 *rpoA* 유전자의 큐어링
- [0155] KJ122 및 *이. 콜리(E. coli)* C 스트레인 사이에서의 비교 게놈 서열 분석은 *rpoC* 유전자에 의하여 코딩되는 RNA 폴리머라제의 β' 서브 유닛의 F 영역에서의 미스센스 포인트 돌연변이(missense point mutation)를 나타내었다. RpoC 단백질은 RNA 합성의 과정에서 DNA 템플레이트에 결합한다. 포지션 747에서의 메티오닌 잔기는 KJ122 내 이소류신으로 대체되었다. KJ122의 RpoC 단백질 내에서의 이러한 포인트 돌연변이의 영향은 새로운 스트레인 RY862A를 제공하기 위하여 *rpoC* 돌연변이의 야생형 대립 형질을 KJ122로 이동시키고, 또한 숙신산 생산에서의 이러한 돌연변이의 영향을 결정하는 것에 의하여 평가되었다. RY862A는 P1 *vir* 형질 도입을 사용하여 제조되었다. 연결된 *thiG::kan* 대립 형질이 제공자로서 케이오 결실 컬렉션(Keio deletion collection)(콜리 유전적 스톡 센터, 예일 대학, 뉴 헤븐, CT, USA(Coli Genetic Stock Center, Yale University, New Haven, CT, USA)로부터 입수 가능한)으로부터의 JW5549를 사용하여 KJ122 내로 형질 도입되었고, 또한 카나마이신 저항성에 대하여 선택하였다. RY862A의 *rpoC* 유전자의 서열이 결정되었고, 또한 야생형인 것으로 보여졌다. 5 mg/리터에서의 티아민 HCl이 RY862A의 성장을 보장하기 위하여 RY862A 및 컨트롤 KJ122에 대한 발효 배지에 더하여졌다. 작은 스케인 발효제들 내 48 시간에서, RY862A의 숙신산염 타이터는 KJ122의 그것, 425 mM 보다 상당히 적은 320 mM 였다. RY862A의 개시 성장 속도는 또한 KJ122의 그것에 비해 느렸고, 0.23 OD550/hr에 대하여 비교되는 0.15 OD550/hr이었다. 발효는 듀플리케이트에서 수행되었고 또한 재생 가능하였다.
- [0156] 실시예 7
- [0157] KJ122 내 돌연변이화된 *gldA* 유전자의 큐어링
- [0158] NAD^+ -의존성 글리세롤 디하이드로지나제(NAD^+ -dependent glycerol dehydrogenase)는 *gldA* 유전자에 의해 인코딩된다. GldA 단백질은 글리세롤의 디하이드록시아세톤으로의 가역적인 산화를 촉매하며, 또한 메틸글록살(methylglyoxal)의 R-락탈알데히드(R-Lactaldehyde)로의 환원을 촉매할 수 있다. 비교 게놈 서열 분석은 KJ122 내 *gldA* 유전자가 프레임 이동 돌연변이를 포함한다는 것을 나타내었다. 숙신산 생산에 대한 *gldA* 유전자의

돌연변이화된 형태의 영향은, 새로운 스트레인 AC6를 제공하기 위해, 상기 기재된 두 단계 유전자 대체 방법을 사용하여, 야생형 *gldA* 유전자를 KJ122 내로 장착하는 것에 의해 평가되었다. 제1 단계에서, *cat-sacB* 카세트가 새로운 스트레인 AC5를 제공하기 위해 스트레인 KJ122의 *gldA* 로커스에 장착되었다. 사용된 프라이머들은 SEQ ID NO. 9 및 10이었다(표 1 참조). 제2 단계에서, *이. 콜리(E. coli)*로부터의 야생형 *gldA* 유전자가 새로운 스트레인 AC6를 제공하기 위하여 AC5 내로 장착되었다. 사용된 프라이머들은 SEQ ID NO. 13 및 14였다(표 1 참조). 작은 스케일 발효제들 내 96 시간에서, AC6의 숙신산염 타이터는, 530 mM이었던 KJ122의 그것에 대해 약간 높은, 580 mM이었다. 그러나 AC6의 개시 성장 속도는 0.20 OD550/hr 였던 KJ122의 그것에 비해 약간 느린, 0.16 OD550/hr였다. 따라서, 돌연변이화된 *gldA*는 KJ122의 대사적 진화의 과정을 위해 선택될 수 있는 약간의 성장 장점을 제공한다.

[0159] 실시예 8

[0160] KJ122 내 돌연변이화된 *dhaM* 유전자의 큐어링

[0161] 디하이드록시아세톤 키나제(Dihydroxyacetone kinase)는, 하나의 오페론, *dhaKLM* 내 세 개의 유전자들에 의해 인코딩되는, 포스포트랜스퍼라제 시스템(phosphotransferase (PTS) system)에 관련된 다중 서브 유닛 단백질이다. DhaM 서브 유닛은 만노스 이송자(mannose transporter)의 IIA 도메인, 포스포릴 캐리어 단백질 Hpr 및 효소 1의 N-터미널 도메인에 대해 서열 유사성을 갖는다. *이. 콜리(E. coli)*에서, DhaM은 디하이드록시아세톤의 인산화에서의 인산염에 대한 제공자로서 ATP 대신 PEP를 사용한다. DhaM은 효소 1 및 포스포릴 캐리어 단백질 HPr을 통해 PEP에 의해 인산화된다. KJ122의 *dhaM* 유전자는 프레임 이동 돌연변이를 포함한다. KJ122 내에서의 *dhaM* 돌연변이의 영향을 평가하기 위하여, 야생형 *dhaM* 유전자가 새로운 스트레인 AC2를 제공하기 위하여 상기 기재된 두 단계 유전자 대체 방법을 사용하여 KJ122 내로 장착되었다. 제1 단계에서, 새로운 스트레인 AC1을 제공하기 위하여 *cat-sacB* 카세트가 스트레인 KJ122의 *dahM* 로커스에 장착되었다. 사용된 프라이머들은 SEQ ID NO. 11 및 12였다(표 1 참조). 제2 단계에서, *이. 콜리(E. coli)*로부터의 야생형 *dahM* 유전자가 새로운 스트레인 AC2를 제공하기 위하여 AC1 내로 장착되었다. 사용된 프라이머들은 SEQ ID NO. 15 및 16이었다(표 1 참조). 작은 스케일 발효제들 내 96 시간에서, AC2의 숙신산염 타이터는, 530 mM이었던 KJ122의 그것에 대해 약간 높은, 560 mM이었다. 그러나 AC6의 개시 성장 속도는 0.20 OD550/hr 였던 KJ122의 그것에 비해 약간 느린, 0.16 OD550/hr였다. 따라서, *dahM*를 돌연변이화 하는 것은 KJ122의 대사적 진화의 과정동안 선택될 수 있는 약간의 성장 장점을 제공한다. DhaKLM은 숙신산 생합성에 대한 기질인 PEP를 소비하기 때문에, PEP의 전환은 KJ122의 성장 장점을 위한 메카니즘이다. *dahKLM* 오페론 내에서의 유전자들을 돌연변이화 하는 것에 더하여, 또는 대신에, 어떤 다른 PTS-의존성 디하이드록시아세톤 키나제의 서브 유닛을 인코딩하는 유전자는 또한 그들의 숙신산 생산에서의 중요성을 테스트하기 위하여 돌연변이화 될 수 있다.

[0162] 실시예 9

[0163] KJ122 내 *ftsI* 유전자 내에서의 미스센스 돌연변이의 큐어링

[0164] *ftsI* 유전자는 세포 벽 및/또는 격벽(septation) 내에 포함된 필수 단백질을 인코딩한다. KJ122의 *ftsI* 유전자는 C로부터 T로의 코딩하는 서열의 염기 619를 변화시키는 미스센스 돌연변이를 포함하며, 이는 아르기닌 207을 시스테인으로 변화시킨다. KJ122의 FtsI 단백질에서의 아르기닌의 시스테인으로의 포인트 돌연변이의 효과는 *ftsI* 돌연변이의 야생형 대립 형질을 새로운 스트레인 RY858G1을 제공하기 위하여 KJ122 내로 이동시키고 또한 숙신산 생산에 대한 이러한 돌연변이의 효과를 결정하는 것에 의해 평가되었다. FtsI이 필수적인 단백질이기 때문에, 두 단계 유전자 대체 방법은 작동하지 않을 것이고, 따라서 두 단계 P1 *vir* 형질 도입이 RY858G1을 제조하는 데에 대신 사용되었다. 제1 단계에서, 연결된 *leuA::kan* 대립 형질이 제공자로서 케이오 결실 컬렉션(Keio deletion collection)(콜리 유전적 스톡 센터, 예일 대학, 뉴 헤븐, CT, USA(Coli Genetic Stock Center, Yale University, New Haven, CT, USA)로부터 입수 가능한)으로부터의 JW0073을 사용하여 KJ122 내로 형질 도입되었고, 또한 새로운 스트레인 RY853G1을 제공하기 위하여, 카나마이신 저항성에 대하여 선택하였다. 제2 단계에 대하여, *이. 콜리(E. coli)* C가 제공자로서 사용되었고, RY853G1이 수여자였으며, 선택은 최소한의 글루코오스 배지 상에서의 성장에 대한 것이었다. RY858G1의 *ftsI* 유전자의 서열이 결정되었고, 또한 야생형인 것으로 보여졌다. 작은 스케인 발효제들 내에서, RY853G1은 KJ122에서와 거의 동일한 개시 성장 속도로 성장하였으나, 72시간에서, RY853G1은 460 mM을 생산하였던 KJ122 보다 약간 적은 숙신산염 440 mM을 생산하였다. 따라서 KJ122의 *ftsI* 유전자 내에서의 미스센스 돌연변이는 숙신산염 타이터에

서의 증가에 대해 약간 기여한다.

[0165] 실시예 10

[0166] 재조합 돌연변이들의 부가적인 효과

[0167] 상기 실시예들에 상세하게 기재된 스트레인 비교들의 결과들은 표 3에 요약되어 있다. 동시에 테스트되었던 돌연변이들의 많은 수가 숙신산 타이터 또는 성장 속도에서의 단지 약간의 증가만을 제공하였다. 그러나, 이러한 증가들은 그것의 조상들에 대하여 상대적으로 KJ122에 대하여 발견되었던 모든 향상(잔타마 외.(Jantama *et al.*), 2008a; 잔타마 외.(Jantama *et al.*), 2008b)을 제공하는데 있어서 부가적인 및/또는 시너지적임에 틀림이 없다. 예로서, *rpoA*에서의 미스센스 돌연변이 및 *rpoC*에서의 미스센스 돌연변이 양자에 대하여 큐어링된 KJ122의 유도체는, 야생형 *rpoC* 유전자를, 야생형 *rpoA* 유전자를 포함하는 RY859A1으로 형질 도입함으로써 제조되었고, 새로운 스트레인 RY860A의 결과를 가져왔다. RY860A의 제조는, 그 수여 스트레인이 KJ122 대신 RY859A1이었던 것을 제외하고, 실시예 6에 상기 기재된 RY862A와 동일한 방법으로 수행되었다. 작은 스케일 발효제들에서 48 시간에서 5 mg/l의 티아민 HCl의 보충으로, RY860A는 KJ122(425 mM) 보다 및 단지 *rpoC* 돌연변이에 대해 큐어링되었던, RY862A(320 mM) 보다 상당히 적은 양의 숙신산염(300 mM)을 생산하였다.

[0168] 실시예 11

[0169] 나이브(naive) 스트레인으로로부터의 새로운 숙신산염 생산자의 제조

[0170] 스트레인 TG128은 유전자형(genotype) Δ frdABCD, Δ adhE, Δ mgsA, 및 Δ ack를 가지고, *이. 콜리(E. coli)* W(ATCC 9637)로부터 유도되었던 D-락테이트 생산자이다. TG128은 2005년 7월 25일 농업 리서치 서비스 배양균 컬렉션, 1815N. 유니버시티 스트리트, 페오리아, 일리노이, 61604, 미국(Agricultural Research Services Culture Collection, 1815 N. University Street, Peoria, Illinois, 61604 U.S. A.)으로 수탁되었고, 스트레인 접근 번호 NRRL B-30962를 갖는다. TG128은 그런 다음 일차적으로 수크로스 상에서의 성장에 대해, 이차적으로 섭씨 39도에서 성장에 대해, 및 삼차적으로 섭씨 40도에서의 성장에 대해 선택함으로써 대사적으로 진화되었다. 결과적인 스트레인은 TR160으로 명명되었다. TG160은 그런 다음 숙신산 생산 스트레인으로 재제작되었다. KJ122로부터의 대립 형질들이, 상기 기재된 두 단계 유전자 대체 방법을 사용하여 하기의 연대순으로 장착되었다: *pck*^{*}, Δ *ldhA*, *frdABCD*⁺, 및 *ptsI*^{*}. *pck*^{*} 및 *ptsI*^{*} 돌연변이들은 KJ122 계통 내에서 발견되었거나 숙신산 생산의 효율을 증가시키는 것으로 보여졌던 양자의 돌연변이들이었다(장 외.(Zhang *et al.*), 2009a). 결과적인 스트레인, WG32b는 그런 다음 작은 미생물 발효제들 내에서, 스트레인 WG32b-T23을 제공하기 위하여, 약 126 세대들(23 이동들)에 대하여 성장되었다. WG32b-T23은 작은 스케일 발효제들 내에서 약 260 mM 숙신산을 제조하였던 적당히 우수한 숙신산 생산자이다. 다음으로, 두 개의 부가적인 돌연변이들이, KJ122의 계통에서의 본 발명의 발견들을 바탕으로 하여, WG32b-T23에 부가되었다. 일차적으로, 자연적인 프로모터 및 터미네이터를 포함하는 서열들 측면에 위치하도록 함께 포함하는 *galP* 유전자의 두 번 째 카피가, 두 단계 유전자 대체 방법을 사용하여, *dnaA* 유전자 근처의 로커스에서 계통 내에 집적되었다. 삽입 사이트에서의 DNA 서열은 다음과 같다:

```
attaaattttccaatatgcggcgtaaatcgtagccgcctcgcggcaggatcggttacacttagcgagttctggaaagtcctgtggataa
atcgggaaaaatctgtgagaaacagaagatct - 삽입 사이트 -
```

cttgcgcagtttaggctatgatccgcggtcccgatcggtttgcaggatcttgactcgggcatataaccgcagacagcggt. 제1 단계에서, *kan-sacB* 카세트가 스트레인 WG62를 제공하기 위하여 *dnaA* 로커스에서 삽입되었다. 사용된 PCR 프라이머들은 SEQ ID NO. 25 및 26이었다(표 1 참조). 제2 단계에서, *galP*의 두 번째 카피가 *dnaA* 로커스에 장착되었다. 사용된 PCR 프라이머들은 SEQ ID NO. 27 및 28이었다(표 1 참조). 결과적인 스트레인, WG74로 명명된, 은 작은 스케일 발효들 내에서 약 320 mM의 숙신산을 생산하였다. 이차적으로, *pykA* 유전자의 완전한 개방 리딩 프레임이 WG74로부터 삭제되었고 또한 새로운 스트레인 WG96을 제공하기 위하여 *cat-sacB* 카세트로 대체되었다. 사용된 PCR 프라이머들은 SEQ ID NO. 29 및 30이었다(표 1 참조). 작은 스케일 발효들 내에서, WG96은 450 mM의 숙신산을 생산하였고, 이것은 전구체 스트레인들 WG32b-T23 및 WG74의 그것에 대하여 분명한 향상이다. 따라서, 본 발명자들은 KJ122 내에서의 돌연변이들의 추출물 성질을 발견하는 것이 다른, 덜 진화된 부모 스트레인들로부터 어떻게 새로운, 향상된 숙신산을 생산하는 스트레인들을 제조할 것이냐에 대한 고유한 정보 및 통찰력을 제공한다는 것을 분명히 확립하였다. 나아가, 본 발명의 발견들은 광범위하게 적용 가능하다. 예로서,

galS 돌연변이에 대한 수득(learning)은 새로운 스트레인 내에 KJ122 안에서 발견되는 동일한 돌연변이를 장착하는 것에 의해서 뿐만 아니라, GalS 억제제의 타겟, *galP* 유전자를 듀플리케이트 하는 것에 의해서도 적용될 수 있다. 또 다른 예로서, *pykA* 돌연변이의 수득은 새로운 스트레인 내에 KJ122 안에서 발견되는 동일한 돌연변이를 장착하는 것에 의해서 뿐만 아니라, *pykA*의 결실을 장착하는 것에 의해서도 적용될 수 있다.

- [0171] 여기 개시된 발명은 숙신산 외 다른 화합물들을 생산하기 위한 스트레인들 및 프로세스들을 향상시키는 것에 대하여 일반화될 수 있다. 제공된 실시예들은, 제한하는 것이 아니라, 예시적인 것으로 의도된다.
- [0172] GalP 단백질은 글루코오스 및 다른 슈거들을, ATP의 대강 1/3의 세포를 소모하는, 프로톤 심포트 메카니즘(proton symport mechanism)에 의해 유입하기 위해 작용한다. 유입 후에, 글루코오스와 같은 슈거는 추가적인 ATP를 소모하는, 약 총 1.33 ATP의 소모에 대하여, 인산화되는 것을 요구한다. 대사적 에너지의 관점에서, 이것은, 에너지적으로 약 2 ATP와 동일한, 하나의 PEP를 소모하는, PTS 시스템에 의해 글루코오스와 같은 슈거를 유입하는 것보다 덜 비싼 것이다. 따라서, GalP에 의한 글루코오스와 같은 슈거를 유입하는 것은 PTS를 사용하는 것에 의하는 것보다 더 효율적이다.
- [0173] 나아가, 많은 발효적 프로세스에서, 특히 그러나 혐기성 및 미세 호기성 프로세스에 제한되는 것이 아닌, PEP는 숙신산 이외의 화합물들, 말산, 푸마르산, 아스파르트산염, 트레오닌, 메티오닌, 라이신, 및 기타와 같은, 의 생합성에 대한 중간체일 수 있다. 따라서 PEP를, PTS를 대신하여 GalP를 사용하는 것에 의해, 또한 일반적으로 피루베이트 키나제 활성을 감소시키는 것에 의해 전환함으로써 생합성의 효율이 향상될 수 있다.
- [0174] 많은 발효들에서, 최소한의 배지 내에서의 더 효과적인 성장이 바람직하며, 또한 KJ122 내에서 발견되는 48 kb 결실은 최소한의 배지에서 *이. 콜리(E. coli)*를 사용하는 어떤 발효에서의 효율을 향상시키는 것을 도울 것이다.
- [0175] KJ122에서 발견되는 *ftsI* 돌연변이는 더 작은 또는 더 큰 구형 세포를 유도하는 것에 의해 효율성을 향상시킬 것이며, 이것은 차례로 더 높은 부피에 대한 표면 영역 비를 유도한다. 이것은, 차례로, 더욱 효율적인 글루코오스, 수크로스, 말토스, 글리세롤 등과 같은 영양분의 유입을 유도할 것이고, 숙신산염, 말산염, 푸마르산염, 아스파르트산염, 트레오닌, 메티오닌, 라이신 등과 같은 생성물의 더 효율적인 유출을 유도할 것이다.
- [0176] *rpoA* 및 *rpoC*에서 발견되는 돌연변이들은 글루코오스에 의한 분해 대사질(catabolite) 억제의 일반적인 손실의 결과를 가져온다. 예로서, 저렴한 슈거들의 바람직한 소스들인 바이오 매스 가수분해물들은 통상적으로 글루코오스를 포함하는 슈거들의 혼합물을 포함한다. 다양한 다른 슈거들이 글루코오스와 함께 동시적으로 사용되는 것이 바람직하며 또한 분해 대사질 억제의 제거는 글루코오스와 함께 비-글루코오스 슈거들을 효율적으로 소비하고 대사하는 것에 대한 세포의 능력을 향상시킨다. 나아가, 분해 대사질의 제거는 바람직한 유전자들, *pck*, *galP*, *mdh*, *fumA*, *xylE*, *lacZ*, *lacY*, *fumB*, 및 *frdABCD*와 같은, 의 발현을 증가시킨다. 따라서, 본 발명의 RNA 폴리머라제 돌연변이들은 다양한 상업적으로 유용한 스트레인들 및 발효 프로세스들에 대한 광범위한 적용 가능성을 갖는다.
- [0177] 따라서, 본 발명에서 개시된 돌연변이들은 숙신산 생산을 증가시키기 위하여 유용하며, 그들은 그들이 숙신산 생산 또는 다른 유기산 생산에 대하여 종래 기재되지 않았다는 점에서 새롭고, 또한 그들은 본 기술 분야에서 숙련된 자가 본 발명들의 개별적인 돌연변이들 또는 이들 돌연변이들의 조합이 숙신산 또는 다른 유기산 생산을 향상시킬 것이라고 기대할 수 없었다는 점에서 진보적이다.
- [0178] KJ122의 게놈 서열 내에서 발견되었던 몇 개의 돌연변이들이 한번에 하나씩 개조될 때에 성장 또는 숙신산 생산에 대하여 단지 약간의 효과만을 생산하였다 하더라도, 성장의 속도에서의 단지 작은 증가는, 시간 주기가 많은 세대들을 커버함에 따라 대사적 진화의 과정 동안 선택될 수 있다는 점이 강조되어야 한다. 나아가, 그 안에서 각 개별적인 돌연변이가 부가적으로 또는 시너지적으로 기여하는 몇 가지 돌연변이들의 조합은, 성장의 효율 및/또는 숙신산 생산에서의 더 큰 증가를 부여할 것이다.
- [0179] 실시예 11에 기재된 바와 같이, 새로운, 이차적 부모 스트레인으로부터의 발효에 의한 바람직한 화합물을 생산하는 목적을 위하여 미생물들 제조하는 것에 대한 방법은 광범위하게 다양한 화합물들 및 스트레인들에 대하여 일반화될 수 있다. 일반적인 방법은 하기를 포함한다: (A) 상기 바람직한 화합물로부터 이격된 발효 생성물에 대한 생합성적 경로에서의 단계를 촉매하는 효소를 인코딩하는 적어도 하나의 유전자를 일차적 부모 스트레인으로부터 삭제하는 것으로, 제1 중간체 스트레인을 유도하는 것, (B) 제2 중간체 스트레인을 유도하기 위해 상기 제1 중간체 스트레인에 대한 대사적 진화를 수행하는 것, (C) 상기 부모 스트레인 및 상기 제2 중간체 스트레인의 게놈 DNA 서열을 결정하는 것, (D) 상기 대사적 진화의 과정을 위해 선택되었던 돌연변이

들을 확인하기 위하여 상기 게놈 DNA 서열들을 비교하는 것, (E) 상기 돌연변이들의 어느 것이 성장 속도 또는 상기 바람직한 화합물의 생산의 효율을 증가시키기 위하여 유리한지 결정하기 위하여 상기 돌연변이들의 적어도 하나를 테스트하는 것, (F) 적어도 하나의 상기 유리한 돌연변이들을 선택하는 것, 및 (G) 적어도 하나의 상기 유리한 돌연변이들, 또는 하나의 상기 유리한 돌연변이들과 기능적으로 유사한 적어도 하나의 돌연변이들을 이차적 부모 스트레인에서 장착하여 상기 미생물을 유도하는 것.

[0180] 이러한 일반적인 방법은 숙신산염 이외의 화합물들, 푸마르산염, 말산염, 아스파르트산염, 글루탐산염 및 이들 화합물들의 어떤 것의 유도체들과 같은, 및 *이. 콜리(E. coli)* 이외의 다른 미생물들, 다른 박테리아, 아르콘들(archons), 이스트들(yeasts), 곰팡이(filamentous fungi), 조류(algae), 및 디노플라젤레이트들(dino flagellates)을 포함하는, 에 대하여 적용될 것이다.

[0181] 출원인들의 발명은 바람직한 실시예들에 대한 특별한 참조로서 상기 상세하게 기재되었다. 상기 상세한 설명에 익숙해진 숙련된 수행자는 후술하는 청구항들의 범위로부터 벗어나는 일 없이 어떤 변형을 이룰 수 있다.

표 1

[0182]

사용된 프라이머들의 서열들		
SEQ ID NO.	프라이머 이름	서열 (5'>3')
9	BY60	caggaaacgctgacgctactggctcggtaccagcagagcggcgtaaacctgtgacggaagatcacttcgcag aataa
10	BY61	gtgagtttgaccgctatctgctgagccaaataaaccgaatatggctattgaagcacttcactgacaccctca tca
11	BY62	tctctggcggtggctataaaaaaccgtaacggcctgcatgtacgtccggcgtgacggaagatcacttcgcag aataaat
12	BY63	aacctgtacggttgaaacgttgcgttttaacgtccagcgttagcgtttctgaagcacttcactgacaccctc atca
13	AC 16	gcatttgtctgttatctacaccgatgagg
14	AC 17	tccactcttgcaggaacgct
15	AC33	gactgggagaaggtgtcggtgaat
16	AC34	catcattaaacagcggcctataaaaata
17	91A	aaacccaatgccagcgctggcaactcaggttagcgacaccattggcgtgggtgtgat ctacgcgatgcattcattta
18	91B	aatcaacgcattacaacgctggcgaattaacacctcaatggcgtgacgctctttttccgc ggccaagaactccagcatga
19	90A	gaataacagcacgctgggtca
20	90B	tcaacgcactcatccagcct
21	60A	atgtccagaaggcttcgcagaacaaaaatcgttaccacgttaggccagcaacagatcgc gacggaagatcacttcgcag
22	60B	ttactctaccgttaaaatacgcgtggatttagtagaacccacggctactcatcacgtcgccaagaataaaaag aaaatgcc
23	59A	tacatgtccagaaggcttcg
24	59B	catccggcaacgtacttact
25	75A	ccaggacgatccttgcgctttaccatcagcccgataatcctccaccggcgcccatgctagcgcatgca tccattta
26	75B	ccgcgctatccgcaccttttcgcagggaaaatgtacgacctcacaccagtggaaaccagc ggccaagaactccagcatga
27	76A	ccaggacgatccttgcgctttaccatcagcccgataatcctccaccggcgcccatgccgattacacca accacaac
28	76B	ccgcgcttttcgcaccttttcgcagggaaaatgtacgacctcacaccagtggaaaccagc ggccaatttcatagctttcc
29	60A	atgtccagaaggcttcgcagaacaaaaatcgttaccacgttaggccagcaacagatcgc gacggaagatcacttcgcag
30	60B	ttactctaccgttaaaatacgcgtggatttagtagaacccacggctactcatcacgtcgccaagaataaaaag aaaatgcc
31	67A	ggctgggacggaagtcgctgtcgttctcaaaatcggtggagctgcatgacaaggtcatcggaacggaagatca cttcgcag

32	67B	atggagcagattagcgggttaaccctgctatttgctgataaatctaaaaccggtaagcaaagaataaaag aaaatgcc
36	87A	atgaaaaagaccaaataatgtttgcaccatcgaccgaaaaccgaatctgaagagatgtctagcgcgcgcac cattta
37	87B	ttacaggacgtgaacagatgcgggtgttagtagtgccgctcggtaccagtgcaccagaaaggcgaagaactcc agcatga
38	78A	acacattccttctgcacgctt
39	78B	aggatgcttccatcggttc

표 2

[0183]

프라이머 쌍들 및 사용된 DNA 템플레이트 및 형성된 박테리아 스트레인들		
PCR 프라이머 쌍	DNA 템플레이트	형성된 스트레인
BY60 / BY61	pCA2	AC5
BY62 / BY63	pCA2	AC1
AC16 / AC17	<i>E. Coli C</i>	AC6
AC33 / AC34	<i>E. Coli C</i>	AC2
91A / 91B	pGW162	WG83
90A / 90B	KJ122	WG86
60A / 60B	pL014151	WG84
59A / 59B	<i>E. Coli C</i>	WG85
75A / 75B	pGW162	WG62
76A / 76B	KJ122	WG74
60A / 60B	pL014151	WG96
67A / 67A	pL014151	WG51
87A / 87B	pGW162	WG87
78A / 78B	pGW191	WG89

표 3

[0184]

전체 게놈 서열화를 통해 검출된 돌연변이들의 상세						
관련 유전자(들)	스트레인 이름	부모 스트레인	대립 형질 타입	숙신산염 타이터	개시 성장 속도 OD550/hr	주어진 타이터에 대한 발효 시간(시간)
<i>pykA</i>	KJ122	-	프레임 이동	550 mM	0.21	96
<i>pykA</i>	WG85a	KJ122	야생형	160 mM	0.022	96
<i>pykA</i>	WG84	KJ122	결실	540 mM	0.21	96
<i>pykA, F</i>	KJ122	-	<i>pykA</i> ^{fs} , <i>pykF</i> ⁺	510 mM	0.20	97
<i>pykA, F</i>	WG89	KJ122	<i>pykA</i> ⁺ , <i>DpykF</i>	115 mM	0.021	97
<i>galS</i>	XZ722	-	야생형	175 mM	0.017	144
<i>galS</i>	WG86a	XZ722	프레임 이동	210 mM	0.019	144
<i>galS</i>	WG83	XZ722	결실	220 mM	0.018	144
<i>rpoA</i>	KJ122	-	미스센스	385 mM	0.24	48
<i>rpoA</i>	RY859A1	KJ122	야생형	370 mM	0.23	48
<i>rpoC</i>	KJ122		미스센스	425 mM	0.23	48
<i>rpoC</i>	RY862A	KJ122	야생형	320 mM	0.15	48
<i>rpoA, rpoC</i>	RY860A	RY859A1	배지에 대한 야생형	300 mM	0.14	48
<i>gldA</i>	KJ122	-	프레임 이동	530 mM	0.20	96

<i>gldA</i>	AC6	KJ122	야생형	580 mM	0.16	96
<i>dhaM</i>	KJ122	-	프레임 이동	530 mM	0.20	96
<i>dhaM</i>	AC2	KJ122	야생형	560 mM	0.16	96
<i>ftsI</i>	KJ122	-	미스센스	460 mM	0.20	72
<i>ftsI</i>	RY858G1	KJ122	야생형	440 mM	0.20	72

[0185]

서열 목록

SEQUENCE LISTING

<110> Myriant Technologie LLC

Gong, Wei

Dole, Sudhanshu

Graber, Tammy

Collard, Andrew

Pero, Janice

Yocum, Rogers

<120> Engineering Microbes for Efficient Production of Chemicals

<130> MT2010_04PCT

<150> US 61/281,481

<151> 2009-11-18

<160> 39

<170> PatentIn version 3.5

<210> SEQ ID NO. 1

<211> 1445

<212> DNA

<213> Escherichia coli KJ122

<220> Feature: pykA gene coding region

<223> Other Information: Frame shift mutation, insertion of GC after base 1023

<400> Sequence

```

atgtccagaa ggcttcgcag aacaaaaatc gttaccacgt taggcccagc aacagatcgc      60
gataataatc ttgaaaaagt tatcgcggcg ggtgccaacg ttgtacgtat gaacttttct      120
cacggctcgc ctgaagatca caaaatgcgc gcggataaag ttcgtgagat tgccgcaaaa      180
ctggggcgctc atgtggctat tctgggtgac ctccaggggc caaaatccg tgtatccacc      240
tttaaagaag gcaaagtttt cctcaatatt ggggataaat tctgctcga cgccaacctg      300

```

ggtaaagtg aaggcgacaa agaaaaagtc ggtatcgact acaaaggcct gcctgctgac 360

gtcgtgcctg gtgacatcct gctgctggac gatggtcgcg tccagttaaa agtactggaa 420

gttcagggca tgaagtggt caccgaagtc accgtcgggtg gtccctctc caacaataaa 480

ggtatcaaca aacttggcgg cggtttgtcg gctgaagcgc tgaccgaaaa agacaaagca 540

gacattaaga ctgcggcggtt gattggcgta gattacctgg ctgtctctt cccacgtgt 600

ggcgaagatc tgaactatgc ccgtcgccig gcacgcgatg caggatgtga tgcgaaaatt 660

gttgccaagg ttgaacgtgc ggaagccgtt tgcagccagg atgcaatgga tgacatcatc 720

ctgcctctg acgtggtaat gggtgcacgt ggcgacctcg gtgtggaaat tggcgaccgc 780

gaactggtcg gcattcagaa agcgttgatc cgtcgtgcgc gtcagctaaa ccgagcggta 840

atcacggcga cccagatgat ggagtcaatg attactaacc cgatgccgac gcgtgcagaa 900

gtcatggacg tagcaaagc cgttctggat ggtactgacg ctgtgatgct gtctgcagaa 960

actgccgtg ggcagtatcc gtcagaaacc gttgcagcca tggcgcgct ttcctgggt 1020

gcgcggaaaa aatcccgagc atcaacgttt ctaaacaccg tctggacgtt cagttcgaca 1080

atgtggaaga agctattgcc atgtcagcaa tgtacgcagc taaccacctg aaaggcgta 1140

cggcgatcat caccatgacc gaatcgggtc gtaccgcgt gatgacctcc cgtatcagct 1200

ctggtctgcc aattttcgcc atgtcgcgcc atgaacgtac gctgaacctg actgctctt 1260

atcgtggcgt tacgccgtg cactttgata gcgctaata cggcgtagca gctgccagcg 1320

aagcggttaa tctgctgcgc gataaagggtt acttgatgtc tggtgacctg gtgattgtca 1380

cccagggcga cgtgatgagt accgtgggtt ctactaatac cacgcgtatt ttaacgtag 1440

agtaa 1445

<210> SEQ ID NO. 2

<211> 1042

<212> DNA

<213> Escherichia coli KJ122

<220> Feature: galS coding sequence

<223> Other Information: frameshift mutation, insertion of A after base 246

<400> Sequence

atgatcacca ttcgtgatgt agcgctcag gctggcgtct ctgtggcaac ggtttcccgg 60

gtgctcaata acagcacgtt ggtcagtgcc gacacgcgtg aagcagtaat gaaagccgtg 120

agtgagctgg attatcgccc aaacccaat gccagcgcgc tggcaactca ggttagcgac 180

accattggcg tgggtggtgat ggacgtttct gatgcgtttt tcggcgcgct ggtaaaagcg 240

gtggatacta gtcgctcagc agcatcagaa atacgtgcta atcggcaata gctatcatga	300
agcggaaaaa gagcgtcagc ccattgaggt gttaattcgc cagcgttgta atgcgttgat	360
tgttcactca aaagcattga gtgacgatga actggcgcaa tttatggata acattcccgg	420
tatggtgtta atcaaccgcg ttgtgccggg gtacgcccac cgttgcgttt gcctggataa	480
tctcagcggg gcccgaaatg cgacgcgcgt gttgctgaat aacggtcac aacgtattgg	540
ttatctttct tccagtcagc gcattgaaga tgacgccatg cgtaaagcag gctggatgag	600
tgcggtgaaa gagcaggata ttattccgcc ggaaagctgg attggcactg gtacgccgga	660
catgccgggc ggtgaggcgg cgatgggtga actgctgggg cgcaatctac aacttaccgc	720
tgtatttgct tataacgaca atatggccgc tggcgactg acagcattaa aagataatgg	780
cattgcgatt ccgttacac tctcaatcat cggtttcgat gatattccca tcgccgtta	840
caccgacccg caattaacga ccgtgcgtta tccattgct tcaatggcta aattagccac	900
cgaactggcc ttgcaggggg cagcaggcaa tattgacct cgtgccagcc actgttttat	960
gccgacgtta gtgcgtcgcc attctgtcgc aacgcgccag aatgcggcgg cgatcactaa	1020
ctcaacaaat caggcgatgt aa	1042

<210> SEQ ID NO. 3

<211> 906

<212> DNA

<213> Escherichia coli KJ122

<220> Feature: 48 kilobase deletion

<223> Other Information: Deletion between bases 448 and 449

Deletion removes genes ydcI, ydcL, ydcO, ydcV

and many others

<400> Sequence

cagtcggtaa cctcgcgcac acagccgggc agtgacgtca tcgtctgcgc ggaaatggac	60
gaacagtggg gctacgtcgg tgctaaatca cgtcagcgt ggctgtttta cgcgtatgac	120
aggatacggg ggacagtgtt ggcgcacgtt ttcggtgaac gcactctggc cacactggag	180
cgtcttctga gcctgtgtc gccctttgag gtcgtggtat ggatgacgga tggctggccg	240
ctgtatgaat cagccctgaa gggaaagctg cagttatca gcaagcgtta cactcagcgc	300
attgagcgac ataattctgaa tctgagacaa catctggcaa ggctgggacg gaagtcgctg	360
tcgtttctaa aatcgggtga gctgcatgac aaggtcatcg ggcatatatc gaacataaaa	420
cactatcagt aagttggagt cactaccctg aattcctgaa tttttgtcct taccgtttta	480

gaattaataa aaaacccggc gctaaatgct taccgggttt tagatttatc aggcaaatag 540
cagggttaac cgctaactctg ctccatcgcc tgcaaaatac gcttatctga aatcggataa 600
ggcgtaccaa gttgttgagc gaagtaacta acgcgcagct cttctatcat ccaacggatc 660
tctttcacgt cticacctc acgacgtgca ggcggcagtt tgttgatcca ttgctgccac 720
gcctgctgga cgttttcgac ttccagcatc tgagcacggt cgcgatgtgg atcaaccgcc 780
agtttttcca gtcgttttcc aatcgctgc aaatatcgca gcgtgctgcc cagccgtttg 840

aagccgttac cagtgacaaa accgcgatat accaaccgc ccatctgcgc tttaatgtca 900
gaaagc 906

<210> SEQ ID NO. 4

<211> 990

<212> DNA

<213> Escherichia coli KJ122

<220> Feature: rpoA gene coding sequence

<223> Other Information: missense mutation C965T; P322L mutation

<400> Sequence

atgcagggtt ctgtgacaga gtttctaaaa ccgcgcctgg ttgatatcga gcaagtgagt 60
tcgacgcacg ccaaggtgac ccttgagcct ttagagcgtg gctttggcca tactctgggt 120

aacgcactgc gccgtattct gctctcatcg atgccgggtt gcgcggtgac cgaggttgag 180
attgatgggt tactacatga gtacagcacc aaagaaggcg ttcaggaaga tatcctggaa 240
atcctgctca acctgaaagg gctggcgggtg agagttcagg gcaaagatga agttattctt 300
accttgaata aatctggcat tggccctgtg actgcagccg atatcaccca cgacggtgat 360
gtcgaatcgc tcaagccgca gcacgtgatc tgccacctga ccgatgagaa cgcgtctatt 420
agcatgcgta tcaaagtcca gcgcggctgt ggttatgtgc cggttctac ccgaattcat 480
tcggaagaag atgagcgccc aatcggccgt ctgctgggtcg acgcatgcta cagccctgtg 540

gagcgtattg cctacaatgt tgaagcagcg cgtgtagaac agcgtaccga cctggacaag 600
ctggtcatcg aaatggaac caacggcaca atcgatcctg aagaggcgat tcgtcgtgcg 660
gcaaccattc tggtgaaca actggaagct ttcgttgact tacgtgatgt acgtcagcct 720
gaagtgaaag aagagaaacc agagttcgat ccgatcctgc tgcgccctgt tgacgatctg 780
gaattgactg tccgctctgc taatgcctt aaagcagaag ctatccacta tatcggtgat 840
ctggtacagc gtaccgaggt tgagctcctt aaaacgccta accttggtaa aaaatctctt 900
actgagatta aagacgtgct ggcttcccgt ggactgtctc tgggcatgcg cctggaaaac 960

tggctaccgg caagcatcgc tgacgagtaa 990
 <210> SEQ ID NO.5
 <211> 4224
 <212> DNA
 <213> Escherichia coli KJ122
 <220> Feature: rpoC coding sequence
 <223> Other Information: missense mutation G2441A; M747I mutation
 <400> Sequence
 gtgaaagatt tattaaagtt tctgaaagcg cagactaaaa ccgaagagtt tgatgcgac 60
 aaaattgctc tggttcgcc agacatgac cgttcattgt ctttcggtga agttaaaaag 120
 ccggaaccca tcaactaccg tacgttcaaa ccagaacgtg acggcctttt ctgcgccgt 180
 atctttgggc cggtaaaaaga ttacgagtgc ctgtgcggtg agtacaagcg cctgaaacac 240
 cgtggcgta tctgtgagaa gtgcggcggt gaagtgaccc agactaaagt acgccgtgag 300
 cgtatgggcc acatcgaact ggcttccccg actgcgcaca tctggttcct gaaatcgctg 360
 ccgtcccgtg tgggtctgct gctcgatatg ccgtgcgcg atatcgaacg cgtactgtac 420
 ttgtaacct atgtggttat cgaaggcggg atgaccaacc tggaacgtca gcagatcctg 480
 actgaagagc agtatctgga cgcgttgga gagttcggg acgaattcga cgcgaagatg 540
 ggggcggaag caatccaggc tctgtgaag agcatggatc tggagcaaga gtgcgaacag 600
 ctgcgtgaag agctgaacga aaccaactcc gaaaccaagc gtaaaaagct gaccaagcgt 660
 atcaaaactgc tggaagcgtt cgttcagtct ggttaacaaac cagagtggat gatcctgacc 720
 gttctgccgg tactgccgcc agatctgcgt ccgtcggttc cgtggatgg tggtcgtttc 780
 gcgacttctg acctgaacga tctgtatcgt cgcgtcatta accgtaacaa ccgtctgaaa 840
 cgtctgctgg atctggctgc gccggacatc atcgtacgta acgaaaaacg tatgctgcag 900
 gaagcggtag acgccctgct ggataacggt cgtcgcggtc gtgcgacac cgtttctaac 960
 aagcgtctc tgaatcttt ggccgacatg atcaaaggta aacagggtcg ttccgtcag 1020
 aacctgctcg gtaagcgtgt tgactactcc ggtcgttctg taatcacctg aggtccatac 1080
 ctgcgtctgc atcagtgcgg tctgccgaag aaaatggcac tggagctgtt caaacgttc 1140
 atctacggca agctggaact gcgtggtctt gctaccacca ttaaagctgc gaagaaaatg 1200
 gttgagcgcg aagaagctgt cgtttgggat atcctggacg aagtatccg cgaacacccg 1260
 gtactgctga accgtgcacc gactctgcac cgtctgggta tccaggcatt tgaaccggtg 1320
 ctgatcgaag gtaaagctat ccagctgcac ccgtgggtt gtgcggcata taacgccgac 1380
 ttcgatggtg accagatggc tgttcacgta ccgtgacgc tggaaagcca gctggaagcg 1440

cgtgcgctga tgatgtctac caacaacatc ctgtccccgg cgaacggcga accaatcatc 1500
 gttccgtctc aggacgttgt actgggtctg tactacatga cccgtgactg tgttaacgcc 1560
 aaaggcgaag gcatggtgct gactggcccc aaagaagcag aacgtctgta tcgctctggt 1620
 ctggcttctc tgcatgcgcg cgttaaagtg cgtatcaccg agtatgaaaa agatgctaac 1680
 ggtgaattag tagcgaatac cagcctgaaa gacacgactg ttggccgtgc cattctgtgg 1740
 atgattgtac cgaagggtct gccttactcc atcgtcaacc aggcgctggg taaaaaagca 1800
 atctccaaaa tgctgaacac ctgctaccgc attctcggtc tgaaaccgac cgttatTTTT 1860

 gcggaccaga tcatgtacac cggcttcgcc tatgcagcgc gttctggtgc atctgttgg 1920
 atcgaatgaca tggcatccc ggagaagaaa caggaatca tctccgagc agaagcagaa 1980
 gttgctgaaa ttcaggagca gttccagtct ggtctggtaa ctgcgggcga acgctacaac 2040
 aaagtattcg atatctgggc tgcggcgaac gatcgtgtat ccaaagcgat gatggataac 2100
 ctgcaaaactg aaaccgtgat taaccgtgac ggtcaggaag agaagcaggt ttccttcaac 2160
 agcatctaca tgatggccga ctccggtgcg cgtggttctg cggcacagat tcgtcagctt 2220
 gctggtatgc gtggtctgat agcgaagccg gatggctcca tcatcgaaac gccaatcacc 2280

 gcgaacttcc gtgaaggctt gaacgtactc cagtacttca tctccacca cgggtgctgt 2340
 aaaggctcgg cggataccgc actgaaaact gcgaactccg gttacctgac tcgtcgtctg 2400
 gttgacgtgg cgcaggacct ggtggttacc gaagacgatt gtggtacca tgaaggtatc 2460
 atgatgactc cggttatcga ggggtgtgac gttaaagagc cgctgcgcga tcgctactg 2520
 ggtcgtgtaa ctgctgaaga cgttctgaag ccgggtactg ctgatatcct cgttccgcgc 2580
 aacacgctgc tgcacgaaca gtggtgtgac ctgctggaag agaactctgt cgacgcggtt 2640
 aaagtacgtt ctgttgtatc ttgtgacacc gactttggtg tatgtgcgca ctgctacggt 2700

 ctgacactgg cgcgtggcca catcatcaac aagggtgaag caatcggtgt tatcgcggca 2760
 cagtccatcg gtgaaccggg tacacagctg accatgcgta cgttccacat cggtggtgcg 2820
 gcatctcgtg cggtgctga atccagcatc caagtgaaaa acaaaggtag catcaagctc 2880
 agcaacgtga agtcggttgt gaactccagc ggtaaacgtg ttatcacttc ccgtaatact 2940
 gaactgaaac tgatcgacga attcggctgt actaaagaaa gctacaaagt accttacggt 3000
 gcgggtactgg cgaaggcgga tggcgaacag gttgctggcg gcgaaaccgt tgcaaacgtg 3060
 gaccgcaca ccatgccggt tatcaccgaa gtaagcgggt ttgtacgctt tactgacatg 3120

 atcgacggcc agaccattac gcgtcagacc gacgaactga ccggtctgtc ttcgctggtg 3180
 gttctggatt ccgcagaacg taccgcaggt ggtaagatc tgcgtccggc actgaaaatc 3240

gttgatgctc agggtaacga cgttctgata ccaggtaccg atatgccagc gcagtacttc 3300
ctgccgggta aagcgattgt tcagctggaa gatggcgtac agatcagctc tggtagacacc 3360
ctggcgcgta ttccgcagga atccggcggt accaaggaca tcaccggtgg tctgccgcgc 3420
gttgcggacc tggtcgaagc acgtcgtccg aaagagccgg caatcctggc tgaaatcagc 3480
ggtatcgttt ctttcggtaa agaaacaaaa ggtaaacgtc gtctggttat caccgccgta 3540

gacggtagcg atccgtacga agagatgatt ccgaaatggc gtcagctcaa cgtgttcgaa 3600
ggtgaacgtg tagaacgtgg tgacgtaatt tccgacggtc cggaagcgcc gcacgacatt 3660
ctgcgtctgc gtggtgttca tgctgttact cgttacatcg ttaacgaagt acaggacgta 3720
taccgtctgc agggcggtta gattaacgat aacacatcg aagttatcgt tcgtcagatg 3780
ctgcgtaaag ctaccatcgt taacgcgggt agctccgact tcctggaagg cgaacaggtt 3840
gaatactctc gcgtcaagat cgcaaaccgc gaactggaag cgaacggcaa agtgggtgca 3900
acttactccc gcgatctgct gggatcacc aaagcgtctc tggcaaccga gtccttcac 3960

tccgcggcat cgttccagga gaccactcgc gtgctgaccg aagcagccgt tgcgggcaaa 4020
cgcgacgaac tgcgcggcct gaaagagaac gttatcgtgg gtcgtctgat cccggcaggt 4080
accggttacg cgtaccacca ggatcgtatg cgtgcctgtg ctgcgggtga agctccggct 4140
gcaccgcagg tgactgcaga agacgcatct gccagcctgg cagaactgct gaacgcaggt 4200
ctgggcgggt ctgataacga gtaa 4224

<210> SEQ ID NO. 6

<211> 1105

<212> DNA

<213> Escherichia coli KJ122

<220> gldA coding sequence

<223> Other Information: frameshift mutation; insertion of C after base 717

<400> Sequence

atggaccgca ttattcaatc accgggtaaa tacatccagg gcctgatgt gattaatcgt 60
ctgggcgaat acctgaagcc gctggcagaa cgctggtag tggtaggtga caaatttgtt 120
ttaggttttg ctcaatccac tgcgagaaa agctttaag atgctggact gtagtagaa 180
attgcgccgt ttggcggtag atgttcgcaa aatgagatcg accgtctgcg tggcatcgcg 240
gagactgcgc agtgtggcgc aattctcggg atcggtagcg gaaaaacct cgatactgcc 300
aaagcactgg cacatttcat ggggtgtccg gtagcgatcg caccgactat cgctctacc 360
gatgcaccgt gcagcgcatg gtctgttatc tacaccgatg agggtagatt tgaccgctat 420

ctgctgttgc caaataaccc gaatatggtc attgtcgaca ccaaaatcgt cgctggcgca 480
 cctgcacgtc tgttagcggc gggatatcggc gatgcgctgg caacctggtt tgaagcgcgt 540
 gcctgtcttc gtagcggcgc gaccacatg gcgggcggca agtgcacca ggctgcgctg 600
 gcactggctg aactgtgcta caacacccitg ctggaagaag gcgaaaaagc gatgcttgct 660
 gccgaacagc atgtagtac tccggcgtg gagcgcgtga ttgaagcgaa cacctatctt 720
 aagcgggtga ggctttgaaa gtgggtggtct ggctgcggcg cacgcagtgc ataacggcct 780
 gaccgctatc ccggacgcgc atcactatta tcacggtgaa aaagtggcat ttggtacgct 840

gacgcagctg gttctggaag acgcaccggt tgaggaaatc gaaaccgtag ctgcacttag 900
 ccatgcggta gggttgccaa taactctcgc tcaactggat attaaagaag atgtcccggc 960
 gaaaaatgca attgtggcag aagcggcatg tgcagaaggt gaaaccatcc acaacatgcc 1020
 tggcggcgcg acgccagatc aggtttacgc cgctctgctg gtagccgacc agtacggtca 1080
 gcgtttcctg caagagtggg aataa 1105

<210> SEQ ID NO. 7

<211> 1417

<212> DNA

<213> Escherichia coli KJ122

<220> Feature: dhaM coding sequence

<223> Other Information: frameshift mutation; deletion of TS between bases

1256 and 1266

<400> Sequence

atggtaaacc tggatcatagt ttacatagc agccgactgg gagaaggtgt cggatgaatta 60
 gcccgtcaga tgttaatgag tgatagttgt aaaatcgcca ttgccgcggg aattgacgat 120
 ccacaaaatc ccattggtac cgatgccgtc aaagtgatgg aggccatcga atctgttgct 180
 gatgccgacc atgtgctggt catgatggat atgggtagcg cattattgag tgctgaaact 240
 gcgctggaat tgctggctcc cgagatcgcc gcaaaagtac gtttgtgtgc tgcgccgttg 300
 gtcgaaggta cactggcagc aacggtcagc gcggcctcgg gggcggtat cgacaaagtt 360
 atctttgacg ccatgcatgc gctggaagcc aaactgaac aactgggttt accgtcctcc 420

gacactgaaa ttcttgacac atgtcctgcg tacgatgaag aagcccgttc tctggcggtg 480
 gtcataaaaa accgtaacgg cctgcatgta cgtccgcct cccggctggt ttatacctta 540
 tcgacattta atgccgatat gttgctggaa aaaaacggca aatgcgtcac accagagagt 600
 attaaccaga ttgcgttact acaagttcgc tataacgata cgctgcgcct gattgcgaaa 660
 gggccagaag ctgaagaggc actgatcgct ttccgtcagc tggctgaaga taactttggt 720

gaaacggagg aagtcgctcc acctactctg cgtcccgttc cgcctgtttc gggtaaagcc 780
 ttttattatc aaccagtttt atgtacggta caggcaaaat caaccctgac cgtggaagaa 840

gaacaagatc gattacgcca ggctattgac ttcacgttat tagatctgat gacgttaaca 900
 gcgaaagcag aagccagcgg gcttgacgat attgccgcaa tcttttctgg tcaccataca 960
 ctgttagatg atccggaact gctggcggcg gcaagcgaac tccttcagca tgaacattgc 1020
 acggcagaat atgcctggca gcaagttctt aaagaactta gccagcaata ccagcaactg 1080
 gatgatgaat atctacaagc tcgctatatt gatgtggacg atcttctgca tcgcaccctg 1140
 gtccacctga cccaaacgaa agaagaactc ccgcagtta actcgccaac tattctactg 1200
 gcggagaaca ttatccttc cacagtactg caactggatc cggcggttgt aaaaggtatc 1260

tgctcgccc ggaagtccgg tateccacag cgcctaata gcccgtaac tggggattgg 1320
 ctggatttgc cagcagggtg agaaactgta tgcgatacaa ccagaagaaa cgtaacgct 1380
 ggacgttaaa acgcaacgtt tcaaccgtca gggttaa 1417

<210> SEQ ID NO. 8
 <211> 1767
 <212> DNA
 <213> Escherichia coli KJ122
 <220> Feature: ftsI coding region
 <223> Other Information: missense mutation C619T; R206C
 <400> Sequence

atgaaagcag cggcgaaaac gcagaaacca aaacgtcagg aagaacatgc caactttatc 60
 agttggcggtt ttgcgttggt atgcggctgt attctcctgg cgttggtttt tctgctcgga 120

cgcgtagcgt ggttacaagt tatctccccg gatatgctgg tgaaagaggg cgacatgcgt 180
 tctcttcgcg ttcagcaagt ttccacctcc cgcggcatga ttactgaccg ttctggtcgc 240
 ccgttagcgg tgagcgtgcc ggtaaaagcg atttgggctg acccgaaaga agtgcattgac 300
 gctggcggta tcagcgtcgg tgaccgctgg aaggcgtgg ctaacgcgct caatattccg 360
 ctggatcagc ttcagccccg cattaacgcc aaccgaaag ggcgctttat ttatctggcg 420
 cgtcaggtga accctgacat ggccgactac atcaaaaaac tgaaactgcc ggggattcat 480
 ctgcgtgaag agtctcgccg ttactatccg tccggcgaag tgactgtca cctcatcggc 540

tttactaacg tcgatagica agggattgag ggcgttgaga agagtttcga taaatggctt 600
 accgggcagc cgggtgagtg cattgtgcgt aaagaccgct atggtcgcgt aattgaagat 660
 atttcttcta ctgacagcca ggcagcgcac aacctggcgc tgagtattga tgaacgcctg 720

caggcgctgg tttatcgca actgaacaac gcggtggcct ttaacaaggc tgaatctggt	780
agcgccgtgc tgggtggaigt caacaccggt gaagtgtggt cgtggctaa cagcccgtca	840
tacaacccta acaatctgag cggcacgccg aaagaggcga tgcgtaacg taccatcacc	900
gacgtgtttg aaccgggctc aacggttaaa ccgatgggtg taatgaccgc gttgcaacgt	960
ggcgtggtgc gggaaaactc ggtactcaat accattcctt atcgaattaa cgcccacgaa	1020
atcaaagacg tggcacgcta cagcgaattia accctgaccg gggtattaca gaagtcgagt	1080
aacgtcggtg tttccaagct ggcgttagcg atgccgtcct cagcgttagt agatacttac	1140
tcacgttttg gactgggaaa agcgaccaat ttgggggttg tgggagaacg cagtggctta	1200
tatcctcaaa aacaacggtg gtctgacata gagaggcca ccttctcttt cggtacggg	1260
ctaatggtaa caccattaca gttagcgca gtctacgaa ctatcggcag ctacggcatt	1320
tatcgccac tgcgattac caaagttgac cccccggtt cgggtgaacg tgtcttccg	1380
gaatccattg tccgcactgt ggtgcatatg atggaaagcg tggcgctacc aggcggcggc	1440
ggcgtgaagg cggcgattaa aggctatcgt atcgccatta aaaccggtac cgcgaaaaag	1500
gtcgggccgg acggtcgcta catcaataaa tatattgctt ataccgcagg cgttgcgcct	1560
gcgagtcagc cgcgcttcgc gctggttgtt gttatcaacg atccgcaggc gggtaaatac	1620
tacggcggcg ccgtttccgc gccggtcttt ggtgccatca tgggcggcgt attgcgtacc	1680
atgaacatcg agccgatgc gctgacaacg ggcgataaaa atgaatttgt gattaatcaa	1740
ggcgagggga caggtggcag atcgtaa	1767

<210> SEQ ID NO. 9

<211> 77

<212> DNA

<213> Synthetic DNA

<220> Feature: PCR Primer BY60

<400> Sequence

caggaacgc tgaccgtact ggtcggctac cagcagagcg gcgtaaacct gtgacggaag	60
atcacttcgc agaataa	77

<210> SEQ ID NO. 10

<211> 76

<212> DNA

<213> Synthetic DNA

<220> Feature: PCR Primer BY61

<400> Sequence

gtgagtttga ccgtatctg ctgttgccaa ataaccgaa tatggtcatt gaagcattc 60
actgacaccc tcatca 76

<210> SEQ ID NO. 11

<211> 79

<212> DNA

<213> Synthetic DNA

<220> Feature: PCR Priemr BY62

<400> Sequence

tctctggcgg tggtcataaa aaaccgtaac ggctgcatg tacgtccggc gtgacggaag 60
atcacttcgc agaataaat 79

<210> SEQ ID NO. 12

<211> 76

<212> DNA

<213> Synthetic DNA

<220> Feature: PCR Primer BY63

<400> Sequence

aaccctgacg gttgaaacgt tgcgttttaa cgtccagcgt tagcgtttct gaagcattc 60
actgacaccc tcatca 76

<210> SEQ ID NO. 13

<211> 28

<212> DNA

<213> Synthetic DNA

<220> Feature: PCR Primer AC16

<400> Sequence

gcattgtctg ttatctacac cgatgagg 28

<210> SEQ I D NO. 14

<211> 22

<212> DNA

<213> Synthetic DNA

<220> Feature: PCR Primer AC17

<400> Sequence

tcccactctt gcaggaaacg ct 22

<210> SEQ ID NO. 15

<211> 24

<212> DNA

<213> Synthetic DNA

<220> Feature: PCR Primer AC33

<400> Sequence

gactgggaga aggtgtcggg gaat 24

<210> SE ID NO. 16

<211> 29

<212> DNA

<213> Synthetic DNA

<220> Feature: PCR Primer AC34

<400> Sequence

catcattaaa cagcggccct aataaaata 29

<210> SEQ ID NO. 17

<211> 80

<212> DNA

<213> Synthetic DNA

<220> Feature: PCR Primer 91A

<400> Sequence

aaacgccaat gccagcgcg tggcaactca ggtagcgac accattggcg tgggtggtgat 60

ctagcgcatg catccattta 80

<210> SEQ ID NO. 18

<211> 80

<212> DNA

<213> Synthetic DNA

<220> Feature: PCR Primer 91B

<400> Sequence

aatcaacgca ttacaacgct ggcaaatata cacctcaatg gcgtgacgct ctttttcgc 60

ggcgaagaac tccagcatga 80

<210> SEQ ID NO. 19

<211> 20

<212> DNA

<213> Synthetic DNA

<220> Feature: PCR Primer 90A

<400> Sequence

gaataacagc acgttggtca 20

<210> SEQ ID NO.20

<211> 20

<212

> DNA

<213> Synthetic DNA

<220> Feature: PCR Primer 90B

<400> Sequence

tcaacgcact catccagcct 20

<210> SEQ ID NO. 21

<211> 80

<212> DNA

<213> Synthetic DNA

<220> Feature: PCR Primer 60A

<400> Sequence

atgtccagaa ggcttcgcag aacaaaaatc gttaccacgt taggcccagc aacagatcgc 60

gacggaagat cacttcgcag 80

<210> SEQ ID NO. 22

<211> 80

<212> DNA

<213> Synthetic DNA

<220> Feature: PCR Primer 60B

<400> Sequence

ttactctacc gttaaaatac gcgtgggtatt agtagaaccc acggtactca tcacgtcgcc 60

aagaaataaa agaaaatgcc 80

<210> SEQ ID NO. 23

<211> 20

<212> DNA

<213> Synthetic DNA

<220> Feature: PCR Primer 59A

<400> Sequence

tacatgtcca gaaggcttcg 20

<210> SEQ ID NO.24

<211> 20

<212> DNA

<213> Synthetic DNA
 <220> Feature: PCR Primer 59B
 <400> Sequence
 catccggcaa cgtacttact 20

<210> SEQ I D NO.25
 <211> 80
 <212> DNA
 <213> Synthetic DNA
 <220> Feature: PCR Primer 75A
 <400> Sequence
 ccaggacgat ccttgcgctt tacccatcag cccgtataat cctccaccg gcgcgccatg 60
 ctagcgcgatg catccattta 80

<210> SEQ ID NO. 26
 <211> 81
 <212> DNA
 <213> Synthetic DNA
 <220> Feature: PCR Primer 75B
 <400> Sequence
 ccgcgctttt ccgcaccttt tcgcaggga aatgtacgac ctacaccag tggaaaccag 60
 cggcgaagaa ctccagcatg a 81

<210> SEQ ID NO,27
 <211> 80
 <212> DNA
 <213> Synthetic DNA
 <220> Feature: PCR Primer 76A
 <400> Sequence
 ccaggacgat ccttgcgctt tacccatcag cccgtataat cctccaccg gcgcgccatg 60
 ccgattacac caaccacaac 80

<210> SEQ ID NO. 28
 <211> 81
 <212> DNA
 <213> Synthetic DNA
 <220> Feature: PCR Primer 76B
 <400> Sequence

ccgcgctttt ccgcaccttt tcgcagggaa aatgtacgac ctacacaccag tggaaaccag 60
cggcgaattt catagctttc c 81

<210> SEQ ID No. 29

<211> 80

<212> DNA

<213> Synthetic DNA

<220> Feature: PCR Primer 60A

<400> Sequence

atgtccagaa ggcttcgcag aacaaaaatc gttaccacgt taggcccagc aacagatcgc 60
gacggaagat cacttcgcag 80

<210> SEQ ID NO.30

<211> 80

<212> DNA

<213> Synthetic DNA

<220> Feature: PCR Primer 60B

<400> Sequence

ttactctacc gttaaaatc gcgttggtatt agtagaacc acggtactca tcacgtcgcc 60
aagaaataaa agaaaatgcc 80

<210> SEQ ID No. 31

<211> 80

<212> DNA

<213> Synthetic DNA

<220> Feature: PCR Primer 67A

<400> Sequence

ggctgggacg gaagtcgctg tcgtttctcaa aatcggtgga gctgcatgac aaggtcatcg 60
gacggaagat cacttcgcag 80

<210> SEQ ID NO. 32

<211> 80

<212> DNA

<213> Synthetic DNA

<220> Feature: PCR Primer 67B

<400> Sequence

atggagcaga ttacgggtta accctgctat ttgcctgata aatctaaaac ccggttaagca 60
aagaaataaa agaaaatgcc 80

<210> SEQ ID NO. 33

<211> 7929

<212> DNA

<213> Escherichia coli

<220> FEATURE: cat sacB, amp; PLasmid template for PCR of cat-sacB cassette

<221> misc_feature

<222> (1545)..(1545)

<223> n is a, c, g, or t

<220><221> misc_feature

<222> (3474)..(3475)

<223> n is a, c, g, or t

<400> Sequence

tcgcgcgttt cggatgatgac ggtgaaaacc tctgacacat gcagctcccg gagacggtea	60
cagcttgtct gtaagcggat gccgggagca gacaagcccg tcagggcgcg tcagcgggtg	120
ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc	180
accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcaggcgcc	240
attcgccatt caggctgcgc aactgttggg aagggcgatc ggtgcgggcc tcttcgctat	300
tacgccagct ggcgaaaggg ggatgtgctg caaggcgatt aagtgggta acgccagggt	360
tttcccagtc acgacgttgt aaaacgacgg ccagtgccaa gcttgcagtc ctgcaggtcg	420
actctagagg atccccgggt accgagctcg aattcccgcg cccgatgaat tgatccgaag	480
ttcctattct ctgaaaagta taggaacttc gaattgtcga caagctagca tgtgacggaa	540
gataccttcg cagaataaat aaatcctggt gtcctgttg ataccgggaa gccctgggcc	600
aacttttggc gaaaatgaga cgttgatcgg cacgtaagag gttccaactt tcaccataat	660
gaaataagat cactaccggg cgtatttttt gagttatcga gattttcagg agctaaggaa	720
gctaaaaatg agaaaaaat cactggatat accaccgttg atatatccca atggcatcgt	780
aaagaacatt ttgaggcatt tcagtcagtt gctcaatgta cctataacca gaccgttcag	840
ctggatatta cggccttttt aaagaccgta aagaaaaata agcacaagtt ttatccggcc	900
tttattcaca ttcttgcccg cctgatgaat gtcacatccg aattccgtat ggcaatgaaa	960
gacggtgagc tggtgatatg ggatagtgtt cacccttgtt acaccgtttt ccatgagcaa	1020
actgaaacgt tttcatcgct ctggagtga taccacgacg atttccggca gtttctacac	1080
atatattcgc aagatgtggc gtgttacggt gaaaacctgg cctatttccc taaagggttt	1140

attgagaata tgtttttcgt ctacagccaat ccttgggtga gtttcaccag ttttgattta	1200
aacgtggcca atatggacaa cttcttcgcc cccgttttca ccatgggcaa atattatacg	1260
caaggcgaca aggtgctgat gccgctggcg attcaggttc atcatgccgt ttgtgatggc	1320
ttccatgtcg gcagaatgct taatgaattt caacagtact gcgatgagtg gcagggcggg	1380
gcgtaatttt ttttaaggcag ttatttgggtgc ccttaaagcg ctggtgctac gcctgaataa	1440
gtgataataa gcggatgaat ggcagaaatt cgaaagcaaa ttcgacccgg tcgtcggttc	1500
agggcagggt cgtaaataag ccgcttatgt ctattgctgg ttantcgggt acccggggat	1560
cgcgcccgcg gaccgatcc catcacatat acctgccgtt cactattatt tagtgaaatg	1620
agatattatg atattttctg aattgtgatt aaaaaggcaa ctttatgccc atgcaacaga	1680
aactataaaa aatacagaga atgaaaagaa acagatagat ttttagttc tttaggcccg	1740
tagtctgcaa atccttttat gattttctat caaacaagag aggaaatag accagttgca	1800
atccaaacga gagtctaata gaatgaggtc gaaaagtaaa tcgcgcgggt ttgttactga	1860
taaagcaggc aagacctaaa atgtgtaaag ggcaaagtgt atactttggc gtcacccctt	1920
acatatatta ggtctttttt tattgtgcgt aactaacttg ccatcttcaa acaggagggc	1980
tggaagaagc agaccgctaa cacagtacat aaaaaaggag acatgaacga tgaacatcaa	2040
aaagtttgca aaacaagcaa cagtattaac ctttactacc gcactgctgg caggaggcgc	2100
aactcaagcg ttgcgaaag aaacgaacca aaagccatat aaggaaacat acggcatttc	2160
ccatattaca cgccatgata tgctgcaaat cctgaacag caaaaaaatg aaaaatatca	2220
agttcctgaa ttcgattcgt ccacaattaa aaatatctct tctgcaaaag gcctggacgt	2280
ttgggacagc tggccattac aaaacgctga cggcactgtc gcaaaactatc acggctacca	2340
catcgtcttt gcattagccg gagatcctaa aaatgcggat gacacatcga ttacatgtt	2400
ctatcaaaaa gtccgcgaaa cttctattga cagctggaaa aacgtggcc gcgtctttaa	2460
agacagcgac aaattcgatg caaatgattc tatcctaaaa gaccaaacac aagaatggtc	2520
aggttcagcc acatttacat ctgacggaaa aatccgttta ttctacactg atttctccgg	2580
taaacattac ggcaaacaaa cactgacaac tgcacaagtt aacgtatcag catcagacag	2640
ctctttgaac atcaacggtg tagaggattt taaatcaatc ttgacgggtg acggaaaaac	2700
gtatcaaaat gtacagcagt tcatcgatga aggcaactac agctcaggcg acaaccatac	2760
gctgagagat cctcactacg tagaagataa aggccacaaa tacttagtat ttgaagcaaa	2820
cactggaact gaagatggct accaaggcga agaattctta ttttaacaaag catactatgg	2880
caaaagcaca tcattcttcc gtcaagaaag tcaaaaactt ctgcaaagcg ataaaaaacg	2940
cacggctgag ttagcaaacg gcgctctcgg tatgattgag ctaaacgatg attacacact	3000

gaaaaaagtg atgaaaccgc tgattgcac taacacagta acagatgaaa ttgaacgcgc	3060
gaacgtcttt aaaaatgaac gcaaatggta cctgttcact gactcccgcg gatcaaaaat	3120
gacgattgac ggcattacgt ctaacgatat ttacatgctt ggttatgttt ctaattcttt	3180
aactggccca tacaagccgc tgaacaaaac tggccttggt ttaaaaatgg atcttgatcc	3240
taacgatgta acctttactt actcacactt cgtgtacct caagcgaaag gaaacaatgt	3300
cgtgattaca agctatatga caaacagagg attctacgca gacaaacaat caacgtttgc	3360
gccaaagctt ctgctgaaca tcaaaggcaa gaaaacatct gttgtcaaag acagcatcct	3420
tgaacaagga caattaacag ttaacaaata aaaacgcaa agaaaatgcc gatnccggt	3480
ttattgacta ccggaagcag tgtgaccgtg tgcttctcaa atgcctcagg ctgtctatgt	3540
gtgactgttg agctgtaaca agttgtctca ggtgttcaat ttcatgttct agttgctttg	3600
ttttactggt ttaccctgtt ctattagggt ttacatgctg ttcatctgtt acattgtcga	3660
tctgttcacg gtgaacagct ttaaatgcac caaaaactcg taaaagctct gatgtatcta	3720
tcttttttac accgttttca tctgtgcata tggacagttt tccctttgat gctagcttgc	3780
atgcctgcag gtcgactcta gaggatcccc gtactatcaa cagggtgaac tgcggatctt	3840
gcgccagct ttaagctgtt aaaccgtttt gtgaaaaaat ttttaaaata aaaaagggga	3900
cctctagggt cccaattaa ttagtaatat aatctattaa aggtcattca aaaggcatc	3960
caccggatca attccccgc tcgcgcaggc tgggtgcca gctctcgggt aacatcaagg	4020
cccgatctt ggagcccttg cctccccga cgatgatcgt gccgtgatcg aaatccagat	4080
ccttgacccg cagttgcaa cctcactga tccggctcac ggtaactgat gccgtatttg	4140
cagtaccagc gtacggccca cagaatgatg tcacgctgaa aatgccggcc tttgaatggg	4200
ttcatgtgca gctccatcag caaaagggga tgataagttt atcaccaccg actatttgca	4260
acagtccgt tgatcgtgct atgatcgact gatgtcatca gcggtggagt gcaatgtcgt	4320
gcaatcgaa tggcgaaaag ccgagctcat cggtcagctt ctcaaccttg gggttacccc	4380
cggcggtgtg ctgctgggcc acagctcctt ccgtagcgtc cggccctcg aagatgggcc	4440
acttggaactg atcgaggccc tgcgtgctgc gctgggtccg ggaggggacgc tcgtcatgcc	4500
ctcgtggtca ggtctggacg acgagccgtt cgatcctgcc acgtcgccc ttacaccgga	4560
ccttgagatt gtctctgaca cattctggcg cctgccaaat gtaaagcgca gcgccatcc	4620
atttgcttt gcggcagcgg ggccacaggc agagcagatc atctctgatc cattgccctt	4680
gccacctcac tcgctgcaa gcccggtcgc ccgtgtccat gaactcgatg ggcaggtact	4740
tctctcggc gtgggacacg atgccaacac gacgtgcat cttgccgagt tgatggcaaa	4800

ggttccctat ggggtgccga gacactgcac cattcttcag gatggcaagt tggtagcgt	4860
cgattatctc gagaatgacc actgctgtga gcgctttgcc ttggcggaca ggtggctcaa	4920
ggagaagagc cttcagaagg aaggfccagt cggatcatgcc tttgctcggt tgatccgctc	4980
ccgcgacatt gtggcgacag ccctgggtca actgggccga gatccgttga tcttctgca	5040
tccgccagag ggcgggatgc gaagaatgcg atgccgctcg ccagtcgatt ggctgagctc	5100
atgagcggag aacgagatga cgttggagg gcaaggctgc gctgattgct ggggcaacac	5160
gtgaaaggcg agatcaccaa ggtagtcggc aaataatgtc taacaattcg ttcaagccga	5220
cgccgcttcg cgccgaggct taactcaagc gttagatgca ctaagcacat aattgctcac	5280
agccaaacta tcaggccaag tctgctttta ttatttttaa gcgtgcataa taagccctac	5340
acaaattggg agatataatc tgaaaggctg gctttttctt gttatcgcaa tagttggcga	5400
agtaatcgca acatccgcat taaaatctag cgagggcttt actaagctga tccggtggat	5460
gaccttttga atgaccttta atagattata ttactaatta attggggacc cttagaggtcc	5520
ccttttttat tttaaaaatt ttttcacaaa acggtttaca agcataaagc tcgatgaatt	5580
gatccgaagt tcctattctc tagaaagtat aggaacttcg aattgtcgac aagctccccg	5640
gggagcttga tctggcttat cgaaattaat acgactcact ataggagagc cggaattcgt	5700
aatcatggtc atagctgttt cctgtgtgaa attgttatcc gtcacaatt ccacacaaca	5760
tacgagccgg aagcataaag tgtaaagcct ggggtgccta atgagtgagc taactcacat	5820
taattgcgtt gcgtcactg cccgctttcc agtcgggaaa cctgtcgtgc cagctgcatt	5880
aatgaatcgg ccaacgcgcg gggagaggcg gtttgcgtat tgggcgctct tccgcttct	5940
cgctcactga ctgcgtgcgc tcggtcgttc ggctgcggcg agcggatatca gctcactcaa	6000
aggcggtaat acggttatcc acagaatcag gggataacgc aggaaagaac atgtgagcaa	6060
aaggccagca aaaggccagg aaccgtaaaa aggccgcgtt gctggcgttt ttccataggc	6120
tccgcccccc tgacgagcat cacaaaaatc gacgtcaag tcagagggtg cgaaacccga	6180
caggactata aagataccag gcgtttcccc ctggaagctc cctcgtgcgc tctcctgttc	6240
cgacctgcc gcttaccgga tacctgtccg cttttctccc ttcggaagc gtggcgcttt	6300
ctcaaagctc acgctgtagg tatctcagtt cgggtgtaggc cgctcgtcc aagctgggct	6360
gtgtgcacga accccccgtt cagcccgacc gctgcgcctt atccggtaac tatcgtcttg	6420
agtccaaccc ggtaagacac gacttatcgc cactggcagc agccactggt aacaggatta	6480
gcagagcgag gtatgtaggc ggtgctacag agttcttgaa gtggtggcct aactacggct	6540
acactagaag aacagtatit ggtatctgcg ctctgctgaa gccagttacc ttcggaaaaa	6600
gagttggtag ctcttgatcc ggcaaaaaa ccaccgctgg tagcgggtgtt tttttgttt	6660

gcaagcagca gattacgcgc agaaaaaag gatctcaaga agatccttg atcttttcta 6720
cggggtctga cgctcagtgg aacgaaaact cacgttaagg gattttggtc atgagattat 6780
caaaaaggat cttcacctag atccttttaa attaaaaatg aagttttaaa tcaatctaaa 6840
gtatatatga gtaaacttgg tctgacagt accaatgctt aatcagttag gcacctatct 6900

cagcgatctg tctatttcgt tcatccatag ttgctgact ccccgctgtg tagataacta 6960
cgatacggga gggcttacca tctggcccca gtgctgcaat gataccgga gacccacgt 7020
caccggtcc agatttatca gcaataaacc agccagccgg aagggccgag cgcagaagtg 7080
gtcctgcaac ttatccgcc tccatccagt ctattaattg ttgccgggaa gctagagtaa 7140
gtagtccgc agttaatagt ttgcgcaacg ttgttgccat tgctacaggc atcgtggtgt 7200
cacgctcgtc gtttggtatg gcttcattca gctccggtc ccaacgatca aggcgagtta 7260
catgatcccc catgttgtgc aaaaagcgg ttagctcctt cggctcctcg atcgttgtca 7320

gaagtaagtt ggccgcagtg ttatcactca tggttatggc agcactgcat aattctctta 7380
ctgtcatgcc atccgtaaga tgcttttctg tgactggtga gtactcaacc aagtcattct 7440
gagaatagtg tatgcggcga ccgagttgct cttgccggc gtcaatacgg gataatacgg 7500
cgccacatag cagaacttta aaagtgtca tcattgaaa acgttcttcg gggcgaaaac 7560
tctcaaggat ctiaccgctg ttgagatcca gtctgatga acccactcgt gcaccaact 7620
gatcttcagc atcttttact ttcaccagcg tttctgggtg agcaaaaaca ggaaggcaaa 7680
atgccgcaaa aaagggaata agggcgacac ggaaatgttg aatactcata ctcttccttt 7740

ttcaatatta ttgaagcatt tatcagggtt attgtctcat gagcggtac atatttgaat 7800
gtatttagaa aaataaaca ataggggttc cgcgcacatt tccccgaaa gtgccacctg 7860
acgtctaaga aaccattatt atgatgacat taacctataa aaataggcgt atcacgaggc 7920
cctttcgtc 7929

<210> SEQ ID NO. 34
<211> 4163
<212> DNA
<213> Escherichia coli KJ122
<220> Feature: kan, sacB
<223> Other Information: Plasmid template fro PCR of kan-sacB
<400> Sequence

tcgagaggcc tgacgggatt taaatcgcta gcggtctgct aaaggaagcg gaacacgtag 60

aaagccagtc cgcagaaacg gtgctgacct cggatgaatg tcagctactg ggctatctgg 120

acaagggaaa acgcaagcgc aaagagaaaag caggtagctt gcagtgggct tacatggcga	180
tagctagact gggcggtttt atggacagca agcgaaccgg aattgccagc tggggcgccc	240
tctggttaagg ttgggaagcc ctgcaaagta aactggatgg ctttcttgcc gccaaaggatc	300
tgatggcgca ggggatcaag atctgatcaa gagacaggat gaggatcggt tcgcatgatt	360
gaacaagatg gattgcacgc aggttctccg gccgcttggg tggagaggct attcggtat	420
gactgggcac aacagacaat cggctgctct gatgccgccg tgttccggct gtcagcgag	480
gggcgcccgg ttctttttgt caagaccgac ctgtccgggtg ccctgaatga actgcaggac	540
gaggcagcgc ggctatcgtg gctggccacg acgggcgttc cttgcgcagc tgtgctcgac	600
gttgtcactg aagcgggaag ggactggctg ctattgggcg aagtgccggg gcaggatctc	660
ctgtcatctc accttgctcc tgccagaaaa gtatccatca tggctgatgc aatgcggcgg	720
ctgcatacgc ttgatccggc tacctgcccc ttcgaccacc aagcgaaaca tcgcatcgag	780
cgagcacgta ctcgatgga agccggtctt gtcgatcagg atgatctgga cgaagagcat	840
caggggctcg cgccagccga actgttcgcc aggcctcaagg cgcgcatgcc cgacggcgag	900
gatctcgtcg tgacctatgg cgatgcctgc ttgccgaata tcatggtgga aaatggccgc	960
ttttctggat tcatcgactg tggccggctg ggtgtggcgg accgctatca ggacatagcg	1020
ttggctaccc gtgatatgc tgaagagctt ggcggcgaat gggctgaccg cttcctcgtg	1080
ctttacggta tcgccgtcc cgattcgag cgcacgcct tctatgcct tcttgacgag	1140
ttctctgag cgggactctg gggttcgaaa tgaccgacca agcgacgcc aacctgccat	1200
cacgagattt cgattccacc gccgccttct atgaaagggt gggcttcgga atcgttttcc	1260
gggacgccgg ctggatgac ctcagcgcg gggatctcat gctggagtgc ttcgccacg	1320
ctagtttaaa caccgcgagt cactgagcg tcagaccct taataagatg atcttcttga	1380
gatcgttttg gtctgcgct aatctcttgc tctgaaaacg aaaaaaccgc ctgacagggc	1440
ggtttttcga aggttctctg agctaccaac tctttgaacc gaggtactg gcttgaggga	1500
gcgcagtcac caaaacttgt ctttcagt tacccttaac cggcgcatga cttcaagact	1560
aactcctcta aatcaattac cagtggctgc tgccagtgg gcttttgcat gtctttccgg	1620
gttgactca agacgatagt taccggataa ggcgcagcgg tcggactgaa cggggggttc	1680
gtgcatacag tccagcttgg agcgaactgc ctaccggaa ctgagtgtca ggcgtggaat	1740
gagacaaaac cgccataac agcggaatga caccggtaaa ccgaaaggca ggaacaggag	1800
agcgacagag ggagccgcca ggggaaacg cctggtatct ttatagtcct gtcgggtttc	1860
gccaccactg attgagcgt cagatttctg gatgcttgc aggggggcgg agcctatgga	1920
aaaacggctt tgcccgggcc ctctcattc cctgttaagt atcttctgg catcttcag	1980

gaaatctccg ccccgttcgt aagccatttc cgctcgccgc agtcgaacga ccgagcgtag	2040
cgagtcagt agcgaggaag cggaatata cctgtatcac atattctgct gacgcaccgg	2100
tgcagccttt tttctcctgc cacatgaagc acttactga caccctcctc agtgccaaca	2160
tagtaagcca gtatacactc cgctagcgca tgcattccatt taaatggaag aaataaaaga	2220
aaatgccaat aggatatgg cattttcttt tgcgttttta tttgttaact gttaatgtc	2280
cttgttcaag gatgctgtct ttgacaacag atgttttctt gcccttgatg ttcagcagga	2340
agctcgccgc aaacgttgat tgtttgtctg cgtagaatcc tctgtttgtc atataccttg	2400
taatcacgac attgtttcct ttcgtttgag gtacagcgaa gtgtgagtaa gtaaaggtta	2460
catcgttagg atcaagatcc atttttaaca caaggccagt tttgttcagc ggcttgtatg	2520
ggccagttaa agaattagaa acataaccaa gcatgtaaat atcgttagac gtaatgccgt	2580
caatcgtcat tttgatccg cgggagtcag tgaacaggta ccatttgccg ttcattttaa	2640
agacgttcgc gcgttcaatt tcatctgtta ctgtgttaga tgcaatcagc ggtttcatca	2700
cttttttcag tgtgtaatca tcgttttagct caatcatacc gagagcgccg tttgctaact	2760
cagccgtgcg ttttttatcg ctttgcagaa gtttttgact ttcctgacgg aagaatgatg	2820
tgcttttgcc atagtatgct ttgttaataa aagattcttc gccttggttag ccattctcag	2880
ttccagtgtt tgcttcaaat actaagtatt tgtggccttt atcttctacg tagtgaggat	2940
ctctcagcgt atggttgtcg cctgagctgt agttgccttc atcgatgaac tgctgtacat	3000
tttgatacgt ttttccgtca ccgtcaaaga ttgatttata atcctctaca ccgttgatgt	3060
tcaaagagct gtctgatgct gatacgttaa ctgtgacagt tgtcagtgtt tgtttgccgt	3120
aatgtttacc ggagaaatca gtgtagaata aacggatttt tccgtcagat gtaaagtgg	3180
ctgaacctga ccattcttgt gtttggcttt ttaggataga atcatttgca tcgaatttgt	3240
cgctgtcttt aaagacgagg ccagcgtttt tccagctgtc aatagaagtt tcgccgactt	3300
tttgatagaa catgtaaatc gatgtgtcat ccgcattttt aggatctccg gctaatacaa	3360
agacgatgtg gtagccgtga tagtttgcga cagtgccgtc agcgttttgt aatggccagc	3420
tgteccaaac gtccaggcct tttgcagaag agatatTTTT aattgtggac gaatcaaatt	3480
cagaaacttg atatTTTTa tttttttgct gttcagggat ttgcagcata tcatggcgtg	3540
taatatggga aatgccgtat gtttccttat atggcttttg gttcgtttct ttcgcaaacg	3600
cttgagttgc gcctcctgcc agcagtgcgg tagtaaaggt taatactgtt gcttgttttg	3660
caaactTTTT gatgttcacg gttcatgtct ctttttttat gtactgtgtt agcggctctgc	3720
ttcttccagc cctcctgttt gaagatggca agttagtac gcacaataaa aaaagaccta	3780
aaatatgtaa ggggtgacgc caaagtatac actttgccct ttacacattt taggtcttgc	3840

ctgctttatc agtaacaaac ccgcgcgatt tacttttcga cctcattcta ttagactctc 3900
 gtttgattg caactggctt attttcctct ttgtttgat agaaaatcat aaaaggattt 3960
 gcagactacg ggccataaga actaaaaaat ctatctgttt cttttcattc tctgtatttt 4020
 ttatagtttc tgitgcatgg gcataaagtt gcctttttta tcacaattca gaaaatatca 4080
 taatatctca ttctactaaa taatagtga cggcaggtat atgtgatggg ttaaaaagga 4140
 tcggcggcgc ctcgatttaa atc 4163

<210> SEQ ID NO. 35

<211> 3350

<212> DNA

<213> Escherichia coli

<220> Feature: pykF coding region

<223> Other Information: Sequence use for deletion of pykF

<400> Sequence

gggcgaattg ggcccacgt cgcatgctcc cggccgccat ggccggccgc ggaattcgat 60
 acattcctct gcacgctttt tcatgtcac ctatccttag agcgaggcac caccactttc 120
 gtaataccgg attcgctttc cggcagtgcg ccagaaaagc aagtttctcc catccttctc 180
 aacttaaaga ctaagactgt catgaaaaag accaaaattg ttgcaccgc atctgttcac 240
 gtctgtaat attgcttttg tgaattaatt tgtatatga agcgccctga tgggcgcttt 300

ttttatttaa tcgataacca gaagcaataa aaaatcaaat cggatttcac tatataatct 360
 cactttatct aagatgaatc cgatggaagc atcctatcac tagtgaattc gcggccgcct 420
 gcaggtcgac catatgggag agctcccaac gcgttgatg catagcttga gtattctata 480
 gtgtcaccta aatagcttgg cgtaatcatg gtcatactg tttcctgtgt gaaattgtta 540
 tccgctcaca attccacaca acatacgagc cggaagcata aagtgtaaag cctgggtgac 600
 ctaatgagtg agctaactca cattaattgc gttgcgtca ctgcccgtt tccagtcggg 660
 aaacctgtcg tgccagctgc attaatgaat cggccaacgc gcggggagag gcggtttgcg 720

tattgggcgc tcttcgctt cctcgtcac tgactcgtg cgctcggtcg ttcggctgcg 780
 gcgagcggta tcagctcact caaaggcggg aataggtta tccacagaat caggggataa 840
 cgcaggaaag aacatgtgag caaaaggcca gcaaaaggcc aggaaccgta aaaaggccgc 900
 gttgctggcg tttttcata ggctccgcc cctgacgag catcacaaaa atcgacgtc 960
 aagtcagagg tggcgaaacc cgacaggact ataaagatac caggcgtttc cccctggaag 1020
 ctccctcgtg cgctctctg ttcgaccct gccgcttacc ggatacctgt ccgcctttct 1080

cccttcggga agcgtggcgc tttctcatag ctcacgtgt aggtatctca gttcgggtga 1140

ggtcgttcgc tccaagctgg gctgtgtgca cgaaccccc gttcagcccg accgtgcgc 1200

cttatccggt aactatcgtc ttgagtccaa cccggttaaga cagacttat cgccactggc 1260

agcagccact ggtaacagga ttagcagagc gaggtatgta ggcggtgcta cagagtctt 1320

gaagtggtag cctaactacg gctacactag aagaacagta tttggtatct gcgtctgct 1380

gaagccagtt accttcggaa aaagagttgg tagctcttga tccggcaaac aaaccaccgc 1440

tggtagcggg ggtttttttg tttgcaagca gcagattacg cgcagaaaaa aaggatctca 1500

agaagatcct ttgatctttt ctacgggggc tgacgctcag tggaaacgaaa actcacgtta 1560

agggattttg gtcatgagat tatcaaaaag gatcttcacc tagatccttt taaattaaaa 1620

atgaagtttt aaatcaatct aaagtatata tgagtaaact tggcttgaca gttaccaatg 1680

cttaatcagt gaggcaccta tctcagcgat ctgtctatct cgttcaccca tagttgcctg 1740

actccccgc gtgtagataa ctacgatacg ggaggggctta ccatctggcc ccagtctgc 1800

aatgataccg cgagaccac gctcaccggc tccagattta tcagcaataa accagccagc 1860

cggaggggcc gagcgcagaa gtggctctgc aactttatcc gcctccatcc agtctattaa 1920

ttgttgccgg gaagctagag taagtagttc gccagttaat agtttgcgca acgtttgtgc 1980

cattgctaca ggcatcgtgg tgtcacgctc gtcgtttggg atggcttcat tcagctccgg 2040

ttccaacga tcaaggcgag ttacatgatc ccccatgttg tgcaaaaaag cggtagctc 2100

cttcggctct ccgatcgttg tcagaagtaa gttggccgca gtgttatcac tcatggttat 2160

ggcagcactg cataattctc ttactgtcat gccatccgta agatgctttt ctgtgactgg 2220

tgagtactca accaagtcat tctgagaata gtgtatgcgg cgaccgagtt gctcttgccc 2280

ggcgtcaata cgggataata ccgcgccaca tagcagaact ttaaaagtgc tcatcattgg 2340

aaaacgttct tcggggcgaa aactctcaag gatcttaccg ctgttgagat ccagttcgat 2400

gtaaccact cgigcaccca actgatcttc agcatctttt actttacca gcgtttctgg 2460

gtgagcaaaa acaggaaggc aaaatgccgc aaaaaaggga ataaggcgca cacggaaatg 2520

ttgaatactc atactcttcc tttttcaata ttattgaagc atttatcagg gttattgtct 2580

catgagcgga tacatatgtg aatgtattta gaaaaataaa caaatagggg ttccgcgcac 2640

atttccccga aaagtgccac ctgatgcggt gtgaaatacc gcacagatgc gtaaggagaa 2700

aataccgcat caggaaattg taagcggtta tttttgtta aaattcgctg taaatttttg 2760

ttaaatacgc tcatttttta accaataggc cgaaatcggc aaaatccctt ataaatcaaa 2820

agaatagacc gagatagggt tgagtgttgt tccagtttgg aacaagagtc cactattaaa 2880

gaacgtggac tccaacgtca aaggcgcaaa aaccgtctat caggcgcatg gccactacg 2940
tgaaccatca ccctaatcaa gttttttggg gtcgaggtgc cgtaaagcac taaatcgga 3000
ccctaaaggg agccccgat ttagagcttg acggggaaag ccggcgaaac tggcgagaaa 3060
ggaagggaag aaagcgaaag gagcgggcgc tagggcgctg gcaagtgtag cggtcacgct 3120
gcgcgtaacc accacacccg ccgcgcttaa tgcgccgcta caggcgcgct ccattcgcca 3180
ttcaggctgc gcaactgttg ggaaggcgga tcggtgcggg cctcttcgct attacgccag 3240

ctggcgaaa ggggatgtgc tgcaaggcga ttaagttggg taacgccagg gttttcccag 3300
tcacgacgtt gtaaacgac ggccagtga ttgtaatac actcactata 3350

<210> SEQ ID NO.36

<211> 78

<212> DNA

<213> Synthetic DNA

<220> Feature: PCR Primer 87A

<400> Sequence

atgaaaaaga ccaaattgt ttgcaccatc ggaccgaaaa ccgaatctga agagatgtct 60

agcgcatgca tccattta 78

<210> SEQ ID NO. 37

<211> 79

<212> DNA

<213> Synthetic DNA

<220> Feature: PCR Primer 87B

<400> Sequence

ttacaggacg tgaacagatg cgggtttagt agtgccgctc ggtaccagtg caccagaaag 60

gcgaagaact ccagcatga 79

<210> SEQ ID NO. 38

<211> 20

<212> DNA

<213> Synthetic DNA

<220> Feature: PCR Primer 78A

<400> Sequence

acacattcct ctgcacgctt 20

<210> SEQ ID NO. 39

<211> 20

<212> DNA

<213> Synthetic DNA

<220> Feature: PCR Primer 78B

<400> Sequence

aggatgcttc catcgattc

20