



(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2013-0119945
(43) 공개일자 2013년11월01일

| | |
|---|---|
| (51) 국제특허분류(Int. Cl.) <i>C12N 1/19</i> (2006.01) <i>C12N 15/81</i> (2006.01) <i>C12P 7/52</i> (2006.01) | (71) 출원인 노보자임스 인코포레이티드 미국 캘리포니아 95616-4880 데이비스 드류 애비뉴 1445 |
| (21) 출원번호 10-2013-7015983 | (72) 발명자 젠센 홀리 미국 미네소타 55317 쉐허슨 브렌덴 코트 6618 |
| (22) 출원일자(국제) 2011년11월21일 심사청구일자 없음 | 러쉬 브라이언 미국 미네소타 55410 미네아폴리스 토마스 애비뉴 에스 4423 |
| (85) 번역문제출일자 2013년06월20일 | (74) 대리인 송봉식, 정삼영 |
| (86) 국제출원번호 PCT/US2011/061717 | |
| (87) 국제공개번호 WO 2012/074818 국제공개일자 2012년06월07일 | |
| (30) 우선권주장 61/416,199 2010년11월22일 미국(US) 61/535,181 2011년09월15일 미국(US) | |

전체 청구항 수 : 총 38 항

(54) 발명의 명칭 3-히드록시프로피온산 생산을 위한 조성물 및 방법

(57) 요약

본 출원은 활성 3-HP 발효 경로를 포함하는 유전적으로 변형된 효모 세포, 및 3-HP를 생산하는 이들 세포의 사용을 개시한다.

(72) 발명자

휴리타 자넷

미국 미네소타 55331 엑셀시오르 63 스트리트 웨스트 1690

마스텔 베스

미국 미네소타 55331 엑셀시오르 63 스트리트 웨스트 1690

베리 앨런

미국 캘리포니아 95746 그라니트 베이 오크 리프웨이 9655

야머 데이

미국 캘리포니아 95618 데이비스 호그 플레이스 5632

캐틀렛 마이클

미국 캘리포니아 95691 웨스트 새크라멘토 파둠 플레이스 1665

바나트 미셸

미국 캘리포니아 95635 새크라멘토 암네스트 웨이 5672

특허청구의 범위

청구항 1

외인성 PPC 유전자;

외인성 PYC 유전자;

외인성 AAT 유전자;

외인성 ADC 유전자;

외인성 BAAT 또는 gabT 유전자; 및

외인성 3-HPDH 유전자

로부터 선택된 하나 이상의 외인성 3-HP 경로 유전자를 포함하는, 활성 3-HP 발효 경로를 포함하는 유전적으로 변형된 효모 세포.

청구항 2

제 1항에 있어서, 외인성 PYC 유전자를 포함하는 것을 특징으로 하는 유전적으로 변형된 효모 세포.

청구항 3

제 2항에 있어서, PYC 유전자는 SEQ ID NO: 2, 3, 4, 5, 6, 7, 및 8로부터 선택된 아미노산 서열과 적어도 60%, 예를 들어, 적어도 65%, 70%, 75%, 80%, 85%, 90%, 95%, 97%, 98%, 99%, 또는 100% 서열 동일성을 갖는 폴리펩티드를 암호화하는 것을 특징으로 하는 유전적으로 변형된 효모 세포.

청구항 4

제 2항 또는 제 3항에 있어서, PYC 유전자가 SEQ ID NO: 1의 뉴클레오티드 서열과 적어도 60%, 예를 들어, 적어도 65%, 70%, 75%, 80%, 85%, 90%, 95%, 97%, 98%, 99%, 또는 100% 서열 동일성을 갖는 뉴클레오티드 서열을 포함하는 것을 특징으로 하는 유전적으로 변형된 효모 세포.

청구항 5

제 1항 내지 제 4항 중 어느 한 항에 있어서, 외인성 AAT 유전자를 포함하는 것을 특징으로 하는 유전적으로 변형된 효모 세포.

청구항 6

제 5항에 있어서, AAT 유전자는 SEQ ID NO: 14, 15, 및 16으로부터 선택된 아미노산 서열과 적어도 60%, 예를 들어, 적어도 65%, 70%, 75%, 80%, 85%, 90%, 95%, 97%, 98%, 99%, 또는 100% 서열 동일성을 갖는 폴리펩티드를 암호화하는 것을 특징으로 하는 유전적으로 변형된 효모 세포.

청구항 7

제 5항 또는 제 6항에 있어서, AAT 유전자는 SEQ ID NO: 13의 뉴클레오티드 서열과 적어도 60%, 예를 들어, 적어도 65%, 70%, 75%, 80%, 85%, 90%, 95%, 97%, 98%, 99%, 또는 100% 서열 동일성을 갖는 뉴클레오티드 서열을 포함하는 것을 특징으로 하는 유전적으로 변형된 효모 세포.

청구항 8

제 1항 내지 제 7항 중 어느 한 항에 있어서, 외인성 ADC 유전자를 포함하는 것을 특징으로 하는 유전적으로 변형된 효모 세포.

청구항 9

제 8항에 있어서, ADC 유전자는 SEQ ID NO: 17, 18, 133, 135, 137, 및 139로부터 선택된 아미노산 서열과 적어도 60%, 예를 들어, 적어도 65%, 70%, 75%, 80%, 85%, 90%, 95%, 97%, 98%, 99%, 또는 100% 서열 동일성을

갖는 폴리펩티드를 암호화하는 것을 특징으로 하는 유전적으로 변형된 효모 세포

청구항 10

제 8항 또는 제 9항에 있어서, ADC 유전자는 SEQ ID NO: 130, 131, 132, 134, 136, 및 138로부터 선택된 뉴클레오티드 서열과 적어도 60%, 예를 들어, 적어도 65%, 70%, 75%, 80%, 85%, 90%, 95%, 97%, 98%, 99%, 또는 100% 서열 동일성을 갖는 뉴클레오티드 서열을 포함하는 것을 특징으로 하는 유전적으로 변형된 효모 세포.

청구항 11

제 1항 내지 제 10항 중 어느 한 항에 있어서, 외인성 BAAT 유전자 또는 외인성 gabT 유전자를 포함하는 것을 특징으로 하는 유전적으로 변형된 효모 세포.

청구항 12

제 11항에 있어서, BAAT 유전자 또는 gabT 유전자는 SEQ ID NO: 20, 21, 22, 23, 및 24로부터 선택된 아미노산 서열과 적어도 60%, 예를 들어, 적어도 65%, 70%, 75%, 80%, 85%, 90%, 95%, 97%, 98%, 99%, 또는 100% 서열 동일성을 갖는 폴리펩티드를 암호화하는 것을 특징으로 하는 유전적으로 변형된 효모 세포.

청구항 13

제 11항 또는 제 12항에 있어서, BAAT 유전자 또는 gabT 유전자는 SEQ ID NO: 19, 140, 141, 및 142로부터 선택된 뉴클레오티드 서열과 적어도 60%, 예를 들어, 적어도 65%, 70%, 75%, 80%, 85%, 90%, 95%, 97%, 98%, 99%, 또는 100% 서열 동일성을 갖는 뉴클레오티드 서열을 포함하는 것을 특징으로 하는 유전적으로 변형된 효모 세포.

청구항 14

제 11항 내지 제 13항 중 어느 한 항에 있어서, 상기 BAAT 유전자 또는 gabT는 BAAT 유전자이며, 이것은 또한 gabT 유전자인 것을 특징으로 하는 유전적으로 변형된 효모 세포.

청구항 15

제 1항 내지 제 14항 중 어느 한 항에 있어서, 외인성 3-HPDH 유전자를 포함하는 것을 특징으로 하는 유전적으로 변형된 효모 세포.

청구항 16

제 15항에 있어서, 3-HPDH 유전자는 SEQ ID NO: 26, 27, 28, 29, 30, 31, 32, 33, 34, 및 129로부터 선택된 아미노산 서열과 적어도 60%, 예를 들어, 적어도 65%, 70%, 75%, 80%, 85%, 90%, 95%, 97%, 98%, 99%, 또는 100% 서열 동일성을 갖는 폴리펩티드를 암호화하는 것을 특징으로 하는 유전적으로 변형된 효모 세포.

청구항 17

제 15항 또는 제 16항에 있어서, 3-HPDH는 SEQ ID NO: 25, 143, 144, 및 343으로부터 선택된 뉴클레오티드 서열과 적어도 60%, 예를 들어, 적어도 65%, 70%, 75%, 80%, 85%, 90%, 95%, 97%, 98%, 99%, 또는 100% 서열 동일성을 갖는 뉴클레오티드 서열을 포함하는 것을 특징으로 하는 유전적으로 변형된 효모 세포.

청구항 18

제 15항 내지 제 17항 중 어느 한 항에 있어서, 3-HPDH 유전자는 또한 HIBADH 유전자인 것을 특징으로 하는 유전적으로 변형된 효모 세포.

청구항 19

제 15항 내지 제 18항 중 어느 한 항에 있어서, 3-HPDH 유전자는 또한 4-히드록시부티레이트 데히드로게나제 유전자인 것을 특징으로 하는 유전적으로 변형된 효모 세포.

청구항 20

제 1항 내지 제 19항 중 어느 한 항에 있어서, 외인성 PPC 유전자를 포함하는 것을 특징으로 하는 유전적으로 변형된 효모 세포.

청구항 21

제 20항에 있어서, PPC 유전자는 SEQ ID NO: 10, 11, 및 12로부터 선택된 아미노산 서열과 적어도 60%, 예를 들어, 적어도 65%, 70%, 75%, 80%, 85%, 90%, 95%, 97%, 98%, 99%, 또는 100% 서열 동일성을 갖는 폴리펩티드를 암호화하는 것을 특징으로 하는 유전적으로 변형된 효모 세포.

청구항 22

제 20항 또는 제 21항 중 어느 한 항에 있어서, PPC 유전자는 SEQ ID NO: 9의 뉴클레오티드 서열과 적어도 60%, 예를 들어, 적어도 65%, 70%, 75%, 80%, 85%, 90%, 95%, 97%, 98%, 99%, 또는 100% 서열 동일성을 갖는 뉴클레오티드 서열을 포함하는 것을 특징으로 하는 유전적으로 변형된 효모 세포.

청구항 23

제 1항 내지 제 22항 중 어느 한 항에 있어서, 상기 효모 세포는 크랩트리-음성인 것을 특징으로 하는 유전적으로 변형된 효모 세포.

청구항 24

제 1항 내지 제 22항 중 어느 한 항에 있어서, 효모 세포는 이사트첸키아, 칸디다, 클루이베로미세스, 피치아, 쉬조사카로미세스, 토를라스포라, 지고사카로미세스, 및 사카로미세스로부터 선택된 속에 속하는 것을 특징으로 하는 유전적으로 변형된 효모 세포.

청구항 25

제 24항에 있어서, 효모 세포는 아이. 오리엔탈리스/피. 페르멘탄스 클레이드 및 사카로미세스 클레이드로부터 선택된 클레이드에 속하는 것을 특징으로 하는 유전적으로 변형된 효모 세포.

청구항 26

제 24항에 있어서, 효모 세포는 아이. 오리엔탈리스, 씨. 람비카, 및 에스. 불테리로부터 선택되는 것을 특징으로 하는 유전적으로 변형된 효모 세포.

청구항 27

제 1항 내지 제 26항 중 어느 한 항에 있어서, 상기 세포는 PDC, ADH, GAL6, CYB2A, CYB2B, GPD, GPP, ALD, 및 PCK 유전자로부터 선택된 네이티브 유전자의 하나 이상의 삭제 또는 붕괴를 추가로 포함하는 것을 특징으로 하는 유전적으로 변형된 효모 세포.

청구항 28

제 27항에 있어서, 삭제 또는 붕괴 중 하나 이상은 외인성 3-HP 경로 유전자로부터 발생하는 것을 특징으로 하는 유전적으로 변형된 효모 세포.

청구항 29

제 1항 내지 제 28항 중 어느 한 항에 있어서, 외인성 3-HP 경로 유전자 중 하나 이상은 하나 이상의 외인성 조절 요소와 작동 가능하게 결합되는 것을 특징으로 하는 유전적으로 변형된 효모 세포.

청구항 30

제 1항 내지 제 29항 중 어느 한 항에 있어서, 세포는 75 g/L 이상의 3-HP를 함유하는 배지에서 4 미만의 pH에서 성장할 수 있는 것을 특징으로 하는 유전적으로 변형된 효모 세포.

청구항 31

제 1항 내지 제 30항 중 어느 한 항에 있어서, 세포는 3-HP-저항성 효모 세포인 것을 특징으로 하는 유전적으로

변형된 효모 세포.

청구항 32

제 1항 내지 제 31항 중 어느 한 항에 있어서, 세포는 돌연변이 및/또는 선택을 겪었으며, 이러한 돌연변이 및/또는 선택된 세포는 같은 종의 야생형 세포보다 3-HP에 대하여 더 높은 정도의 저항성을 소유하는 것을 특징으로 하는 유전적으로 변형된 효모 세포.

청구항 33

제 32항에 있어서, 세포는 하나 이상의 외인성 3-HP 경로 유전자를 가지고 유전적으로 변형되기 전에 돌연변이 및/또는 선택을 겪는 것을 특징으로 하는 유전적으로 변형된 효모 세포.

청구항 34

제 32항 또는 제 33항에 있어서, 세포는 젖산 또는 3-HP의 존재시 선택을 겪는 것을 특징으로 하는 유전적으로 변형된 효모 세포.

청구항 35

제 34항에 있어서, 선택은 케모스타트 선택인 것을 특징으로 하는 유전적으로 변형된 효모 세포.

청구항 36

(i) 적어도 하나의 탄소 공급원을 포함하는 배지의 존재시 제 1항 내지 제 35항 중 어느 한 항의 유전적으로 변형된 효모 세포를 배양하는 단계

(ii) 배양물로부터 3-HP를 분리하는 단계

를 포함하는, 3-HP를 생산하는 방법.

청구항 37

제 36항에 있어서, 상기 탄소 공급원은 글루코스, 자일로스, 아라비노스, 수크로스, 프럭토스, 셀룰로스, 글루코스 올리고머, 및 글리세롤로부터 선택되는 것을 특징으로 하는 방법.

청구항 38

제 36항 또는 제 37항에 있어서, 배지는 5 미만, 예를 들어, 약 1.5 내지 약 4.5, 약 2.0 내지 약 4.0, 또는 약 2.0 내지 약 3.5의 범위의 pH인 것을 특징으로 하는 방법.

명세서

기술분야

[0001] 서열 목록에 대한 참고

[0002] 본 출원은 컴퓨터 판독 가능 형태로 서열 목록을 함유하며, 이것은 본원에 참고로 포함된다.

배경 기술

[0003] 3-히드록시프로피온산 (3-HP)은 발효에 의해 만들어질 수 있는 상위 12개의 고잠재성 빌딩 블록 (building block) 화합물 중 하나로서 미국 에너지부 (U.S. Department of Energy)에 의해 확인된 세 개의 탄소 카르복실산이다. 젖산 (2-히드록시프로피온산)의 이소머 (isomer)인, 3-HP에 대한 다른 이름은 에틸렌 젖산 및 3-히드록시프로피오네이트를 포함한다. 3-HP는 글루코스로부터 이론상 100% 수득율, 다양한 화학 반응에 참여하는 것을 허용하는 다수의 작용기, 및 낮은 독성을 갖는, 매력적인 재생 가능한 플랫폼 화학물질이다. 3-HP는 1,3-프로판디올, 말론산, 아크릴아미드, 및 아크릴산과 같은, 여러 화학 상품을 형성하는 기질로서 사용될 수 있다. 아크릴산은 아크릴레이트 에스테르 및 초흡수성 폴리머를 만들기 위해 사용된 대량의 화학물질이고 (> 70억 lbs/년), 현재 프로필렌의 촉매 산화로부터 유도된다. 3-HP의 발효성 생산은 이들 상업적으로 중요한 화학물질에 대한 공급원료로서 석유화학물질에 대한 지속 가능한 대체물질을 제공할 것이고, 따라서 에너지 소비, 석유에 대

한 미국 의존, 및 온실 가스의 생산을 감소시킨다.

[0004] 박테리아는 당을 유기산으로 발효시키는데 사용될 수 있다. 하지만, 박테리아는 대규모 유기산 생산에 대해 특정 문제점을 제공한다. 유기산이 생산되면, 발효 배지는 점점 산성이 된다. 낮은 pH 조건이 실제로 바람직한데, 결과물이 부분적으로 또는 전체적으로 산성 형태이기 때문이다. 하지만, 유기산을 생산하는 대부분의 박테리아는 강한 산성 환경에서 잘 수행하지 못하고, 그러므로 배지가 더 산성이 되면서 죽거나 그것들이 경제적으로 살 수 없도록 천천히 생산하기 시작한다. 이것을 방지하기 위해, 더 높은 pH를 유지하기 위해 배지에 버퍼를 처리하는 것이 필요하다. 하지만, 이것은 유기산 생성물의 회수를 더 어렵고 비싸게 만든다.

[0005] 최근 수년 동안 당을 유기산으로 발효시키는 효모의 사용에 관한 흥미가 증가하였다. 효모는 많은 발효 공업에서 생체 촉매로서 사용되고, 박테리아를 통해 여러 이점을 제공한다. 많은 박테리아가 성장 및 당의 효과적인 대사에 필요한 특정 아미노산 또는 단백질을 합성할 수 없는 한편, 대부분의 효모 종은 무기 질소 화합물로부터 그것들의 필수 아미노산 또는 단백질을 합성할 수 있다. 효모는 또한 박테리오파지 감염에 민감하지 않으며, 이것은 박테리아의 생산성 또는 전체 발효 실행의 손실로 이어질 수 있다.

[0006] 효모가 유기산 생산에 대한 매력적인 후보물질이지만, 그것들은 여러 문제점을 제공한다. 첫 번째로, 효모에서 경로 조작은 전형적으로 박테리아에서 보다 더 어렵다. 효모에서 효소는 세포질, 미토콘드리아, 또는 퍼옥시좀에서 분류되는 반면, 박테리아에서 그것들은 세포질에 모인다. 이것은 모든 효소의 생합성 경로가 단일 세포 내의 같은 구획에 동시 존재한다는 것을 보장하도록 표적화 신호가 제거될 필요가 있다는 것을 의미한다. 경로의 수송의 조절은 구획 사이를 중재하고 또한 원하는 생성물에 대한 탄소 흐름을 최대화하는 것이 필요할 수도 있다. 두 번째로, 모든 효모 종이 대규모로 경제적 발효에 대하여 필요한 기준을 만족하는 것은 아니다. 사실은, 작은 퍼센트의 효모만이 충분히 높은 부피 및 낮은 pH 조건 하에 강하게 성장하는 능력을 가진 특이적 당 이용의 조합을 보유한다. 미국 에너지부는 3-HP를 포함하는, 여러 유기산의 경제적인 발효에 대하여 대략 2.5 g/L/시간의 생산율이 필요하다는 것을 추정하였다 (<http://www1.eere.energy.gov/biomass/pdfs/35523.pdf>).

[0007] 많은 효모 종이 헥소스 당을 에탄올로 발효시키지만, 자연적으로 유기산의 큰 수득을 생산하는 것은 적다. 이것은 유기산을 생산하는 다양한 효모 종을 유전적으로 변형시키기 위한 노력으로 이어졌다. 젖산을 생산하는 유전적으로 변형된 효모 균주는 이전에 네이티브 (native) 피루베이트 데카르복실라제 (PDC) 유전자를 방해하고 삭제하고 락테이트 데히드로게나제 (LDH)를 삽입하여 에탄올 생산을 제거하는 단계에 의해 개발되었다 (예를 들어, W099/14335, W000/71738, W002/42471, W003/049525, W003/102152 및 W003/102201). 이 대안은 당 대사 작용을 에탄올 생산에서 젖산 생산으로 전환한다. 효모에 대한 발효 생성물 및 경로는 박테리아의 것들과 다르고, 따라서 다른 조작 접근법이 수득율을 최대화하기 위해 필요하다. 유기산 생성물 수득율 또는 순도를 향상시키기 위해 제거 또는 감소가 필요할 수도 있는 다른 네이티브 생성물은 글리세롤, 아세테이트, 및 디올이다. 유전적으로 변화된 효모 균주에서 글리세롤의 감소는 예를 들어, W007/106524에 설명된다.

[0008] 젖산과 달리 3-HP는 자연에 알려진 어떤 경로의 주요 최종 생성물이 아니며, 일부 박테리아 및 균류에서 미량으로만 발견된다. 따라서, 유전공학의 큰 거래가 3-HP를 생산하는 효모를 생성하기 위해 필요하다. 사카로미세스 세레비시애 (*Saccharomyces cerevisiae*) 균주는 이전에 락테이트 중재를 통해 3-HP를 생산하도록 조작되었다 (W002/042418 참조). 하지만, 야생형 에스. 세레비시애 (*S. cerevisiae*)의 오차 수준은 그것을 3-HP 생산을 위한 최적의 숙주로 만드는데 불충분하다. 그러므로, 산업상 규모에서 더 비용-효율적 방식으로 3-HP를 생성하는 향상된 효모 균주가 필요하다.

발명의 내용

해결하려는 과제

[0009] 특정 구체예에서 여기에 제공된 것은 PEP, 피루베이트, 및/또는 글리세롤에서 3-HP로 활성 3-HP 발효 경로를 포함하는 유전적으로 변형된 효모 세포이다. 특정 구체예에서, 여기에 제공된 세포는 PPC, PYC, AAT, ADC, BAAT, gabT, 3-HPDH, HIBADH, 4-히드록시부티레이트 데히드로게나제, ACC, AAM, 알라닌 데히드로게나제, 알데히드 데히드로게나제, BCKA, KGD, 4-아미노부티레이트 아미노트랜스퍼라제, β -알라닐-CoA 암모니아 리아제, Co-A 아실화 말로네이트 세미알데히드 데히드로게나제, CoA 신테타제, CoA 트랜스퍼라제, 글리세롤 데히드라타제, IPDA, LDH, 락틸-CoA 데히드라타제, 말레이트 데카르복실라제, 말레이트 데히드로게나제, 말로닐-CoA 리덕타제, OAA 포르메이트리아제, OAA 데히드로게나제, 피루베이트/알라닌 아미노트랜스퍼라제, PDH, 2-케토산 데카르복실라제, 3-HP-CoA 데히드라타제, 3-HP-CoA 히드롤라제, 또는 3-히드록시이소부티릴-CoA 히드롤라제 활

성과 함께 효소를 암호화하는 하나 이상의 3-HP 경로 유전자를 함유한다.

- [0010] 특정 구체예에서 여기에 제공된 것은 PEP 또는 피루베이트에서 3-HP로 활성 3-HP 발효 경로를 포함하는 유전적으로 변형된 효모 세포이며, 세포는 PPC, PYC, AAT, ADC, BAAT, gabT, 3-HPDH, HIBADH, 및 4-히드록시부티레이트 데히드로게나제 활성을 갖는 효소를 암호화하는 하나 이상의 유전자를 함유한다. 특정 구체예에서, 3-HP 경로 유전자 중 하나 이상은 외인성이고, 이들 구체예에서 유전자는 효모, 균류, 박테리아, 식물, 곤충, 또는 포유동물 근원으로부터 유도될 수도 있다. 예를 들어, 세포는 아이. 오리엔탈리스 (*I. orientalis*)로부터 유도된 효모 PYC 유전자 또는 알. 스페로이데스 (*R. sphaeroides*), 알. 에틀리 (*R. etli*), 피. 플루오레센스 (*P. fluorescens*), 씨. 글루타미쿰 (*C. glutamicum*), 또는 에스. 멜리로티 (*S. meliloti*)로부터 유도된 박테리아 PYC 유전자; 이. 콜리 (*E. coli*), 엠. 씨모아우토포피쿰 (*M. thermoautotrophicum*), 또는 씨. 페르프링겐스 (*C. perfringens*)로부터 유도된 박테리아 PPC 유전자, 아이. 오리엔탈리스 또는 에스. 세레비시아로부터 유도된 효모 AAT 유전자 또는 이. 콜리로부터 유도된 박테리아 AAT 유전자; 에스. 아베르미틸리스 (*S. avermitilis*), 씨. 아세토부틸리쿰 (*C. acetobutylicum*), 에이치. 필로리 (*H. pylori*), 비. 리체니포르미스 (*B. lichenim*), 또는 씨. 글루타미쿰으로부터 유도된 박테리아 ADC 유전자; 아이. 오리엔탈리스 또는 에스. 클루이베리 (*S. kluyveri*)로부터 유도된 효모 BAAT 유전자 또는 에스. 아베르미틸리스로부터 유도된 박테리아 BAAT; 에스. 세레비시아로부터 유도된 효모 gabT 유전자 또는 에스. 아베르미틸리스로부터 유도된 박테리아 gabT 유전자; 아이. 오리엔탈리스 또는 에스. 세레비시아로부터 유도된 효모 3-HPDH 유전자 또는 이. 콜리 또는 엠. 세둘라 (*M. sedula*)로부터 유도된 박테리아 3-HPDH 유전자; 에이. 패칼리스 (*A. faecalis*), 피. 푸티다 (*P. putida*), 또는 피. 애루기노사 (*P. aeruginosa*)로부터 유도된 박테리아 HIBADH 유전자; 및/또는 씨. 클루이베리 (*C. kluyveri*)로부터 유도된 효모 4-히드록시부티레이트 데히드로게나제 유전자 또는 알. 유트로파 (*R. eutropha*)로부터 유도된 박테리아 4-히드록시부티레이트 데히드로게나제 유전자를 함유한다.
- [0011] 특정 구체예에서 여기에 제공된 것은 PEP 또는 피루베이트에서 3-HP로 활성 3-HP 발효 경로를 포함하는 유전적으로 변형된 효모 세포이며, 세포는 PPC, 말레이트 데히드로게나제, 및 말레이트 데카르복실라제 활성을 갖는 효소를 암호화하는 하나 이상의 유전자를 함유한다. 특정 구체예에서, 3-HP 경로 유전자 중 하나 이상은 외인성이고, 이들 구체예에서, 유전자는 효모, 균류, 박테리아, 식물, 곤충, 또는 포유동물 근원으로부터 유도될 수도 있다.
- [0012] 특정 구체예에서 여기에 제공된 것은 PEP 또는 피루베이트에서 3-HP로 활성 3-HP 발효 경로를 포함하는 유전적으로 변형된 효모 세포이며, 세포는 PPC, 2-케토산 데카르복실라제, KGD, BCKA, 인돌피루베이트 데카르복실라제, 3-HPDH, HIBADH, 또는 4-히드록시부티레이트 데히드로게나제 활성을 갖는 효소를 암호화하는 하나 이상의 유전자를 함유한다. 특정 구체예에서, 3-HP 경로 유전자 중 하나 이상은 외인성이고, 이들 구체예에서, 유전자는 효모, 균류, 박테리아, 식물, 곤충, 또는 포유동물 근원으로부터 유도된다.
- [0013] 특정 구체예에서 여기에 제공된 것은 PEP 또는 피루베이트에서 3-HP로 활성 3-HP 발효 경로를 포함하는 유전적으로 변형된 효모 세포이며, 세포는 PPC, OAA 포르메이트리아제, 말로닐-CoA 리덕타제, Co-A 아실화 말로네이트 세미알데히드 데히드로게나제, 3-HPDH, HIBADH, 및 4-히드록시부티레이트 데히드로게나제 활성을 갖는 효소를 암호화하는 하나 이상의 유전자이다. 특정 구체예에서, 3-HP 경로 유전자 중 하나 이상은 외인성이고, 이들 구체예에서, 유전자는 효모, 균류, 박테리아, 식물, 곤충, 또는 포유동물 근원으로부터 유도된다.
- [0014] 특정 구체예에서 여기에 제공된 것은 피루베이트에서 3-HP로 활성 3-HP 발효 경로를 포함하는 유전적으로 변형된 효모 세포이며, 세포는 PDH, 아세틸-CoA 카르복실라제, 말로닐-CoA 리덕타제, CoA 아실화 말로네이트 세미알데히드 데히드로게나제, 3-HPDH, HIBADH, 및 4-히드록시부티레이트 데히드로게나제 활성을 갖는 효소를 암호화하는 하나 이상의 유전자이다. 특정 구체예에서, 3-HP 경로 유전자 중 하나 이상은 외인성이고, 이들 구체예에서, 유전자는 효모, 균류, 박테리아, 식물, 곤충, 또는 포유동물 근원으로부터 유도된다.
- [0015] 특정 구체예에서 여기에 제공된 것은 피루베이트에서 3-HP로 활성 3-HP 발효 경로를 포함하는 유전적으로 변형된 효모 세포이며, 세포는 알라닌 데히드로게나제, 피루베이트/알라닌 아미노트랜스퍼라제, 알라닌 2,3 아미노뮤타제, CoA 트랜스퍼라제, CoA 신데타제, β -알라닐-CoA 암모니아 리아제, 3-HP-CoA 데히드라타제, 3-HP-CoA 히드롤라제, 3-히드록시이소부티릴-CoA 히드롤라제, BAAT, 3-HPDH, HIBADH, 및 4-히드록시부티레이트 데히드로게나제 활성을 갖는 효소를 암호화하는 하나 이상의 유전자를 함유한다. 특정 구체예에서, 3-HP 경로 유전자 중 하나 이상은 외인성이고, 이들 구체예에서, 유전자는 효모, 균류, 박테리아, 식물, 곤충, 또는 포유동물 근원으로부터 유도된다.
- [0016] 특정 구체예에서 여기에 제공된 것은 피루베이트에서 3-HP로 활성 3-HP 발효 경로를 포함하는 유전적으로 변형

된 효모 세포이며, 세포는 LDH, CoA 트랜스퍼라제, 락틸-CoA 데히드라타제, 3-HP-CoA 데히드라타제, 3-HP-CoA 히드롤라제, 및 3-히드록시이소부티릴-CoA 히드롤라제 활성을 갖는 효소를 암호화하는 하나 이상의 유전자를 함유한다. 특정 구체예에서, 3-HP 경로 유전자 중 하나 이상은 외인성이고, 이들 구체예에서, 유전자는 효모, 균류, 박테리아, 식물, 곤충, 또는 포유동물 근원으로부터 유도된다.

[0017] 특정 구체예에서 여기에 제공된 것은 PEP 또는 피루베이트에서 3-HP로 활성 3-HP 발효 경로를 포함하는 유전적으로 변형된 효모 세포이며, 세포는 글리세롤 데히드라타제 및 알데히드 데히드로게나제 활성을 갖는 효소를 암호화하는 하나 이상의 유전자를 함유한다. 특정 구체예에서, 3-HP 경로 유전자 중 하나 이상은 외인성이고, 이들 구체예에서, 유전자는 효모, 균류, 박테리아, 식물, 곤충, 또는 포유동물 근원으로부터 유도된다.

[0018] 여기에 제공된 유전적으로 변형된 효모 세포는 어떤 효소 종도 될 수 있다. 특정 구체예에서, 세포는 크랩트리(Crabtree)-음성이고, 이들 중 특정 구체예에서, 그것들은 속(genus) 이사트첸키아(Issatchenkia), 칸디다(Candida), 클루이베로미세스(Kluyveromyces), 피치아(Pichia), 쉬조사카로미세스(Schizosaccharomyces), 토룰라스포라(Torulaspora), 지고사카로미세스(Zygosaccharomyces), 또는 사카로미세스(Saccharomyces)에 속한다. 이들 중 특정 구체예에서, 세포는 아이. 오리엔탈리스/피. 페르멘탄스(P. fermentans) 클레이드(clade) 또는 사카로미세스 클레이드에 속할 수도 있고, 이들 구체예에서, 그것들은 아이. 오리엔탈리스, 씨. 람비카(C. lambica), 또는 에스. 불데리(S. bulderi)일 수도 있다. 특정 구체예에서, 효모 세포는 3-HP 저항성 효모 세포일 수도 있다. 3-HP 저항성은 세포의 네이티브 특성일 수도 있고 그것은 활성 3-HP 발효 경로와 관련된 유전적 변형, 또는 이들의 조합의 도입 전에, 중에, 또는 후에 돌연변이 및/또는 선택을 겪은 세포로부터 발생한다. 특정 구체예에서, 효모 세포는 3-HP, 다른 발효 제품 또는 부산물, 및/또는 같은 종의 야생형 효모 세포에 의해 나타나는 것보다 더 큰 다양한 배지 성분 외에 유기산에 대한 내성의 정도를 나타낼 수도 있다. 특정 구체예에서, 효모 세포는 돌연변이 및/또는 선택된 세포는 같은 종의 야생형 세포보다 3-HP에 대하여 더 높은 정도의 저항을 소유하도록, 돌연변이 및/또는 선택을 겪었다. 이들 중 일부 구체예에서, 세포는 하나 이상의 외인성 3-HP 경로 유전자로 유전적으로 변형되기 전에 돌연변이 및/또는 선택을 겪었다. 일부 구체예에서, 세포는 젖산 또는 3-HP의 존재시 선택을 겪었다. 일부 구체예에서, 선택은 케모스태트(chemostat) 선택이다.

[0019] 활성 3-HP 발효 경로에 관련된 변형에 더하여, 여기에 제공된 세포는 하나 이상의 네이티브 유전자의 삭제 또는 붕괴를 함유할 수도 있다. 예를 들어, 세포는 하나 이상의 PDC, ADH, GAL6, CYB2A, CYB2B, GPD, GPP, ALD, 또는 PCK 유전자의 삭제 또는 붕괴를 함유할 수도 있다. 특정 구체예에서, 이들 삭제 또는 붕괴는 활성 3-HP 발효 경로와 관련된 하나 이상의 유전자의 도입에 결합될 수도 있다.

[0020] 특정 구체예에서 여기에 제공된 것은 적어도 하나의 탄소 근원의 존재시 세포를 배양하고 배양 배지로부터 3-HP를 분리함으로써 여기에 제공된 유전적으로 변형된 효모 세포를 사용하여 3-HP를 생산하는 방법이다. 이들 중 특정 구체예에서, 탄소 근원은 글루코스, 자일로스, 아라비노스, 수크로스, 프럭토스, 셀룰로스, 글루코스 올리고머, 및 글리세롤 중 하나 이상으로부터 선택될 수도 있다.

도면의 간단한 설명

[0021] 도 1: 선택 3-HP 발효 경로의 개요

도 2: 플라스미드 pM1Ba107

도 3: 표적화된 통합 기술의 개략적 표현

도 4: 플라스미드 pGMEr125(a)

도 5: 플라스미드 pGMR 25(b)

도 6: 플라스미드 pGMEr121

도 7: 플라스미드 pMhCt074

도 8: 플라스미드 pMhCt083

도 9: 플라스미드 pMhCt087

도 10: 플라스미드 pMhCt075

도 11: 플라스미드 pMhCt077

도 12: 플라스미드 pMhCt095
 도 13: 플라스미드 pMhCt096
 도 14: 플라스미드 pMeJi310-2
 도 15: 플라스미드 pMeJi312-2
 도 16: 플라스미드 pGMEr126
 도 17: 플라스미드 pGMEr130
 도 18: 플라스미드 pGMEr137
 도 19: 플라스미드 pACN5
 도 20: 플라스미드 pACN23
 도 21: 플라스미드 pHJJ27
 도 22: 플라스미드 pACN43
 도 23: 플라스미드 pHJJ75
 도 24: 플라스미드 pHJJ76
 도 25: 플라스미드 pJLJ49
 도 26: 플라스미드 pJLJ62
 도 27: 플라스미드 pMI458
 도 28: 플라스미드 pCM208
 도 29: 플라스미드 pJY39
 도 30: 플라스미드 pMcTs64
 도 31: 플라스미드 pMcTs65
 도 32: 플라스미드 pJLJ8.

발명을 실시하기 위한 구체적인 내용

- [0022] 본 발명의 다음 설명은 단지 본 발명의 다양한 구체예를 설명하기 위한 것이다. 이와 같이, 논의된 특정 변형은 본 발명의 범위에 대한 제한으로 생각되는 것이 아니다. 당업자에게 다양한 등가물, 변화, 및 변형은 본 발명의 범위로부터 벗어나지 않고 이루어질 수도 있다는 것이 명백하고, 이러한 동등한 구체예는 여기에 포함되는 것으로 생각된다.
- [0023] 여기에 인용된 모든 참고문헌은 전문이 참고로 포함된다.
- [0024] 약어
- [0025] 3-HP, 3-히드록시프로피온산; 3-HPA, 3-히드록시프로피온알데히드; 3-HPDH, 3-히드록시프로피온산데히드로게나제; AAM, 알라닌 2,3 아미노뮤타제; AAT, 아스파르테이트 아미노트랜스퍼라제; ACC, 아세틸-CoA 카르복실라제; ADC, 아스파르테이트 1-데카르복실라제; AKG, 알파-케토글루타레이트; ALD, 알데히드 데히드로게나제; BAAT, β -알라닌 아미노트랜스퍼라제; BCKA, 분지형-사슬 알파-케토산 데카르복실라제; bp, 염기쌍; CYB2, L-(+)-락테이트-시토크롬 c 옥시도리덕타제; CYC, 이소-2-시토크롬 c; EMS, 에탄 메틸 술포나제; ENO, 에놀라제; gabT, 4-아미노부티레이트 아미노트랜스퍼라제; GAPDH, 글리세르알데히드-3-포스페이트 데히드로게나제 3; GPD, 글리세롤 3-포스페이트 데히드로게나제; GPP, 글리세롤 3-포스페이트 포스파타제; HIBADH, 3-히드록시이소부티레이트 데히드로게나제; IPDA, 인돌피루베이트 데카르복실라제; KGD, 알파-케토글루타레이트 데카르복실라제; LDH, 락테이트 데히드로게나제; MAE, 말산 효소; OAA, 옥살로아세테이트; PCK, 포스포에놀피루베이트 카르복시키나제; PDC, 피루베이트 데카르복실라제; PDH, 피루베이트 데히드로게나제; PEP, 포스포에놀피루베이트; PGK, 포스포글리세레이트 키나제; PPC, 포스포에놀피루베이트 카르복실라제; PYC, 피루베이트 카르복실라제; RKI, 리보스 5-포스페이트 케톨-이소머라제; TAL, 트랜스알돌라제; TEF1, 번역 신장 인자-1; TEF2, 번역 신장 인자-2; TKL, 트

랜스케톨라제, XDH, 자일리톨 데히드로게나제; XR, 자일로스 리덕타제, YP, 효모 추출물/펩톤.

[0026] 설명

[0027] 여기에 제공된 것은 3-HP의 생산을 위한 유전적으로 변형된 효모 세포, 이들 효모 세포를 만드는 방법, 및 3-HP를 생산하는 이들 세포를 사용하는 방법이다. 여기에 사용된 바와 같이 "3-HP"는 3-히드록시프로피온산의 염 및 산성 형태를 포함한다.

[0028] 많은 3-HP 발효 경로가 업계에 알려져 있다 (예컨대, 미국 특허 번호 6,852,517; 미국 특허 번호 7,309,597; 미국 공개 번호 2001/0021978; 미국 공개 번호 2008/0199926; W002/42418; 및 W010/031083이며, 모든 것은 본원에 참조로 포함됨). 3-HP 발효 경로는 포스포에놀피루베이트 (PEP), 피루베이트, 옥살로아세트레이트 (OAA), 아스파르테이트, β -알라닌, 말로네이트 세미알데히드, 말레이트, 말로닐-CoA, 아세틸-CoA, 알라닌, 락테이트, 락틸-CoA, 아크릴일-CoA, 글리세롤, 3-히드록시프로피온알데히드 (3-HPA), β -알라닐-CoA, 3-HP-CoA, 및 글리세레이트를 포함할 수도 있는, 일련의 중간물을 통해 작동된다. 여러 알려진 3-HP 발효 경로의 개요는 도 1에 설명된다.

[0029] 여기에 개시된 바와 같이, 한 세트의 다양한 종의 효모 세포는 3-HP 저항성에 대하여 테스트되었다. 3-HP 저항을 나타내는 세포는 다양한 농도의 3-HP를 함유하는 배지에서 그것들의 성장률 및 글루코스 소비 속도에 기초하여 추가로 평가되었다. 이 실험에 기초하여, 한 세트의 이상적인 3-HP 생산용 숙주 세포를 확인하였다. 이들 숙주 세포들은 활성 3-HP 발효 경로를 함유하도록 유전적으로 변형되었으며, 낮은 pH 조건 하에 3-HP를 생산하는 유전적으로 변형된 효모 세포를 발생시킨다.

[0030] 특정 구체예에서 여기에 제공된 것은 적어도 하나의 활성 3-HP 발효 경로를 갖는, PEP, 피루베이트, 및/또는 글리세롤에서 3-HP로 유전적으로 변형된 효모 세포이다. "활성 3-HP 발효 경로"를 갖는 효모 세포는 여기에 사용된 바와 같이 3-HP 발효 경로에서 각각의 반응을 촉진하는데 필요한 활성 효소를 생산하고, 그러므로 적어도 하나의 발효 가능한 당의 존재시 발효 조건 하에 배양될 때 측정 가능한 수득율로 3-HP를 생산할 수 있다. 활성 3-HP 발효 경로를 갖는 효모 세포는 하나 이상의 3-HP 경로 유전자를 포함한다. "3-HP 경로 유전자"는 여기에 사용된 바와 같이 3-HP 발효 경로에 수반되는 효소를 암호화하는 뉴클레오티드 서열의 암호화 영역을 나타낸다.

[0031] 특정 구체예에서, 여기에 제공된 효모 세포는 PEP 또는 피루베이트, OAA, 아스파르테이트, β -알라닌, 및 말로네이트 세미알데히드 중간물 (예컨대, 미국 공개 번호 2010/0021978, 도 1)를 통해 진행되는 활성 3-HP 발효 경로를 갖는다. 이들 구체예에서, 효모 세포는 피루베이트 카르복실라제 (PYC), PEP 카르복실라제 (PPC), 아스파르테이트 아미노트랜스퍼라제 (AAT), 아스파르테이트 1-데카르복실라제 (ADC), β -알라닌 아미노트랜스퍼라제 (BAAT), 아미노부티레이트 아미노트랜스퍼라제 (gabT), 3-HP 데히드로게나제 (3-HPDH), 3-히드록시이소부티레이트 데히드로게나제 (HIBADH), 및 4-히드록시부티레이트 데히드로게나제 유전자 중 하나 이상을 포함하는, 한 세트의 3-HP 발효 경로 유전자를 포함한다. 3-HP 발효 경로 유전자는 또한 바람직하게 PEP의 OAA로 전환을 촉진하는 폴리펩티드를 생산하도록 변형된 PEP 카르복시키나제 (PCK) 유전자를 포함할 수도 있다 (네이티브 PCK 유전자는 유전적으로 바람직하게 OAA의 PEP로 역반응을 촉진하는 폴리펩티드를 생산한다).

[0032] 특정 구체예에서, 여기에 제공된 효모 세포는 PEP 또는 피루베이트, OAA, 및 말레이트 중간물을 통해 진행되는 활성 3-HP 발효 경로를 갖는다 (예컨대, 미국 공개 번호 2010/0021978, 도 4). 이들 구체예에서, 효모 세포는 PPC, PYC, 말레이트 데히드로게나제, 및 말레이트 데카르복실라제 유전자 중 하나 이상을 포함하는, 한 세트의 3-HP 발효 경로 유전자를 포함한다. 3-HP 발효 경로 유전자는 또한 바람직하게 PEP의 OAA로 전환을 촉진하는 폴리펩티드를 생산하도록 변형된 PCK 유전자를 포함할 수도 있다.

[0033] 특정 구체예에서, 여기에 제공된 효모 세포는 PEP 또는 피루베이트, OAA, 및 말로네이트 세미알데히드 중간물을 통해 진행되는 활성 3-HP 발효 경로를 갖는다 (예컨대, 미국 공개 번호 2010/0021978, 도 1). 이들 구체예에서, 효모 세포는 PPC, PYC, 2-케토산 데카르복실라제, 알파-케토글루타레이트 (AKG) 데카르복실라제 (KGD), 분지형-사슬 알파-케토산 데카르복실라제 (BCKA), 인돌피루베이트 데카르복실라제 (IPDA), 3-HPDH, HIBADH, 및 4-히드록시부티레이트 데히드로게나제 유전자 중 하나 이상을 포함하는, 한 세트의 3-HP 발효 경로를 포함한다. 3-HP 발효 경로 유전자는 또한 바람직하게 PEP의 OAA로 전환을 촉진하는 폴리펩티드를 생산하도록 변형된 PCK 유전자를 포함할 수도 있다. 게다가, 3-HP 발효 경로 유전자는 OAA의 말로네이트 세미알데히드로 전환을 촉진할 수 있는 폴리펩티드를 암호화하도록 변형된, PDC 유전자 및/또는 벤조일포르메이트 데카르복실라제 유전자를 포함할 수도 있다.

[0034] 특정 구체예에서, 여기에 제공된 효모 세포는 PEP 또는 피루베이트, OAA, 말로닐-CoA, 및 말로네이트 세미알데

히드 중간물을 통해 진행되는 활성 3-HP 발효 경로를 가지며, 말로네이트 세미알데히드 중간물은 선택적이다 (예컨대, 미국 공개 번호 2010/0021978, 도 2). 이들 구체예에서, 효모 세포는 PPC, PYC, OAA 포르메이트리아제, 말로닐-CoA 리덕타제, CoA 아실화 말로네이트 세미알데히드 데히드로게나제, 3-HPDH, HIBADH, 및 4-히드록시부티레이트 데히드로게나제 유전자 중 하나 이상을 포함하는, 한 세트의 3-HP 발효 경로 유전자를 포함한다. 3-HP 발효 경로 유전자는 또한 바람직하게 PEP의 OAA로 전환을 촉진하는 폴리펩티드를 생산하도록 변형된 PCK 유전자를 포함할 수도 있다. 게다가, 3-HP 발효 경로 유전자는 OAA의 말로닐 CoA로 전환을 촉진하는 폴리펩티드를 생산하기 위해 2-케토산 데히드로게나제 유전자를 변형함으로써 유도된 OAA 데히드로게나제 유전자를 포함할 수도 있다.

[0035] 특정 구체예에서, 여기에 제공된 효모 세포는 피루베이트, 아세틸-CoA, 말로닐-CoA, 및 말로네이트 세미알데히드 중간물을 통해서 진행되는 활성 3-HP 발효 경로를 가지며, 말로네이트 세미알데히드 중간물은 선택적이다 (예컨대, W002/042418, 도 44). 이들 구체예에서, 효모 세포는 피루베이트 데히드로게나제 (PDH), 아세틸-CoA 카르복실라제 (ACC), 말로닐-CoA 리덕타제, CoA 아실화 말로네이트 세미알데히드 데히드로게나제, 3-HPDH, HIBADH, 및 4-히드록시부티레이트 데히드로게나제 유전자 중 하나 이상을 포함하는, 한 세트의 3-HP 발효 경로 유전자를 포함한다.

[0036] 특정 구체예에서, 여기에 제공된 효모 세포는 피루베이트, 알라닌, β -알라닌, β -알라닐-CoA, 아크릴일-CoA, 3-HP-CoA, 및 말로네이트 세미알데히드 중간물을 통해서 진행되는 활성 3-HP 발효 경로를 가지며, β -알라닐-CoA, 아크릴일-CoA, 3-HP-CoA, 및 말로네이트 세미알데히드 중간물은 선택적이다 (β -알라닌은 말로네이트 세미알데히드 중간물을 통해 또는 β -알라닐-CoA, 아크릴일-CoA, 및 3-HP-CoA 중간물을 통해 3-HP로 전환될 수 있다 (예컨대, 미국 특허 7,309,597, 도 1)). 이들 구체예에서, 효모 세포는 알라닌 데히드로게나제, 피루베이트/알라닌 아미노트랜스퍼라제, 알라닌 2,3 아미노뮤타제, CoA 트랜스퍼라제, CoA 신테타제, β -알라닐-CoA 암모니아 리아제, 3-HP-CoA 데히드라타제, 3-HP-CoA 히드롤라제, 3-히드록시이소부티릴-CoA 히드롤라제, BAAT, 3-HPDH, HIBADH, 및 4-히드록시부티레이트 데히드로게나제 유전자 중 하나 이상을 포함하는, 한 세트의 3-HP 발효 경로를 포함한다.

[0037] 특정 구체예에서, 여기에 제공된 효모 세포는 피루베이트, 락테이트, 락틸-CoA, 아크릴일-CoA, 및 3-HP-CoA 중간물을 통해서 진행되는 활성 3-HP 발효 경로를 갖는다 (예컨대, W002/042418, 도 1). 이들 구체예에서, 효모 세포는 LDH, CoA 트랜스퍼라제, CoA 신테타제, 락틸-CoA 데히드라타제, 3-HP-CoA 데히드라타제, 3-HP-CoA 히드롤라제, 및 3-히드록시이소부티릴-CoA 히드롤라제 유전자 중 하나 이상을 포함하는, 한 세트의 3-HP 발효 경로를 포함한다.

[0038] 특정 구체예에서, 여기에 제공된 효모 세포는 글리세롤 및 3-HPA 중간물을 통해서 진행되는 활성 3-HP 발효 경로를 갖는다 (예컨대, 미국 특허 6,852,517). 이들 구체예에서, 효모 세포는 글리세롤 데히드라타제 및 알데히드 데히드로게나제 유전자 중 하나 이상을 포함하는, 한 세트의 3-HP 발효 경로를 포함한다.

[0039] 특정 구체예에서, 여기에 제공된 효모 세포는 PEP 또는 피루베이트, OAA, 아스파르테이트, β -알라닌, β -알라닐-CoA, 아크릴일-CoA, 3-HP-CoA, 및 알라닌 중간물을 통해서 진행되는 활성 3-HP 발효 경로를 가지며, OAA, 아스파르테이트, 및 알라닌 중간물은 선택적이다 (PEP 또는 피루베이트는 OAA 및 아스파르테이트를 통해 또는 알라닌을 통해 β -알라닌으로 전환될 수 있다) (W002/042418, 도 54; 미국 특허 7,309,597, 도 1 참조). 이들 구체예에서, 효모 세포는 PPC, PYC, AAT, ADC, CoA 트랜스퍼라제, CoA 신테타제, β -알라닐-CoA 암모니아 리아제, 3-HP-CoA 데히드라타제, 3-HP-CoA 히드롤라제, 3-히드록시이소부티릴-CoA 히드롤라제, 알라닌 데히드로게나제, 피루베이트/알라닌 아미노트랜스퍼라제, 및 AAM 유전자 중 하나 이상을 포함하는, 한 세트의 3-HP 발효 경로를 포함한다. 3-HP 발효 경로 유전자는 또한 바람직하게 PEP의 OAA로 전환을 촉진하는 폴리펩티드를 생산하도록 변형된 PCK 유전자를 포함할 수도 있다.

[0040] 여기에 제공된 효모 세포에서 3-HP 발효 경로 유전자는 내인성 또는 외인성일 수도 있다. "내인성"은 유전자, 프로모터, 및 종결 서열과 같은 유전적 구성요소에 관하여 여기에 사용된 바와 같이 유전적 구성요소가 특정 효모 세포의 네이티브 형태의 게놈의 특정 위치에 존재한다는 것을 의미한다. "외인성"은 유전적 구성요소에 관하여 여기에 사용된 바와 같이 유전적 구성요소가 특정 효모 세포의 네이티브 형태의 게놈의 특정 위치에 존재하지 않는다는 것을 의미한다. "네이티브"는 효모 세포에 관하여 여기에 사용된 바와 같이 특정 효모 종의 야생형 효모 세포를 나타낸다. "네이티브"는 대사 경로에 관하여 여기에 사용된 바와 같이 네이티브 효모 세포에 존재하고 활성인 대사 경로를 나타낸다.

[0041] 외인성 유전적 구성요소는 네이티브 또는 비-네이티브 서열을 가질 수도 있다. 네이티브 서열을 갖는 외인성 유

전적 구성요소는 네이티브 세포의 계놈에 존재하는 유전적 구성요소와 동일한 서열을 포함한다 (기능에 영향을 미치지 않는 개체 간 돌연변이를 제외) (즉, 외인성 유전적 구성요소는 내인성 유전적 구성요소와 동일하다). 하지만, 외인성 구성요소는 내인성 구성요소에 비해 숙주 세포 계놈의 다른 위치에 존재한다. 예를 들어, 내인성 PYC 유전자와 동일한 외인성 PYC 유전자는 효모 세포 내에 삽입될 수도 있으며, 비-네이티브 (증가된) 수의 PYC 유전자 카피를 갖는 변형된 세포를 발생시킨다. 비-네이티브 서열을 갖는 외인성 유전적 구성요소는 네이티브 세포의 계놈에서 발견되지 않는 서열을 포함한다. 예를 들어, 특정 종의 외인성 PYC 유전자는 또 다른 종의 효모 세포 내에 삽입될 수도 있다. 외인성 유전자는 바람직하게 작동 방식에서 숙주 세포 계놈로 통합되며, 숙주 세포에서 활성 단백질을 생산할 수 있다는 것을 의미한다. 하지만, 특정 구체예에서, 외인성 유전자는 숙주 세포질에 안정적으로 유지되는 벡터의 일부로서 세포 내로 도입될 수도 있다.

[0042] 특정 구체예에서, 여기에 제공된 효모 세포는 하나 이상의 외인성 3-HP 발효 경로 유전자를 포함한다. 특정 구체예에서, 여기에 개시된 유전적으로 변형된 효모 세포는 단일 외인성 유전자를 포함한다. 다른 구체예에서, 효모 세포는 다수의 외인성 유전자를 포함한다. 이들 구체예에서, 효모 세포는 단일 외인성 유전자의 다수의 카피 및/또는 둘 이상의 다른 외인성 유전자의 카피를 포함할 수도 있다. 다수의 외인성 유전자를 포함하는 효모 세포는 어떤 수의 외인성 유전자도 포함할 수 있다. 예를 들어, 이들 효모 세포는 1 내지 20개의 외인성 유전자를 포함할 수도 있고, 특정 바람직한 구체예에서, 그것들은 1 내지 7개의 외인성 유전자를 포함할 수도 있다. 외인성 유전자의 다수의 카피는 그것들이 서로 인접한 단일 위치에서 통합될 수도 있다. 대안으로, 그것들은 숙주 세포의 계놈 내의 여러 위치에서 통합될 수도 있다.

[0043] 특정 구체예에서, 여기에 제공된 효모 세포는 하나 이상의 내인성 3-HP 발효 경로 유전자를 포함한다. 이들 중 특정 구체예에서, 세포는 이들 내인성 유전자 중 하나 이상을 과발현하도록 조작될 수도 있으며, 변형된 세포가 적어도 일부 조건 하에서 네이티브 세포보다 더 높은 수준으로 내인성 유전자를 발현한다는 것을 의미한다. 이들 중 특정 구체예에서, 발현되는 내인성 유전자는 하나 이상의 외인성 조절 인자에 작동가능하게 결합될 수도 있다. 예를 들어, 하나 이상의 네이티브 또는 비-네이티브 외인성 강한 프로모터는 그것들이 하나 이상의 내인성 3-HP 경로 유전자에 작동 가능하게 결합되는 세포 내로 도입될 수도 있다.

[0044] 유전자 여기에 제공된 변형된 효모 세포에서 3-HP 발효 경로는 프로모터 또는 종결자와 같은 하나 이상의 조절 인자에 작동 가능하게 결합될 수도 있다. 여기에 사용된 바와 같이, 용어 "프로모터"는 유전자의 전사의 시작을 조절하는, 유전자의 번역 개시 코돈 (일반적으로 약 1 내지 1000개의 염기쌍 (bp) 이내, 바람직하게 약 1 내지 500 bp 이내)에 대하여 업스트림 (즉, 5')에 위치한 번역되지 않은 서열을 나타낸다. 용어 "종결자"는 여기에 사용된 바와 같이 유전자의 전사의 끝을 조절하는, 유전자의 번역 종결 코돈 (일반적으로 약 1 내지 1000 bp 이내, 바람직하게 약 1 내지 500 bp 이내, 및 특히 약 1 내지 100 bp 이내)에 대하여 다운스트림 (즉, 3')에 위치한 번역되지 않은 서열을 나타낸다. 프로모터 또는 종결자는 유전자의 그것에 관하여 계놈에서 그것의 위치가 프로모터 또는 종결자가, 경우에 따라, 그것의 전사 조절 기능을 수행하는 위치인 경우 유전자와 "작동 가능하게 결합" 된다. 적합한 프로모터 및 종결자는, 예를 들어, W099/14335, W000/71738, W002/42471, W003/102201, W003/102152 및 W003/049525에 설명된다 (모두 전문이 본원에 참고로 포함됨).

[0045] 여기에 제공된 세포의 3-HP 발효 경로 유전자에 결합된 조절 요소는 내인성 또는 외인성일 수도 있다. 예를 들어, 외인성 3-HP 발효 경로 유전자는 내인성 프로모터 및/또는 종결자에 의해 조절되는 효모 세포 내에 삽입될 수도 있다. 대안으로, 외인성 3-HP 발효 경로 유전자는 하나 이상의 외인성 조절 요소에 결합될 수도 있다. 예를 들어, 외인성 유전자는 하나 이상의 외인성 조절 요소를 포함하는 유전자 발현 구조물의 일부로서 세포 내로 도입될 수도 있다. 특정 구체예에서, 외인성 조절 요소, 또는 외인성 조절 요소의 기능적인 부분은 네이티브 서열을 포함할 수도 있다. 다른 구체예에서, 외인성 조절 요소는 비-네이티브 서열을 포함할 수도 있다. 이들 구체예에서, 외인성 조절 요소는 네이티브 조절 요소에 대한 상대적으로 높은 정도의 서열 동일성을 갖는 서열을 포함할 수도 있다. 예를 들어, 외인성 유전자는 네이티브 프로모터 또는 종결자에 대한 적어도 50%, 적어도 60%, 적어도 70%, 적어도 80%, 또는 적어도 90% 서열 동일성을 갖는 외인성 프로모터 또는 종결자에 결합될 수도 있다. 뉴클레오타이드 또는 아미노산 서열에 대한 서열 동일성 퍼센트는 업계에 알려진 방법, 예를 들어, 기본 파라미터를 갖는 BLAST (미국 국립 생물 정보 센터 (National Center Biological Information; NCBI) 기본 지역 정렬 검색 도구) 버전 2.2.1 소프트웨어를 사용하여 계산될 수 있다. 예를 들어, 기본 파라미터를 갖는 BLAST 버전 2.2.1 알고리즘을 사용하여, 적어도 90%의 동일성 점수를 갖는 서열은 적어도 90% 서열 동일성을 갖는 것으로 생각된다. BLAST 소프트웨어는 NCBI, Bethesda, Maryland에서 이용 가능하다.

[0046] 특정 양태에서, 여기에 제공된 세포에서 3-HP 발효 경로 유전자에 결합된 조절 요소 (예를 들어, 프로모터)는 경로 유전자에 이질적일 수도 있다. 경로 유전자에 이질적인 조절 요소는 그것의 자연적인 형태로 유전자에 결

합되지 않는 조절 요소이다. 당업자는 경로 유전자에 이질적인 조절 요소가 내인성 또는 외인성일 수도 있다는 것을 인정할 수 있으며, 경로 유전자 및 효모 세포와 그것의 관계에 의존적이다. 일부 예에서, 내인성 3-HP 발효 경로 유전자는 경로 유전자에 이질적인 조절 요소 (예를 들어, 프로모터)에 작동 가능하게 결합된다. 다른 예에서, 외인성 3-HP 발효 경로 유전자는 경로 유전자에 이질적인 외인성 조절 요소 (예를 들어, 프로모터)에 작동 가능하게 결합된다.

[0047] 다수의 외인성 유전자가 숙주 세포 내에 삽입되는 이들 구체예에서, 각각의 외인성 유전자는 다른 조절 요소에 의해 조절될 수도 있거나, 둘 이상의 외인성 유전자는 같은 조절 요소에 의해 조절될 수도 있다. 예를 들어, 첫 번째 외인성 유전자가 첫 번째 조절 요소에 결합되는 경우, 두 번째 외인성 유전자는 또한 첫 번째 조절 요소에 결합될 수도 있거나, 그것은 두 번째 조절 요소에 결합될 수도 있다. 첫 번째 및 두 번째 조절 요소는 동일하거나 높은 정도의 서열 동일성을 공유할 수도 있거나, 그것들은 전혀 관련이 없을 수도 있다.

[0048] 여기에 제공된 효모 세포에서 하나 이상의 3-HP 발효 경로 유전자에 작동 가능하게 결합되는 프로모터의 예는 PDC1, 포스포글리세레이트 키나제 (PGK), 자일로스 리덕타제 (XR), 자일리톨 데히드로게나제 (XDH), L-(+)-락테이트-시토크롬 c 옥시도리덕타제 (CYB2), 번역 신장 인자-1 (TEF1), 번역 신장 인자-2 (TEF2), 에놀라제 (ENO1), 글리세르알데히드-3-포스페이트 데히드로게나제 (GAPDH), 및 오로티딘 5'-포스페이트 데카복실라제 (URA3) 유전자에 대한 프로모터를 포함하지만, 이에 제한되지 않는다.

[0049] 이들 예에서, 3-HP 발효 경로 유전자는 PDC1, PGK, XR, XDH, CYB2, TEF1, TEF2, ENO1, GAPDH, 또는 URA3 유전자에 대한 내인성 또는 외인성 프로모터에 결합될 수도 있다. 프로모터가 외인성인 경우, 그것들은 PDC1, PGK, XR, XDH, CYB2, TEF1, TEF2, ENO1, GAPDH, 또는 URA3 유전자에 대한 네이티브 프로모터와 동일하거나 높은 정도의 서열 동일성 (즉, 적어도 약 80%, 적어도 약 85%, 적어도 약 90%, 적어도 약 95%, 또는 적어도 약 99%)을 공유할 수도 있다.

[0050] 여기에 제공된 효모 세포에서 하나 이상의 3-HP 발효 경로 유전자에 결합될 수도 있는 종결자의 예는 PDC1, XR, XDH, 트랜스알돌라제 (TAL), 트랜스케톨라제 (TKL), 리보스 5-포스페이트 케톨-이소머라제 (RKI), CYB2, 또는 이소-2-시토크롬 c (CYC) 유전자 또는 유전자의 갈락토스 패밀리에 (특히 GAL10 종결자)에 대한 종결자를 포함하지만, 이에 제한되지 않는다. 이들 예에서, 3-HP 발효 경로 유전자는 PDC1, XR, XDH, TAL, TKL, RKI, CYB2, 또는 CYC 유전자 또는 갈락토스 패밀리에 유전자에 대한 내인성 또는 외인성 종결자에 결합될 수도 있다. 종결자가 외인성인 경우, 그것들은 PDC1, XR, XDH, TAL, TKL, RKI, CYB2, 또는 CYC 유전자 또는 갈락토스 패밀리에 유전자에 대한 네이티브 종결자와 동일하거나 높은 정도의 서열 동일성 (즉, 적어도 약 80%, 적어도 약 85%, 적어도 약 90%, 적어도 약 95%, 또는 적어도 약 99%)을 공유할 수도 있다. 특정 구체예에서, 3-HP 발효 경로 유전자는 숙주 세포에 네이티브인 네이티브 GAL10 유전자의 기능적 부분 또는 네이티브 GAL10 종결자와 적어도 80%, 적어도 85%, 적어도 90%, 또는 적어도 95% 서열 동일성을 공유하는 서열을 포함하는 종결자에 결합된다.

[0051] 외인성 유전자는 업계에 알려진 어떤 방법을 통해서도 효모 숙주 세포 내에 삽입될 수도 있다. 바람직한 구체예에서, 유전자는 숙주 세포 게놈으로 통합된다. 외인성 유전자는 표적화된 또는 무작위의 방식으로 게놈으로 통합될 수도 있다. 이들 구체예에서, 유전자가 표적화된 방식으로 통합될 수도 있는 경우, 그것은 외인성 유전자의 통합이 네이티브 유전자의 삭제 또는 붕괴와 결합되는, 특정 유전자에 대한 위치로 통합될 수도 있다. 예를 들어, 외인성 3-HP 경로 유전자의 도입은 다른 발효 생산 경로에 수반된 효소를 암호화하는 하나 이상의 유전자의 삭제 또는 붕괴와 결합될 수도 있다. 대안으로, 외인성 유전자는 유전자에 해당하지 않는 게놈의 일부로 통합될 수도 있다.

[0052] 표적화된 통합 및/또는 삭제는 통합 구조물을 이용할 수도 있다. 용어 "구조물"은 여기에 사용된 바와 같이 숙주 세포를 형질전환하기 위해 사용되는 DNA 서열을 나타낸다. 구조물은, 예를 들어, 원형 플라스미드 또는 벡터, 원형 플라스미드 또는 벡터의 일부 (예를 들어, 제한 효소 분해 생성물), 선형 플라스미드 또는 벡터, 또는 주형으로 플라스미드 또는 게놈 DNA를 사용하여 제조된 PCR 생성물일 수도 있다. 외인성 구조물을 갖는 효모 세포를 형질전환하는 방법은, 예를 들어, W099/14335, W000/71738, W002/42471, W003/102201, W003/102152, 및 W003/049525에 설명된다. 통합 구조물은 삽입 부위 표적의 두 개의 클로닝된 표적 DNA 서열을 사용하여 조립될 수 있다. 두 개의 표적 DNA 서열은 네이티브 숙주 게놈에서 인접하거나 비-인접할 수도 있다. 이 맥락에서, "비-인접"은 DNA 서열이 네이티브 게놈에서 서로 바로 인접하지 않지만, 대신 삭제되는 영역에 의해 분리된다. "인접한" 서열은 여기에 사용된 바와 같이 네이티브 게놈에서 서로 직접적으로 인접하다. 표적화된 통합이 표적 유전자의 삭제 또는 붕괴에 결합되는 경우, 통합 구조물은 또한 삭제 구조물로서 나타날 수도 있다. 삭제 구조물에서, 표적 서열 중 하나는 표적 유전자의 프로모터, 프로모터 영역의 모두 또는 일부, 표적 유전자의 모두 또는

는 일부를 암호화하는 서열, 또는 이들의 일부 조합에 대한 영역 5'을 포함할 수도 있다. 다른 표적 서열은 표적 유전자의 종결자, 종결 영역의 모두 또는 일부, 및/또는 표적 유전자의 모두 또는 일부를 암호화하는 서열에 대한 영역 3'을 포함할 수도 있다. 표적화된 통합이 네이티브 유전자의 삭제 또는 붕괴에 결합되지 않을 경우, 간섭 서열 (intervening sequence)의 삽입이 네이티브 유전자의 발현을 붕괴시키지 않는 표적 서열이 선택된다. 두 개의 표적 서열이 원래 숙주 세포의 게놈에 나타나기 때문에 서로 같은 방향인 통합 또는 삭제 구조물이 제조된다. 통합 또는 삭제 구조물이 외인성 유전자를 숙주 세포 내로 도입하기 위해 사용되는 경우, 유전자 발현 카세트는 외인성 유전자의 발현을 허용하는 두 개의 표적 유전자 서열 사이의 구조물로 클로닝된다. 유전자 발현 카세트는 외인성 유전자를 함유하고, 외인성 유전자에 작동 가능하게 결합된 프로모터 또는 종결자와 같은, 하나 이상의 조절 서열을 추가로 포함할 수도 있다. 유전자 발현 카세트를 함유하지 않는 삭제 구조물이 또한 구성될 수 있다. 이러한 구조물은 외인성 유전자의 삽입 없이 유전자 서열을 삭제 또는 붕괴하도록 설계된다.

[0053] 통합 또는 삭제 구조물은 두 개의 표적 유전자 서열 사이의 구조물로 클로닝된 하나 이상의 선택 마커 카세트를 포함할 수도 있다. 선택 마커 카세트는 형질전환체 (형질전환체)의 선택을 허용하는 적어도 하나의 선택 마커 유전자를 함유한다. "선택 마커 유전자"는 선택 배양 배지에서 형질 전환된 세포의 생존 및/또는 성장에 필요한 단백질을 암호화하는 유전자이고, 그러므로 세포에 대한 선택 압력을 적용하기 위해 사용될 수도 있다. 성공적인 형질전환체는 선택 마커 유전자를 함유할 것이며, 이것은 성공적인 형질전환된 세포에 선택을 위한 근거를 제공하는 적어도 하나의 특징을 전달한다. 전형적인 선택 마커 유전자는 (a) 항생제 또는 다른 독소에 대한 저항성 (예를 들어, 블레오마이신 또는 제오마이신 (예를 들어, 스트렙토알로테이쿠스 힌두스타누스 블레 (Streptoalloteichus hindustanus ble) 유전자), 아미노글리코사이드, 예를 들어, G418 또는 카나마이신 (예를 들어, 트랜스포손 Tn903의 카나마이신 저항 유전자), 또는 히그로마이신 (예를 들어, 이. 콜리의 아미노글리코사이드 항생제 저항 유전자)에 대한 저항성)을 부여하고, (b) 세포의 영양 요구성 (auxotroph) 결핍 (예를 들어, 루신 (예를 들어, 케이. 마르시아누스 (K. marxianus) LEU2 유전자), 우라실 (예를 들어, 케이. 마르시아누스, 에스. 세레비시애, 또는 아이. 오리엔탈리스 URA3 유전자), 또는 트립토판 (예를 들어, 케이. 마르시아누스, 에스. 세레비시애, 또는 아이. 오리엔탈리스 TRP 유전자)의 결핍)을 보충하고, 세포가 단순 배지에서 이용할 수 없는 임계 양분의 합성을 가능하게 하거나, (d) 세포의 특정 탄소 근원에서 성장하는 능력을 부여하는 (예를 들어, 에스. 세레비시애의 MEL5 유전자, 이것은 알파-갈락토시다제 (멜리비아제) 효소를 암호화하고 독특한 탄소 근원으로서 멜리비오제에서 성장하는 능력을 부여한다) 단백질을 암호화한다. 바람직한 선택 마커는 URA3 유전자, 제오신 저항 유전자, G418 저항 유전자, MEL5 유전자, 및 히그로마이신 저항 유전자를 포함한다. 또 다른 바람직한 선택 마커는 L-락테이트:페리시토크롬 c 옥시도리덕타제 (CYB2) 유전자 카세트이며, 숙주 세포는 원래 이러한 유전자가 부족하거나 그것의 네이티브 CYB2 유전자(들)이 먼저 삭제되거나 붕괴된다는 것을 제공한다. 선택 마커 유전자는 숙주 세포에서 작동 가능한 하나 이상의 프로모터 및/또는 종결 서열에 작동 가능하게 결합된다. 특정 구체예에서, 이들 프로모터 및/또는 종결 서열은 선택 마커 카세트에 포함되는 외인성 프로모터 및/또는 종결 서열이다. 적합한 프로모터 및 종결자는 여기에 설명된 바와 같다.

[0054] 통합 또는 삭제 구조물은 숙주 세포를 형질전환하기 위해 사용된다. 형질전환은, 예를 들어, 전기천공법 (electroporation) 및/또는 화학적 형질전환 (예를 들어, 염화 칼슘, 아세트산 리튬-기반, 등) 방법을 사용하여 달성될 수도 있다. 선택 마커의 존재 또는 부재에 기초한 선택 또는 스크리닝은 성공적인 형질전환체를 확인하기 위해 수행될 수도 있다. 성공적인 형질전환체에서, 표적 부위의 위치에서 상동 재조합 이벤트는 표적 부위 서열의 붕괴 또는 삭제를 일으킨다. 구조물이 삭제 또는 붕괴되는 네이티브 유전자를 표적화하는 경우, 네이티브 표적 유전자의 모두 또는 일부, 그것의 프로모터, 및/또는 그것의 종결자는 그것의 재조합 이벤트 중에 삭제될 수도 있다. 통합 구조물의 표적 서열 사이의 발현 카세트, 선택 마커 카세트, 및 어떤 다른 유전물질이 표적 서열과 일치하는 위치에서 숙주 세포 내에 삽입된다. PCR 또는 서던 (Southern) 분석법에 의한 분석은 원하는 삽입/삭제가 일어났는지 확인하기 위해서 수행될 수 있다.

[0055] 일부 구체예에서, 세포 형질전환은 단일 구조물 또는 PCR 생성물 대신, 둘 이상의 구조물, PCR 생성물, 또는 이들의 조합의 DNA를 사용하여 수행될 수도 있다. 이들 구체예에서, 한 통합 단편의 3' 말단은 또 다른 통합 단편의 5' 말단과 중복되지 않는다. 한 예에서, 한 구조물은 표적 서열의 위치 및 마커 유전자 카세트의 비-기능적 부분으로부터 첫 번째 서열을 함유하는 한편, 다른 것은 표적 서열의 위치의 두 번째 서열 및 마커 유전자 카세트의 두 번째 비-기능적 부분을 함유할 것이다. 완벽한 카세트를 형성하도록 결합될 수 있는 마커 유전자 카세트의 일부가 선택된다. 세포는 이들 조각으로 동시에 형질전환되며, 완벽한, 기능적 마커 또는 구조물적 유전자 카세트의 형성을 일으킨다. 성공적인 형질전환체는 선택 마커에 의해 전달된 특징에 기초하여 선택될 수 있다. 또 다른 예에서, 한 단편 상에 있지만 표적 서열에 없는 선택 마커 잔기는 원하는 부위에서의 높은 통합 확률을 갖도록, 별도의 단편 상에 있다. 다른 구체예에서, 세 개의 선형 DNA로부터 형질전환은 외인성 유전물질을 통합

하기 위해 사용될 수 있다. 이들 구체예에서, 한 단편은 두 번째 단편이 있는 5' 말단 상에서 및 세 번째 단편이 있는 3' 말단 상에서 중복된다.

[0056] 선택 마커 유전자 및 그것의 조절 요소의 일부 또는 모두가 이후의 상동 재조합 이벤트의 결과로서 자발적으로 삭제될 수 있는 통합 또는 삭제 구조물이 설계될 수도 있다. 이것을 달성하는 편리한 방법은 선택 마커 유전자 및/또는 조절 요소가 반복 서열에 의해 플랭크 (flank) 되는 구조물을 설계하는 것이다. 반복 서열은 동일한 DNA 서열이고, 숙주 세포에 네이티브 또는 비-네이티브이고, 서로 같은 또는 다른 방향에 있는 구조물의 방향이다. 반복 서열은 유리하게 길이가 약 50 내지 1500 bp이고, 어떤 것에 대해서도 암호화할 필요가 없다. 반복 서열의 내용물은 상동 재조합 이벤트가 발생하는 것을 허용하고, 이것은 선택 마커 유전자 및 반복 서열 중 하나의 삭제를 일으킨다. 상동 재조합이 상대적으로 낮은 빈도로 발생하기 때문에, 자발적 상동 재조합이 세포 중 일부에서 나타나는 것을 허용하기 위해 비-선택 배지에서 여러 라운드 동안 형질전환체를 성장시키는 것이 필요할 수도 있다. 선택 마커 유전자가 자발적으로 삭제되는 세포는 선택 마커 유전자에 의해 전달된 그것들의 선택 특징의 손실에 기초하여 선택되거나 스크리닝 될 수 있다. 특정 경우에, 리콤비나제 효소의 발현은 반복 부위 사이의 재조합을 향상시킬 수도 있다.

[0057] 여기에 제공된 변형된 효모 세포에서 외인성 3-HP 발효 경로 유전자는 어떤 적합한 근원의 근원 유전자로부터 유도될 수도 있다. 예를 들어, 외인성 유전자는 효모, 균류, 박테리아, 식물, 곤충, 또는 포유동물 근원으로부터 유도될 수도 있다. 여기에 사용된 바와 같이, 외인성 유전자는 네이티브 근원 유전자로부터 "유도된 (derived from)" 외인성 유전자는 1) 네이티브 유전자에 의해 암호화된 폴리펩티드와 동일하고, 2) 네이티브 유전자에 의해 암호화된 폴리펩티드와 적어도 50%, 적어도 60%, 적어도 70%, 적어도 80%, 적어도 85%, 적어도 90%, 적어도 95%, 적어도 97%, 또는 적어도 99% 서열 동일성을 공유하고, 및/또는 3-HP 발효 경로에서 네이티브 유전자에 의해 암호화된 폴리펩티드로서 같은 기능을 갖는 폴리펩티드를 암호화한다. 예를 들어, 아이. 오리엔탈리스 PYC 유전자로부터 유도된 PYC 유전자는 SEQ ID NO: 2의 아미노산 서열을 포함하는 폴리펩티드, SEQ ID NO: 2의 아미노산 서열에 대하여 적어도 50%, 적어도 60%, 적어도 70%, 적어도 80%, 적어도 85%, 적어도 90%, 적어도 95%, 적어도 97%, 또는 적어도 99% 서열 동일성을 갖는 폴리펩티드, 및/또는 피루베이트의 OAA로 전환을 촉진하는 능력을 가진 폴리펩티드를 암호화할 수도 있다. 네이티브 유전자로부터 유도된 유전자는 네이티브 유전자의 암호화 영역에 대하여 적어도 50%, 적어도 60%, 적어도 70%, 적어도 80%, 적어도 85%, 적어도 90%, 적어도 95%, 적어도 97%, 또는 적어도 99% 서열 동일성을 갖는 뉴클레오티드 서열을 포함할 수도 있다. 특정 구체예에서, 네이티브 유전자로부터 유도된 유전자는 근원 유전자의 암호화 영역과 동일한 뉴클레오티드 서열을 포함할 수도 있다. 예를 들어, 아이. 오리엔탈리스 PYC 유전자로부터 유도된 PYC 유전자는 SEQ ID NO: 1의 뉴클레오티드 서열 또는 SEQ ID NO: 1의 뉴클레오티드에 대하여 적어도 50%, 적어도 60%, 적어도 70%, 적어도 80%, 적어도 85%, 적어도 90%, 적어도 95%, 적어도 97%, 또는 적어도 99% 서열 동일성을 갖는 뉴클레오티드 서열을 포함할 수도 있다.

[0058] 여기에 제공된 변형된 효모 세포의 특정 구체예에서, 외인성 3-HP 발효 경로 유전자가 유발되는 네이티브 근원 유전자는 3-HP 발효 경로에 수반되는 폴리펩티드를 생산한다. 하지만, 다른 구체예에서, 네이티브 근원 유전자는 3-HP 발효 경로에 수반되지 않거나 3-HP 발효 경로에서 역반응을 촉진하는 폴리펩티드를 암호화할 수도 있다. 이들 구체예에서, 외인성 3-HP 경로 유전자는 변형된 활성 및/또는 기질 선호도를 발생시키는, 하나 이상의 표적화된 또는 무작위의 돌연변이 대 네이티브 근원 유전자를 겪었을 것이다. 예를 들어, 네이티브 근원 유전자는 3-HP 발효 경로에서 원하는 반응 방향의 증가된 활성 및/또는 원하는 방향의 감소된 활성으로 폴리펩티드를 암호화하는 유전자를 발생시키도록 돌연변이될 수도 있다. 예를 들어, 3-HP 발효 경로에서 네이티브 근원 유전자가 정반응 및 역반응을 촉진할 수 있는 폴리펩티드를 암호화하는 경우, 결과의 외인성 유전자가 정방향에서 증가된 활성 및 역방향에서 감소된 활성을 갖는 유전자는 변형될 수도 있다. 유사하게, 네이티브 근원 유전자는 네이티브 폴리펩티드와 다른 기질 선호도를 갖는 폴리펩티드를 암호화하는 유전자를 생산하도록 돌연변이될 수도 있다. 예를 들어, 3-HP 경로 유전자는 바람직하지 않거나 네이티브 폴리펩티드에 의해 전혀 작용하지 않는 기질에 대하여 작용하는 능력을 가진 폴리펩티드를 생산하도록 돌연변이 될 수도 있다. 이들 구체예에서, 외인성 3-HP 경로 유전자에 의해 암호화된 폴리펩티드는 네이티브 근원 유전자에 의해 암호화된 폴리펩티드가 완벽히 촉진할 수 없는 반응을 촉진할 수도 있다. 결과의 3-HP 경로 유전자가 하나 이상의 다운스트림 3-HP 경로 중간물 또는 부산물의 존재시 DNA, RNA, 또는 단백질 수준에서 감소된 피드백 (feedback) 역제를 나타내는 네이티브 근원 유전자가 또한 돌연변이될 수도 있다.

[0059] 여기에 제공된 변형된 효모 세포의 특정 구체예에서, 외인성 3-HP 경로 유전자는 숙주 효모 종으로부터 유도될 수도 있다. 예를 들어, 숙주 세포가 아이. 오리엔탈리스인 경우, 외인성 유전자는 아이. 오리엔탈리스 유전자로

부터 유도될 수도 있다. 이들 구체예에서, 외인성 유전자는 야생형 세포의 유전자의 발현을 구동하는 프로모터와 다른 프로모터에 의해 조절되는 경우 외인성 유전자의 숙주 세포로 통합이 네이티브 유전자 서열의 카피 수를 증가시키고 및/또는 유전자의 조절 또는 발현 수준을 변화시키는, 네이티브 유전자의 암호화 영역과 동일한 뉴클레오타이드 서열을 포함할 수도 있다. 다른 구체예에서, 외인성 3-HP 경로 유전자는 네이티브 3-HP 경로 유전자의 암호화 영역과 다른 뉴클레오타이드 서열을 포함할 수도 있지만, 그럼에도 불구하고, 네이티브 3-HP 경로 유전자에 의해 암호화되는 폴리펩티드와 동일한 폴리펩티드를 암호화한다. 다른 구체예에서, 외인성 3-HP 경로 유전자는 하나 이상의 네이티브 3-HP 경로 유전자에 의해 암호화된 폴리펩티드에 대하여 적어도 50%, 적어도 60%, 적어도 70%, 적어도 80%, 적어도 85%, 적어도 90%, 적어도 95%, 적어도 97%, 또는 적어도 99% 서열 동일성을 갖는 폴리펩티드를 암호화하는 뉴클레오타이드 서열을 포함할 수도 있다. 이들 중 특정 구체예에서, 외인성 유전자는 하나 이상의 네이티브 유전자에 대하여 적어도 50%, 적어도 60%, 적어도 70%, 적어도 80%, 적어도 85%, 적어도 90%, 적어도 95%, 적어도 97%, 또는 적어도 99% 서열 동일성을 갖는 뉴클레오타이드 서열을 포함한다. 다른 구체예에서, 외인성 3-HP 유전자는 네이티브 3-HP 경로 유전자에 의해 암호화된 폴리펩티드에 대하여 50%보다 작은 서열 동일성을 갖지만 그럼에도 불구하고, 3-HP 발효 경로의 네이티브 폴리펩티드와 같은 기능 (즉, 같은 반응을 촉진하는 능력)을 갖는 폴리펩티드를 암호화할 수도 있다. 네이티브 근원 유전자는 보통의 진행생물 개시 코돈 (ATG)으로 시작하는 암호화 서열을 제공하기 위해 필요한 경우, 또는 다른 목적으로 돌연변이 유발을 받을 수도 있다.

[0060] 다른 구체예에서, 외인성 3-HP 경로 유전자는 숙주 효모 세포의 그것과 다른 종으로부터 유도될 수도 있다. 이들 중 특정 구체예에서, 외인성 3-HP 경로 유전자는 숙주 세포와 다른 효모 종으로부터 유도될 수도 있다. 예를 들어, 숙주 세포가 아이. 오리엔탈리스인 경우, 외인성 유전자는 에스. 세레비시에로부터 유도될 수도 있다. 다른 구체예에서, 외인성 3-HP 경로 유전자는 균류, 박테리아, 식물, 곤충, 또는 포유동물 근원으로부터 유도될 수도 있다. 예를 들어, 숙주 세포가 아이. 오리엔탈리스인 경우, 외인성 유전자는 박테리아 근원, 예를 들어, 이. 콜리로부터 유도될 수도 있다. 이들 구체예에서, 외인성 3-HP 경로 유전자가 비-효모 근원으로부터 유도되는 경우, 외인성 유전자 서열은 효모 숙주 세포의 발현에 코돈-최적화될 수도 있다.

[0061] 이들 구체예에서, 외인성 3-HP 경로 유전자가 숙주 세포 중 이외의 종으로부터 유도되는 경우 외인성 유전자는 근원 유기체의 네이티브 3-HP 경로 유전자에 의해 암호화된 폴리펩티드와 동일한 폴리펩티드를 암호화할 수도 있다. 이들 중 특정 구체예에서, 외인성 3-HP 경로 유전자는 근원 유기체의 네이티브 3-HP 경로 유전자와 동일할 수도 있다. 다른 구체예에서, 외인성 유전자는 근원 유기체의 네이티브 3-HP 경로 유전자에 대하여 적어도 50%, 적어도 60%, 적어도 70%, 적어도 80%, 적어도 85%, 적어도 90%, 적어도 95%, 적어도 97%, 또는 적어도 99% 서열 동일성을 공유할 수도 있다. 다른 구체예에서, 외인성 3-HP 경로 유전자는 근원 유기체의 네이티브 3-HP 경로 유전자에 의해 암호화된 폴리펩티드와 적어도 50%, 적어도 60%, 적어도 70%, 적어도 80%, 적어도 85%, 적어도 90%, 적어도 95%, 적어도 97%, 또는 적어도 99% 서열 동일성을 공유하는 폴리펩티드를 암호화할 수도 있다. 이들 중 특정 구체예에서, 외인성 유전자는 근원 유기체의 하나 이상의 네이티브 3-HP 경로 유전자에 대하여 적어도 50%, 적어도 60%, 적어도 70%, 적어도 80%, 적어도 85%, 적어도 90%, 적어도 95%, 적어도 97%, 또는 적어도 99% 서열 동일성을 갖는 뉴클레오타이드 서열을 포함할 수도 있다. 다른 구체예에서, 외인성 3-HP 유전자는 근원 유기체의 네이티브 3-HP 경로 유전자에 의해 암호화된 폴리펩티드에 대하여 50%보다 작은 서열 동일성을 갖지만, 그럼에도 불구하고 3-HP 발효 경로에서 근원 유기체의 네이티브 폴리펩티드와 같은 기능을 갖는 폴리펩티드를 암호화할 수도 있다.

[0062] 특정 구체예에서, 여기에 제공된 효모 세포는 ACC (아세틸-CoA의 말로닐-CoA로 전환을 촉진), 알라닌 2,3 아미노무타제 (AAM, 알라닌의 β -알라닌으로 전환을 촉진), 알라닌 데히드로게나제 (피루베이트의 알라닌으로 전환을 촉진), 알데히드 데히드로게나제 (3-HPA의 3-HP로 전환을 촉진), KGD (OAA의 말로네이트 세미알데히드로 전환을 촉진), AAT (OAA의 아스파르테이트로 전환을 촉진), ADC (아스파르테이트의 β -알라닌으로 전환을 촉진), BCKA (OAA의 말로네이트 세미알데히드로 전환을 촉진), BAAT (β -알라닌의 말로네이트 세미알데히드로 전환을 촉진), 4-아미노부티레이트 아미노트랜스퍼라제 (gabT, β -알라닌의 말로네이트 세미알데히드로 전환을 촉진), β -알라닐-CoA 암모니아 리아제 (β -알라닐-CoA의 아크릴일-CoA로 전환을 촉진), Co-A 아실화 말로네이트 세미알데히드 데히드로게나제 (말로닐-CoA의 말로네이트 세미알데히드로 전환을 촉진), CoA 신테타제 (β -알라닌의 β -알라닐-CoA로 전환 또는 락테이트의 락틸-CoA로 전환을 촉진), CoA 트랜스퍼라제 (β -알라닌의 β -알라닐-CoA 및/또는 락테이트의 락틸-CoA로 전환을 촉진), 글리세롤 데히드라타제 (글리세롤의 3-HPA로 전환을 촉진), IPDA (OAA의 말로네이트 세미알데히드로 전환을 촉진), LDH (피루베이트의 락테이트로 전환을 촉진), 락틸-CoA 데히드라타제 (락틸-CoA의 아크릴일-CoA로 전환을 촉진), 말레이트 데카복실라제 (말레이트의 3-HP로 전환을 촉진), 말레이트 데히드로게나제 (OAA의 말레이트로 전환을 촉진), 말로닐-CoA 리덕타제 (말로닐-CoA의 말로네

이트 세미알데히드 또는 3-HP로 전환을 촉진), OAA 포르메이트리아제 (피루베이트-포르메이트 리아제 및 케토산 포르메이트-리아제로도 알려짐, OAA의 말로닐 CoA로 전환을 촉진), OAA 데히드로게나제 (OAA의 말로닐 CoA로 전환을 촉진); PPC (PEP의 OAA로 전환을 촉진), 피루베이트/알라닌 아미노트랜스퍼라제 (피루베이트의 알라닌으로 전환을 촉진), PYC (피루베이트의 OAA로 전환을 촉진), PDH (피루베이트의 아세틸-CoA로 전환을 촉진), 2-케토산 데카르복실라제 (OAA의 말로네이트 세미알데히드로 전환을 촉진), 3-HP-CoA 데히드라타제 (아크릴일-CoA 히드라타제로도 알려짐, 아크릴일-CoA의 3-HP-CoA로 전환을 촉진), 3-HPDH (말로네이트 세미알데히드의 3-HP로 전환을 촉진), 3-HP-CoA 히드롤라제 (3-HP-CoA의 3-HP로 전환을 촉진), HIBADH (말로네이트 세미알데히드의 3-HP로 전환을 촉진), 3-히드록시이소부티릴-CoA 히드롤라제 (3-HP-CoA의 3-HP로 전환을 촉진), 및 4-히드록시부티레이트 데히드로게나제 (말로네이트 세미알데히드의 3-HP로 전환을 촉진)로 구성된 그룹으로부터 선택되는 하나 이상의 3-HP 경로 효소를 암호화하는 유전자를 발현한다. 이들 효소 활성 각각에 대하여, 팔호의 원하는 반응은 네이티브 또는 비-네이티브 활성의 결과일 수도 있다.

[0063] "피루베이트 카르복실라제 유전자" 또는 "PYC 유전자"는 여기에 사용된 바와 같이 피루베이트 카르복실라제 활성을 갖는 폴리펩티드를 암호화하는 어떤 유전자도 나타내며, 피루베이트, CO₂, 및 ATP의 OAA, ADP, 및 포스페이트로 전환을 촉진하는 능력을 의미한다. 특정 구체예에서, PYC 유전자는 효모 근원으로부터 유도될 수도 있다. 예를 들어, PYC 유전자는 SEQ ID NO: 2에 설명된 아미노산 서열을 암호화하는 아이. 오리엔탈리스 PYC 유전자로부터 유도될 수도 있다. 다른 구체예에서, 유전자는 SEQ ID NO: 2의 아미노산 서열에 대하여 적어도 50%, 적어도 60%, 적어도 70%, 적어도 80%, 적어도 85%, 적어도 90%, 적어도 95%, 적어도 97%, 또는 적어도 99% 서열 동일성을 갖는 아미노산 서열을 암호화할 수도 있다. 특정 구체예에서, 아이. 오리엔탈리스-유도 PYC 유전자는 SEQ ID NO: 1에 설명된 뉴클레오티드 서열 또는 서열 SEQ ID NO: 1에 설명된 뉴클레오티드에 대하여 적어도 50%, 적어도 60%, 적어도 70%, 적어도 80%, 적어도 85%, 적어도 90%, 적어도 95%, 적어도 97%, 또는 적어도 99% 서열 동일성을 갖는 뉴클레오티드 서열을 포함할 수도 있다. 다른 구체예에서, PYC 유전자는 박테리아 근원으로부터 유도될 수도 있다. 예를 들어, PYC 유전자는 보전 (anaplerosis)을 위해 PYC만 사용하고 PPC는 사용하지 않는 (하기 참조) 소수의 박테리아 종 중 하나, 예를 들어, 알. 스페로이데스, 또는 PYC 및 PPC 둘 다를 소유하는 박테리아 종, 예를 들어, 알. 에틀리로부터 유도될 수도 있다. 알. 스페로이데스 및 알. 에틀리의 PYC 유전자에 의해 암호화된 아미노산 서열은 SEQ ID NO: 3 및 4에 각각 설명된다. PYC 유전자는 SEQ ID NO: 3 또는 4의 아미노산 서열을 암호화하는 유전자, 또는 SEQ ID NO: 3 또는 4의 아미노산 서열에 대하여 적어도 50%, 적어도 60%, 적어도 70%, 적어도 80%, 적어도 85%, 적어도 90%, 적어도 95%, 적어도 97%, 또는 적어도 99% 서열 동일성을 갖는 아미노산 서열을 암호화하는 유전자로부터 유도될 수도 있다. 대안으로, PYC 유전자는 활성화를 위해 아세틸-CoA에 대한 의존성을 갖지 않는 효소를 암호화하는 PYC 유전자, 예를 들어, SEQ ID NO: 5 (카르복시 트랜스퍼라제 서브유닛) 또는 SEQ ID NO: 6 (비오틴 카르복실라제 서브유닛)에 설명된 아미노산 서열을 암호화하는 피. 플루오레센스 PYC 유전자, SEQ ID NO: 7에 설명된 아미노산 서열을 암호화하는 씨. 글루타미쿰 PYC 유전자, 또는 SEQ ID NO: 5, 6, 또는 7의 아미노산 서열에 대하여 적어도 50%, 적어도 60%, 적어도 70%, 적어도 80%, 적어도 85%, 적어도 90%, 적어도 95%, 적어도 97%, 또는 적어도 99% 서열 동일성을 갖는 아미노산 서열을 암호화하는 유전자로부터 유도될 수도 있다. PYC 유전자는 또한 아스파르테이트에 의해 억제되지 않는 효소를 암호화하는 PYC 유전자, 예를 들어, SEQ ID NO: 8 (Sauer FEMS Microbiol Rev 29:765 (2005))에 설명된 아미노산 서열을 암호화하는 에스. 멜리로티 PYC 유전자, 또는 SEQ ID NO: 8의 아미노산 서열에 대하여 적어도 50%, 적어도 60%, 적어도 70%, 적어도 80%, 적어도 85%, 적어도 90%, 적어도 95%, 적어도 97%, 또는 적어도 99% 서열 동일성을 갖는 아미노산을 암호화하는 유전자로부터 유도될 수도 있다.

[0064] "PEP 카르복실라제 유전자" 또는 "PPC 유전자"는 여기에 사용된 바와 같이 PEP 카르복실라제 활성을 갖는 폴리펩티드를 암호화하는 어떤 유전자도 나타내며, PEP 및 CO₂의 OAA 및 포스페이트로 전환을 촉진하는 능력을 의미한다. 특정 구체예에서, PPC 유전자는 박테리아 PPC 유전자로부터 유도될 수도 있다. 예를 들어, PPC 유전자는 SEQ ID NO: 10에 설명된 아미노산 서열 또는 SEQ ID NO: 10의 아미노산 서열에 대하여 적어도 50%, 적어도 60%, 적어도 70%, 적어도 80%, 적어도 85%, 적어도 90%, 적어도 95%, 적어도 97%, 또는 적어도 99% 서열 동일성을 갖는 아미노산 서열을 암호화하는 이. 콜리 PPC 유전자로부터 유도될 수도 있다. 특정 구체예에서, 이. 콜리-유도 PPC 유전자는 SEQ ID NO: 9에 설명된 뉴클레오티드 서열 또는 SEQ ID NO: 9에 설명된 뉴클레오티드 서열에 대하여 적어도 50%, 적어도 60%, 적어도 70%, 적어도 80%, 적어도 85%, 적어도 90%, 적어도 95%, 적어도 97%, 또는 적어도 99% 서열 동일성을 갖는 뉴클레오티드 서열을 포함할 수도 있다. 다른 구체예에서, PPC 유전자는 "A" 타입 PPC로부터 유도될 수도 있으며, 많은 고세균류 (archaea) 및 제한된 수의 박테리아에서 발견되고, 이것은 아세틸 CoA에 의해 활성화되며 아스파르테이트에 의해 덜 억제된다. 예를 들어, PPC 유전자는 SEQ ID NO: 11에 설명된 아미노산 서열을 암호화하는 엠. 씨모아우트로피쿰 PPC A 유전자, 예를 들어, 서열

SEQ ID NO: 12에 설명된 아미노산 서열을 암호화하는 씨. 페르프링겐스 PPC A 유전자, 또는 SEQ ID NO: 11 또는 12의 아미노산 서열에 대하여 적어도 50%, 적어도 60%, 적어도 70%, 적어도 80%, 적어도 85%, 적어도 90%, 적어도 95%, 적어도 97%, 또는 적어도 99% 서열 동일성을 갖는 아미노산을 암호화하는 유전자로부터 유도될 수도 있다. 이들 중 특정 구체예에서, 유전자는 네이티브 유전자에 비해 개선된 특징을 갖는 효소를 발생시키기 위하여 하나 이상의 돌연변이를 겪을 수도 있다. 예를 들어, 유전자는 네이티브 폴리펩티드에 비해 아스파르테이트 피드백에 대한 증가된 저항성을 갖는 PPC 폴리펩티드를 암호화하도록 돌연변이 될 수도 있다. 다른 구체예에서, PPC 유전자는 식물 근원으로부터 유도될 수도 있다.

[0065] "아스파르테이트 아미노트랜스퍼라제 유전자" 또는 "AAT 유전자"는 여기에 사용된 바와 같이 아스파르테이트 아미노트랜스퍼라제 활성을 갖는 폴리펩티드를 암호화하는 어떤 유전자도 나타내며, OAA의 아스파르테이트로 전환을 촉진하는 능력을 의미한다. 아스파르테이트 아미노트랜스퍼라제 활성을 갖는 효소는 EC 2.6.1.1로서 분류된다. 특정 구체예에서, AT 유전자는 아이. 오리엔탈리스 또는 에스. 세레비시아와 같은 효모 근원으로부터 유도될 수도 있다. 예를 들어, AAT 유전자는 SEQ ID NO: 14에 설명된 아미노산 서열을 암호화하는 아이. 오리엔탈리스 AAT 유전자 또는 SEQ ID NO: 15에 설명된 아미노산 서열을 암호화하는 에스. 세레비시아 AAT2 유전자로부터 유도될 수도 있다. 다른 구체예에서, 유전자는 SEQ ID NO: 14 또는 15의 아미노산 서열에 대하여 적어도 50%, 적어도 60%, 적어도 70%, 적어도 80%, 적어도 85%, 적어도 90%, 적어도 95%, 적어도 97%, 또는 적어도 99% 서열 동일성을 갖는 아미노산 서열을 암호화할 수도 있다. 특정 구체예에서, 아이. 오리엔탈리스-유도 AAT 유전자는 SEQ ID NO: 13에 설명된 뉴클레오티드 서열 또는 SEQ ID NO: 13에 설명된 뉴클레오티드 서열에 대하여 적어도 50%, 적어도 60%, 적어도 70%, 적어도 80%, 적어도 85%, 적어도 90%, 적어도 95%, 적어도 97%, 또는 적어도 99% 서열 동일성을 갖는 뉴클레오티드 서열을 포함할 수도 있다. 다른 구체예에서, AT 유전자는 박테리아 근원으로부터 유도될 수도 있다. 예를 들어, AAT 유전자는 SEQ ID NO: 16에 설명된 아미노산 서열을 포함하는 폴리펩티드를 암호화하는 이. 콜리 aspC 유전자로부터 유도될 수도 있다. 다른 구체예에서, 유전자는 SEQ ID NO: 16의 아미노산 서열에 대하여 적어도 50%, 적어도 60%, 적어도 70%, 적어도 80%, 적어도 85%, 적어도 90%, 적어도 95%, 적어도 97%, 또는 적어도 99% 서열 동일성을 갖는 아미노산 서열을 암호화할 수도 있다.

[0066] "아스파르테이트 데카르복실라제 유전자" 또는 "ADC 유전자"는 여기에 사용된 바와 같이 아스파르테이트 데카르복실라제 활성을 갖는 폴리펩티드를 암호화하는 어떤 유전자도 나타내며, 아스파르테이트의 β -알라닌으로 전환을 촉진하는 능력을 의미한다. 아스파르테이트 데카르복실라제 활성을 가진 효소는 EC 4.1.1.11로서 분류된다. 특정 구체예에서, DC 유전자는 박테리아 근원으로부터 유도될 수도 있다. 활성 아스파르테이트 데카르복실라제는 불활성화 프로효소의 단백질 가수분해 과정이 필요할 수도 있기 때문에, 이들 구체예에서, 효모 숙주 세포는 박테리아 ADC 유전자에 의해 암호화된 활성 효소의 형성을 지지하도록 선택되어야 한다.

[0067] 일부 구체예에서, DC 유전자는 SEQ ID NO: 17에 설명된 아미노산 서열을 암호화하는 에스. 아베르미틸리스 panD 유전자로부터 유도될 수도 있다. 일부 구체예에서, DC 유전자는 SEQ ID NO: 17의 아미노산 서열에 대하여 적어도 50%, 적어도 60%, 적어도 70%, 적어도 80%, 적어도 85%, 적어도 90%, 적어도 95%, 적어도 97%, 또는 적어도 99% 서열 동일성을 갖는 아미노산 서열을 암호화할 수도 있다. 특정 구체예에서, 에스. 아베르미틸리스-유도 ADC 유전자는 SEQ ID NO: 130, 145, 146, 또는 147 중 어느 하나에 설명된 뉴클레오티드 서열; 또는 SEQ ID NO: 130, 145, 146, 또는 147 중 어느 하나에 설명된 뉴클레오티드 서열에 대하여 적어도 50%, 적어도 60%, 적어도 70%, 적어도 80%, 적어도 85%, 적어도 90%, 적어도 95%, 적어도 97%, 또는 적어도 99% 서열 동일성을 갖는 뉴클레오티드 서열을 포함할 수도 있다.

[0068] 다른 구체예에서, DC 유전자는 SEQ ID NO: 18에 설명된 아미노산 서열을 암호화하는 씨. 아세토부틸리쿰 panD 유전자로부터 유도될 수도 있다. 일부 구체예에서, DC 유전자는 SEQ ID NO: 18의 아미노산 서열에 대하여 적어도 50%, 적어도 60%, 적어도 70%, 적어도 80%, 적어도 85%, 적어도 90%, 적어도 95%, 적어도 97%, 또는 적어도 99% 서열 동일성을 갖는 아미노산 서열을 암호화할 수도 있다. 특정 구체예에서, 씨. 아세토부틸리쿰-유도 ADC 유전자는 SEQ ID NO: 131에 설명된 뉴클레오티드 서열, 또는 SEQ ID NO: 131에 설명된 뉴클레오티드 서열에 대하여 적어도 50%, 적어도 60%, 적어도 70%, 적어도 80%, 적어도 85%, 적어도 90%, 적어도 95%, 적어도 97%, 또는 적어도 99% 서열 동일성을 갖는 뉴클레오티드 서열을 포함할 수도 있다.

[0069] 다른 구체예에서, DC 유전자는 SEQ ID NO: 133에 설명된 아미노산 서열을 암호화하는 에이치. 필로리 ADC 유전자로부터 유도될 수도 있다. 일부 구체예에서, DC 유전자는 SEQ ID NO: 133의 아미노산 서열에 대하여 적어도 50%, 적어도 60%, 적어도 70%, 적어도 80%, 적어도 85%, 적어도 90%, 적어도 95%, 적어도 97%, 또는 적어도 99% 서열 동일성을 갖는 아미노산 서열을 암호화할 수도 있다. 특정 구체예에서, 에이치. 필로리-유도 ADC 유전자는 SEQ ID NO: 133에 설명된 뉴클레오티드 서열, 또는 SEQ ID NO: 133에 설명된 뉴클레오티드 서열에 대하여

적어도 50%, 적어도 60%, 적어도 70%, 적어도 80%, 적어도 85%, 적어도 90%, 적어도 95%, 적어도 97%, 또는 적어도 99% 서열 동일성을 갖는 뉴클레오티드 서열을 포함할 수도 있다.

[0070] 다른 구체예에서, DC 유전자는 SEQ ID NO: 135에 설명된 아미노산 서열을 암호화하는 바실루스 종 TS25 ADC 유전자로부터 유도될 수도 있다. 일부 구체예에서, DC 유전자는 SEQ ID NO: 135의 아미노산 서열에 대하여 적어도 50%, 적어도 60%, 적어도 70%, 적어도 80%, 적어도 85%, 적어도 90%, 적어도 95%, 적어도 97%, 또는 적어도 99% 서열 동일성을 갖는 아미노산 서열을 암호화할 수도 있다. 특정 구체예에서, 바실루스 종 TS25-유도 ADC 유전자는 SEQ ID NO: 134에 설명된 뉴클레오티드 서열, 또는 SEQ ID NO: 134에 설명된 뉴클레오티드 서열에 대하여 적어도 50%, 적어도 60%, 적어도 70%, 적어도 80%, 적어도 85%, 적어도 90%, 적어도 95%, 적어도 97%, 또는 적어도 99% 서열 동일성을 갖는 뉴클레오티드 서열을 포함할 수도 있다.

[0071] 다른 구체예에서, DC 유전자는 SEQ ID NO: 137에 설명된 아미노산 서열을 암호화하는 씨. 글루타미쿰 ADC 유전자로부터 유도될 수도 있다. 일부 구체예에서, DC 유전자는 SEQ ID NO: 137의 아미노산 서열에 대하여 적어도 50%, 적어도 60%, 적어도 70%, 적어도 80%, 적어도 85%, 적어도 90%, 적어도 95%, 적어도 97%, 또는 적어도 99% 서열 동일성을 갖는 아미노산 서열을 암호화할 수도 있다. 특정 구체예에서, 씨. 글루타미쿰-유도 ADC 유전자는 SEQ ID NO: 136에 설명된 뉴클레오티드 서열, 또는 SEQ ID NO: 136에 설명된 뉴클레오티드 서열에 대하여 적어도 50%, 적어도 60%, 적어도 70%, 적어도 80%, 적어도 85%, 적어도 90%, 적어도 95%, 적어도 97%, 또는 적어도 99% 서열 동일성을 갖는 뉴클레오티드 서열을 포함할 수도 있다.

[0072] 다른 구체예에서, DC 유전자는 SEQ ID NO: 139에 설명된 아미노산 서열을 암호화하는 비. 리체니포르미스 ADC 유전자로부터 유도될 수도 있다. 일부 구체예에서, DC 유전자는 SEQ ID NO: 139의 아미노산 서열에 대하여 적어도 50%, 적어도 60%, 적어도 70%, 적어도 80%, 적어도 85%, 적어도 90%, 적어도 95%, 적어도 97%, 또는 적어도 99% 서열 동일성을 갖는 아미노산 서열을 암호화할 수도 있다. 특정 구체예에서, 비. 리체니포르미스-유도 ADC 유전자는 SEQ ID NO: 138, 148, 149, 150, 또는 151 중 어느 하나에 설명된 뉴클레오티드 서열, 또는 SEQ ID NO: 138, 148, 149, 150, 또는 151 중 어느 하나에 설명된 뉴클레오티드 서열에 대하여 적어도 50%, 적어도 60%, 적어도 70%, 적어도 80%, 적어도 85%, 적어도 90%, 적어도 95%, 적어도 97%, 또는 적어도 99% 서열 동일성을 갖는 뉴클레오티드 서열을 포함할 수도 있다.

[0073] "β-알라닌 아미노트랜스퍼라제 유전자" 또는 "BAAT 유전자"는 여기에 사용된 바와 같이 β-알라닌 아미노트랜스퍼라제 활성을 갖는 폴리펩티드를 암호화하는 어느 유전자도 나타내며, β-알라닌의 말로네이트 세미알데히드로 전환을 촉진하는 능력을 의미한다. β-알라닌 아미노트랜스퍼라제 활성을 가진 효소는 EC 2.6.1.19로서 분류된다. 특정 구체예에서, BAAT 유전자는 효모 근원으로부터 유도될 수도 있다. 예를 들어, BAAT 유전자는 SEQ ID NO: 20에 설명된 아미노산 서열을 암호화하는 pyd4 유전자에 대한 아이. 오리엔탈리스 상동체로부터 유도될 수도 있다. 일부 구체예에서, BAAT 유전자는 SEQ ID NO: 20의 아미노산에 대하여 적어도 50%, 적어도 60%, 적어도 70%, 적어도 80%, 적어도 85%, 적어도 90%, 적어도 95%, 적어도 97%, 또는 적어도 99% 서열 동일성을 갖는 아미노산 서열을 암호화할 수도 있다. 특정 구체예에서, 아이. 오리엔탈리스-유도 BAAT 유전자는 SEQ ID NO: 19에 설명된 뉴클레오티드 서열 또는 SEQ ID NO: 19에 설명된 뉴클레오티드 서열에 대하여 적어도 50%, 적어도 60%, 적어도 70%, 적어도 80%, 적어도 85%, 적어도 90%, 적어도 95%, 적어도 97%, 또는 적어도 99% 서열 동일성을 갖는 뉴클레오티드 서열을 포함할 수도 있다. 다른 구체예에서, BAAT 유전자는 SEQ ID NO: 21에 설명된 아미노산 서열을 암호화하는 에스. 클루이베리 pyd4 유전자로부터 유도될 수도 있다. 일부 구체예에서, BAAT 유전자는 SEQ ID NO: 21의 아미노산 서열에 대하여 적어도 50%, 적어도 60%, 적어도 70%, 적어도 80%, 적어도 85%, 적어도 90%, 적어도 95%, 적어도 97%, 또는 적어도 99% 서열 동일성을 갖는 아미노산 서열을 암호화할 수도 있다. 특정 구체예에서, 에스. 클루이베리-유도 BAAT 유전자는 SEQ ID NO: 142에 설명된 뉴클레오티드 서열 또는 SEQ ID NO: 142에 설명된 뉴클레오티드 서열에 대하여 적어도 50%, 적어도 60%, 적어도 70%, 적어도 80%, 적어도 85%, 적어도 90%, 적어도 95%, 적어도 97%, 또는 적어도 99% 서열 동일성을 갖는 뉴클레오티드 서열을 포함할 수도 있다. 다른 구체예에서, BAAT 유전자는 박테리아 근원으로부터 유도될 수도 있다. 예를 들어, BAAT 유전자는 SEQ ID NO: 22에 설명된 아미노산 서열을 암호화하는 에스. 아베르미틸리스 BAAT 유전자로부터 유도될 수도 있다. 일부 구체예에서, BAAT 유전자는 SEQ ID NO: 22의 아미노산 서열에 대하여 적어도 50%, 적어도 60%, 적어도 70%, 적어도 80%, 적어도 85%, 적어도 90%, 적어도 95%, 적어도 97%, 또는 적어도 99% 서열 동일성을 갖는 아미노산 서열을 암호화할 수도 있다. 특정 구체예에서, 에스. 아베르미틸리스-유도 BAAT 유전자는 서열 SEQ ID NO: 140에 설명된 뉴클레오티드 또는 SEQ ID NO: 140에 설명된 뉴클레오티드 서열에 대하여 적어도 50%, 적어도 60%, 적어도 70%, 적어도 80%, 적어도 85%, 적어도 90%, 적어도 95%, 적어도 97%, 또는 적어도 99% 서열 동일성을 갖는 뉴클레오티드 서열을 포함할 수도 있다.

- [0074] BAAT 유전자는 또한 "4-아미노부티레이트 아미노트랜스퍼라제" 또는 "gabT 유전자"일 수도 있고, 그것이 4-아미노부티레이트 뿐만 아니라 β -알라닌에 대한 네이티브 활성을 갖는다는 것을 의미한다. 대안으로, BAAT 유전자는 박테리아 또는 효모 근원으로부터 무작위로 유도될 수도 있거나 BAAT 활성을 갖는 폴리펩티드를 암호화하도록 박테리아 또는 효모 근원의 네이티브 gabT 유전자를 조작하도록 지시될 수도 있다. 예를 들어, BAAT 유전자는 SEQ ID NO: 23에 설명된 아미노산 서열을 암호화하는 에스. 아베르미틸리스 gabT로부터 유도될 수도 있다. 일부 구체예에서, 에스. 아베르미틸리스-유도 BAAT 유전자는 SEQ ID NO: 23의 아미노산 서열에 대하여 적어도 50%, 적어도 60%, 적어도 70%, 적어도 80%, 적어도 85%, 적어도 90%, 적어도 95%, 적어도 97%, 또는 적어도 99% 서열 동일성을 갖는 아미노산 서열을 암호화할 수도 있다. 다른 구체예에서, BAAT 유전자는 SEQ ID NO: 24에 설명된 아미노산 서열을 암호화하는 에스. 세레비시아 gabT 유전자 UGA1로부터 유도될 수도 있다. 일부 구체예에서, 에스. 세레비시아-유도 BAAT 유전자는 SEQ ID NO: 24의 아미노산 서열에 대하여 적어도 50%, 적어도 60%, 적어도 70%, 적어도 80%, 적어도 85%, 적어도 90%, 적어도 95%, 적어도 97%, 또는 적어도 99% 서열 동일성을 갖는 아미노산 서열을 암호화할 수도 있다. 특정 구체예에서, 에스. 세레비시아-유도 BAAT 유전자는 SEQ ID NO: 141에 설명된 뉴클레오티드 서열 또는 서열 SEQ ID NO: 141에 설명된 뉴클레오티드에 대하여 적어도 50%, 적어도 60%, 적어도 70%, 적어도 80%, 적어도 85%, 적어도 90%, 적어도 95%, 적어도 97%, 또는 적어도 99% 서열 동일성을 갖는 뉴클레오티드 서열을 포함할 수도 있다.
- [0075] "3-HP 테히드로게나제 유전자" 또는 "3-HPDH 유전자"는 여기에 사용된 바와 같이 3-HP 테히드로게나제 활성을 암호화하는 어떤 유전자도 나타내며, 말로네이트 세미알데히드의 3-HP로 전환을 촉진하는 능력을 의미한다. 3-HP 테히드로게나제 활성을 가진 효소는 그것들이 NAD(H) 공동인자를 이용하면 EC 1.1.1.59로서, 및 그것들이 NADP(H) 공동인자를 이용하면 EC 1.1.1.298로서 분류된다. EC 1.1.1.298로서 분류된 효소는 대안으로, 말로네이트 세미알데히드 리덕타제로 나타난다.
- [0076] 특정 구체예에서, 3-HPDH 유전자는 효모 근원으로부터 유도될 수도 있다. 예를 들어, 3-HPDH 유전자는 SEQ ID NO: 26에 설명된 아미노산 서열을 암호화하는 YMR226C 유전자에 대한 아이. 오리엔탈리스 상동체로부터 유도될 수도 있다. 일부 구체예에서, 3-HPDH 유전자는 SEQ ID NO: 26의 아미노산 서열에 대하여 적어도 50%, 적어도 60%, 적어도 70%, 적어도 80%, 적어도 85%, 적어도 90%, 적어도 95%, 적어도 97%, 또는 적어도 99% 서열 동일성을 갖는 아미노산 서열을 암호화할 수도 있다. 특정 구체예에서, 아이. 오리엔탈리스-유도 3-HPDH 유전자는 SEQ ID NO: 25에 설명된 뉴클레오티드 서열 또는 SEQ ID NO: 25에 설명된 뉴클레오티드 서열에 대하여 적어도 50%, 적어도 60%, 적어도 70%, 적어도 80%, 적어도 85%, 적어도 90%, 적어도 95%, 적어도 97%, 또는 적어도 99% 서열 동일성을 갖는 뉴클레오티드 서열을 포함할 수도 있다. 다른 구체예에서, 3-HPDH 유전자는 SEQ ID NO: 129에 설명된 아미노산 서열을 암호화하는 에스. 세레비시아 YMR226C 유전자로부터 유도될 수도 있다. 일부 구체예에서, 3-HPDH 유전자는 SEQ ID NO: 129의 아미노산 서열에 대하여 적어도 50%, 적어도 60%, 적어도 70%, 적어도 80%, 적어도 85%, 적어도 90%, 적어도 95%, 적어도 97%, 또는 적어도 99% 서열 동일성을 갖는 아미노산 서열을 암호화할 수도 있다. 특정 구체예에서, 에스. 세레비시아-유도 3-HPDH 유전자는 SEQ ID NO: 144에 설명된 뉴클레오티드 서열 또는 SEQ ID NO: 144에 설명된 뉴클레오티드 서열에 대하여 적어도 50%, 적어도 60%, 적어도 70%, 적어도 80%, 적어도 85%, 적어도 90%, 적어도 95%, 적어도 97%, 또는 적어도 99% 서열 동일성을 갖는 뉴클레오티드 서열을 포함할 수도 있다.
- [0077] 다른 구체예에서, 3-HPDH 유전자는 박테리아 근원으로부터 유도될 수도 있다. 예를 들어, 3-HPDH 유전자는 SEQ ID NO: 27의 아미노산 서열을 암호화하는 이. 콜리 ydfG 유전자로부터 유도될 수도 있다. 일부 구체예에서, 유전자는 SEQ ID NO: 27의 아미노산 서열에 대하여 적어도 50%, 적어도 60%, 적어도 70%, 적어도 80%, 적어도 85%, 적어도 90%, 적어도 95%, 적어도 97%, 또는 적어도 99% 서열 동일성을 갖는 아미노산 서열을 암호화할 수도 있다. 특정 구체예에서, 이. 콜리-유도 3-HPDH 유전자는 SEQ ID NO: 143에 설명된 뉴클레오티드 서열 또는 SEQ ID NO: 143에 설명된 뉴클레오티드 서열에 대하여 적어도 50%, 적어도 60%, 적어도 70%, 적어도 80%, 적어도 85%, 적어도 90%, 적어도 95%, 적어도 97%, 또는 적어도 99% 서열 동일성을 갖는 뉴클레오티드 서열을 포함할 수도 있다. 다른 구체예에서, 3-HPDH 유전자는 SEQ ID NO: 29에 설명된 아미노산 서열을 암호화하는 엠. 세둘라 말로네이트 세미알데히드 리덕타제 유전자로부터 유도될 수도 있다. 일부 구체예에서, 3-HPDH 유전자는 SEQ ID NO: 29에 설명된 아미노산 서열에 대하여 적어도 50%, 적어도 60%, 적어도 70%, 적어도 80%, 적어도 85%, 적어도 90%, 적어도 95%, 적어도 97%, 또는 적어도 99% 서열 동일성을 갖는 아미노산 서열을 암호화할 수도 있다. 특정 구체예에서, 엠. 세둘라-유도 3-HPDH 유전자는 SEQ ID NO: 343에 설명된 뉴클레오티드 서열 또는 SEQ ID NO: 343에 설명된 뉴클레오티드 서열에 대하여 적어도 50%, 적어도 60%, 적어도 70%, 적어도 80%, 적어도 85%, 적어도 90%, 적어도 95%, 적어도 97%, 또는 적어도 99% 서열 동일성을 갖는 뉴클레오티드 서열을 포함

할 수도 있다.

[0078] "3-히드록시이소부티레이트 데히드로게나제 유전자" 또는 "HIBADH 유전자"는 여기에 사용된 바와 같이 3-히드록시이소부티레이트 데히드로게나제 활성을 갖는 폴리펩티드를 암호화하는 어떤 유전자도 나타내며, 3-히드록시이소부티레이트의 메틸말로네이트 세미알데히드로 전환을 촉진하는 능력을 의미한다. 3-히드록시이소부티레이트 데히드로게나제 활성을 가진 효소는 EC 1.1.1.31로서 분류된다. 일부 3-히드록시이소부티레이트 데히드로게나제는 또한 3-HPDH 활성을 갖는다. 특정 구체예에서, HIBADH 유전자는 박테리아 근원으로부터 유도될 수도 있다. 예를 들어, HIBADH 유전자는 SEQ ID NO: 28에 설명된 아미노산 서열을 암호화하는 에이. 패칼리스 M3A 유전자, SEQ ID NO: 30 또는 SEQ ID NO: 31에 설명된 아미노산 서열을 각각 암호화하는 피. 푸티다 KT2440 또는 E23440 mmsB 유전자, 또는 SEQ ID NO: 32에 설명된 아미노산 서열을 암호화하는 피. 애루기노사 PA01 mmsB 유전자로부터 유도될 수도 있다. 특정 구체예에서, HIBADH 유전자는 SEQ ID NO: 28, 30, 31, 또는 32에 설명된 아미노산 서열에 대하여 적어도 50%, 적어도 60%, 적어도 70%, 적어도 80%, 적어도 85%, 적어도 90%, 적어도 95%, 적어도 97%, 또는 적어도 99% 서열 동일성을 갖는 아미노산 서열을 암호화한다.

[0079] "4-히드록시부티레이트 데히드로게나제 유전자"는 여기에 사용된 바와 같이 4-히드록시부티레이트 데히드로게나제 활성을 갖는 폴리펩티드를 암호화하는 어떤 유전자도 나타내며, 4-히드록시부타노에이트의 숙시네이트 세미알데히드로 전환을 촉진하는 능력을 의미한다. 4-히드록시부티레이트 데히드로게나제 활성을 갖는 효소는 EC 1.1.1.61로서 분류된다. 일부 4-히드록시부티레이트 데히드로게나제는 또한 3-HPDH 활성을 갖는다. 특정 구체예에서, 4-히드록시부티레이트 데히드로게나제 유전자는 박테리아 근원으로부터 유도될 수도 있다. 예를 들어, 4-히드록시부티레이트 데히드로게나제 유전자는 SEQ ID NO: 33에 설명된 아미노산 서열을 암호화하는 알. 유트로파 H16 4hbd 유전자 또는 SEQ ID NO: 34에 설명된 아미노산 서열을 암호화하는 씨. 클루이베리 DSM 555 hbd 유전자로부터 유도될 수도 있다. 다른 구체예에서, 유전자는 SEQ ID NO: 33 또는 34에 설명된 아미노산 서열에 대하여 적어도 50%, 적어도 60%, 적어도 70%, 적어도 80%, 적어도 85%, 적어도 90%, 적어도 95%, 적어도 97%, 또는 적어도 99% 서열 동일성을 갖는 아미노산 서열을 암호화할 수도 있다.

[0080] "PEP 카르복시키나제 유전자" 또는 "PCK 유전자"는 여기에 사용된 바와 같이 PEP 카르복시키나제 활성을 갖는 폴리펩티드를 암호화하는 어떤 유전자도 나타내며, PEP, CO₂, 및 ADP 또는 GDP의 OAA 및 ATP 또는 GTP로, 또는 그 반대의 전환을 촉진하는 능력을 의미한다. PEP 카르복시키나제 활성을 가진 효소는 EC 4.1.1.32로서 (GTP/GDP 이용) 및 EC 4.1.1.49 (ATP/ADP 이용)로서 분류된다. 특정 구체예에서, PCK 유전자는 효모 근원으로부터 유도될 수도 있다. 다른 구체예에서, PCK 유전자는 박테리아 근원으로부터 유도될 수도 있고, 이들 중 특정 구체예에서, 유전자는 데카르복실화가 우세한 경우 PCK 반응이 반응의 더 일반적인 형태 대신에 OAA의 생산을 선호하는 박테리아로부터 유도될 수도 있다. 예를 들어, PCK 유전자는 SEQ ID NO: 35에 설명된 아미노산 서열을 암호화하는 엠. 숙시니시프로듀센스 (*M. succiniciproducens*) PCK 유전자, SEQ ID NO: 36에 설명된 아미노산 서열을 암호화하는 에이. 숙시니시프로듀센스 (*A. succiniciproducens*) PCK 유전자, SEQ ID NO: 37에 설명된 아미노산 서열을 암호화하는 알. 숙시노게네스 (*A. succinogenes*) PCK 유전자, 또는 SEQ ID NO: 38에 설명된 아미노산 서열을 암호화하는 알. 유트로파 PCK 유전자로부터 유도될 수도 있다. 다른 구체예에서, PCK 유전자는 유도되는 네이티브 유전자에 비해, 결과의 유전자가 바람직하게 PEP의 OAA로 전환을 촉진하는 폴리펩티드를 암호화하는 하나 이상의 돌연변이를 겪었다. 예를 들어, PCK 유전자는 SEQ ID NO: 39에 설명된 아미노산 서열을 암호화하는 이. 콜리 K12 균주 PCK 유전자로부터 유도될 수도 있으며, 유전자는 바람직하게 PEP의 OAA로 전환을 촉진하도록 돌연변이되었다. 다른 구체예에서, PEP의 OAA로 전환은, 예를 들어, ATP/ADP 또는 GTP/GDP 대신에 무기 포스페이트 및 디포스페이트를 사용하는 프로피온산 박테리아 (예를 들어, 피. 셰르마니이 (*P. shermanii*)), 에이. 우디이 (*A. woodii*))에서 발견되는 PEP 카르복시트랜스포스포릴라제에 의해 촉진된다.

[0081] "말레이트 데히드로게나제 유전자"는 여기에 사용된 바와 같이 말레이트 데히드로게나제 활성을 갖는 폴리펩티드를 암호화하는 어떤 유전자도 나타내며, OAA의 말레이트로 전환을 촉진하는 능력을 의미한다. 특정 구체예에서, 말레이트 데히드로게나제 유전자는 박테리아 또는 효모 근원으로부터 유도될 수도 있다.

[0082] "말레이트 데카르복실라제 유전자"는 여기에 사용된 바와 같이 말레이트 데카르복실라제 활성을 갖는 폴리펩티드를 암호화하는 어떤 유전자도 나타내며, 말레이트의 3-HP로 전환을 촉진하는 능력을 의미한다. 말레이트 데카르복실라제 활성은 자연 발생하는 것으로 알려져 있지 않다. 그러므로, 말레이트 데카르복실라제 유전자는 하나 이상의 돌연 변이를 아세트락테이트 데카르복실라제 활성을 갖는 폴리펩티드를 암호화하는 네이티브 근원 유전자로 통합함으로써 유도될 수도 있다. 아세트락테이트 데카르복실라제 활성을 갖는 폴리펩티드는 2-히드록시-2-메틸-3-옥소부타노에이트의 2-아세트인으로 전환을 촉진하고, EC 4.1.1.5로서 분류된다. 특정 구체예에서, 말레이트 데카르복실라제 유전자는 박테리아 근원으로부터 유도될 수도 있다. 예를 들어, 말레이트 데카르복실라제

유전자는 SEQ ID NO: 40에 설명된 아미노산 서열을 암호화하는 엘. 락티스 (*L. lactis*) aldB 유전자, SEQ ID NO: 41에 설명된 아미노산 서열을 암호화하는 에스. 썬모필루스 (*S. thermophilus*) aldB 유전자, SEQ ID NO: 42에 설명된 아미노산 서열을 암호화하는 비. 브레비스 (*B. brevis*) aldB 유전자, 또는 SEQ ID NO: 43에 설명된 아미노산 서열을 암호화하는 이. 애로게네스 (*E. aerogenes*) budA 유전자로부터 유도될 수도 있다.

[0083] "알파-케토글루타레이트 (AKG) 데카르복실라제 유전자" 또는 "KGD 유전자"는 여기에 사용된 바와 같이 알파-케토글루타레이트 데카르복실라제 활성을 갖는 폴리펩티드를 암호화하는 어떤 유전자도 나타내며, 알파-케토글루타레이트 (2-옥소글루타레이트)의 숙시네이트 세미알데히드 전환을 촉진하는 능력을 의미한다. AKG 데카르복실라제 활성을 가진 효소는 EC 4.1.1.71로서 분류된다. KGD 유전자는 OAA의 말로네이트 세미알데히드로 전환을 촉진할 수 있는 폴리펩티드를 암호화하는 유전자를 유도하기 위해 사용될 수도 있다. 이 활성은 네이티브 KGD 유전자에서 발견될 수도 있거나, 그것은 하나 이상의 돌연변이를 네이티브 KGD 유전자로 통합함으로써 유도될 수도 있다. 특정 구체예에서, KGD 유전자는 박테리아 근원으로부터 유도될 수도 있다. 예를 들어, KGD 유전자는 SEQ ID NO: 44에 설명된 아미노산 서열을 암호화하는 엠. 튜버쿨로시스 (*M. tuberculosis*) KGD 유전자, SEQ ID NO: 45에 설명된 아미노산 서열을 암호화하는 비. 자포니쿰 (*B. japonicum*) KGD 유전자, 또는 SEQ ID NO: 46에 설명된 아미노산 서열을 암호화하는 엠. 로티 (*M. loti*; 아카 리조비움 로티 (aka *Rhizobium loti*)) KGD 유전자로부터 유도될 수도 있다.

[0084] "분지형-사슬 알파-케토산 데카르복실라제 유전자" 또는 "BCKA 유전자"는 여기에 사용된 바와 같이 분지형-사슬 알파-케토산 데카르복실라제 활성을 갖는 폴리펩티드를 암호화하는 어떤 유전자도 나타내며, 이것은 데카르복실레이트에 길이가 3 내지 6개의 탄소의 다양한 알파-케토산을 제공할 수 있다. BCKA 활성을 가진 효소는 EC 4.1.1.72로서 분류된다. BCKA 유전자는 OAA의 말로네이트 세미알데히드로 전환을 촉진할 수 있는 폴리펩티드를 암호화하는 유전자를 유도하기 위해 사용될 수도 있다. 이 활성은 네이티브 BCKA 유전자에서 발견될 수도 있거나, 그것은 하나 이상의 돌연변이를 네이티브 BCKA 유전자로 통합함으로써 유도될 수도 있다. 특정 구체예에서, BCKA 유전자는 박테리아 근원으로부터 유도될 수도 있다. 예를 들어, BCKA 유전자는 SEQ ID NO: 47에 설명된 아미노산 서열을 암호화하는 엘. 락티스 *kdcA* 유전자로부터 유도될 수도 있다.

[0085] "인돌피루베이트 데카르복실라제 유전자" 또는 "IPDA 유전자"는 여기에 사용된 바와 같이 인돌피루베이트 데카르복실라제 활성을 갖는 폴리펩티드를 암호화하는 어떤 유전자도 나타내며, 인돌피루베이트의 인돌아세트알데히드로 전환을 촉진하는 능력을 의미한다. IPDA 활성을 가진 효소는 EC 4.1.1.74로서 분류된다. IPDA 유전자는 OAA의 말로네이트 세미알데히드로 전환을 촉진할 수 있는 폴리펩티드를 암호화하는 유전자를 유도하기 위해 사용될 수도 있다. 이 활성은 네이티브 IPDA 유전자에서 발견될 수도 있거나, 그것은 하나 이상의 돌연변이를 네이티브 IPDA 유전자로 통합함으로써 유도될 수도 있다. 특정 구체예에서, 인돌피루베이트 데카르복실라제 유전자는 효모, 박테리아, 또는 식물 근원으로부터 유도될 수도 있다.

[0086] "피루베이트 데카르복실라제 유전자" 또는 "PDC 유전자"는 여기에 사용된 바와 같이 피루베이트 데카르복실라제 활성을 갖는 폴리펩티드를 암호화하는 어떤 유전자도 나타내며, 피루베이트의 아세트알데히드로 전환을 촉진하는 능력을 의미한다. PDC 활성을 가진 효소는 EC 4.1.1.1로서 분류된다. 바람직한 구체예에서, 여기에 제공된 바와 같이 변형된 효모 세포로 통합되는 PDC 유전자는 네이티브 유전자에 비해 결과의 유전자가 OAA의 말로네이트 세미알데히드로 전환을 촉진할 수 있는 폴리펩티드를 암호화하는 것으로부터 유도된 하나 이상의 돌연변이를 겪었다. 특정 구체예에서, PDC 유전자는 효모 근원으로부터 유도될 수도 있다. 예를 들어, PDC 유전자는 SEQ ID NO: 49에 설명된 아미노산 서열을 암호화하는 이. 오리엔탈리스 PDC 유전자, SEQ ID NO: 50에 설명된 아미노산 서열을 암호화하는 스. 세레비시애 PDC1 유전자, 또는 SEQ ID NO: 51에 설명된 아미노산 서열을 암호화하는 케이. 락티스 (*K. lactis*) PDC로부터 유도될 수도 있다. 특정 구체예에서, PDC 유전자는 SEQ ID NO: 48에 설명된 뉴클레오티드 서열 또는 SEQ ID NO: 48에 설명된 뉴클레오티드 서열에 대하여 적어도 50%, 적어도 60%, 적어도 70%, 적어도 80%, 적어도 85%, 적어도 90%, 적어도 95%, 적어도 97%, 또는 적어도 99% 서열 동일성을 갖는 뉴클레오티드 서열을 포함할 수도 있는 아이. 오리엔탈리스 PDC 유전자로부터 유도된다. 다른 구체예에서, PDC 유전자는 박테리아 근원으로부터 유도될 수도 있다. 예를 들어, PDC 유전자는 SEQ ID NO: 52에 설명된 아미노산 서열을 암호화하는 제트. 모빌리스 (*Z. mobilis*) PDC 유전자 또는 SEQ ID NO: 53에 설명된 아미노산 서열을 암호화하는 에이. 파스투리아누스 (*A. pasteurianus*) PDC 유전자로부터 유도될 수도 있다.

[0087] "벤조일포르메이트 데카르복실라제" 유전자 여기에 사용된 바와 같이 벤조일포르메이트 데카르복실라제 활성을 갖는 폴리펩티드를 암호화하는 어떤 유전자도 나타내며, 벤조일포르메이트의 벤즈알데히드로 전환을 촉진하는 능력을 의미한다. 벤조일포르메이트 데카르복실라제 활성을 가진 효소는 EC 4.1.1.7로서 분류된다. 바람직한 구체예에서, 여기에 제공된 바와 같이 변형된 효모 세포로 통합되는 벤조일포르메이트 데카르복실라제 유전자는

네이티브 유전자에 비해 결과의 유전자가 OAA의 말로네이트 세미알데히드로 전환을 촉진할 수 있는 폴리펩티드를 암호화하는 것으로부터 유도된 하나 이상의 돌연변이를 겪었다. 특정 구체예에서, 벤조일포르메이트 데카르복실라제 유전자는 박테리아 근원으로부터 유도될 수도 있다. 예를 들어, 벤조일포르메이트 데카르복실라제 유전자는 SEQ ID NO: 54에 설명된 아미노산 서열을 암호화하는 피. 푸티다 mdIC 유전자, SEQ ID NO: 55에 설명된 아미노산 서열을 암호화하는 피. 애루기노사 mdIC 유전자, SEQ ID NO: 56에 설명된 아미노산 서열을 암호화하는 피. 스트루트제리 (*P. stutzeri*) dpgB 유전자, 또는 SEQ ID NO: 57에 설명된 아미노산 서열을 암호화하는 피. 플루오레센스 ilvB-1 유전자로부터 유도될 수도 있다.

[0088] "OAA 포르메이트리아제 유전자"는 여기에 사용된 바와 같이 OAA 포르메이트리아제 활성을 갖는 폴리펩티드를 암호화하는 어떤 유전자도 나타내며, 아세틸레이트 케토산의 그것의 해당 CoA 유도체로 전환을 촉진하는 능력을 의미한다. OAA 포르메이트리아제 유전자에 의해 암호화된 폴리펩티드는 피루베이트 또는 또 다른 케토산에 대한 활성을 가질 수도 있다. 특정 구체예에서, OAA 포르메이트리아제 유전자는 OAA를 말로닐-CoA로 전환시키는 폴리펩티드를 암호화한다.

[0089] "말로닐-CoA 리덕타제 유전자"는 여기에 사용된 바와 같이 말로닐-CoA 리덕타제 활성을 갖는 폴리펩티드를 암호화하는 어떤 유전자도 나타내며, 말로닐-CoA의 말로네이트 세미알데히드로 전환을 촉진하는 능력을 의미한다 (Co-A 아실화 말로네이트 세미알데히드 데히드로게나제 활성으로도 나타남). 특정 구체예에서, 말로닐-CoA 리덕타제 유전자는 말로네이트 세미알데히드의 3-HP로 전환을 촉진하는 능력도 갖는 이기능성 말로닐-CoA 리덕타제 유전자로부터 유도될 수도 있다. 이들 중 특정 구체예에서, 말로닐-CoA 리덕타제 유전자는 박테리아 근원으로부터 유도될 수도 있다. 예를 들어, 말로닐-CoA 리덕타제 유전자는 SEQ ID NO: 58에 설명된 아미노산 서열을 암호화하는 씨. 아우란티아쿠스 (*C. aurantiacus*) 말로닐-CoA 리덕타제 유전자, SEQ ID NO: 59에 설명된 아미노산 서열을 암호화하는 알. 카스텐홀지이 (*R. castenholzii*) 말로닐-CoA 리덕타제 유전자, 또는 SEQ ID NO: 60에 설명된 아미노산 서열을 암호화하는 에리트로박터 종 NAP1 말로닐-CoA 리덕타제 유전자로부터 유도될 수도 있다. 다른 구체예에서, 말로닐-CoA 리덕타제 유전자는 말로닐-CoA의 말로네이트 세미알데히드로 전환을 촉진하는 폴리펩티드를 암호화하는 말로닐-CoA 리덕타제 유전자로부터 유도될 수도 있다. 예를 들어, 말로닐-CoA 리덕타제 유전자는 SEQ ID NO: 61에 설명된 아미노산 서열을 암호화하는 엠. 세둘라 Msed_0709 유전자 또는 SEQ ID NO: 62에 설명된 아미노산 서열을 암호화하는 에스. 토크다이어 (*S. tokodaii*) 말로닐-CoA 리덕타제로부터 유도될 수도 있다.

[0090] "피루베이트 데히드로게나제 유전자" 또는 "PDH 유전자"는 여기에 사용된 바와 같이 피루베이트 데히드로게나제 활성을 갖는 폴리펩티드를 암호화하는 어떤 유전자도 나타내며, 피루베이트의 아세틸-CoA로 전환을 촉진하는 능력을 의미한다. 특정 구체예에서, PDH 유전자는 효모 근원으로부터 유도될 수도 있다. 예를 들어, PDH 유전자는 각각 SEQ ID NO: 63-66에 설명된 아미노산 서열을 암호화하는 에스. 세레비시아 LAT1, PDA1, PDB1, 또는 LPD 유전자로부터 유도될 수도 있다. 다른 구체예에서, PDH 유전자는 박테리아 근원으로부터 유도될 수도 있다. 예를 들어, PDH 유전자는 각각 SEQ ID NO: 67-69에 설명된 아미노산 서열을 암호화하는 이. 콜리 균주 K12 substr. MG1655 aceE, aceF, 또는 Ipd 유전자, 또는 각각 SEQ ID NO: 70-73에 설명된 아미노산 서열을 암호화하는 비. 서브틸리스 (*B. subtilis*) pdhA, pdhB, pdhC, 또는 pdhD 유전자로부터 유도될 수도 있다.

[0091] "아세틸-CoA 카르복실라제 유전자" 또는 "ACC 유전자"는 여기에 사용된 바와 같이 아세틸-CoA 카르복실라제 활성을 갖는 폴리펩티드를 암호화하는 어떤 유전자도 나타내며, 아세틸-CoA의 말로닐-CoA로 전환을 촉진하는 능력을 의미한다. 아세틸-CoA 카르복실라제 활성을 가진 효소는 EC 6.4.1.2로서 분류된다. 특정 구체예에서, 아세틸-CoA 카르복실라제 유전자는 효모 근원으로부터 유도될 수도 있다. 예를 들어, 아세틸-CoA 카르복실라제 유전자는 SEQ ID NO: 74에 설명된 아미노산 서열을 암호화하는 에스. 세레비시아 ACC1 유전자로부터 유도될 수도 있다. 다른 구체예에서, 아세틸-CoA 카르복실라제 유전자는 박테리아 근원으로부터 유도될 수도 있다. 예를 들어, 아세틸-CoA 카르복실라제 유전자는 각각 SEQ ID NO: 75-78에 설명된 아미노산 서열을 암호화하는 이. 콜리 accA, accB, accC, 또는 accD 유전자, 또는 각각 SEQ ID NO: 79-82에 설명된 아미노산 서열을 암호화하는 씨. 아우란티아쿠스 accA, accB, accC, 또는 accD 유전자로부터 유도될 수도 있다.

[0092] "알라닌 데히드로게나제 유전자"는 여기에 사용된 바와 같이 알라닌 데히드로게나제 활성을 갖는 폴리펩티드를 암호화하는 어떤 유전자도 나타내며, 피루베이트의 알라닌으로 NAD-의존적 환원성 아미노화를 촉진하는 능력을 의미한다. 알라닌 데히드로게나제 활성을 가진 효소는 EC 1.4.1.1로서 분류된다. 특정 구체예에서, 알라닌 데히드로게나제 유전자는 박테리아 근원으로부터 유도될 수도 있다. 예를 들어, 알라닌 데히드로게나제 유전자는 SEQ ID NO: 83에 설명된 아미노산 서열을 암호화하는 비. 서브틸리스 알라닌 데히드로게나제 유전자로부터 유도

될 수도 있다.

- [0093] "피루베이트/알라닌 아미노트랜스퍼라제 유전자"는 여기에 사용된 바와 같이 피루베이트/알라닌 아미노트랜스퍼라제 활성을 갖는 폴리펩티드를 암호화하는 어떤 유전자도 나타내며, 피루베이트 및 L-글루타메이트의 알라닌 및 2-옥소글루타레이트로 전환을 촉진하는 능력을 의미한다. 특정 구체예에서, 피루베이트/알라닌 아미노트랜스퍼라제 유전자는 효모 근원으로부터 유도된다. 예를 들어, 피루베이트/알라닌 아미노트랜스퍼라제 유전자는 SEQ ID NO: 84에 설명된 아미노산 서열을 암호화하는 에스. 폼베 피루베이트/알라닌 아미노트랜스퍼라제 유전자 또는 SEQ ID NO: 85에 설명된 아미노산 서열을 암호화하는 에스. 세레비시에 ALT2 유전자로부터 유도될 수도 있다.
- [0094] "알라닌 2,3 아미노뮤타제 유전자" 또는 "AAM 유전자"는 여기에 사용된 바와 같이 알라닌 2,3 아미노뮤타제 활성을 갖는 폴리펩티드를 암호화하는 유전자를 나타내며, 알라닌의 β -알라닌으로 전환을 촉진하는 능력을 의미한다. 알라닌 2,3 아미노뮤타제 활성은 자연 발생하는 것으로 알려져 있지 않다. 그러므로, 알라닌 2,3 아미노뮤타제 유전자는 하나 이상의 돌연변이를 유사한 활성, 예를 들어, 리신 2,3 아미노뮤타제 활성을 갖는 폴리펩티드를 암호화하는 네이티브 근원 유전자로 통합함으로써 유도될 수 있다 (예컨대, 미국 특허 7,309,597). 특정 구체예에서, 네이티브 근원 유전자는 SEQ ID NO: 86에 설명된 아미노산 서열을 암호화하는 비. 서브틸리스 리신 2,3 아미노뮤타제 유전자, SEQ ID NO: 87에 설명된 아미노산 서열을 암호화하는 피. 강기발리스 (*P. gingivalis*) 리신 2,3 아미노뮤타제 유전자, 또는 SEQ ID NO: 88에 설명된 아미노산 서열을 암호화하는 에프. 뉴클레아툼 (*F. nucleatum*) (ATCC-10953) 리신 2,3 아미노뮤타제 유전자일 수도 있다.
- [0095] "CoA 트랜스퍼라제 유전자"는 여기에 사용된 바와 같이 CoA 트랜스퍼라제 활성을 갖는 폴리펩티드를 암호화하는 어떤 유전자도 나타내며, 이것은 예를 들어, β -알라닌의 β -알라닐-CoA으로 전환 및/또는 락테이트의 락틸-CoA로 전환을 촉진하는 능력을 포함한다. 특정 구체예에서, CoA 트랜스퍼라제 유전자는 효모 근원으로부터 유도될 수도 있다. 다른 구체예에서, CoA 트랜스퍼라제 유전자는 박테리아 근원으로부터 유도될 수도 있다. 예를 들어, CoA 트랜스퍼라제 유전자는 SEQ ID NO: 89에 설명된 아미노산 서열을 암호화하는 엠. 엘스데니 (*M. elsdenii*) CoA 트랜스퍼라제 유전자로부터 유도될 수도 있다.
- [0096] "CoA 신테타제 유전자"는 여기에 사용된 바와 같이 CoA 신테타제 활성을 갖는 폴리펩티드를 암호화하는 어떤 유전자도 나타낸다. 한 예에서, 이것은 β -알라닌의 β -알라닐-CoA로 전환을 촉진하는 능력을 포함한다. 또 다른 예에서, 이것은 락테이트의 락틸-CoA로 전환을 촉진하는 능력을 포함한다. 특정 구체예에서, CoA 신테타제 유전자는 효모 근원으로부터 유도될 수도 있다. 예를 들어, CoA 신테타제 유전자는 에스. 세레비시에 CoA 신테타제 유전자로부터 유도될 수도 있다. 다른 구체예에서, CoA 신테타제 유전자는 박테리아 근원으로부터 유도될 수도 있다. 예를 들어, CoA 신테타제 유전자는 이. 콜리 CoA 신테타제, 알. 스페로이데스, 또는 에스. 엔테리카 (*S. enterica*) CoA 신테타제 유전자로부터 유도될 수도 있다.
- [0097] " β -알라닐-CoA 암모니아 리아제 유전자"는 여기에 사용된 바와 같이 β -알라닐-CoA 암모니아 리아제 활성을 갖는 폴리펩티드를 암호화하는 어떤 유전자도 나타내며, β -알라닐-CoA의 아크릴일-CoA로 전환을 촉진하는 능력을 의미한다. 특정 구체예에서, β -알라닐-CoA 암모니아 리아제 유전자는 박테리아 근원, 예를 들어, SEQ ID NO: 90에 설명된 아미노산 서열을 암호화하는 씨. 프로피오니쿰 (*C. propionicum*) β -알라닐-CoA 암모니아 리아제 유전자로부터 유도될 수도 있다.
- [0098] "3-HP-CoA 데히드라타제 유전자" 또는 "아크릴일-CoA 히드라타제 유전자"는 여기에 사용된 바와 같이 3-HP-CoA 데히드라타제 유전자 활성을 갖는 폴리펩티드를 암호화하는 어떤 유전자도 나타내며, 아크릴일-CoA의 3-HP-CoA로 전환을 촉진하는 능력을 의미한다. 3-HP-CoA 데히드라타제 활성을 가진 효소는 EC 4.2.1.116로서 분류된다. 특정 구체예에서, 3-HP-CoA 데히드라타제 유전자는 효모 또는 균류 근원, 예를 들어, SEQ ID NO: 91에 설명된 아미노산 서열을 암호화하는 피. 소재 (*P. sojae*) 3-HP-CoA 데히드라타제 유전자로부터 유도될 수도 있다. 다른 구체예에서, 3-HP-CoA 데히드라타제 유전자는 박테리아 근원으로부터 유도될 수도 있다. 예를 들어, 3-HP-CoA 데히드라타제 유전자는 SEQ ID NO: 92에 설명된 아미노산 서열을 암호화하는 씨. 아우란티아쿠스 3-HP-CoA 데히드라타제 유전자, SEQ ID NO: 93에 설명된 아미노산 서열을 암호화하는 알. 루브룸 (*R. rubrum*) 3-HP-CoA 데히드라타제 유전자, 또는 SEQ ID NO: 94에 설명된 아미노산 서열을 암호화하는 알. 캡슐라테스 (*R. capsulatus*) 3-HP-CoA 데히드라타제 유전자로부터 유도될 수도 있다. 다른 구체예에서, 3-HP-CoA 데히드라타제 유전자는 포유동물 근원으로부터 유도될 수도 있다. 예를 들어, 3-HP-CoA 데히드라타제 유전자는 SEQ ID NO: 95에 설명된 아미노산 서열을 암호화하는 에이치. 사피엔스 (*H. sapiens*) 3-HP-CoA 데히드라타제 유전자로부터 유도될 수도 있다.

- [0099] "3-HP-CoA 히드롤라제 유전자"는 여기에 사용된 바와 같이 3-HP-CoA 히드롤라제 활성을 갖는 폴리펩티드를 암호화하는 어떤 유전자도 나타내며, 3-HP-CoA의 3-HP로 전환을 촉진하는 능력을 의미한다. 특정 구체예에서, 3-HP-CoA 유전자는 효모 또는 균류 근원으로부터 유도될 수도 있다. 다른 구체예에서, 3-HP-CoA 유전자는 박테리아 또는 포유동물 근원으로부터 유도될 수도 있다.
- [0100] "3-히드록시이소부티릴-CoA 히드롤라제 유전자"는 여기에 사용된 바와 같이 3-히드록시이소부티릴-CoA 히드롤라제 활성을 갖는 폴리펩티드를 암호화하는 어떤 유전자도 나타내며, 이것은 예를 들어, 3-HP-CoA의 3-HP로 전환을 촉진하는 능력을 포함한다. 특정 구체예에서, 3-히드록시이소부티릴-CoA 히드롤라제 유전자는 박테리아 근원, 예를 들어, SEQ ID NO: 96에 설명된 아미노산 서열을 암호화하는 피. 플루오레센스 3-히드록시이소부티릴-CoA 히드롤라제 유전자 또는 SEQ ID NO: 97에 설명된 아미노산 서열을 암호화하는 비. 세레우스 (*B. cereus*) 3-히드록시이소부티릴-CoA 히드롤라제 유전자로부터 유도될 수도 있다. 다른 구체예에서, 3-히드록시이소부티릴-CoA 히드롤라제 유전자는 SEQ ID NO: 98에 설명된 아미노산 서열을 암호화하는 포유동물 근원, 예를 들어, 에 이치. 사피엔스 3-히드록시이소부티릴-CoA 히드롤라제 유전자로부터 유도될 수도 있다.
- [0101] "락테이트 데히드로게나제 유전자" 또는 "LDH 유전자"는 여기에 사용된 바와 같이 락테이트 데히드로게나제 활성을 갖는 폴리펩티드를 암호화하는 어떤 유전자도 나타내며, 피루베이트의 락테이트로 전환을 촉진하는 능력을 의미한다. 특정 구체예에서, LDH 유전자는 균류, 박테리아, 또는 포유동물 근원으로부터 유도될 수도 있다.
- [0102] "락틸-CoA 데히드라타제 유전자"는 여기에 사용된 바와 같이 락틸-CoA 데히드라타제 활성을 갖는 폴리펩티드를 암호화하는 어떤 유전자도 나타내며, 락틸-CoA의 아크릴일-CoA로 전환을 촉진하는 능력을 의미한다. 특정 구체예에서, 락틸-CoA 데히드라타제 유전자는 박테리아 근원으로부터 유도될 수도 있다. 예를 들어, 락틸-CoA 데히드라타제 유전자는 SEQ ID NO: 99-101에 설명된 아미노산 서열을 암호화하는 엠. 엘스테니이 락틸-CoA 데히드라타제 E1, E11a, 또는 E11b 서브유닛 유전자로부터 유도될 수도 있다.
- [0103] "알데히드 데히드로게나제 유전자"는 여기에 사용된 바와 같이 알데히드 데히드로게나제 활성을 갖는 폴리펩티드를 암호화하는 어떤 유전자도 나타내며, 이것은 예를 들어, 3-HPA의 3-HP로 및 그 반대의 전환을 촉진하는 능력을 포함한다. 특정 구체예에서, 알데히드 데히드로게나제 유전자는 효모 근원, 예를 들어, SEQ ID NO: 102에 설명된 아미노산 서열을 암호화하는 에스. 세레비시에 알데히드 데히드로게나제 유전자 또는 SEQ ID NO: 122, 124, 또는 126에 설명된 아미노산 서열을 암호화하는 아이. 오리엔탈리스 알데히드 데히드로게나제 유전자로부터 유도될 수도 있다. 다른 구체예에서, 알데히드 데히드로게나제는 박테리아 근원, 예를 들어, SEQ ID NO: 103에 설명된 아미노산 서열을 암호화하는 이. 콜리 aldH 유전자 또는 SEQ ID NO: 104에 설명된 아미노산 서열을 암호화하는 케이. 뉴모니아 (*K. pneumoniae*) 알데히드 데히드로게나제 유전자로부터 유도될 수도 있다.
- [0104] "글리세롤 데히드라타제 유전자"는 여기에 사용된 바와 같이 글리세롤 데히드라타제 활성을 갖는 폴리펩티드를 암호화하는 어떤 유전자도 나타내며, 글리세롤의 3-HPA로 전환을 촉진하는 능력을 의미한다. 특정 구체예에서, 글리세롤 데히드라타제 유전자는 박테리아 근원, 예를 들어, 케이. 뉴모니아 (*K. pneumonia*) 또는 씨. 프레운디이 (*C. freundii*) 글리세롤 데히드라타제 유전자로부터 유도될 수도 있다.
- [0105] 특정 구체예에서, 여기에 제공된 유전적으로 변형된 효모 세포는 하나 이상의 네이티브 유전자의 삭제 또는 붕괴를 더 포함한다. 네이티브 유전자에 관한 "삭제 또는 붕괴"는 유전자의 전체 암호화 영역이 제거 (삭제)되거나 유전자가 더 이상 활성 효소를 생산하지 않거나, 크게 감소된 양 (적어도 75% 감소, 바람직하게 적어도 90% 감소)의 활성 효소를 생산하거나, 크게 감소된 (적어도 75% 감소, 바람직하게 적어도 90% 감소) 활성을 갖는 효소를 생산하는 유전자의 암호화 영역, 그것의 프로모터, 및/또는 그것의 종결 영역이 변형된다는 것을 의미한다.
- [0106] 특정 구체예에서, 하나 이상의 네이티브 유전자의 삭제 또는 붕괴는 하나 이상의 네이티브 대사 경로의 삭제 또는 붕괴를 일으킨다. 대사 경로에 관한 "삭제 또는 붕괴"는 경로가 작동 불가능하거나 네이티브 경로에 비해 적어도 75%, 적어도 85%, 또는 적어도 95% 감소된 활성을 나타낸다는 것을 의미한다. 특정 구체예에서, 네이티브 대사 경로의 삭제 또는 붕괴는 3-HP 생산을 감소시키는 하나 이상의 네이티브 유전자의 감소된 발현을 일으키는 하나 이상의 유전적 변형을 통합함으로써 달성된다.
- [0107] 특정 구체예에서, 네이티브 유전자의 삭제 또는 붕괴는 강제 진화 (ced evolution), 돌연변이 유발, 또는 유전적 조작 방법에 이어서, 원하는 돌연변이를 확인하기 위해 적절한 선택 또는 스크리닝에 의해 달성될 수 있다. 특정 구체예에서, 네이티브 숙주 세포 유전자의 삭제 또는 붕괴는 하나 이상의 외인성 유전자의 숙주 세포로 통합과 결합될 수도 있다, 즉, 외인성 유전자는 또한 삭제 구조물인 유전자 발현 통합 구조물을 사용하여 통합될

수도 있다. 다른 구체예에서, 삭제 또는 붕괴는 외인성 유전자를 함유하지 않는 삭제 구조물을 사용하여 또는 업계에 알려진 다른 방법으로 달성될 수도 있다.

[0108] 특정 구체예에서, 여기에 제공된 유전적으로 변형된 효모 세포는 예를 들어, 피루베이트 데카르복실라제 (PDC, 피루베이트를 아세트알데히드로 전환) 및/또는 알콜 데히드로게나제 (ADH, 아세트알데히드를 에탄올로 전환) 유전자를 포함하는, 에탄올 발효에 수반되는 효소를 암호화하는 하나 이상의 네이티브 유전자의 삭제 또는 붕괴를 포함한다. 이들 변형은 효모 세포의 에탄올을 생산하는 능력을 감소시키고, 이로 인해 3-HP 생산을 최대화한다. 하지만, 특정 구체예에서, 여기에 제공된 유전적으로 변형된 효모 세포는 3-HP 및 에탄올을 동시 생산하도록 조작될 수도 있다. 이들 구체예에서, 에탄올 발효에 수반되는 효소를 암호화하는 네이티브 유전자는 바람직하게 삭제되거나 붕괴되고, 특정 구체예에서, 효모 세포는 에탄올 생산을 증가시키는 하나 이상의 외인성 유전자를 포함할 수도 있다.

[0109] 특정 구체예에서, 여기에 제공된 유전적으로 변형된 효모 세포는 대체 발효 생성물, 예를 들어, 글리세롤 또는 다른 부산물, 예를 들어, 아세테이트 또는 디올의 생산에 수반되는 효소를 암호화하는 하나 이상의 네이티브 유전자의 삭제 또는 붕괴를 포함한다. 예를 들어, 여기에 제공된 세포는 글리세롤 3-포스페이트 데히드로게나제 (GPD, 디히드록시아세톤 포스페이트의 글리세롤 3-포스페이트로 반응을 촉진함), 글리세롤 3-포스파타제 (GPP, 글리세롤-3 포스페이트의 글리세롤로 전환을 촉진함), 글리세롤 키나제 (글리세롤 3-포스페이트의 글리세롤로 전환을 촉진함), 디히드록시아세톤 키나제 (디히드록시아세톤 포스페이트의 디히드록시아세톤로 전환을 촉진함), 글리세롤 데히드로게나제 (디히드록시아세톤의 글리세롤로 전환을 촉진함), 알데히드 데히드로게나제 (ALD, 예를 들어, 아세트알데히드를 아세테이트로 또는 3-HP를 3-HPA로 전환함), 또는 부탄디올 데히드로게나제 (부탄디올의 아세토인으로 및 그 반대의 전환을 촉진함) 유전자 중 하나 이상의 삭제 또는 붕괴를 포함할 수도 있다.

[0110] 특정 구체예에서, 여기에 제공된 유전적으로 변형된 효모 세포는

[0111] 3-HP 발효 경로에서 역반응을 촉진하는 효소, 예를 들어, PEP 카르복시키나제 (PCK), OAA 데카르복실라제 활성을 갖는 효소, 또는 CYB2A 또는 CYB2B (락테이트의 피루베이트로 전환을 촉진함)를 암호화하는 하나 이상의 네이티브 유전자의 삭제 또는 붕괴를 포함한다. PCK는 PEP의 OAA로 및 그 반대의 전환을 촉진하지만, PEP 반응에 대한 OAA에 대하여 선호도를 나타낸다. OAA의 PEP로 전환을 감소시키기 위해, 네이티브 PCK 유전자의 하나 이상의 카피가 삭제되거나 붕괴될 수도 있다. 특정 구체예에서, 하나 이상의 네이티브 PCK 유전자가 삭제되거나 붕괴되는 효모 세포는 PEP의 OAA로 전환을 선호하는 폴리펩티드를 암호화하도록 돌연변이된 하나 이상의 외인성 PCK 유전자를 발현할 수도 있다. OAA 데카르복실라제는 OAA의 피루베이트로 전환을 촉진한다. OAA 데카르복실라제 활성을 갖는 효소, 예를 들어, 이. 콜리의 eda 유전자에 의해 암호화된 것들 및 효모 및 균류의 말산 효소 (MAE)가 확인되었다. OAA 데카르복실라제 활성을 감소시키기 위해, OAA 데카르복실라제 활성을 갖는 효소를 암호화하는 네이티브 유전자의 하나 이상의 카피는 삭제되거나 붕괴될 수도 있다. 특정 구체예에서, 하나 이상의 네이티브 OAA 데카르복실화 유전자가 삭제되거나 붕괴되는 효모 세포는 피루베이트의 OAA로 전환을 촉진하는 폴리펩티드를 암호화하도록 돌연변이된 하나 이상의 외인성 OAA 데카르복실화 유전자를 발현할 수도 있다.

[0112] 특정 구체예에서, 여기에 제공된 유전적으로 변형된 효모 세포는 3-HP 발효 경로 생성물 또는 중간물로 원하지 않는 반응에 수반되는 효소를 암호화하는 하나 이상의 네이티브 유전자의 삭제 또는 붕괴를 포함한다. 이러한 유전자의 예는 3HP를 3HP의 알데히드로 전환시키는 효소를 암호화하는 것들을 포함하며, 이것은 박테리아 세포에 해로운 것으로 알려져 있다.

[0113] 특정 구체예에서, 여기에 제공된 유전적으로 변형된 효모 세포는, 예를 들어, GAL6 (갈락토스를 글루코스로 전환하는 GAL 시스템의 음성 조절기)를 포함하는, 3-HP 발효 경로에 대하여 중성 효과를 갖는 효소를 암호화하는 하나 이상의 네이티브 유전자의 삭제 또는 붕괴를 포함한다. 중성 유전자의 삭제 또는 붕괴는 네이티브 발효 경로에 영향을 미치지 않고 하나 이상의 외인성 유전자의 삽입을 허용한다.

[0114] 특정 구체예에서, 여기에 제공된 효모 세포는 3-HP 저항성 효모 세포이다. "3-HP-저항성 효모 세포"는 여기에 사용된 바와 같이 4.0 미만의 pH에서 75 g/L 이상의 3-HP를 함유하는 배지에 적어도 2.5 g/L/hr의 평균 당분해 속도를 나타내는 효모 세포를 나타낸다. 이러한 속도 및 조건은 3-HP를 생산하는 경제적인 공정을 대표한다. 이들 중 특정 구체예에서, 효모 세포는 그것들의 네이티브 형태로 3-HP 저항성을 나타낼 수도 있다. 다른 구체예에서, 세포는 활성 3-HP 발효 경로와 관련된 유전적 변형의 도입 전에, 중에, 또는 후에 돌연변이 및/또는 선택된 세포는 같은 종의 야생형 세포보다 3-HP에 대하여 더 높은 정도의 저항성을 소유하는, 돌연변이 및/또는 선택 (예를 들어, 케모스타트 선택 또는 반복된 일련의 계대 배양)을 겪을 수도 있다. 예를 들어, 일부 구체예에

서, 세포는 하나 이상의 외인성 3-HP 경로 유전자를 가지고 유전적으로 변형되기 전에 3-HP 또는 젖산의 존재시 돌연변이 및/또는 선택을 겪었다. 특정 구체예에서, 돌연변이 및/또는 선택은 그것들의 네이티브 형태로 3-HP 저항성을 나타내는 세포에서 수행될 수도 있다. 돌연변이 및/또는 선택을 겪는 세포는 3-HP 생산을 위한 산업상 숙주로서 그것들의 가능성을 결정하기 위해 다양한 수준의 3-HP의 존재시 당 소비 및 다른 특징에 대하여 테스트될 수도 있다. 3-HP 저항성 이외에, 여기에 제공된 효모 세포는 하나 이상의 추가적 유기산 (예를 들어, 젖산) 또는 다른 발효 생성물, 부산물, 또는 배지 성분에 대한 저항성의 돌연변이 및/또는 선택을 겪을 수도 있다.

[0115] 선택, 예를 들어, 3-HP에 대한 또는 다른 화합물에 대한 저항성의 선택은 업계에 잘 알려진 방법을 사용하여 달성될 수도 있다. 예를 들어, 상기 언급된 바와 같이, 선택은 케모스타트 선택일 수도 있다. 케모스타트 선택은 미생물 (예를 들어, 효모)의 연속 배양을 허용하는 케모스타트를 사용하며, 특이적 성장 속도 및 세포 수가 독립적으로 조절될 수 있다. 연속 배양은 본질적으로 배지가 연속적으로 추가되고 어떤 오버플로우 (overflow)의 연속적인 제거도 일어날 수 있는 일정한 부피의 플로우 시스템 (flow system)이다. 세포 수 및 영양 상태는 일정하게 유지되고, 시스템은 정상 상태이다. 케모스타트는 희석 속도 및 제한된 영양, 예를 들어, 탄소 또는 질소 근원의 농도의 변화를 통해 배양의 군집 밀도 및 특이적 성장 속도 둘 다의 조절을 허용한다. 배양이 성장함에 따라 조건을 변화시킴으로써 (다른 것들 중에, 예를 들어, 접종물 균주의 성장에 필요한 두 번째 탄소 근원의 농도를 감소시킴), 변화된 조건에서 더 빨리 성장할 수 있는 집단의 미생물은 선택될 것이고 새로운 조건 하에 작용하지 않는 미생물보다 커질 것이다. 전형적으로 이러한 선택은 케모스타트 배양의 성장의 과정을 통해 적어도 하나의 배양 성분의 점진적 증가 또는 감소가 필요할 것이다. 미생물의 유도 진화에서 케모스타트의 작동 및 그것들의 사용은 업계에 잘 알려져 있다 (예컨대, Novick Proc Natl Acad Sci USA 36:708-719 (1950), Harder J Appl Bacteriol 43:1-24 (1977)). 선택되는 다른 방법은, 예를 들어, 미국 특허 번호 7,629,162에 설명된 바와 같이 선택적 조건 하에 반복된 연속적 계대 배양이다. 이러한 방법은 상기 설명된 글루코스 제한 케모스타트 방법을 대신해서, 또는 이에 더하여 사용될 수 있다.

[0116] 하기 실시예에 개시된 바와 같이 3-HP 배지에서 성장 및 글루코스 소비의 최고의 조합을 나타내는 효모 균주는 3-HP 발효 경로에 관한 다양한 유전적 변형에 대하여 바람직한 숙주 세포이다. 상대적으로 높은 정도의 3-HP 저항성에 대한 가능성을 소유하는 효모 속은 4 미만의 pH에서 75 g/L 이상의 3-HP의 존재시 성장에 의해 지시된 바와 같이, 예를 들어, 칸디다, 클루이베로미세스, 이사트첸키아, 사카로미세스, 피치아, 쉬조사카로미세스, 토룰라스포라, 및 지고사카로미세스를 포함한다. 3-HP 저항성을 나타내는 종은 아이. 오리엔탈리스 (씨. 크루세이 (*C. krusei*))로도 알려짐), 씨. 람비카 (피치아 페르멘탄스로도 알려짐), 및 에스. 불데리 (카자흐스탄리아 불데리 (*Kazachstania bulderi*))로도 알려짐)를 포함한다. 아이. 오리엔탈리스 및 씨. 람비카는 아이. 오리엔탈리스/피. 페르멘탄스 클레이드의 것인 한편, 에스. 불데리는 사카로미세스 클레이드의 것이다. 3-HP 저항성을 나타내는 특이적 균주는 아이. 오리엔탈리스 균주 24210, PTA-6658, 60585, 및 CD1822, 에스. 불데리 균주 MYA-402 및 MYA-404, 및 씨. 람비카 균주 ATCC 38617를 포함하였다.

[0117] 다른 야생형 효모 또는 균류는 여기에 설명된 바와 같이 높은 수준의 3-HP의 존재시 성장 및 글루코스 이용의 허용 가능한 수준을 갖는지 유사한 방식으로 테스트되고 확인될 수도 있다. 예를 들어, Gross 및 Robbins (*Hydrobiologia* 433(103):91-109)는 잠재적 생산 숙주로서 테스트하기에 적절한 낮은 pH (<4) 환경에서 81개의 균류 종의 목록을 편집하였다.

[0118] 특정 구체예에서, 여기에 제공된 변형된 효모 세포는 하나 이상의 유전적 변형을 크랩트리-음성 숙주 효모 세포로 통합함으로써 발생된다. 이들 중 특정 구체예에서, 숙주 효모 세포는 속 이사트첸키아, 칸디다, 또는 사카로미세스에 속하고, 이들 중 특정 구체예에서, 숙주 세포는 아이. 오리엔탈리스/피. 페르멘탄스 또는 사카로미세스 클레이드에 속한다. 특정 구체예에서, 숙주 세포는 아이. 오리엔탈리스 또는 씨. 람비카, 또는 에스. 불데리이다.

[0119] 아이. 오리엔탈리스/피. 페르멘탄스 클레이드는 적어도 종 아이. 오리엔탈리스, 피. 갈레이포르미스 (*P. galei* mis), 피. 종 YB-4149 (NRRL 지명), 씨. 에타놀리카 (*C. 에타놀리카*), 피. 데세르티콜라 (*P. deserticola*), 피. 멤브라니파시엔스 (*P. membranifaciens*), 및 피. 페르멘탄스를 함유하는 가장 말단의 클레이드이다. 아이. 오리엔탈리스/피. 페르멘탄스 클레이드의 멤버는 "Identification and Phylogeny of Ascomycetous Yeasts from Analysis of Nuclear Large Subunit (26S) Ribosomal DNA Partial Sequences," Antonie van Leeuwenhoek 73:331-371, 1998에서 Kurtzman 및 Robnett에 의해 설명된 방법을 사용하여, 효모 종의 26S 리보솜 DNA의 가변 D1/D2 도메인의 분석에 의해 확인되며, 본원에 참고로 포함된다 (특히 p. 349 참조). 수백개의 자낭균류 (ascomycetes)의 26S 리보솜 DNA의 가변 D1/D2 도메인의 분석은 아이. 오리엔탈리스/피. 페르멘탄스 클레이드가

아주 밀접하게 관련된 종을 함유한다는 것을 나타낸다. 아이. 오리엔탈리스/피. 페르멘탄스 클레이드의 멤버는 26S 리보솜 DNA의 가변 D1/D2 도메인에서 클레이드 외부의 효모 종보다 클레이드의 다른 멤버에 대하여 더 큰 유사성을 나타낸다. 그러므로, Kurtzman 및 Robnett의 방법을 사용하여, 아이. 오리엔탈리스/피. 페르멘탄스 클레이드의 다른 멤버는 그것들 각각의 리보솜 DNA의 D1/D2 도메인의 비교에 의해 및 클레이드의 다른 멤버의 그것과 클레이드 외부의 밀접하게 관련된 종을 비교함으로써 확인될 수 있다.

- [0120] 특정 구체예에서, 여기에 제공된 유전적으로 변형된 효모 세포는 속 이사트첸키아에 속하고, 이들 중 특정 구체예에서, 효모 세포는 아이. 오리엔탈리스이다. 첫 번째 특징으로 할 때, 종 아이. 오리엔탈리스는 명칭 피치아 쿠드리야브제비 (Pichia kudriavzevii)로 지정되었다. 아이. 오리엔탈리스의 무성생식 (무성 형태)는 칸디다 크루세이로서 알려져 있다. 종 아이. 오리엔탈리스에 대한 많은 추가적 동의어가 별도로 나열되었다 (Kurtzman and Fell, The 효모s, a Taxonomic Study. 섹션 35. 이사트첸키아 쿠드리야브트세브 (Issatchenkia kudryavtsev), pp 222-223 (1998)).
- [0121] 3-HP 생산을 위한 이상적인 효모 세포는 이상적으로 낮은 pH 수준에서 성장할 수 있다. 낮은 pH에서 발효를 수행하는 능력은 다운스트림 회수 비용을 감소시키며, 더 경제적인 생산을 일으킨다. 그러므로, 특정 구체예에서, 효모 숙주 세포는 낮은 pH 수준 (예를 들어, 7, 6, 5, 4, 또는 3 미만의 pH 수준)에서 성장할 수 있다.
- [0122] 적합한 숙주 세포는 3-HP 저항성 및/또는 낮은 pH 성장 능력 외에 하나 이상의 바람직한 특징을 소유할 수도 있다. 예를 들어, 3-HP 저항성을 나타내는 잠재적 숙주 세포는 당분해 속도, 열 내성, 바이오매스 가수분해 생성물 억제제에 대한 내성, 전체적인 과정의 견고성, 등에 기초하여 추가로 선택될 수도 있다. 이들 기준은 3-HP 발효 경로와 관련된 어떤 유전적 변형 이전에 평가될 수도 있으며, 그것들은 하나 이상의 이러한 변형이 일어나고 난 후에 평가될 수도 있다.
- [0123] 대부분 효모가 에탄올의 네이티브 생산자이기 때문에, 피루베이트 (PDC)로부터 에탄올 생산의 첫 번째 단계를 촉진하는 효소의 제거 또는 극심한 감소는 대체 생성물의 충분한 수득율에 필요하다. 크랩트리-양성 효모, 예를 들어, 사카로미세스에서, 삭제된 또는 봉쇄된 PDC 유전자는 숙주가 에탄올 또는 아세테이트와 같은 이-탄소 화합물에 대한 영양 요구성을 얻는 것을 유발하고, 글루코스를 함유한 배지에서 성장의 결핍을 유발한다. 이들 한계를 극복할 수 있는 돌연변이는 아세테이트 독립성 및 글루코스 내성에 대하여 점진적 선택을 사용하여 얻어질 수 있다 (예컨대, van Maris Appl Environ Microbiol 70:159 (2004)). 그러므로, 특정 구체예에서, 바람직한 효모 숙주 세포는 크랩트리-음성 효모 세포이며, PDC 삭제 균주는 글루코스에서 성장할 수 있고 C2 자가 영양성 (prototrophy)을 유지한다.
- [0124] 특정 세포에서 이용되는 유전자 발현의 수준 및/또는 외인성 유전자의 수는 선택된 효모 종에 따라 다를 것이다. 완전히 게놈-나열된 효모에 대하여, 전체-게놈 화학양론적 모델은 효소가 원하는 경로를 3-HP 발효 경로로 발달시키기 위해 발현되어야 하는지 결정하기 위해 사용될 수도 있다. 전체-게놈 화학양론적 모델은, 예를 들어, Hjersted et al, "게놈-scale analysis of Saccharomyces cerevisiae metabolism and ethanol production in fed-batch culture," Biotechnol. Bioeng. 2007; 및 Famili et al, "Saccharomyces cerevisiae phenotypes can be predicted by using con균t-based analysis of a 게놈-scale reconstructed metabolic network," Proc. Natl. Acad. Sci. 2003, 100(23): 13134-9에 설명된다.
- [0125] 알려진 게놈 서열이 없는 효모에 대하여, 원하는 유전자에 대한 서열 (과발현 후보물질로서 또는 삽입 부위로서)은 전형적으로 실시예 2A에 하기 설명된 것들과 같은 기술을 사용하여 얻어질 수 있다. 일상적인 실험 설계는 3-HP 경로에서 작용하는 유전자 및 효소를 포함하는, 다양한 유전자의 발현 및 다양한 효소의 활성을 테스트하기 위해 이용될 수 있다. 각각의 효소가 효모에서 개별적으로 및 효소의 최대 차단시 발현되고 개선된 3-HP 생산위해 필요한지 (또는 요구되는지) 확립하기 위해, 바람직하게 모든 경로 효소를 포함하는, 실험이 수행될 수도 있다. 하나의 예시적 실험 설계는 각각의 개별적 효소 뿐만 아니라 효소의 각각의 독특한 쌍의 발현을 테스트하고, 모든 필요한 효소의 발현, 또는 효소의 각각의 독특한 조합을 추가로 테스트할 수 있다. 많은 접근법이 인정되는 것만큼, 취해질 수 있다.
- [0126] 특정 구체예에서, 방법은 여기에 제공된 바와 같이 유전적으로 변형된 효모 세포로부터 3-HP를 생산하기 위해 제공된다. 특정 구체예에서, 이들 방법은 적어도 하나의 탄소 근원의 존재시 여기에 제공된 바와 같이 유전적으로 변형된 효모 세포를 배양하는 단계를 포함하며, 일정 기간 동안 세포가 3-HP를 생산하는 것을 허용하고, 배지로부터 세포에 의해 생산된 3-HP를 분리한다. 탄소 근원은 제공된 효모에 의해 발효될 수 있는 어떤 탄소 근원일 수도 있다. 탄소 근원은 12개의 탄소 당, 예를 들어, 수크로스, 헥소스 당, 예를 들어, 글루코스 또는 프럭토스, 글리칸 또는 글루코스의 다른 폴리머, 글루코스 올리고머, 예를 들어, 말토스, 말토트리오스 및 이소말

토트리오스, 판노스, 및 프럭토스 올리고머일 수도 있다. 세포가 펜토스 당을 발효시키는 능력을 전달하기 위해 변형되면, 발효 배지는 펜토스 당, 예를 들어, 자일로스, 자일란 또는 자일로스의 다른 올리고머, 및/또는 아라비노스를 포함할 수도 있다. 이러한 펜토스 당은 적합하게 헤미셀룰로스-함유 바이오매스의 가수분해물이다. 올리고머 당의 경우에서, 세포에 의한 발효를 위하여 이들을 해당하는 모노머 당으로 분해하기 위해 효소를 발효 브로스에 추가하는 것이 필요할 수도 있다. 특정 구체예에서, 하나 이상의 타입의 유전적으로 변형된 효모 세포가 배지에 존재할 수도 있다. 유사하게, 특정 구체예에서, 유전적으로 변형된 효모 세포보다 같은 또는 다른 종의 하나 이상의 네이티브 효모 세포가 배지에 존재할 수도 있다.

[0127] 특정 구체예에서, 3-HP를 생산하기 위해 여기에 제공된 세포의 배양은 단계별로 나누어질 수도 있다. 예를 들어, 세포 배양 과정은 배양 단계, 생산 단계 및 회수 단계로 나누어질 수도 있다. 업계의 보통의 기술 중 하나는 이들 단계에 사용된 조건이 사용되는 효모의 종, 효모에 의해 이용된 특이적 3-HP 발효 경로, 원하는 수득율, 또는 다른 인자와 같은 인자에 기초하여 달라질 수도 있다.

[0128] 배지는 전형적으로 특정 세포에 필요한 만큼, 질소의 근원 (예를 들어, 아미노산, 단백질, 암모니아 또는 암모늄 염, 등과 같은 무기 질소 근원), 및 다양한 비타민, 미네랄 등을 포함하는, 영양소를 함유할 것이다. 일부 구체예에서, 본 발명의 세포는 화학적 한정 배지에서 배양될 수 있다. 한 예에서, 배지는 약 5 g/L 암모늄 술페이트, 약 3 g/L 칼륨 디히드로젠 포스페이트, 약 0.5 g/L 마그네슘 술페이트, 미량의 원소, 비타민 및 약 150 g/L 글루코스를 함유한다. pH는 배양 중에 자유롭게 다양한 것이 허용될 수도 있거나 pH가 사전 결정된 수준 이하로 떨어지거나 그 이상으로 증가하는 것을 방지하는데 필요한 경우 완충될 수도 있다. 특정 구체예에서, 발효 배지는 약 1.0의 OD600을 생산하기 위해 평가의 대상이 되는 충분한 효모 세포로 접종된다. 달리 명백하게 지시되지 않으면, OD600은 여기에 사용된 바와 같이 모델 DU600 분광광도계 (Beckman Coulter)를 사용하여 1 cm 경로길이의 600 nm의 파장에서 측정된 최적의 밀도를 나타낸다. 배양 온도는 약 30-40 °C로 다양할 수도 있고, 배양 시간은 최대 약 120 시간일 수도 있다.

[0129] 한 예에서, 발효 배지에서 세포의 농도는 전형적으로 생산 단계 중에 약 0.1 내지 20, 바람직하게 0.1 내지 5, 더 바람직하게 1 내지 3 g 건조 세포/발효 배지의 리터의 범위에 있다. 발효는 경로 필요조건에 의존적으로 호기성으로, 미세호기성으로, 또는 혐기성으로 수행될 수도 있다. 원하면, 산소 흡수율 (OUR)은 과정 조절로서 발효를 통해 달라질 수 있다 (예컨대, W003/102200). 일부 구체예에서, 여기에 제공된 변형된 효모 세포는 2 내지 45 mmol/L/hr, 예를 들어, 2 내지 25, 2 내지 20, 2 내지 15, 2 내지 10, 10 내지 45, 15 내지 40, 20 내지 35, 또는 25 내지 35 mmol/L/hr의 산소 흡수율을 특징으로 하는 미세호기성 조건 하에서 배양된다. 특정 구체예에서, 여기에 제공된 변형된 효모 세포는 2 내지 25 mmol/L/hr의 산소 흡수율을 특징으로 하는 미세호기성 조건 하에 배양될 때 특히 잘 수행될 수도 있다. 배지는 생산 단계 중에 pH가 약 3.0 내지 약 7.0, 또는 약 4.0 내지 약 6.0의 범위에서 유지되도록 완충될 수도 있다. 적합한 완충제는 형성되는 산을 중화하는 염기성 물질이고, 예를 들어, 칼슘 히드록시드, 칼슘 카르보네이트, 나트륨 히드록시드, 칼륨 히드록시드, 칼륨 카르보네이트, 나트륨 카르보네이트, 암모늄 카르보네이트, 암모니아, 암모늄 히드록시드 등을 포함한다. 일반적으로, 통상적인 발효 과정에 사용되는 이들 완충제가 또한 여기에 적합하다.

[0130] 이들 구체예에서, 완충된 발효가 이용되는 경우, 산성 발효 생성물은 형성되는 해당하는 염에 대하여 중화될 수도 있다. 이들 구체예에서, 산의 회수는 자유 산의 재생을 수반한다. 이것은 세포를 제거함으로써 및 발효 브로스를 황산과 같은 강산으로 산성화함으로써 수행될 수도 있다. 이것은 염 부산물의 형성을 일으킨다. 예를 들어, 칼슘 염이 중화제로서 이용되고 황산이 산성화제로서 이용되는 경우, 석고가 염 부산물로서 생산된다. 이 부산물은 브로스로부터 분리되고, 산은 액체-액체 추출, 증류, 흡수, 등과 같은 기술을 사용하여 회수된다 (예컨대, T.B. Vickroy, Vol. 3, Chapter 38 of Comprehensive Biotechnology, (ed. M. Moo-Young), Pergamon, Oxford, 1985; R. Datta, et al, FEMS Microbiol Rev, 1995, 16:221-231; 미국 특허 번호 4,275,234, 4,771,001, 5,132,456, 5,420,304, 5,510,526, 5,641,406, 및 5,831,122, 및 W093/00440).

[0131] 다른 구체예에서, 발효 브로스의 pH는 배양 중에 3-HP의 pKa 이상, 전형적으로 4.5 이상인 시작 pH로부터 산성 발효 생성물의 pKa 이하, 예를 들어, 4.5 또는 4.0 미만, 예를 들어, 약 1.5 내지 약 4.5의 범위, 약 2.0 내지 약 4.0의 범위, 또는 약 2.0 내지 약 3.5의 범위로 떨어지는 것이 허용될 수도 있다.

[0132] 다른 구체예에서, 발효는 발효 과정의 시작 전에 또는 시작할 때 발효 배지의 pH를 생성물 산의 pKa 이하로 조정함으로써 생성물 산을 생산하기 위해 수행될 수도 있다. 그 이후 pH는 배양을 통해 생성물 산의 pKa 이하로 유지될 수도 있다. 특정 구체예에서, pH는 4.5 또는 4.0 미만, 예를 들어, 약 1.5 내지 약 4.5의 범위, 약 2.0 내지 약 4.0의 범위, 또는 약 2.0 내지 약 3.5의 범위로 유지될 수도 있다.

- [0133] 여기에 제공된 방법의 특정 구체예에서, 유전적으로 변형된 효모 세포는 상대적으로 낮은 수준의 에탄올을 생산한다. 특정 구체예에서, 에탄올은 10% 이하의 수득율, 바람직하게 2% 이하의 수득율로 생산될 수도 있다. 이들 중 특정 구체예에서, 에탄올은 검출 가능하게 생산되지 않는다. 하지만, 다른 구체예에서, 3-HP 및 에탄올은 동시 생산될 수도 있다. 이들 구체예에서, 에탄올은 10% 이상, 25% 이상, 또는 50% 이상의 수득율로 생산될 수도 있다.
- [0134] 여기에 제공된 방법의 특정 구체예에서, 탄소 근원에서 3-HP의 최종 수득율은 이론상 수득율의 적어도 10%, 적어도 20%, 적어도 30%, 적어도 40%, 적어도 50%, 또는 50% 이상이다. 3-HP의 농도, 또는 역가는 탄소 근원의 수득율 뿐만 아니라 농도의 함수일 것이다. 특정 구체예에서, 역가는 발효 중의 일부 시점에서, 및 바람직하게 발효의 끝에 적어도 1-3, 적어도 5, 적어도 10, 적어도 20, 적어도 30, 적어도 40, 적어도 50, 또는 50 g/L 이상에 도달할 수도 있다. 특정 구체예에서, 3-HP의 최종 수득율은, 특히 생산 단계 중에, 발효 배지의 온도를 변화시킴으로써 증가될 수도 있다.
- [0135] 생산되면, 업계에 알려진 어떤 방법도 발효 배지로부터 3-HP를 분리하기 위해 사용될 수 있다. 예를 들어, 일반적인 분리 기술은 브로스로부터 바이오매스를 제거하기 위해 사용될 수 있고, 일반적인 분리 과정 (예를 들어, 추출, 증류, 및 이온-교환 과정)은 미생물이 없는 브로스로부터 3-HP를 얻기 위해 사용될 수 있다. 게다가, 3-HP는 그것이 생산되는 동안 분리될 수 있거나, 그것은 생성물 생산 단계가 종결된 후에 브로스로부터 분리될 수 있다.
- [0136] 여기에 개시된 방법을 사용하여 생산된 3-HP는 화학적으로 다른 유기 화합물로 전환될 수 있다. 예를 들어, 3-HP는 1,3 프로판디올, 가치있는 폴리에스테르 모노머를 형성하기 위해 수소화될 수 있다. 프로판디올은 또한 시험관 내 또는 생체 내에서 옥시도리덕타제 활성을 가진 폴리펩티드를 사용하여 3-HP로부터 생성될 수 있다. 3-HP와 같은 유기산을 수소화하는 것은 어떤 방법, 예를 들어, 숙신산 및/또는 젖산을 수소화하기 위해 사용된 것들을 사용해서도 수행될 수 있다. 예를 들어, 3-HP는 금속 촉매를 사용하여 수소화될 수 있다. 또 다른 예에서, 3-HP는 탈수 반응을 수행하기 위해 어떤 알려진 방법을 사용하여 아크릴산을 형성하도록 탈수화될 수 있다. 예를 들어, 3-HP는 아크릴산을 형성하기 위해 촉매 (예를 들어, 금속 또는 미네랄 산 촉매)의 존재시 가열될 수 있다.
- [0137] 다음 예는 본 발명을 더 잘 설명하기 위해 제공되고 본 발명의 범위를 제한하는 것으로 이해되지 않는다. 특이적 물질이 언급된 규모에 대하여, 그것은 단지 설명의 목적을 위한 것이고 본 발명을 제한하려는 것이 아니다. 당업자는 창의적인 능력의 사용 없이 및 본 발명의 범위를 제한하지 않고 동등한 수단 또는 반응물을 개발할 수도 있다. 많은 변형이 여기에 설명된 과정으로 만들어질 수 있지만, 본 발명의 범위 내에 남아있다고 생각될 것이다. 이러한 변형이 본 발명의 범위 내에 포함되는 것이 발명자의 의도이다.
- [0138] 실시예
- [0139] 배지 및 용액
- [0140] **TE**는 10 mM 염기성 트리스 및 1 mM EDTA, pH 8.0로 구성되었다.
- [0141] **2X YT + amp 플레이트**는 16 g/L 트립톤, 10 g/L 효모 추출물, 5 g/L NaCl, 100 mg/L 앰피실린, 및 15 g/L 박토아가로 구성되었다.
- [0142] **ura 선택 플레이트**는 암모늄 술페이트가 있는 6.7 g 효모 질소 염기, 5 g 카사미노산, 100 mL 0.5 M 숙신산 pH 5, 20 g 노블 아가 (Noble agar), 및 855 mL 탈염수로 구성되었다. 가압 증기 멸균 후, 40 mL 멸균 50% 글루코스 및 2 mL 10 mg/mL 클로라페니콜을 추가하였고 플레이트에 부었다.
- [0143] **ura 선택 배지**는 암모늄 술페이트가 있는 6.7 g 효모 질소 염기, 5 g 카사미노산, 100 mL 0.5 M 숙신산 pH 5, 및 855 mL 탈염수로 구성되었다. 가압 증기 멸균 후, 40 mL 멸균 50% 글루코스 및 2 mL 10 mg/mL 클로라페니콜을 추가하였다.
- [0144] **YP + 10% 글루코스 배지**는 500 mL YP 브로스 및 100 mL 멸균 50% 글루코스로 구성되었다.
- [0145] **YP 브로스**는 10 g/L의 효모 추출물, 20 g/L의 펩톤으로 구성되었다.
- [0146] **YPD 플레이트**는 10 g의 효모 추출물, 20 g의 펩톤, 20 g 박토 아가, 및 960 mL까지 탈염수로 구성되었다. 가압 증기 멸균 후, 40 mL 멸균 50% 글루코스를 추가하였고 플레이트에 부었다.

- [0147] TAE는 4.84 g/L의 염기성 트리스, 1.14 mL/L의 빙초산, 및 2 mL/L의 0.5 M EDTA pH 8.0으로 구성되었다.
- [0148] TBE는 10.8 g/L의 염기성 트리스, 5.5 g/L 붕산, 및 4 mL/L의 0.5 M EDTA pH 8.0으로 구성되었다.
- [0149] LiOAc/TE 용액은 멸균수 8, 1 M LiOAc 1, 및 10X TE 1로 구성되었다.
- [0150] 10X TE (200 mL)는 2.42 g 염기성 트리스, 4 mL 0.5M EDTA, pH 8.0. 5 M HCl을 pH를 7.5로 조정하기 위해 사용하였고 용액을 가압 증기 멸균으로 멸균하였다.
- [0151] PEG/LiOAc/TE 용액은 50% PEG3350 8, 1 M LiOAc 1, 및 10X TE 1로 구성되었다.
- [0152] 50% PEG3350를 150 mL 물에 100 g PEG3350를 추가하고 용해될 때까지 가열하고 교반함으로써 제조하였다. 부피는 물로 최대 200 mL까지 채워지고 가압 증기 멸균으로 멸균하였다.
- [0153] ScD FOA 플레이트는 275 mL 2X-ScD 2X FOA 액체 배지 및 275 mL 2X-ScD 2X FOA 플레이트 배지로 구성되었고, 녹였고 65 °C로 냉각하였다.
- [0154] 2X-ScD 2X FOA 액체 배지는 아미노산이 없는 6.66 g 효모 질소 염기, 1.54 g ura-D0 보충물 (Clontech, Mountain View, CA, USA), 20 g 텍스트로스, 50 mg 우라실, 2 mg 유리딘, 및 2 g 5-FOA (5-플루오로오로트산, 1수화물; Toronto Research Chemicals, North York, ON, Canada) 및 1 L까지 물로 구성되었다. 결과의 용액을 멸균하기 위해 여과하였다.
- [0155] 2X-ScD 2X FOA 플레이트 배지는 11 g 박토 아가 및 275 mL 물로 구성되었다. 결과의 용액을 멸균하기 위해 여과하였다.
- [0156] DM2 배지는 암모늄 술페이트 (5.0 g/L), 마그네슘 술페이트 7수화물 (0.5 g/L), 1염기 칼륨 포스페이트 (3.0 g/L), 미량의 원소 용액 (1 mL/L) 및 비타민 용액 (1 mL/L)으로 구성되었다. 모든 배지 성분이 용해된 후, 배지의 pH를 적절한 염기 (예를 들어, KOH)를 사용하여 원하는 초기 pH로 조정하였다.
- [0157] 미량의 원소 용액은 EDTA (15.0 g/L), 아연 술페이트 7수화물 (4.5 g/L), 망간 클로라이드 탈수화물 (1.0 g/L), 코발트(II)클로라이드 6수화물 (0.3 g/L), 구리(II)술페이트 5수화물 (0.3 g/L), 이나트륨 몰리브덴 탈수화물 (0.4 g/L), 염화 칼슘 탈수화물 (4.5 g/L), 철 술페이트 7수화물 (3 g/L), 붕산 (1.0 g/L), 및 칼륨 이오다이드 (0.1 g/L)로 구성되었다.
- [0158] 비타민 용액은 비오틴 (D-; 0.05 g/L), 칼슘 판토텐네이트 (D+; 1 g/L), 니코틴산 (5 g/L), 미오-이노시톨 (25 g/L), 피리독신 히드로클로라이드 (1 g/L), p-아미노벤조산 (0.2 g/L), 및 티아민 히드로클로라이드 (1 g/L)로 구성되었다.
- [0159] DM1 X-a-gal 플레이트는 DM1 염, 20 g/L 글루코스, 미량의 원소 용액, 비타민 용액, 2 mL/L X-a-gal (16 mg/mL), 20 g/L 아가로 구성되었다.
- [0160] DM1 염 용액은 2.28 g/L 유레아, 3 g/L 1염기 칼륨 포스페이트, 및 0.5 g/L 마그네슘 술페이트 7수화물로 구성되었다.
- [0161] 버터필드 포스페이트 버퍼는 1.25 mL/L의 스톡 용액 (26.22 g/L 칼륨 디히드로젠 포스페이트 및 7.78 g/L 나트륨 카르보네이트) 및 5 mL/L의 마그네슘 클로라이드 용액 (81.1 g/L $MgCl_2 \cdot 6H_2O$)로 구성되었다. 결과의 용액을 멸균하기 위해 가압 증기 멸균하였고, pH를 7.2로 조정하였다.
- [0162] CNB1 셰이크 플라스크 배지는 유레아 (2.3 g/L), 마그네슘 술페이트 7수화물 (0.5 g/L), 1염기 칼륨 포스페이트 (3.0 g/L), 미량의 원소 용액 (1 mL/L) 및 비타민 용액 (1 mL/L), 글루코스 (120.0 g/L), 2-(N-모폴리노)에탄술폰산 (MES) (97.6 g/L)으로 구성되었다. 모든 배지 성분이 용해된 후, 배지의 pH를 적절한 염기 (예를 들어, KOH)를 사용하여 5.8의 초기 pH로 조정하였다.

[0163] 표 0: 프라이머 서열

| 식별자 | SEQ ID NO: | 서열 (5'-3') |
|---------|------------|---|
| 0611166 | 152 | TAAACGACGGCCAGTGAATTCGCGGCGGCCGCGAGTCCATCGGTTCTGTCA |
| 0611167 | 153 | CATAAGAAAATCAAAGACAGAAGGCGCGCCTTTGCTAGCATTTTTGTGTTTGCT GTGT |
| 0611168 | 154 | ACACAGCAAAACACAAAAATGCTAGCAAAGGCGCGCCTTCTGTCTTTGATTTCTTATG |
| 0611169 | 155 | GGGGGAAAGAACTACCAATAGGCCTCCTTTAATTCGGAGAAAATC |
| 0611170 | 156 | GATTTTCTCCGAATTAAAGGAGGCCTATTGGTAGTTCTTTCCCCC |
| 0611171 | 157 | AAAATAAACTAGTAAATAAATTAATTAATTATCTAGAGAGGGGTTATAT |
| 0611172 | 158 | ATATAACCCCTCTCTAGATAATTAATTAATTTATTTTACTAGTTTATTTT |
| 0611173 | 159 | GACCATGATTACGCCAAGCTCCGCGGCGGCCGCCAGTCAAAACCTTCTTCTC |
| 0611174 | 160 | TAAACGACGGCCAGTGAATTCGCGGCGGCCGCTTTGAAGGAGCTTGCCA |
| 0611175 | 161 | CTATTCTTCTCAAATTGCTGTTTAAACGCGTTGAAGATCTATTCTCC |
| 0611176 | 162 | GGAGAATAGATCTTCAACGCGTTTAAACAGCAATTTGAGGAAGGAATAG |
| 0611177 | 163 | AATGTTCATTTTACATTCAGATGTTAATTAAGGTCTAGATGTGTTTGTGTTGTG |
| 0611178 | 164 | CACACAAACAAACACATCTAGACCTTAATTAACATCTGAATGTAATGAAATT |
| 0611179 | 165 | GACCATGATTACGCCAAGCTCCGCGGCGGCCGCAATGCCAAGAGTTATGGGGC |
| 0611184 | 166 | GGCTACCCTATATATGGTGAGCG |
| 0611185 | 167 | GGGTCCAAGTTATCCAAGCAG |
| 0611186 | 168 | CCTTAATTAACCGTAAAGTTGTCTCAATG |
| 0611189 | 169 | CAGCAAAACACAAAAATTCTAGAAAATTTAATTAACATCTGAATGTAATGAAC |
| 0611191 | 346 | CCTTAATTAATTATCCACGGAAGATATGATG |
| 0611195 | 170 | GTTCATTTTACATTCAGATGTTAATTAATTTTCTAGAATTTTTGTGTTTTGCTG |
| 0611196 | 171 | GCTCTAGATAAAATGCCCTCTTATTCTGTCCG |
| 0611199 | 347 | GCTCTAGATAAAATGTCCCAAGGTAGAAAAGC |
| 0611225 | 172 | GACTGGATCATATGACTCC |
| 0611245 | 272 | CGAACCAATTCAAGAAAACCAAC |
| 0611250 | 173 | ACGCCTTGCCAAATGCGGCCGCGAGTCCATCGGTTCTCTGTCAGA |
| 0611251 | 174 | TCAAAGACAGAATTAATTAAGCTCTAGAATTTTTGTGTTTTGCTGTGTT |
| 0611252 | 175 | AACACAAAAATTCTAGAGCTTAATTAATTCTGTCTTTGATTTCTTATG |
| 0611253 | 176 | GTTTAAACCTTTAATTCGGAGAAAATCTGATC |
| 0611254 | 177 | GTTAACGGTACCGAGCTCTAAGTAGTGGTG |
| 0611255 | 178 | AACCGATGGACTCGCGGCCGCATTTGGCAAGGCGTATCTAT |
| 0611256 | 179 | GGCGCCAGCAATTTGAGGAAGGAATAGGAG |
| 0611257 | 180 | AAACTATTAGATTAATTAAGCTCGCGATGTGTTTGTGTTGTGTTTGTGTG |
| 0611258 | 181 | CAACAAACACATCGCGAGCTTAATTAATCTAATAGTTAATCACAGCTTAT |
| 0611259 | 182 | TACATTATGGTAGCGGCCGCGTGTGACATTTTGATCCACTCG |
| 0611260 | 183 | TGGATCAAAATGTCACACGCGGCCGCTACCATAAATGTATGCGTTGAG |
| 0611261 | 184 | GGGCCCTAAAAGTGTTGGTGATTAGA |
| 0611263 | 185 | GCTAGCTCAACAACTCTTTATCAGATTTAGCA |
| 0611264 | 186 | GCTAGCGAGGAAAAAGAGTCTAACCTTTGT |
| 0611266 | 187 | TCGCGATAAAATGTCAACTGTGGAAGATCACTCCT |
| 0611267 | 188 | TTAATTAAGCTGCTGGCGCTTCATCTT |
| 0611268 | 189 | TCGCGATAAAATGTCCAGAGGCTTCTTTACTG |

[0164]

| | | |
|---------|-----|--|
| 0611269 | 190 | TTAATTAACATAAGGTCTCTCACGACAGAG |
| 0611283 | 191 | GTTAACCCGGTTTAAACATAGCCTCATGAAATCAGCCATT |
| 0611284 | 192 | GGGCCCATATGGCGCCCGGGCGTTGAAGATCTATTCTCCAGCA |
| 0611295 | 193 | GTTTAAACGATTGGTAGTTCTTTCCCCCTC |
| 0611296 | 194 | AATAAATTAAGGGCCCTTTATCGCGAGAGGGGGTTATATGTGTAAA |
| 0611297 | 195 | TATAACCCCTCTCGCGATAAAGGGCCCTTAATTTATTTACTAGTTTAT |
| 0611298 | 196 | GTTTAAACTTTTGTAGCACCTCCTGGT |
| 0611376 | 197 | CAGCAAAACACAAAAAGCTAGCTAAAATGTTACGTACCATGTTCAAAA |
| 0611377 | 198 | GAAATCAAAGACAGAAAGCGCGCCTTATGCTGTAACAGCCTGCGG |
| 0611378 | 199 | CAGCAAAACACAAAAAGCTAGC TAAA ATGTTAAGAACCATGTTCAAATC |
| 0611379 | 200 | GAAATCAAAGACAGAAAGCGCGCC TTATGCAGTAACAGCTTGTGGG |
| 0611552 | 273 | TCTGTCCCTTGGCGACGC |
| 0611553 | 274 | CTTTTCAAACAGATAAGTACC |
| 0611554 | 201 | GCATGGTGGTGCAGCGACG |
| 0611555 | 202 | GGTGTGCATTTGCTGCTG |
| 0611622 | 275 | ATGGGCTGACCTGAAAATTC |
| 0611631 | 203 | TGTATACAGGATCGAAGAATAGAAG |
| 0611632 | 204 | GAACGTCTACAACGAGGTGAAC |
| 0611661 | 205 | GGCGCGCCTCGCGATAAAATG |
| 0611662 | 206 | AGGGCCCTTAATTAATTATGCAGTAACAGCTT |
| 0611717 | 207 | CGCTACGATACGCTACGATA |
| 0611718 | 208 | CTCCCTTCCCTGATAGAAGG |
| 0611814 | 209 | GGGGAGCAATTTGCCACCAGG |
| 0611815 | 210 | CTCCTTCATTTAACTATACCAGACG |
| 0611816 | 211 | GACAGATGTAAGGACCAATAGTGTC |
| 0611817 | 212 | CCATATCGAAATCTAGCCCGTCC |
| 0611828 | 213 | CATAAGAAAATCAAAGACAGAAAGCGCGCCTTTGCTAGCTTTTGTGTTT TGCTGTGT |
| 0611954 | 348 | GCTAGCTAAAAATGTTTGGTAATATTTCCCA |
| 0611957 | 349 | TTAATTAACATTTATCTAATGATCCTC |
| 0611997 | 350 | TCTAGATAAAATGTCTATTAGTGAAAAATATTTTCTCAAG |
| 0611998 | 351 | TTAATTAACTTTTAAATTTTGGAAAAAGCTTGATCAATAATGG |
| 0612055 | 214 | ATTGGACACAACATTATAT |
| 0612150 | 238 | ACGCGTCGACTCGACATTTGCTGCAACGGC |
| 0612151 | 239 | CTAGTCTAGATGTTGTTGTTGTTGTCGTT |
| 0612271 | 215 | GTAAAACGACGGCCAGTGAATTGTTAACATAGGCTCCAACATCTCG |
| 0612272 | 216 | GAACCCGATGGACTCGCGGCCGC GTGGGATATTGGAAG |
| 0612273 | 217 | GGAGGTGCTACAAAAGGAATTC |
| 0612274 | 218 | GACCATGATTACGCCAAGCTCCGCGGAGTCAAACCTTCTCTACC |
| 0612275 | 219 | GTAAAACGACGGCCAGTGAATTC TTTGAAGGAGCTTGCCAAGAAAC |
| 0612276 | 220 | CCTATTCTTCTCAAATTGCTG |
| 0612277 | 221 | ATAACTCTTGGCATTGCGGCCGCAAAGTTAGTTAGAGC |
| 0612278 | 222 | ATGACCATGATTACGCCAAGCTCCGCGGCAAAGACGGTGTATTAGTGCT TG |
| 0612356 | 223 | CACAAAACACAAAAACACACAGCTAGCAAAGCGCGCCATCTAATAG TTAATCACAGCTTA |
| 0612357 | 224 | GCCGTTGCAGCAAATGTCGAGGCCTGTGTGACATTTTGATCCACTCG |
| 0612358 | 225 | CGAGTGGATCAAAATGTCACACAGGCCTCGACATTTGCTGCAACGGC |
| 0612359 | 226 | CATTTTACATTGATGTTAATTAATTATCTAGATGTTGTTGTTGTTGTCG |
| 0612360 | 227 | CGACAACAACAACATCTAGATAATTAATTAACATCTGAATGTAAAATG |
| 0612361 | 228 | GCTCTAACTAACTTGGCGGCCGCTTTATTATAAAATTATATATTCTT |
| 0612366 | 276 | GCTGAAAATATCATTGAGCAT |
| 0612367 | 277 | ACTGTTGATGTCGATGCC |
| 0612378 | 229 | AACACACAAACAAACACAGCTAGCTAAAATGTTAAGAACTATGTTTA |
| 0612379 | 230 | GATTAACTATTAGATGGCGCGCCTTATGCAGTAAGTCTTGTGGA |
| 0612470 | 231 | CGACGGCCAGTGAATTCGTTAACCCGTTTCGATGGGATTCCC |
| 0612471 | 232 | CAGGAACCGATGGAATCGCGGCCGCTCCCTTCTCTAAATGGACTGC |
| 0612472 | 233 | TATATAATTTTATAATAAAAGCGGCCGCACCAGGGGTTAGTGAAGTC |

[0165]

| | | |
|---------|-----|---|
| 0612473 | 234 | CATGATTACGCCAAGCTCCGCGGCCATAACTGACATTTATGGTAAGG |
| 0612579 | 278 | TCTGAATGCAGTACGAGTTG |
| 0612695 | 240 | CAAAACACAGCAAAACACAAAAGCTAGCATGTATAGAACCTTGATGAG |
| 0612698 | 242 | CAAAACACACAAACACACAGCTAGCATGTACAGAACGTTAATGTCTGC |
| 0612724 | 241 | CAAAGACAGAAGGCGCGCCTTATAAGATGGTTCTCGCTGG |
| 0612725 | 243 | GTGATTAACTATTAGATGGCGCGCCTTACAGGATGGTTCTGGCAGG |
| 0612794 | 279 | GAAGGGGGTCCAAGTTATCC |
| 0612795 | 235 | CCAACAATCTTAATTGGTGAC |
| 0612891 | 236 | GGCTGTTACCGCCTAATTAA |
| 0612893 | 237 | GTTCTTAACATTTTAGCTAGCTG |
| 0612908 | 280 | GATATGGGCGGTAGAGAAGA |
| 0612909 | 281 | GCTCCTTCAAAGGCAACACA |
| 0612910 | 282 | TGAACTATCACATGAACGTA |
| 0612911 | 283 | TCAAGGTAGGGTCACTTAAC |
| 0613034 | 284 | TGATTCCTTCAATCACAGGT |
| 0613035 | 285 | AACCGACCTATCGAATGCCT |
| 0613178 | 286 | ACCATGATTACGCCAAGCTTGGTACCTTGGGGTTTACGCTTACAGCGTA CT |
| 0613179 | 287 | TCAACGCCCCCGGGGGATCTGGATCCGCGGCCGCAAGAAATTCCTTTCTT TTCCCTTTATA |
| 0613180 | 288 | AACCGATGGACTCCTCGAGGGATCCGCGGCCGCGCATAACAAAATTGTG CCTAACCCA |
| 0613181 | 289 | ACGACTCACTATAGGGCGAATTGGGCCCGGAAAGGAGAGAGAAAA GGAGA |
| 0613183 | 290 | TATGACCATGATTACGCCAAGCTTGGTACCTCCTACAGAAGTAGCACCA GCACCA |
| 0613184 | 291 | GGAACCGATGGACTCCTCGAGGGATCCGCGGCCGCGAGACTACCGTGTA AATAAGAGTACC |
| 0613185 | 292 | TCTTCAACGCCCCGGGGATCTGGATCCGCGGCCGCTTTGATATAAAC GCTTCTATAATA |
| 0613186 | 293 | GTAAACGACTCACTATAGGGCGAATTGGGCCCAACATCTGCTGCTGTA ATATATTCA |
| 0613241 | 294 | GCATGTCTGTAACTCTCAAAACCAT |
| 0613242 | 295 | TCCATAATCCAATAGATCATCCC |
| 0613243 | 296 | AACACAATGGAACCAACCTAGT |
| 0613416 | 297 | ATTTACACGGTAGTCTGCGGCCGCCATAGCCTCATGAAATCAGCC |
| 0613417 | 298 | CTCTAGGTTCACTGGTTGTTTCTTGGGCTGCCTCCTTCAA |
| 0613418 | 299 | TTGAAGGAGGCAGCCCAAGAAACAACCAAGTGAACCTAGAG |
| 0613419 | 300 | TATTATAGAAGCGTTTATATCAAATGCGGCCGCGGATCCAGATCCCCCG GGGCGTT |
| 0613550 | 301 | AATGATCAACTTGAGAGGTA |
| 0613551 | 302 | CAGGTCTGTTACATAAAGCA |
| 0613688 | 303 | GTTTACGCTCAAACTCCTCC |
| 0613689 | 304 | GGTACATGAAGCAGGCTTTGAAGG |
| 0613695 | 305 | GATTGTGTCCGTCAGCCTTTGCTC |
| 0613746 | 306 | TACCATATTTTCAGAGGATCA |
| 0613747 | 307 | AGGATGTTCTTGCCTGCAAGT |
| 0614233 | 308 | GATGATATAGTTGATGCTTTCCAAAG |
| 0614626 | 344 | AGGGTACCTTAGTACGAAGG |
| 0614627 | 345 | CTATTCTTACGATGAAGGCG |
| CM647 | 262 | AATGATCCATGGTCCGAGAG |
| oACN48 | 309 | GGGCCCTTCATTTACGAAATAAAGTGCCGCGG |
| oACN49 | 310 | GCGGCCGCAATAAATTTAAAAATAACGATATCAAAATTC |
| oACN50 | 311 | CGACGCCAAAGAAGGGCTCAGCCGAAAAAG |
| oACN51 | 312 | CCATTTCTTTTTCGGCTGAGCCCTTCTTTGGC |
| oACN52 | 313 | GCGGCCGCAATAACCTCAGGGAGAACTTTGGC |
| oACN53 | 314 | GAGCTCCCAACAAGGTCTCAGTAGGATCG |
| oACN54 | 315 | CGCCATAAGGAGAGGCTGTAGATTTGTC |

[0166]

| | | |
|---------|-----|--|
| oACN55 | 316 | CCAGGACATCTTGCTTGAATGTCG |
| oANN1 | 317 | GTTCCATCGGGCCCTAAAGGTCTCTCAGCAG |
| oANN2 | 318 | CAAACACA TCGCGA TAAATGTCCAGAGGCTTC |
| oANN5 | 319 | GCAAGACCTTGGATCTGAAGGG |
| oANN6 | 320 | CGAACCAATTCAAGAAAACCAACAG |
| oANN7 | 321 | GGGCCC GTCCCTTGGCGACGCCCTGATC |
| oANN8 | 322 | GCGGCCGCTATTTTGTGTTTGTGCTGTTTTG |
| oANN9 | 323 | GCGGCCGCC ATCTGAATGTAAATGAACATTAAATG |
| oANN10 | 324 | GAGCTCCCCAGTTGTTGTTGCAATTAAC |
| oANN11 | 325 | GAAGAGACGTACAAGATCCGCC |
| oANN12 | 326 | TAGGAATGGTGCATCATCCAAC |
| oANN13 | 327 | TTCTTATCTGAAAACCTCCGAGTTCGCAAGAAGGTTGAAG |
| oANN14 | 328 | TTATTGAAATTAATCAAGGATTCAAGTTGAACATACAATTACTG |
| oANN15 | 329 | TTCCAAAAAAGTTTCGTGAGTTGATGGTTGTATGATTATGG |
| oANN16 | 330 | CCTCTACCACCACCACCAATG |
| oANN20 | 331 | GGGAAGAACTAAGAAGAAGTATG |
| oCA405 | 260 | GCAACTGATGTTACGAATGCG |
| oCA406 | 261 | TTGCCGTTGCAGCAAATCTC |
| oHJ1 | 332 | CAATCCTTCTAGAGAGGGGGTTATATGTGTAAATATAGAGTTTG |
| oHJ2 | 333 | TTCTACCCCTCGAGATTGGTTCTTTCCCCCTCTCAAG |
| oHJJ116 | 334 | GTGGTCAATCTAGAAAATATGACTGACAAAATCTCCCTAGGTAC |
| oHJJ117 | 335 | CAATTTTGGAGCTGATTCCAAATCGTAAAC |
| oLJ28 | 264 | GCACCAAGAGCAGTTTTCCCATCTATTG |
| oLJ29 | 265 | CCATATAGTTCTTTTCTAACATCAACATCACACTTC |
| oLJ30 | 266 | GGAGAATAGATCTTCAACGCTTTAATAAAGTAGTTTG |
| oLJ31 | 267 | CGTGTGCGCAATAAAACCAATGAC |
| oLJ43 | 336 | CAAGAGTATCCCATCTGACAGGAACCGATGG |
| oLJ44 | 337 | GCTGGAGAATAGATCTTCAACGCCCCG |
| oLJ45 | 338 | CGGAGAAGGCGTATAAAAAGGACACGGAG |
| oLJ46 | 339 | GGATAAAAGTATTACATACGTACAGGATTGTGTATTAGTGATTTCG |
| oLJ57 | 340 | CCTCCAGTGTTTTCTCTCTGTCTCTTTGTTTTTTTTTC |
| oJY11 | 268 | CCTCGAAGAGCTTGAATTTG |
| oJY12 | 269 | GTAGTGAATGTCCGGATAAG |
| oJY13 | 270 | GCAAGGTCATGAGGTTAAAG |
| oJY14 | 271 | AACACTTATGGCGTCTCCTC |
| oJY44 | 341 | GGTTAATTAATTTATTTGTACATAAAACCACATAAATGTAAAAGC |
| oJY45 | 342 | GAATTCCTTTAATTCGGAGAAAATCTGATCAAGAG |
| WG26 | 263 | ACGGCAGTATACCCTATCAGG |

[0167]

[0168] 잡다한 서열

[0169] 프로모터: 실시예에 설명된 PDC, TDH3, EN01, PGK1 프로모터는 SEQ ID NO: 244, 245, 246, 및 247의 아이. 오리엔탈리스 서열로부터 각각 유도된다. 종결자: 실시예에 설명된 The TAL, TKL, RKI, 및 PDC 종결자는 SEQ ID NO: 248, 249, 250, 및 251의 아이. 오리엔탈리스 서열로부터 각각 유도된다. URA3 프로모터, ORF, 및 실시예에 설명된 종결자는 SEQ ID NO: 252의 아이. 오리엔탈리스 서열로부터 유도된다.

[0170] 실시예 1A: 3-HP에 기초한 숙주 효모 세포의 선택

[0171] 한 세트의 야생형 효모 균주를 3-HP의 존재시 그것들의 성장하는 능력에 대해 테스트하였다.

[0172] 1차 스크리닝 과정에 이용되는 3-HP 농도의 범위를 7개의 야생형 효모 균주의 다양한 수준의 3-HP를 함유하는 배지에서 성장하는 능력을 평가함으로써 결정하였다. 이 실험에 사용된 3-HP를 W004/076398에 설명된 바와 같이 200 psi 및 175 °C에서 수성 아크릴산 용액 (30%) 및 CO₂ 촉매로부터 화학적으로 합성하였다. 잔여 아크릴산을 상온에서 이소프로필로 역류 추출을 사용하여 제거하였다 (W005/021470 참조).

[0173] 세포를 YPD 플레이트에 스트리킹 (streak)하였고 하룻밤 동안 키웠다. 약 4의 OD600으로 세포 슬러리 (slurry)를 YPD 배지, pH 3.0에서 만들었고 이 슬러리를 0.05의 OD600으로 다양한 농도의 3-HP (pH 3.9)를 함유하는 미세액가 웰을 접종하기 위해 사용하였다. 플레이트를 가스 투과성 막으로 덮었고 30 °C/300 RPM 셰이커에서 하룻밤 동안 배양하였다. 각각의 웰에 대한 최적의 밀도를 GENios 모델 플레이트 리더에서 600 nm의 파장으로 측정하였고, 플레이트를 성장에 대하여 시각적으로 관찰하였다. (125 g/L)로 자란 균주 중 하나 이상이 선택되는 가장 높은 3-HP 농도를 1차 스크리닝 과정에 대한 상한 범위로서 선택하였다.

[0174] 1차 스크리닝

[0175] 1차 스크리닝 과정을 위하여, 89개의 야생형 효모 균주를 찾는 범위에 대하여 사용된 같은 프로토콜을 사용하여 0 g/L, 75 g/L, 100 g/L, 또는 125 g/L 3-HP (pH 3.9)의 미세액가 플레이트에서 성장에 대하여 스크리닝하였다. 신선한 YPD 플레이트를 각각의 균주를 위해 사용하였고, 약 4의 OD600의 슬러리를 YPD 배지, pH 3.0에서 만들었다. 슬러리를

[0176] 슬러리를 0.05의 OD600으로 각각의 웰을 접종하기 위해 사용하였다. 플레이트를 가스 투과성 막으로 덮었고 30

℃/300 RPM 셰이커에서 하룻밤 동안 배양하였다. 각각의 웰에 대한 최적의 밀도를 GENios 모델 플레이트 리더에서 600 nm의 파장으로 측정하였고, 플레이트를 성장에 대하여 시각적으로 관찰하였다. 유사한 프로토콜을 0 g/L, 30 g/L, 45 g/L, 및 60 g/L의 젖산 농도에서 성장을 평가하기 위해 실행하였다. 표 1은 성장이 관찰된 3-HP 또는 젖산의 가장 높은 농도를 요약한다.

[0177] 100 g/L 3-HP에서 성장할 수 있고 75 g/L 3-HP에서 잘 자란 15개의 균주를 확인하였다 (표 1, 세트 B). 균주를 더 줄이기 위해, 두 번째 미세역가 플레이트 테스트를 수행하였다. 이 테스트는 첫 번째와 유사하지만, 100 g/L, 112.5 g/L, 및 125 g/L의 3-HP 농도 (pH 3.9)를 이용한다. 이 테스트로부터, 75 g/L 3-HP에서 잘 자라거나 75 및 112.5 g/L 둘 다에서 일부 성장을 보이는 3-HP 11개의 균주를 확인하였다 (표 1, 세트 C). 이들 11개의 균주를 2차 스크리닝으로 진행하였다. 75 g/L 3-HP에서 나쁜 성장을 갖고 112.5 g/L 3-HP에서 성장하지 않는 4개의 균주는 2차 스크리닝으로 진행하지 않았다. 2차 스크리닝으로 진행하지 않는 균주는 상업적 발효 공정에서 경제적으로 낮은 성능을 나타내는 것이 예상된다. 하지만, 하나 이상의 이러한 균주가 그럼에도 불구하고 상업적으로 실행 가능한 발효 공정에 대한 최소의 필요조건을 만족시키는 것은 가능하다.

[0178] 표 1: 3-HP 또는 젖산에서 성장

| 호모 균주 | 균주 수집 # | 세트 | 1 차 스크리닝: 3-HP (g/L) | 1 차 스크리닝: 젖산 (g/L) |
|---|-------------------|------|-------------------------------|-----------------------------|
| <i>불레로미세스 알부스</i> (<i>Bulleromyces albus</i>) | ATCC 96272 | | 0 g/L 에서 성장 안함 (NG) | NG |
| <i>칸디다 블랑키이</i> (<i>Candida blankii</i>) | ATCC 18735 | | 0 | 30 |
| <i>칸디다 보이디니이</i> (<i>Candida boidinii</i>) | PYCC 70-104 | | 0 | 30 |
| <i>칸디다 카테눌라타</i> (<i>Candida catenulata</i>) | ATCC 20117 | | 0 | NG |
| <i>칸디다 에첼시이</i> (<i>Candida etchellsii</i>) | PYCC 60-8 | | 0 | 30 |
| <i>칸디다 파마타</i> (<i>Candida famata</i>) | ATCC 20284 | | 0 | 0 |
| <i>칸디다 플루비아틸리스</i> (<i>Candida fluvialilis</i>) | ATCC 38623 | | 0 | 0 |
| <i>칸디다 지오차레스</i> (<i>Candida geochares</i>) | 카길 (Cargill) 1978 | | 75 | 60 |
| <i>칸디다 굴리에르몽디이</i> (<i>Candida guilliermondii</i>) | ATCC 20118 | | 0 | 0 |
| <i>칸디다 인테르메디아</i> (<i>Candida intermedia</i>) | ATCC 20178 | | 0 | 0 |
| <i>칸디다 케피르</i> (<i>Candida kefir</i>) | ATCC 44691 | | 75 | 30 |
| <i>칸디다 락티스콘덴시</i> (<i>Candida lactiscondensi</i>) | ATCC 96927 | | 0 | NG |
| <i>칸디다 람비카</i> (<i>Candida lambica</i>) | ATCC 38617 | B, C | 75 | 45 |
| <i>칸디다 메타노소르보사</i> (<i>Candida methanosorbosa</i>) | ATCC 20361 | | 0 | NG |
| <i>칸디다 밀레리</i> (<i>Candida milleri</i>) | ATCC 60591 | | 0 | 30 |
| <i>칸디다 밀레리</i> | ATCC 60592 | | 0 | 45 |
| <i>칸디다 나오덴드라</i> (<i>Candida naeodendra</i>) | ATCC 56465 | | 0 | NG |
| <i>칸디다 파랍실로시스</i> | ATCC 20179 | | 0 | 0 |

[0179]

| | | | | |
|--|---------------------------|-----------|-----|----|
| (<i>Candida parapsilosis</i>) | | | | |
| 칸디다 피그날리아 (<i>Candida pignaliae</i>) | ATCC 36592 | | 0 | NG |
| 칸디다 슈도람비카 (<i>Candida pseudolambica</i>) | ATCC 96309 | | 0 | 0 |
| 칸디다 루고사 (<i>Candida rugosa</i>) | ATCC 20306 | | 0 | 45 |
| 칸디다 쉬하타에 (<i>Candida shehatae</i>) | NCYC 2389 | | 0 | 0 |
| 칸디다 소노렌시스 (<i>Candida sonorensis</i>) | ATCC 32109 | A | 0 | 30 |
| 칸디다 소르보필라 (<i>Candida sorbophila</i>) | 카길 1973 | | 75 | 45 |
| 칸디다 소르복실로사 (<i>Candida sorboxylosa</i>) | ATCC 24120 | | 0 | 30 |
| 칸디다 소로시보란스 (<i>Candida sorosivorans</i>) | ATCC 38619 | B | 100 | 60 |
| 칸디다 테누이스 (<i>Candida tenuis</i>) | MUCL 47216 또는 MUCL 31253B | | 0 | NG |
| 칸디다 발리다 (<i>Candida valida</i>) | ATCC 28525 | B | 75 | 45 |
| 칸디다 반더왈티이 (<i>Candida vanderwaltii</i>) | MUCL 300000 | | 0 | 45 |
| 칸디다 켄플리니나 (<i>Candida zemplinina</i>) | PYCC 04-501 | A,B, C | 100 | 60 |
| 칸디다 제이라노이데스 (<i>Candida zeylanoides</i>) | ATCC 20347 | | 75 | 45 |
| 시테로미세스 매트리티텐시스 (<i>Citeromyces matritensis</i>) | ATCC 34087 | | 0 | NG |
| 데바리오미세스카스텔리이 (<i>Debaryomyces castellii</i>) | PYCC 70-1022 | | 0 | 30 |
| 데바리오미세스 한센니이 (<i>Debaryomyces hansenii</i>) | ATCC 90624 | | 0 | NG |
| 데바리오미세스 폴리모르푸스 (<i>Debaryomyces polymorphus</i>) | ATCC 20280 | | 0 | 0 |
| 데케라 아노말라 (<i>Dekkera anomala</i>) | ATCC 20277 | | 0 | 0 |
| 데케라 람비카 (<i>Dekkera lambica</i>) | ATCC 10563 | | 0 | 0 |
| 히포피치아 부르토니이 (<i>Hypocichia burtonii</i>) | ATCC 20030 | | 0 | 0 |

[0180]

| | | | | |
|---|----------------------------|-----------|-----|----|
| (<i>Hyphopichia burtonii</i>) | | | | |
| 이샤트첸키아 오리엔탈리스 (<i>Issatchenkia orientalis</i>) | ATCC 24210 | B, C | 100 | 45 |
| 이샤트첸키아 오리엔탈리스 | ATCC 60585 | B, C | 100 | 45 |
| 이샤트첸키아 오리엔탈리스 | ATCC PTA-6658 | A,B, C | 100 | 60 |
| 이샤트첸키아 오리엔탈리스 | CD1822 (카길 수집-하기 설명 참조) | A,B, C | 100 | 60 |
| 클루이베로미세스 락티스 (<i>Kluyveromyces lactis</i>) | ATCC 8585 | A | 0 | 0 |
| 클루이베로미세스 막시아누스 (<i>Kluyveromyces marxianus</i>) | ATCC 52486 | A | 75 | 45 |
| 클루이베로미세스 써모톨레란스 (<i>Kluyveromyces thermotolerans</i>) | ATCC 52709 | | 0 | 30 |
| 코다메아 오메리 (<i>Kodamaea ohmeri</i>) | ATCC 20282 | | 0 | 45 |
| 클루이베로미세스 야로위이 (<i>Kluyveromyces yarrowii</i>) | ATCC 36591 | | NG | NG |
| 리포미세스 스타르케이 (<i>Lipomyces starkeyi</i>) | ATCC 12659 | | 0 | 0 |
| 리포미세스 테트라스포루스 (<i>Lipomyces tetrasporus</i>) | ATCC 56306 | | NG | NG |
| 메치니코비아 풀체리마 (<i>Metschnikowia pulcherrima</i>) | ATCC 9889 | | 0 | 0 |
| 믹소지마 클루이베리 (<i>Myxozyma kluyveri</i>) | ATCC 76214 | | NG | NG |
| 네마토스포라 코릴리 (<i>Nematospora coryli</i>) | ATCC 20292 | | 0 | NG |
| 파치솔렌 탄노필루스 (<i>Pachysolen tannophilus</i>) | NCYC 614 | | 0 | 30 |
| 피치아 아노말라 (<i>Pichia anomala</i>) | ATCC 2102 | | 0 | 0 |
| 피치아 페르멘탄스 (<i>Pichia fermentans</i>) | ATCC 28526 | | 0 | 30 |
| 피치아 플럭숨 (<i>Pichia fluxuum</i>) | ATCC 28778 | | NG | NG |
| 피치아 자다니이 (<i>Pichia jadinii</i>) | ATCC 9950 | | 0 | 0 |

[0181]

| | | | | |
|---|---------------|------|-----|----|
| 피치아 멤브라니파시엔스 (<i>Pichia membranifaciens</i>) | NCYC 2696 | B,C | 125 | 60 |
| 피치아 나카세이 (<i>Pichia nakasei</i>) | ATCC 24116 | | 0 | 0 |
| 피치아 실비콜라 (<i>Pichia silvicola</i>) | ATCC 16768 | | 0 | 0 |
| 피치아 스티피티스 (<i>Pichia stipites</i>) | CBS 6054 | | 0 | 0 |
| 피치아 스트라스부르겐시스 (<i>Pichia strasburgensis</i>) | ATCC 34024 | | 0 | 0 |
| 피치아 탄니콜라 (<i>Pichia tannicola</i>) | ATCC 2261 | | 0 | 0 |
| 피치아 톨레타나 (<i>Pichia toletana</i>) | ATCC 58362 | | NG | NG |
| 사카로미세스 세레비시애 (<i>Saccharomyces cerevisiae</i>) | ATCC 96784 | | 75 | 30 |
| 사카로미세스 바야누스 (<i>Saccharomyces bayanus</i>) | ATCC 90739 | | 0 | 45 |
| 사카로미세스 불데리 (<i>Saccharomyces bulderi</i>) | MYA-402 | B, C | 100 | 60 |
| 사카로미세스 불데리 | MYA-404 | B,C | 100 | 45 |
| 사카로미세스 캡슐라리스 (<i>Saccharomyces capsularis</i>) | ATCC 20033 | | 0 | 0 |
| 사카로미세스 세레비시애 | CEN-PK 113-7D | A,B | 100 | 45 |
| 사카로미세스 루드윅기이 (<i>Saccharomyces ludwigii</i>) | NCYC 734 | | 75 | 45 |
| 사카로미콕시스 크라테겐시스 (<i>Saccharomycopsis crataegensis</i>) | MUCL 44417 | | 0 | 0 |
| 사카로미콕시스 자벤시스 (<i>Saccharomycopsis javensis</i>) | MUCL 31237 | | 0 | 45 |
| 사카로미콕시스 비니 (<i>Saccharomycopsis vini</i>) | NRRL Y-7290 | | 0 | 0 |
| 사카로미세스 우바룸 (<i>Saccharomyces uvarum</i>) | ATCC 76514 | | 0 | 0 |
| 쉬조사카로미세스 자포니쿠스 (<i>Schizosaccharomyces japonicas</i>) | ATCC 10660 | | 0 | 0 |

[0182]

| | | | | |
|---|-------------|-----|-----|----|
| 쉬조사카로미세스 폼베 (<i>Schizosaccharomyces pombe</i>) | NCYC 535 | B,C | 100 | 60 |
| 톨루아스포라 델브루에키이 (<i>Torulaspora delbrueckii</i>) | ATCC 52714 | | 75 | 0 |
| 톨루아스포라 프레토리엔시스 (<i>Torulaspora pretoriensis</i>) | ATCC 42479 | B | 75 | 0 |
| 위커라미엘라 옥시덴탈리스 (<i>Wickerhamiella occidentalis</i>) | CBS 8452 | | 0 | 30 |
| 야마다지마 굴리에르몽디이 (<i>Yamadazyma guilliermondii</i>) | ATCC 90197 | | 0 | 0 |
| 야마다지마 하플로필라 (<i>Yamadazyma haplophila</i>) | ATCC 20321 | | 0 | 0 |
| 야마다지마 스티피티스 (<i>Yamadazyma stipites</i>) | ATCC 201225 | | 0 | 0 |
| 야로위아 리폴리티카 (<i>Yarrowia lipolytica</i>) | ATCC 46330 | | 0 | 45 |
| 지고사카로미세스 바일리이 (<i>Zygosaccharomyces bailii</i>) | ATCC 36946 | | 75 | 45 |
| 지고사카로미세스 비스포루스 (<i>Zygosaccharomyces bisporus</i>) | NCYC 3134 | | 0 | 45 |
| 지고사카로미세스 콤부첸시스 (<i>Zygosaccharomyces kombuchaensis</i>) | NCYC2897 | | 0 | 60 |
| 지고사카로미세스 렌투스 (<i>Zygosaccharomyces lentus</i>) | NCYC 2928 | B,C | 100 | 45 |
| 지고사카로미세스 로옥시이 (<i>Zygosaccharomyces rouxii</i>) | ATCC 34890 | | 75 | 30 |

[0183]

[0184]

상기 테스트된 아이. 오리엔탈리스 균주 CD1822를 글루코스 제한 케모스태트에서 91일 동안 아이. 오리엔탈리스 ATCC PTA-6658를 발달시킴으로써 생성하였다. 시스템은 DM 배지에서 15 g/L 텍스트로스가 제공되고, 공급 배지에서 추가된 젖산을 가지고 pH=3에서 0.06 h^{-1} 의 회석 속도로 작동된다. 약 $2 \text{ mmol L}^{-1} \text{ h}^{-1}$ 의 낮은 산소 전달 속도, 및 공기 포화도 0%로 일정하게 남아있는 용존 산소 농도를 갖는 조건을 유지하였다. 최종 시점의 단일 콜로니 분리체는 두 개의 셰이크 플라스크 검정으로 특징화된다. 첫 번째 검정에서, 균주는 DM1 한정 배지에서 pH 조정 없이 25 g/L 총 젖산의 존재시 그것들의 글루코스를 에탄올로 발효시키는 능력을 특징으로 한다. 두 번째 검정에서, 분리체의 성장 속도를 DM1 한정 배지에서 pH 조정 없이 25, 32 및 45 g/L의 총 젖산의 존재시 측정하였다. 균주 CD1822는 측정된 발효 속도 및 성장 속도에 기초하여 선택된 단일 분리체였다. 아이. 오리엔탈리스를 발달시키는 다른 방법은 예를 들어, 미국 특허 번호 7,629,162에 설명된 바와 같이 선택적 조건 하에 반복된 연속적 계대 배양을 포함하지만, 이에 제한되지 않는다. 이러한 방법은 상기 설명된 글루코스 제한 케모스태트 방법의 사용을 대신해서, 또는 이에 더하여 사용될 수 있다. 당업자에 의해 인정될 수 있는 바와 같이, 균주는 개선된 3-HP 내성을 발달시키기 위해 젖산 대신에 추가된 3-HP의 존재시 유사한 진화 과정을 사용하여 발생할 수 있다. 추가적으로, 균주는 여기에 설명된 바와 같이, 선택 전에 돌연변이될 수 있다 (예컨대, 실시예 1B).

[0185]

2차 스크리닝

[0186]

2차 스크리닝의 첫 번째 부분에 대하여, 성장 속도를 0 g/L, 35 g/L, 또는 75 g/L 3-HP를 함유하는 YPD 배지에서 pH 3.9에서 측정하였다. 셰이크 플라스크를 6 내지 10의 OD600으로 하룻밤 동안 성장된 시드 플라스크로부터 수확된 바이오매스로 접종하였다. 250 mL 배플 (baffled) 성장 속도 플라스크 (50 mL 작업 부피)를 0.1의 OD600으로 접종하였고 250 rpm 및 30 °C에서 키웠다. 샘플을 시간 단위의 검정을 통해 채취하였고 OD600을 통해 바이오매스 성장에 대하여 분석하였다. 결과의 OD600 데이터를 플로팅하였고 성장 속도를 확립하였다.

[0187] 표 2: 3-HP에서 성장 속도 (μ)

| 균주 | 0 g/L 3-HP | 35 g/L 3-HP | 75 g/L 3-HP |
|--------------------------|------------|-------------|-------------|
| 이스트첸키아 오리엔탈리스 ATCC 60585 | 0.56 | 0.51 | 0.29 |
| 이스트첸키아 오리엔탈리스 CD1822 | 0.62 | 0.52 | 0.28 |
| 칸디다 람비카 ATCC 38617 | 0.65 | 0.53 | 0.30 |
| 칸디다 발리다 | 0.51 | 0.32 | 0.14 |
| 이스트첸키아 오리엔탈리스 PTA-6658 | 0.61 | 0.53 | 0.32 |
| 이스트첸키아 오리엔탈리스 24210 | 0.58 | 0.51 | 0.26 |
| 사카로미세스 볼데리 MYA 402 | 0.53 | 0.45 | 0.28 |
| 피치아 멤브라니파시엔스 | 0.41 | 0.39 | 0.32 |
| 사카로미세스 볼데리 MYA 404 | 0.55 | 0.44 | 0.27 |
| 쉬조사카로미세스 폼베 | 0.41 | 0.35 | 0.21 |

[0188]

[0189] 2차 스크리닝의 두 번째 부분에 대하여, 글루코스 소비를 100 g/L 및 0 g/L, 35 g/L, 또는 75 g/L 3-HP 글루코스를 함유하는 YPD 배지에서 pH 3.9에서 같은 10개의 균주에 대하여 측정하였다. 셰이크 플라스크를 6 내지 10의 OD600으로 하룻밤 동안 성장된 시드 플라스크로부터 수확된 바이오매스로 접종하였다. 250 mL 배플 당분해 검정 플라스크 (50 mL 작업 부피)를 0.1의 OD600으로 접종하였고 250 rpm 및 30 °C에서 키웠다.

[0190] 샘플을 시간 단위의 검정을 통해 채취하였고 Yellow Springs Instruments (YSI)의 2700 생화학 분석기를 사용하여 글루코스 소비에 대하여 분석하였다. 결과의 데이터를 플로팅하였고 글루코스 소비 속도를 확립하였다.

[0191] 표 3: 글루코스 소비 속도 (g/L/hr) in 3-HP

| 균주 | 0 g/L 3-HP | 35 g/L 3-HP | 75 g/L 3-HP |
|--------------------------|------------|-------------|-------------|
| 이스트첸키아 오리엔탈리스 ATCC 60585 | 5.5 | 4.2 | 3.3 |
| 이스트첸키아 오리엔탈리스 CD1822 | 5.5 | 4.2 | 4.2 |
| 칸디다 람비카 ATCC 38617 | 4.2 | 4.2 | 3.5 |
| 칸디다 발리다 | 5.5 | 2.2 | 2.1 |
| 이스트첸키아 오리엔탈리스 PTA-6658 | 5.5 | 4.2 | 4.1 |
| 이스트첸키아 오리엔탈리스 24210 | 4.2 | 4.2 | 3.8 |
| 사카로미세스 볼데리 MYA 402 | 4.2 | 4.2 | 4.0 |
| 피치아 멤브라니파시엔스 | 0.4 | 2.1 | 1.2 |
| 사카로미세스 볼데리 MYA 404 | 4.2 | 4.2 | 3.8 |
| 쉬조사카로미세스 폼베 | 2.5 | 3.1 | 2.0 |
| 지고사카로미세스 렌투스 | 3.4 | 0.8 | 0.3 |

[0192]

[0193] 4개의 균주 (피. 멤브라니파시엔스, 에스. 폼베 (S. pombe), 씨. 발리다 (C. valida), 및 제트. 렌투스 (Z. lentus))는 경제적 발효 공정에 필요한 75 g/L 3-HP (pH 3.9) 조건 하에 2.5 g/L/hr 글루코스 이용 속도를 달성하지 않았다.

[0194] 3-HP에서 선도 균주를 확인하기 위해, 균주 성능을 3개의 범주로 분류하였다. 이들 범주 중 둘은 성장 속도의 다른 양태에 기초하였다: 1) 가장 높은 산 농도에서 성장 속도 및 2) 산 농도에 대하여 플로팅된 성장 속도의 기울기. 세 번째 범주는 가장 높은 산 농도에서 당분해 속도였다. 이 분류는 표준화된 경계로서 각각의 속도에 대하여 최고값 및 최저값을 사용하여 표준화된 규모로 이루어졌다. 따라서 각각의 균주는 각각의 범주에 대하여 0 내지 1의 등급을 받았으며, 1은 최고 가능 점수이다. 균주의 전체 속도는 3개의 범주에 대한 표준화 값의 합계였다. 성장 속도 및 당분해 속도가 똑같이 가중되는 가중점수가 되었다. 이 경우에, 가장 높은 산 농도에서 당분해 속도는 50% 가중되는 한편, 두 개의 성장 속도는 각각 25% 가중되었다. 범주 당 표준화 값 및 합계 및 가중 점수는 표 4에 요약된다.

[0195] 표 4: 3-HP에서 균주 등급

| 균주 | 성장 속도 @ 75 g/L 3-HP | 성장 속도 기울기 | 글리콜화 속도 | 합계 점수 | 가중 점수 |
|-----------------------------------|------------------------|--------------|------------|-------|-------|
| 이스트첸키아 오리엔탈리스 ATCC PTA-6658 | 1.00 | 0.37 | 0.97 | 2.34 | 0.83 |
| 사카로미세스 볼데리 MYA 402 | 0.78 | 0.49 | 0.94 | 2.21 | 0.79 |
| 이스트첸키아 오리엔탈리스 CD1822 | 0.78 | 0.21 | 1.00 | 1.99 | 0.75 |
| 사카로미세스 볼데리 MYA 404 | 0.72 | 0.42 | 0.90 | 2.04 | 0.74 |
| 이스트첸키아 오리엔탈리스 ATCC 60585 | 0.83 | 0.44 | 0.77 | 2.04 | 0.70 |
| 이스트첸키아 오리엔탈리스 24210 | 0.67 | 0.28 | 0.90 | 1.85 | 0.69 |
| 칸디다 람비카 ATCC 38617 | 0.89 | 0.19 | 0.82 | 1.90 | 0.68 |
| 피치아 멤브라니파시엔스 | 1.00 | 1.00 | 0.23 | 2.23 | 0.62 |
| 쉬조사카로미세스 폼네 | 0.39 | 0.65 | 0.47 | 1.51 | 0.50 |
| 칸디다 발리다 | 0 | 0.14 | 0.40 | 0.54 | 0.24 |
| 지고사카로미세스 렌투스 | 0.33 | 0 | 0 | 0.33 | 0.08 |

[0196]

[0197] 테스트된 균주 중에, 종 아이. 오리엔탈리스, 씨. 람비카, 및 에스. 볼데리의 균주는 낮은 pH에서 3-HP에 대한 생산 숙주로서 가장 큰 가능성을 나타냈다.

[0198] 같은 과정을 pH 2.85 (~80% 자유 산)에서 0, 25, 및 50 g/L 젖산을 함유하는 배지로 1차 스크리닝의 원래의 91 개의 야생형 효모 균주를 스크리닝하고, 평가하고, 점수를 매기기 위해 이용하였다. 표준화 값 및 가중 및 합계 점수는 2차 스크리닝으로 진행된 13개의 균주에 대하여 유도되었다.

[0199] 표 5: 젖산에서 균주 등급

| 균주 | 성장 속도 @ 50 g/L 젖산 | 성장 속도 기울기 | 글리콜화 속도 | 합계 점수 | 가중 점수 |
|---|----------------------|--------------|------------|-------|-------|
| 칸디다 람비카 ATCC 38617 | 0.92 | 1.00 | 1.00 | 2.92 | 0.98 |
| 이스트헨키아 오리엔탈리스 ATCC PTA-6658 | 0.94 | 0.95 | 1.00 | 2.89 | 0.97 |
| 이스트헨키아 오리엔탈리스 CD1822 | 1.00 | 0.86 | 1.00 | 2.86 | 0.97 |
| 이스트헨키아 오리엔탈리스 24210 | 0.89 | 0.73 | 1.00 | 2.62 | 0.91 |
| 칸디다젠폴리니나 | 0.22 | 0.95 | 1.00 | 2.17 | 0.79 |
| 사카로미세스 불데리 MYA 404 | 0.47 | 0.45 | 1.00 | 1.92 | 0.73 |
| 사카로미세스 바야누스 | 0.08 | 0.91 | 0.96 | 1.95 | 0.73 |
| 사카로미세스 불데리 MYA 402 | 0.50 | 0.23 | 1.00 | 1.73 | 0.68 |
| 칸디다 밀레리 | 0 | 0.64 | 0.92 | 1.56 | 0.62 |
| 칸디다 소로시보란스 | 0.28 | 0.95 | 0.59 | 1.82 | 0.60 |
| 코다매아 오메리 | 0.42 | 0 | 0.76 | 1.18 | 0.49 |
| 칸디다 지오차레스 | 0.17 | 0.27 | 0.69 | 1.13 | 0.46 |
| 사카로미세스 자벤시스 (<i>Saccharomyces javensis</i>) | 0.11 | 0.68 | 0 | 0.79 | 0.20 |

[0200]

[0201]

젖산에 대하여, 에스. 자벤시스 (*S. javensis*)만이 50 g/L 젖산이 있는 배지에서 pH 2.85에서 2.5 g/L/hr 글루코스 이용 속도를 달성하지 않았다. 아이. 오리엔탈리스, 씨. 람비카, 및 에스. 불데리는 3-HP 및 젖산에 대하여 산 내성을 나타내는 한편, 산 중 하나에 대해서만 내성인 많은 균주가 있었다. 이것은 또한 1차 스크리닝의 결과에서 보여질 수 있다 (표 1). 예를 들어, 씨. 밀레리 (*C. milleri*), 씨. 루고사 (*C. rugosa*), 씨. 반테르왈티 (*C. vanderwaltii*), 케이. 오메리 (*K. ohmeri*), 에스. 바야누스 (*S. bayanus*), 에스. 자벤시스, 와이. 리폴리티카 (*Y. lipolytica*), 제트. 비스포루스 (*Z. bisporus*), 및 제트. 콤부첸시스 (*Z. kombuchaensis*)는 모두 45-60 g/L 젖산에서 성장을 입증하였지만 테스트된 가장 낮은 농도의 3-HP (35 g/L)에서 성장은 입증하지 않았다. 따라서, 균주의 한 유기산에 대한 내성은 명확하게 다른 산에 대하여 그것의 내성의 예측기로서 사용될 수 없다. 이것은 상기 3-HP 저항성을 나타내는 균주와 W003/049525에서 유기산 생성을 위해 바람직한 숙주로서 확인된 8개의 균주의 목록을 비교함으로써 추가로 강조된다. 이들 균주 중 둘 (씨. 디덴시아 (*C. diddensiae*) 및 씨. 엔토모필라 (*C. entomophila*))은 테스트를 위해 얻어질 수 있고, 다른 6개는 상기 설명된 1차 스크리닝에 포함되었다. 이들 6개 중에서, 씨. 크루세이 (아이. 오리엔탈리스로서 테스트됨)만이 35 g/L 3-HP의 존재시 성장할 수 있다.

[0202]

실시에 1B: 돌연변이 유발 및 3-HP에 대한 저항성을 가진 돌연변이 균주의 선택

[0203]

실시에 1A에 선택된 효모 세포를 돌연변이 유발을 받고 높은 3-HP 내성을 갖는 돌연변이를 확인하기 위해 선택 압력에 노출하였다.

[0204]

예를 들어, 신선한 YP (효모 추출물/펩톤) + 20 g/L 글루코스 플레이트 또는 액체 배양 (OD600 1-4)의 효모 세포를 약 10의 OD600으로 멸균수에 재현탁한다. 이 세포 현탁액의 200 μ L 엘리퀼트를 개개의 튜브로 피펫팅하였고 약 1시간 동안 3 μ L 에탄 메틸 술포네이트 (EMS)에 노출하였으며, 이것은 세포 중 약 65%를 살해한다. 더 높은 EMS 농도가 또한 살해율을 증가시키기 위해 사용될 수 있다. 노출 후, 세포는 5% 나트륨 티오솔페이트로 중화되고, PBS 버퍼로 세척되고, 약 4시간 동안 풍부 배지에서 회수되고, 선택 배지에서 배양된다. 조건이 선택적이라는 것을 확인하기 위해 mock 샘플 (EMS 없음)을 또한 실행하였다. 대안으로, 세포는 UV 조사를 사용하여 돌연변이가 일어날 수 있다.

- [0205] 3-HP 저항 돌연변이 균주를 선택하기 위해, EMS 처리된 세포 현탁액의 엘리퀼트 (돌연변이된 세포 중 약 2×10^8 개)를 감자 텍스트로스 아가 (PDA) 또는 모체 균주가 성장하지 않거나 매우 느리게 성장하는 수준으로 3-HP를 함유하는 또 다른 배지에 도말하였다. 이들 플레이트를 콜로니가 나타날 때까지 수 일 동안 배양한다. 단일 콜로니를 정제하고, 선택의 어떤 적합한 효과를 제거하기 위해 비-선택 배지에 스트리킹하고, 증가된 저항성을 확인하기 위해 선택 배지에서 다시 테스트한다. 저항성 균주를 생성물 및 기질의 HPLC 분석에 대한 주기적인 샘플링을 갖는 셰이크 플라스크 포맷으로 테스트한다. 대안으로, 3-HP 내성에 대한 선택을 케모스테트 또는 연속적 셰이크 플라스크 진화에 의해 수행할 수도 있다. 추가적 라운드의 돌연변이 유발 및 선택을 수행할 수 있다. 돌연변이 유발을 상업적 3-HP 생산을 위해 필수적인 속성을 갖도록 원래 3-HP 생산 필요조건을 만족시키지 않는 숙주의 저항성을 증가시키기 위해 사용할 수도 있다.
- [0206] 실시예 2A: DNA의 숙주 게놈에 형질전환 과정
- [0207] 다음 실시예에 설명된 변형된 효모 균주를 생성하기 위해 효모 숙주 게놈으로 DNA 형질전환을 하기 특이적 과정에 기초하여 수행하였다.
- [0208] 4 mL의 YP + 10% 글루코스 배지를 14 mL 팔콘 (Falcon) 튜브에 추가하였고 원하는 균주를 멸균 루프를 사용하여 이 배지로 접종하였다. 배양물을 37 °C에서 하룻밤 동안 (~16 시간) 250 rpm으로 흔들면서 키웠다. 배양물의 작은 엘리퀼트를 약 시간 간격으로 회수하였고 OD600을 측정하였다. 배양물을 OD600이 0.6-1.0일 때까지 키웠다.
- [0209] 세포를 상온에서 2279X g로 원심분리에 의해 수확하였고, 펠렛을 25 mL 멸균수에 재현탁하였고, 상온에서 2279X g로 원심분리되었다. 펠렛을 1 mL 멸균수에 재현탁하였고, 재현탁된 세포를 1.5 mL 튜브로 옮겼고 16,100X g로 펠렛화하였다. 세포를 1 mL LiOAc/TE 용액에 재현탁하였고 16,100X g로 펠렛화하였다. 세포 펠렛을 500 μ L LiOAc/TE 용액에 재현탁하였다.
- [0210] 다음 성분을 1.5 mL 튜브에 추가하였다: 100 μ L의 상기 세포, 10 μ L 갓 끓인 후 냉각된 연어 정자 DNA (Agilent Technologies, Santa Clara, CA, USA), 및 10 μ L의 원하는, 선형화 형질전환 DNA. DNA 대신에 물로 대조군 반응을 또한 제조하였다. 각각의 형질전환 반응에 대하여, 600 μ L의 PEG/LiOAc/TE 용액을 추가하였고 반응물을 30 °C에서 30분 동안 250 rpm 셰이커 플랫폼 상에서 그것의 측면에서 배양하였다. 40 μ L DMSO를 각각의 반응에 추가하였고 42 °C 항온조에서 5분 동안 배양하였다. 세포를 5,400X g로 1분 동안 펠렛화하였다. 세포를 물에 재현탁하였고, 둘로 나누었고, 형질전환 반응물의 각 절반을 ura 선택 배지 플레이트에 도말하였다. 플레이트를 37 °C에 배치하였다. 콜로니는 일반적으로 18 내지 24시간 성장 후 볼 수 있고, 균주 배경에 의존적이다.
- [0211] 멸균 루프를 소량의 효모를 페트리 디쉬에서 300 μ L 효모 용해 용액 (EPICENTRE® Biotechnologies, Madison, WI, USA)을 함유한 1.5 mL 튜브로 옮기기 위해 사용하였고 게놈 DNA를 MasterPure™ Yeast DNA Purification Kit (EPICENTRE® Biotechnologies)를 사용하여 제조사의 제시에 따라 추출하였다.
- [0212] MasterPure™ Yeast DNA Purification Kit (EPICENTRE® Biotechnologies)를 사용하여 제조된 게놈 DNA를 분리된 형질전환체에서 올바른 통합 이벤트가 발생하는지 결정하기 위한 PCR 반응에 사용하였다. PCR 반응물 (25 μ L)은 스크리닝 되는 균주에 대한 0.5 μ L 게놈 DNA, 1X Crimson Taq™ 반응 버퍼 (New England Biolabs, Ipswich, MA, USA), 각각 25 pmol의 센스 및 안티센스 프라이머, 각각 200 μ M의 dATP, dCTP, dGTP, 및 dTTP, 및 0.625 유닛의 Crimson Taq™ DNA Polymerase (New England Biolabs)를 함유하였다. PCR을 가장 큰 것으로 예상되는 생성물의 kbp 당 95 °C에서 30초 동안 1 사이클 후 이어서, 95 °C에서 30초 동안, 50 °C에서 30초 동안, 및 68 °C에서 1분 동안 30 사이클; 68 °C에서 10분 동안 최종 연장으로 프로그램된 EPPENDORF® MASTERCYCLER® (Eppendorf Scientific, Westbury, New York, USA)로 수행하였다. 열 순환 (thermocycling) 후, PCR 반응 생성물을 TAE 또는 TBE 버퍼 중의 1% 아가로스 겔 전기영동으로 분리하였고 밴드의 크기를 시각화하였고 설명된 바와 같이 특이적 프라이머 세트에 대하여 설명하였다.
- [0213] 실시예 2B: 삽입 부위의 선택
- [0214] 외인성 유전자를 숙주 효모 세포로 통합하는데 적합한 삽입 부위는 효모 숙주 세포에서 삭제될 때 3-HP 생산에 대하여 이롭거나 중성화 효과를 갖는 유전자에 대한 위치일 수도 있다. 선택된 효모 균주에 대하여 적합한 삽입 부위의 비-제한 예는 여기에 작용하는 예에서 설명된다. 당업자는 쉽게 이것들 및 다른 삽입 부위, 예를 들어, 하나 이상의 PDC (예를 들어, SEQ ID NO: 49에 설명된 아미노산 서열을 암호화하는 및/또는 SEQ ID NO: 48에 설명된 뉴클레오티드 서열의 암호화 영역을 포함하는 아이. 오리엔탈리스 PDC 유전자), ADH (예를 들어, SEQ ID NO: 106, 108, 또는 110에 설명된 아미노산 서열을 암호화하는 및/또는 SEQ ID NO: 105, 107, 또는 109에 설명

된 뉴클레오타이드 서열의 암호화 영역을 포함하는 아이. 오리엔탈리스 ADH 유전자), GAL6 (예를 들어, SEQ ID NO: 112에 설명된 아미노산 서열을 암호화하는 및/또는 서열 SEQ ID NO: 111에 설명된 뉴클레오타이드의 암호화 영역을 포함하는 아이. 오리엔탈리스 GAL6 유전자), CYB2A (예를 들어, 서열 SEQ ID NO: 114에 설명된 아미노산을 암호화하는 및/또는 SEQ ID NO: 113에 설명된 뉴클레오타이드 서열의 암호화 영역을 포함하는 아이. 오리엔탈리스 CYB2A 유전자), CYB2B (예를 들어, 서열 SEQ ID NO: 116에 설명된 아미노산을 암호화하는 및/또는 SEQ ID NO: 115에 설명된 뉴클레오타이드 서열의 암호화 영역을 포함하는 아이. 오리엔탈리스 CYB2B 유전자), GPD (예를 들어, 서열 SEQ ID NO: 118에 설명된 아미노산을 암호화하는 및/또는 SEQ ID NO: 117에 설명된 뉴클레오타이드 서열의 암호화 영역을 포함하는 아이. 오리엔탈리스 GPD 유전자), ALD (예를 들어, 서열 SEQ ID NO: 120에 설명된 아미노산을 암호화하는 및/또는 SEQ ID NO: 119에 설명된 뉴클레오타이드 서열의 암호화 영역을 포함하는 아이. 오리엔탈리스 ALD 상동체 5680 유전자, 서열 SEQ ID NO: 122에 설명된 아미노산을 암호화하는 및/또는 SEQ ID NO: 121에 설명된 뉴클레오타이드 서열의 암호화 영역을 포함하는 아이. 오리엔탈리스 ALD 상동체 42026 유전자, 서열 SEQ ID NO: 124에 설명된 아미노산을 암호화하는 및/또는 SEQ ID NO: 123에 설명된 뉴클레오타이드 서열의 암호화 영역을 포함하는 아이. 오리엔탈리스 ALD 상동체 42426 유전자, 또는 서열 SEQ ID NO: 126에 설명된 아미노산을 암호화하는 및/또는 SEQ ID NO: 125에 설명된 뉴클레오타이드 서열의 암호화 영역을 포함하는 아이. 오리엔탈리스 ALD 상동체 42727 유전자), 또는 PCK (예를 들어, 서열 SEQ ID NO: 128에 설명된 아미노산을 암호화하는 및/또는 SEQ ID NO: 127에 설명된 뉴클레오타이드 서열의 암호화 영역을 포함하는 아이. 오리엔탈리스 PCK 유전자) 유전자 또는 이들의 상동체에 대한 위치)의 사용에 대한 교시 내용을 적용할 수 있다. 이들 유전자에 대한 서열이 발표되지 않은 경우, 그것들은 표준 과정, 예를 들어, 예를 들어, 게놈 시퀀싱, 게놈 또는 cDNA 라이브러리의 프로브 잡종화, 또는 알려진 상동 서열에 기초하는 퇴화 프라이머를 사용하여 유전자 단편의 증폭 후 이어서, 전체 서열을 얻기 위한 게놈 위킹 (게놈 walking) 을 사용하여 얻어질 수 있다. 삽입 부위에 대한 다른 적합한 위치는 오픈 리딩 프레임에 함유하지 않는 유전자 간 영역을 포함한다.

[0215] 실시예 2C: 삽입 벡터, 선택 마커 카세트, 유전자 발현 카세트, 및 통합 구조물을 위한 기술

[0216] 삽입 부위 벡터를 하나 이상의 외인성 유전자를 숙주 효모 세포로 통합하기 위해 생성하였다. 숙주 효모 세포는 실시예 1에 설명된 바와 같이 선택 공정을 겪은 세포일 수도 있거나, 그것들은 돌연변이 유발 및/또는 선택을 겪지 않은 세포일 수도 있다.

[0217] 삽입 부위 벡터를 생성하기 위해, 원하는 삽입 부위의 영역 업스트림 (5') 및 영역 다운 스트림 (3')을 주형으로서 숙주 게놈 DNA를 사용하여 둘 다 증폭시켰다. 업스트림 영역은 바람직하게 70 bp 이상이고 1.5 kbp 이하이다. 결과의 표적 서열은 단편이 인접하거나 거의 인접하도록 각각의 단편의 하나의 카피를 가진 벡터를 얻기 위해서 클로닝 벡터에 동시에 또는 순차적으로 결합된다. 유전자 발현 카세트 및/또는 선택 마커 카세트의 삽입을 허용하기 위해 단편들 사이에 독특한 제한 부위가 포함될 수도 있다. 독특한 제한 효소 부위는 또한 클로닝 벡터로부터 이들 부위 사이에서 DNA의 이후의 제거를 허용하기 위해 업스트림 단편의 5' 말단에 또는 근처에 또는 다운스트림 단편의 3' 말단에 또는 근처로 통합될 수도 있다.

[0218] 삽입 부위 벡터로 통합을 위한 선택 마커 카세트를 표준 클로닝 기술을 사용하여 생성하였다. 이들 선택 마커 카세트는 선택 가능 마커를 위한 유전자를 함유할 수도 있고, 또한 업스트림 프로모터 및/또는 다운스트림 종결 서열을 함유할 수도 있다. 적합한 선택 마커 유전자의 예는 URA3, TRP1, HIS, MEL5, CYB2A, LEU2, 및 G418 유전자를 포함한다. 플랭킹 (플랭킹) 서열은 재조합을 통해 마커의 미래의 손실을 허용하기 위해 프로모터/마커 유전자/종결 서열의 측면 상의 카세트에 통합될 수도 있다. 이들 플랭킹 서열은 직접적인 또는 역방향 반복 서열 (기능적 또는 비기능적 서열) 또는 하나 이상의 loxP 부위를 포함할 수도 있다.

[0219] 유전자 발현 카세트를 표준 클로닝 기술을 사용하여 생성하였다. 이들 유전자 발현 카세트는 과발현되는 유전자를 함유하고, 또한 업스트림 프로모터 및/또는 다운스트림 종결 서열을 함유할 수도 있다. 특정 구체예에서, 이들 프로모터/유전자/종결자 조합의 둘 이상의 카피는 단일 유전자-발현 카세트에 통합될 수도 있다. 이중 기원 유전자는 숙주 효모 균주에서 개선된 발현에 코돈-최적화될 수도 있다. 선택된 마커 카세트는 여기에 설명된 바와 같이 과발현되는 유전자 및 어떤 연관된 프로모터 및/또는 종결자와 인접하거나 거의 인접한 유전자 발현 카세트에 클로닝될 수 있다.

[0220] 대안으로, 네이티브 프로모터의 외인성 프로모터와 대체를 위해, 발현 카세트는, 표적화 서열 사이에서, 통합되는 프로모터의 선택 카세트 업스트림을 가질 수도 있다.

[0221] 유전자 발현 카세트는 유전자 발현 통합 구조물을 생성하기 위해 표준 클로닝 기술을 사용하여 여기에 설명된 삽입 부위 벡터에서 두 개의 표적 부위 서열 사이에 삽입될 수 있다. 하나 이상의 선택 마커 카세트는 유전자

발현 카세트의 일부로서 또는 별도로 또한 표적 서열 사이에 삽입될 수도 있다. 특정 변화에서, 유전자 발현 카세트의 조각은 통합 단편 사이에 일반적인 중복 단편이 있도록 다른 삽입 부위 벡터로 클로닝될 수 있다. 예를 들어, 하나의 벡터는 업스트림 삽입 단편, 프로모터, 유전자, 및 종결자를 함유할 수도 있고, 두 번째 벡터는 종결자, 선택 마커 카세트, 및 다운스트림 삽입 단편을 함유할 수도 있다. 또 다른 예에서, 두 개의 유전자의 동시 삽입을 허용하기 위해, 하나의 벡터는 업스트림 삽입 단편, 프로모터, 유전자, 종결자 및 선택 마커 카세트의 모두 또는 일부를 함유할 수도 있고, 두 번째 벡터는 선택 마커 카세트의 모두 또는 일부, 두 번째 프로모터, 유전자, 종결자, 및 다운스트림 삽입 단편을 함유할 수도 있다.

[0222] 유전자 녹아웃 (knockout) 구조물을 생성하기 위해, 삽입 부위 벡터를 업스트림 및 다운스트림 플랭킹 영역으로부터 유도된 삭제되거나 붕괴되는 유전자의 표적 DNA 서열을 사용하여 만들었다. 선택된 표적 서열은 표적 유전자의 업스트림 및 다운스트림 플랭킹 영역 및/또는 표적 유전자 암호화 서열의 모두 또는 일부 또는 그것의 조절 요소 (예를 들어, 프로모터 또는 종결자)를 포함할 수도 있다. 하나 이상의 선택 마커 카세트는 두 개의 표적 서열 사이의 삽입 부위 벡터로 통합될 수도 있다. 녹아웃이 외인성 유전자의 발현과 커플링되는 경우, 하나 이상의 유전자 발현 카세트는 또한 삽입 부위 벡터로 통합된다.

[0223] 숙주 효모 계놈으로 통합되는 DNA 단편을 클로닝 벡터의 단편의 및 다수의 벡터의 중복 단편의 제한 효소 분해로 선형화할 수 있다. 대안으로, 선형 통합 단편을 PCR, 또는 PCR 및 제한 효소 분해의 조합을 사용하여 생성할 수 있다. 삽입 부위 플랭킹 영역은 벡터 주형에서 그것들의 존재에 의해 또는 증폭 프라이머로 통합에 의해 통합 단편으로 통합될 수 있다. 후자의 경우에, 플랭킹 영역의 뉴클레오티드 중 최소 70개는 바람직하게 프라이머로 통합된다.

[0224] 선택된 효모 균주를 위해 적합한 삽입 벡터, 선택 마커 카세트, 유전자 발현 카세트, 및 통합 구조물의 비-제한 예는 다음 작용 예에서 설명된다. 당업자는 3-HP를 생산하는, 대체 변형된 효모 균주를 생성하기 위해 이들 실시예의 교시 내용 및 이전의 설명을 쉽게 적용할 수 있다

[0225] 실시예 2D: adh1202 위치에서 외인성 유전자를 발현하는 삽입 벡터의 구조물

[0226] 선택 가능 마커로서 URA3을 사용하여 플라스미드 pM1Ba107를 PDC 프로모터 및 종결자에 의해 조절되는 아이. 오리엔탈리스 adh1202 위치에서 단일 유전자의 통합을 허용하기 위해 생성하였다. ura 선택 가능 마커와 함께 PDC 프로모터 및 종결자를 PCR 증폭하였고 하기 설명된 바와 같이 pCR4™4BLUNT TOPO® (Invitrogen, La Jolla, CA, USA)로 클로닝하였다. PDC 프로모터, 종결자 및 URA3 선택 가능 마커 PCR 단편을 SOE PCR로 구성하였다. PDC 프로모터를 PCR 생성물의 3' 말단 상의 PDC 종결자에 대하여 상동성을 함유하는 프라이머로 증폭하였고 PDC 종결자 및 URA3 선택 가능 마커를 생성물의 5' 말단 상의 PDC 프로모터에 대하여 상동성을 갖는 프라이머를 사용하여 증폭하였다. 이들 두 개의 단편을 SOE PCR을 통해 함께 넣었다.

[0227] PDC 프로모터를 프라이머 0611184 및 0611195를 사용하여 pACN5로부터 증폭하였다 (도 19). 프라이머 0611184는 NotI 제한 부위를 PCR 생성물의 5' 말단으로 도입한다. 프라이머 0611195는 PDC 프로모터 이후에 XbaI 제한 부위를 도입하고 PCR 생성물의 3' 말단 상의 PDC 종결자에 대하여 상동성을 도입한다.

[0228] 증폭 반응을 제조사의 지시에 따라 Platinum® Pfx DNA Polymerase (Invitrogen)를 사용하여 수행하였다. 각각의 PCR 반응물은 50 μ L의 최종 부피에 0.5 μ L의 희석된 pACN5 (도 19), 25 pM 각각의 프라이머 0611184 및 0611195, 1X Pfx 증폭 버퍼 (Invitrogen), 2 mM MgSO₄, 0.2 mM dNTP 혼합물, 1.25 유닛 Platinum® Pfx DNA Polymerase (Invitrogen)를 함유하였다. 증폭 반응물을 95 °C에서 2분 동안 1 사이클; 각각 95 °C에서 30초 동안, 55 °C에서 30초 동안, 72 °C에서 2분 동안 30 사이클; 72 °C에서 3분 동안 최종 연장으로 프로그램된 EPPENDORF® MASTERCYCLER® (Eppendorf Scientific)에서 배양하였다.

[0229] PCR 생성물을 89 mM 염기성 트리스-89 mM 붕산-2 mM 이나트륨 EDTA (TBE) 버퍼를 사용하는 1% 아가로스 겔 전기영동으로 정제하였다. 약 700 bp의 단편을 겔에서 잘라냈고 QIAQUICK® Gel Extraction Kit (Qiagen, Valencia, CA, USA)를 사용하여 아가로스로부터 추출하였다.

[0230] PDC 종결자 및 URA3 선택 가능 마커를 프라이머 0611189 및 0611185를 사용하여 pHJJ76으로부터 증폭하였다 (도 24). 프라이머 0611189는 PCR 생성물의 5' 말단 상의 PDC 프로모터에 대하여 상동성 및 PDC 종결자 바로 앞에 Pad 제한 부위를 도입한다. 프라이머 0611185는 PCR 생성물의 3' 말단에 NotI 제한 부위를 도입한다. 증폭 반응을 제조사의 지시에 따라 Platinum® Pfx DNA Polymerase (Invitrogen)를 사용하여 수행하였다. 각각의 PCR 반응물은 50 μ L의 최종 부피에 0.5 μ L의 희석된 pHJJ76 (도 19), 25 pM 각각의 프라이머 0611189 및 0611185,

1X Pfx 증폭 버퍼 (Invitrogen), 2 mM MgSO₄, 0.2 mM dNTP 혼합물, 1.25 유닛 Platinum® Pfx DNA Polymerase (Invitrogen)를 함유하였다. 증폭 반응물을 95 °C에서 2분 동안 1 사이클; 각각 95 °C에서 30초 동안, 55 °C에서 30초 동안, 72 °C에서 2분 동안 30 사이클; 72 °C에서 3분 동안 최종 연장으로 프로그램된 EPPENDORF® MASTERCYCLER® (Eppendorf Scientific)에서 배양하였다.

[0231] PCR 생성물을 89 mM 염기성 트리스-89 mM 붕산-2 mM 이나트륨 EDTA (TBE) 버퍼를 사용하는 1% 아가로스 겔 전기영동으로 정제하였다. 약 2000 bp의 단편을 겔에서 잘라냈고 QIAQUICK® Gel Extraction Kit (Qiagen, Valencia, CA, USA)를 사용하여 아가로스로부터 추출하였다.

[0232] 2000 bp PDC 종결자 및 URA3 선택 가능 마커 PCR 생성물 및 700 bp PDC 프로모터 PCR 생성물을 SOE-PCR을 사용하여 융합하였다. 증폭 반응을 제조사의 지시에 따라 Platinum® Pfx DNA Polymerase (Invitrogen)를 사용하여 수행하였다. 각각의 PCR 반응물은 100 µL의 최종 부피에 8 ng의 2000 bp PDC 종결자 및 URA3 선택 가능 마커 PCR 생성물, 24 ng의 700 bp PDC 프로모터 PCR 생성물, 50 pM 각각의 프라이머 0611184 및 0611185, 1X Pfx amplification buffer (Invitrogen), 2 mM MgSO₄, 0.2 mM dNTP 혼합물, 2.5 유닛 Platinum® Pfx DNA Polymerase (Invitrogen)를 함유하였다. 증폭 반응물을 95 °C에서 2분 동안 1 사이클; 각각 95 °C에서 30초 동안, 55 °C에서 30초 동안, 68 °C에서 3분 동안 30 사이클; 68 °C에서 3분 동안 1 사이클로 프로그램된 EPPENDORF® MASTERCYCLER® (Eppendorf Scientific)에서 배양하였다.

[0233] 2700 bp PCR 생성물을 89 mM 염기성 트리스-89 mM 붕산-2 mM 이나트륨 EDTA (TBE) 버퍼를 사용하는 1% 아가로스 겔 전기영동으로 정제하였다. 약 2700 bp의 단편을 겔에서 잘라냈고 QIAQUICK® Gel Extraction Kit (Qiagen, Valencia, CA, USA)를 사용하여 아가로스로부터 추출하였다.

[0234] 2700 bp PCR 생성물을 제조사의 지시에 따라 시퀀싱용 Zero Blunt® TOPO® PCR Kit를 사용하여 pCR™4BLUNT TOPO® (Invitrogen) 벡터로 클로닝하였다. 6 µL의 총 반응량에서, 1 또는 4 µL의 2700 bp PCR 생성물, 1 µL 염 용액 (Invitrogen) 및 1 µL pCR™4BLUNT TOPO® (Invitrogen)을 상온에서 15분 동안 함께 배양하였다. 2 µL의 각 클로닝 반응물을 제조사의 지시에 따라 One Shot® TOP10 화학적 컴피턴트 (competent) 이. 콜리 (Invitrogen) 세포에 형질전환하였다. 형질전환체를 2X YT + amp 플레이트에 도말하였고 37 °C에서 하룻밤 동안 배양하였다. 여러 결과의 형질전환체를 NotI 분해에 의해 원하는 PCR 생성물의 적절한 삽입으로 스크리닝하였다. 원하는 밴드 크기를 수득하는 클론을 DNA 시퀀싱에 의해 올바른 것으로 확인하였고 pM1Ba100으로 지정하였다.

[0235] 플라스미드 pHJJ76 (도 24)는 adh1202 위치에서 유전자 통합을 허용하는 상동성을 함유한다. 플라스미드 pHJJ76를 adh1202 상동 서열의 내부에 존재하는 URA3 선택 가능 마커를 제거하기 위해 NotI로 분해하였다. 분해된 pHJJ76를 89 mM 염기성 트리스-89 mM 붕산-2 mM 이나트륨 EDTA (TBE) 버퍼를 사용하는 1% 아가로스 겔 전기영동으로 정제하였다. 5.2 kbp 단편을 QIAQUICK® Gel Extraction Kit (Qiagen)를 사용하여 아가로스로부터 추출하였고, T4 DNA 리가제를 사용하여 함께 다시 결합하였다. 결합 생성물을 제조사의 지시에 따라 One Shot® TOP10 화학적 컴피턴트 이. 콜리 (Invitrogen) 세포에 형질전환하였다. 형질전환체를 2X YT + amp 플레이트에 도말하였고 37 °C에서 하룻밤 동안 배양하였다. 여러 결과의 형질전환체를 ApaI 및 SacI 분해로 스크리닝하였다. 원하는 분해 생성물을 수득하는 클론을 pHJJ76-no ura로 분해하였다.

[0236] 유전자가 adh1202에서 통합을 위해 PDC 프로모터 및 종결자에 의해 조절되는 배치될 수 있는 경우, pM1Ba100 (상기)의 PDC 프로모터 및 종결자 및 URA3 선택 가능 마커를 플라스미드를 생성하기 위해 pHJJ76-no ura로 클로닝하였다. pHJJ76-no ura를 NotI으로 분해한 후 이어서 CIP를 처리하였다. 선형 5.2 kbp 단편을 QIAQUICK® PCR Purification Kit (Qiagen)를 사용하여 분해하였다. pM1Ba100을 NotI으로 분해하였고 89 mM 염기성 트리스-89 mM 붕산-2 mM 이나트륨 EDTA (TBE) 버퍼를 사용하는 1% 아가로스 겔에서 실행하였다. 2742 bp 단편을 겔에서 잘라냈고, QIAQUICK® Gel Extraction Kit (Qiagen)를 사용하여 추출하였고, T4 DNA 리가제를 사용하여 pHJJ76-no ura의 5.2 kbp 단편으로 결합하였다. 결합 생성물을 제조사의 지시에 따라 One Shot® TOP10 화학적 컴피턴트 이. 콜리 (Invitrogen) 세포에 형질전환하였다. 형질전환체를 2X YT + amp 플레이트에 도말하였고 37 °C에서 하룻밤 동안 배양하였다. 여러 결과의 형질전환체를 KpnI 및 XbaI 분해로 스크리닝하였다. 원하는 분해 생성물을 수득하는 클론을 pM1Ba107로 지정하였다 (도 2).

[0237] 실시예 2E: PDC 위치에서 다수의 외인성 유전자를 발현하는 삽입 벡터 단편의 구조물

[0238] 다음 삽입 벡터 단편을 내인성 아이. 오리엔탈리스 PDC 유전자를 다수의 유전자, 예를 들어, PDC, ENO1, 및 TDH3 프로모터로부터 발현되는 것으로 여기에 설명된 세 개의 유전자를 발현하는 카세트에 대체하는 설계된 DNA

구조물을 생성하기 위해 사용할 수 있다.

- [0239] TDH3 프로모터. 왼쪽 구조물 (pMhCt068 및 유도체) 및 오른쪽 구조물 (pMhCt069 및 유도체) 사이의 상동 재조합은 URA3 단백질의 발현을 일으키며, 균주의 우라실 영양 요구성에서 우라실 자가 영양성으로 전환을 일으키고, 원하는 성분의 선택을 허용한다. 각각의 왼쪽 구조물의 5' 말단은 PDC 위치의 DNA 업스트림에 대하여 상동성인 한편, 각각의 오른쪽 구조물의 3' 말단은 PDC 위치의 DNA 다운스트림에 대하여 상동성이다. 이들 상동 영역은 발현 카세트를 표적화하기 위해 PDC 위치에 제공한다. 이 표적화 접근법은 도 3에 개략적으로 나타나고, 어떤 적합한 위치, 예를 들어, 상기 설명된 어떤 위치, 예를 들어, ADH 위치 (하기 실시예 참조) 또는 ALD 위치를 표적화하기 위해 여기에 설명된 다수의 유전자의 어떤 조합도 사용하도록 변형될 수 있다.
- [0240] 왼쪽 단편의 구조물
- [0241] 빈 벡터 왼쪽 구조물, pMhCt068을 하기 설명된 바와 같이 다수의 단계로 클로닝하였다.
- [0242] PDC 프로모터 영역 및 원하는 추가의 제한 부위를 함유하는 PCR 단편 및 플랭킹 DNA를 프라이머 0611166 및 0611167을 사용하여 게놈 아이. 오리엔탈리스 DNA로부터 증폭하였다. PCR 반응물 (50 μ L)은 100 ng의 게놈 아이. 오리엔탈리스 DNA (바람직하게, 예를 들어, EPICENTRE® Biotechnologies의 MasterPure™ Yeast DNA Purification Kit를 사용하여), 1X ThermoPol 반응 버퍼 (New England Biolabs), 100 pmol 각각의 프라이머 0611166 및 0611167, 200 μ M 각각의 dATP, dCTP, dGTP, 및 dTTP, 2 μ L 100 mM MgSO₄, 및 2 유닛 VentR® (엑소-) DNA Polymerase (New England Biolabs)를 함유하였다. PCR을 94 °C에서 2분 동안 1 사이클 후 이어서 각각 94 °C에서 30초 동안, 54 °C에서 30초 동안, 및 72 °C에서 1분 동안 34 사이클, 72 °C에서 10분 동안 최종 연장으로 프로그램된 EPPENDORF® MASTERCYCLER® (Eppendorf Scientific)로 수행하였다. 열 순환 후, 약 780개의 염기쌍 PCR 생성물을 겔에서 잘라냈고 제조사의 지시에 따라 NUCLEOSPIN® Extract II Kit (Macherey-Nagel, Bethlehem, PA, USA)를 사용하여 정제한 경우, PCR 반응 생성물을 TAE 버퍼 중의 0.9% 아가로스 겔 전기영동으로 분리하였다.
- [0243] TAL 종결 영역 및 원하는 추가의 제한 부위를 함유하는 PCR 단편 및 플랭킹 DNA를 프라이머 0611168 및 0611169를 사용하여 pACN5로부터 증폭시켰다 (도 19). PCR 반응물 (50 μ L)은 1 μ L의 pACN5 미니-프랩 (Mini-prep) 플라스미드 DNA, 1X iProof™ HF 버퍼 (Bio-Rad Laboratories, Hercules, CA, USA), 100 pmol 각각의 프라이머 0611168 및 0611169, 200 μ M 각각의 dATP, dCTP, dGTP, 및 dTTP, 1.5 μ L DMSO 및 1 유닛 iProof™ High-Fidelity DNA Polymerase (Bio-Rad Laboratories)를 함유하였다. PCR을 98 °C에서 30초 동안 1 사이클 후 이어서 각각 98 °C에서 10초 동안, 59 °C에서 20초 동안, 72 °C에서 45초 동안 34 사이클, 72 °C에서 10분 동안 최종 연장으로 프로그램된 EPPENDORF® MASTERCYCLER® (Eppendorf Scientific)로 수행하였다. 열 순환 후, 약 460개의 염기쌍 PCR 생성물을 겔에서 잘라냈고 제조사의 지시에 따라 NUCLEOSPIN® Extract II Kit (Macherey-Nagel, Bethlehem, PA, USA)를 사용하여 정제한 경우, PCR 반응 생성물을 TAE 버퍼 중의 0.9% 아가로스 겔 전기영동으로 분리하였다.
- [0244] PCR을 상기 생성물 둘 다를 융합하는 단일 증폭 생성물을 생성하기 위해 사용하였다. PCR 반응물 (50 μ L)은 125 ng의 PDC 프로모터 함유 PCR 생성물, 76 ng의 TAL 종결자 함유 PCR 생성물, 1X ThermoPol 반응 버퍼 (New England Biolabs), 100 pmol 각각의 프라이머 0611166 및 0611169, 200 μ M 각각의 dATP, dCTP, dGTP, 및 dTTP, 2 μ L 100 mM MgSO₄, 및 2 유닛 VentR® (엑소-) DNA Polymerase (New England Biolabs)를 함유하였다. PCR을 94 °C에서 2분 동안 1 사이클 후 이어서 각각 94 °C에서 30초 동안, 54 °C에서 30초 동안, 72 °C에서 1분 동안 34 사이클, 72 °C에서 10분 동안 최종 연장으로 프로그램된 EPPENDORF® MASTERCYCLER® (Eppendorf Scientific)로 수행하였다. 열 순환 후, 약 1250개의 염기쌍 PCR 생성물을 겔에서 잘라냈고 제조사의 지시에 따라 NUCLEOSPIN® Extract II Kit (Macherey-Nagel, Bethlehem, PA, USA)를 사용하여 정제한 경우, PCR 반응 생성물을 TAE 버퍼 중의 0.9% 아가로스 겔 전기영동으로 분리하였다.
- [0245] EN01 프로모터 영역 및 원하는 추가의 제한 부위를 함유하는 PCR 단편 및 플랭킹 DNA를 프라이머 0611170 및 0611171을 사용하여 pACN43으로부터 증폭하였다 (도 22). PCR 반응물 (50 μ L)은 1 μ L의 pACN43 미니-프랩 플라스미드 DNA, 1X Phusion HF 버퍼 (New England Biolabs), 100 pmol 각각의 프라이머 0611170 및 0611171, 200 μ M 각각의 dATP, dCTP, dGTP, 및 dTTP, 및 1 유닛 Phusion™ High-Fidelity DNA Polymerase (New England Biolabs)를 함유하였다. PCR을 98 °C에서 30초 동안 1 사이클 후 이어서 각각 98 °C에서 10초 동안, 59 °C에서 20초 동안, 72 °C에서 45초 동안 34 사이클, 72 °C에서 10분 동안 최종 연장으로 프로그램된 EPPENDORF® MASTERCYCLER® (Eppendorf Scientific)로 수행하였다. 열 순환 후, 약 1050개의 염기쌍 PCR 생성

물을 겔에서 잘라냈고 제조사의 지시에 따라 NUCLEOSPIN® Extract II Kit (Macherey-Nagel, Bethlehem, PA, USA)를 사용하여 정제한 경우, PCR 반응 생성물을 TAE 버퍼 중의 0.9% 아가로스 겔 전기영동으로 분리하였다.

[0246] 원하는 추가의 제한 부위 및 플랜킹 DNA와 함께, RKI 종결 영역에 이어서 URA3 프로모터 영역 및 URA3 ORF의 5' 말단을 함유하는 PCR 단편을 프라이머 0611172 및 0611173을 사용하여 pACN43으로부터 증폭하였다 (도 22). PCR 반응물 (50 μ L)은 1 μ L의 pACN43 미니-프랩 플라스미드 DNA, 1X iProof™ HF 버퍼 (Bio-Rad Laboratories), 100 pmol 각각의 프라이머 0611172 및 0611173, 200 μ M 각각의 dATP, dCTP, dGTP, 및 dTTP, 1.5 μ L DMSO 및 1 유닛 iProof™ High-Fidelity DNA Polymerase (Bio-Rad Laboratories)를 함유하였다. PCR을 98 °C에서 30초 동안 1 사이클 후 이어서 각각 98 °C에서 10초 동안, 59 °C에서 20초 동안, 72 °C에서 45초 동안 34 사이클, 72 °C에서 10분 동안 최종 연장으로 프로그램된 EPPENDORF® MASTERCYCLER® (Eppendorf Scientific)로 수행하였다. 열 순환 후, 약 1400개의 염기쌍 PCR 생성물을 겔에서 잘라냈고 제조사의 지시에 따라 NUCLEOSPIN® Extract II Kit (Macherey-Nagel, Bethlehem, PA, USA)를 사용하여 정제한 경우, PCR 반응 생성물을 TAE 버퍼 중의 0.9% 아가로스 겔 전기영동으로 분리하였다.

[0247] PCR을 상기 생성물 둘 다를 융합하는 단일 증폭 생성물을 생성하기 위해 사용하였다. PCR 반응물 (50 μ L)은 93 ng의 EN01 프로모터 함유 PCR 생성물 (상기); 125 ng의 RKI 종결자, URA3 프로모터 및 부분적 ORF 함유 PCR 생성물 (상기); 1X Phusion HF 버퍼 (New England Biolabs); 100 pmol 각각의 프라이머 0611170 및 0611173; 200 μ M 각각의 dATP, dCTP, dGTP, 및 dTTP; 1.5 μ L DMSO; 및 1 유닛 Phusion™ High-Fidelity DNA Polymerase (New England Biolabs)를 함유하였다. PCR을 98 °C에서 30초 동안 1 사이클 후 이어서 각각 98 °C에서 10초 동안, 56 °C에서 20초 동안, 72 °C에서 2분 동안 34 사이클, 72 °C에서 10분 동안 최종 연장으로 프로그램된 EPPENDORF® MASTERCYCLER® (Eppendorf Scientific)로 수행하였다. 열 순환 후, 약 2460개의 염기쌍 PCR 생성물을 겔에서 잘라냈고 제조사의 지시에 따라 NUCLEOSPIN® Extract II Kit (Macherey-Nagel, Bethlehem, PA, USA)를 사용하여 정제한 경우, PCR 반응 생성물을 TAE 버퍼 중의 0.9% 아가로스 겔 전기영동으로 분리하였다.

[0248] PCR 생성물에 대한 수령자 벡터를 생성하기 위해, 플라스미드 pMhCt017 (관련없는 인서트 (insert)를 갖는 표준 클로닝 벡터 pUC19)을 HindIII 및 EcoRI로 분해하였고, 37 °C에서 30분 동안 10 유닛 소 위장 포스파타제 (New England Biolabs)를 처리하였고, TAE에서 0.9% 아가로스 겔 전기영동으로 정제하였고, 약 2.6 kbp 밴드를 겔에서 잘라냈고 제조사의 지시에 따라 NUCLEOSPIN® Extract II Kit (Macherey-Nagel)를 사용하여 정제하였다. 결과의 HindIII 내지 EcoRI 정제된 단편은 pUC18에서 발견된 것과 동일하였다 (Yanisch-Perron, C, Vieira, J. and Messing, J. (1985) 유전자, 33, 103-119).

[0249] 상기 정제된 1250 bp 및 2460 bp PCR 생성물을 125 ng pMhCt017 HindIII 내지 EcoRI 벡터 단편, 92 ng의 PDC 프로모터 및 TAL 종결자 PCR 생성물, 165 ng의 EN01 프로모터 및 URA3 프로모터 및 부분적 ORF 함유 PCR 생성물, 1X In-Fusion 반응 버퍼 (Clontech) 및 1 μ L의 IN-FUSION™ 효소 (Clontech)로 구성된 10 μ L의 총 반응량에서 IN-FUSION™ Advantage PCR Cloning Kit (Clontech)를 사용하여 분해된 pMhCt017 단편에 삽입하였다. 반응물을 37 °C에서 15분 동안, 50 °C에서 15분 동안 배양하였고, 얼음 위에 두었다. 반응물을 40 μ L의 TE 버퍼로 희석하였고 2.5 μ L을 SoloPack Gold SuperCompetent 세포 (Agilent Technologies)를 제조사의 지시에 따라 형질전환하기 위해 사용하였다. 형질전환체를 2X YT + amp 플레이트에 도말하였고 37 °C에서 하룻밤 동안 배양하였다. 여러 결과의 형질전환체를 ApaI 분해에 의해 원하는 PCR 생성물의 적절한 삽입을 위해 스크리닝하였다. 원하는 밴드 크기를 수득하는 클론은 DNA 시퀀싱으로 올바른 것으로 확인하였고 pMhCt068로 지정하였다.

[0250] 플라스미드 pMhCt068은 여기에 설명된 원하는 이소성 유전자의 추가를 위해 PDC 프로모터 영역에 이어서 NheI 및 AscI 제한 부위, TAL 종결자, 여기에 설명된 원하는 두 번째 이소성 유전자의 클로닝을 위해 EN01 프로모터 영역에 이어서 XbaI 및 PacI 제한 부위, RKI 종결자, 아이. 오리엔탈리스 URA3 프로모터 및 아이. 오리엔탈리스 URA3 ORF의 5' 말단을 함유한다. 플라스미드 pMhCt068은 PDC 프로모터 내로 약 200 bp에서 A에서 T로 뉴클레오티드 변화, PDC 프로모터로 방법의 2/3에서 G에서 T로 뉴클레오티드 변화, 및 NheI 제한 부위의 5' 측면 상에 존재하는 조숙한 시작 코돈 (ATG)을 갖는 것으로 발견되었다. 따라서, 올바른 버전의 pMhCt068를 하기 설명된 바와 같이 구성하였다.

[0251] PDC 프로모터 영역을 pACN5로부터 프라이머 0611166 및 0611828로 PCR 증폭하였으며 (도 19), 상기 원하는 시작 코돈을 도입하지 않는다. PCR 반응물 (50 μ L)은 1 내지 50배 희석된 pACN5의 미니-프랩 중 1 μ L, 1X ThermoPol 반응 버퍼 (New England Biolabs), 100 pmol 각각의 프라이머 0611166 및 0611828, 200 μ M 각각의 dATP, dCTP, dGTP, 및 dTTP, 2 μ L 100 mM MgSO₄, 및 2 유닛 VentR® (엑소-) DNA Polymerase (New England Biolabs)를 함유하였다. PCR을 94 °C에서 2분 동안 1 사이클 후 이어서 각각 94 °C에서 30초 동안, 54 °C에서

30초 동안, 72 °C에서 1분 동안 34 사이클, 72 °C에서 10분 동안 최종 연장으로 프로그램된 EPPENDORF® MASTERCYCLER® (Eppendorf Scientific)로 수행하였다. 열 순환 후, 약 780개의 염기쌍 PCR 생성물을 겔에서 잘라냈고 제조사의 지시에 따라 NUCLEOSPIN® Extract II Kit (Macherey-Nagel, Bethlehem, PA, USA)를 사용하여 정제한 경우, PCR 반응 생성물을 TAE 버퍼 중의 0.9% 아가로스 겔 전기영동으로 분리하였다.

[0252] PDC 프로모터 함유 PCR 생성물을 상기 설명된 바와 같이 TAL 종결자 함유 PCR 생성물에 융합하였다. TAL 종결자 PCR 단편을 0611168 프라이머로 만들었고, 결과의 PCR 융합 생성물은 조속한 시작 코돈이 없고 원하지 않는 시작 코돈을 포함하는 생성물을 갖는, 혼합물일 것이다. 결과의 1250 bp PCR 생성물을 정제하였고 상기 설명된 바와 같이 IN-FUSION™ Advantage PCR Cloning Kit (Clontech)를 통해 RKI 종결자, URA3 프로모터 및 부분적 ORF 함유 융합 PCR 생성물 및 pUC18와 결합시켰다. 예상된 ApaLI 분해를 수득하는 클론은 DNA 시퀀싱에 의해 올바른 것으로 나타나며, PDC 프로모터에서 돌연변이의 원하는 부재 및 NheI 제한 부위의 조속한 ATG 5'의 결핍을 포함하고, pMhCt082로 지정하였다.

[0253] 오른쪽 단편의 구조물

[0254] 빈 벡터 오른쪽 구조물, pMhCt068을 하기 설명된 바와 같이 다수의 단계로 클로닝하였다.

[0255] pACN43으로부터 아이. 오리엔탈리스 URA3 ORF, URA3 종결자 (URA3 종결 코돈의 275 bp 다운스트림), URA3 프로모터 (효모 숙주로 통합 후 마커 밖으로 루프-아웃 (loop-out) 되는 반복 영역의 역할을 함)의 3' 말단 및 원하는 추가의 제한 부위를 함유하는 PCR 단편 및 플랭킹 DNA를 프라이머 0611174 및 0611175를 사용하여 증폭하였다 pACN43 (도 22). PCR 반응물 (50 µL)은 1 µL의 pACN43 미니-프랩 플라스미드 DNA, 1X Phusion HF 버퍼 (New England Biolabs), 100 pmol 각각의 프라이머 0611174 및 0611175, 200 µM 각각의 dATP, dCTP, dGTP, 및 dTTP, 및 1 유닛 Phusion™ High-Fidelity DNA Polymerase (New England Biolabs)를 함유하였다. PCR을 98 °C에서 30초 동안 1 사이클 후 이어서 각각 98 °C에서 10초 동안, 59 °C에서 20초 동안, 72 °C에서 45초 동안 34 사이클, 72 °C에서 10분 동안 최종 연장으로 프로그램된 EPPENDORF® MASTERCYCLER® (Eppendorf Scientific)로 수행하였다. 열 순환 후, 약 1210 bp PCR 생성물을 겔에서 잘라냈고 제조사의 지시에 따라 NUCLEOSPIN® Extract II Kit (Macherey-Nagel, Bethlehem, PA, USA)를 사용하여 정제한 경우, PCR 반응 생성물을 TAE 버퍼 중의 0.9% 아가로스 겔 전기영동으로 분리하였다.

[0256] TDH3 프로모터 영역 및 원하는 추가의 제한 부위를 함유하는 PCR 단편 및 플랭킹 DNA를 프라이머 0611176 및 0611177를 사용하여 pACN23으로부터 증폭하였다 (도 20). PCR 반응물 (50 µL)은 1 µL의 pACN23 미니-프랩 플라스미드 DNA, 1X Phusion HF 버퍼 (New England Biolabs), 100 pmol 각각의 프라이머 0611176 및 0611177, 200 µM 각각의 dATP, dCTP, dGTP, 및 dTTP, 및 1 유닛 Phusion™ High-Fidelity DNA Polymerase (New England Biolabs)를 함유하였다. PCR을 98 °C에서 30초 동안 1 사이클 후 이어서 각각 98 °C에서 10초 동안, 59 °C에서 20초 동안, 72 °C에서 45초 동안 34 사이클, 72 °C에서 10분 동안 최종 연장으로 프로그램된 EPPENDORF® MASTERCYCLER® (Eppendorf Scientific)로 수행하였다. 열 순환 후, 약 1028 bp PCR 생성물을 겔에서 잘라냈고 제조사의 지시에 따라 NUCLEOSPIN® Extract II Kit (Macherey-Nagel, Bethlehem, PA, USA)를 사용하여 정제한 경우, PCR 반응 생성물을 TAE 버퍼 중의 0.9% 아가로스 겔 전기영동으로 분리하였다.

[0257] 아이. 오리엔탈리스 PDC 유전자 영역 (PDC 종결 영역)의 종결 코돈의 영역 3' 및 및 원하는 추가의 제한 부위를 함유하는 PCR 단편 및 플랭킹 DNA를 프라이머 0611178 및 0611179를 사용하여 아이. 오리엔탈리스 게놈 DNA로부터 증폭하였다. PCR 반응물 (50 µL)은 100 ng의 아이. 오리엔탈리스 게놈 DNA, 1X iProof™ HF 버퍼 (Bio-Rad Laboratories), 100 pmol 각각의 프라이머 0611178 및 0611179, 200 µM 각각의 dATP, dCTP, dGTP, 및 dTTP, 1.5 µL DMSO 및 1 유닛 iProof™ High-Fidelity DNA Polymerase (Bio-Rad Laboratories)를 함유한다. PCR을 98 °C에서 30초 동안 1 사이클 후 이어서 각각 98 °C에서 10초 동안, 59 °C에서 20초 동안, 72 °C에서 45초 동안 34 사이클, 72 °C에서 10분 동안 최종 연장으로 프로그램된 EPPENDORF® MASTERCYCLER® (Eppendorf Scientific)로 수행하였다. 열 순환 후, 약 938개의 염기쌍 PCR 생성물을 겔에서 잘라냈고 제조사의 지시에 따라 NUCLEOSPIN® Extract II Kit (Macherey-Nagel, Bethlehem, PA, USA)를 사용하여 정제한 경우, PCR 반응 생성물을 TAE 버퍼 중의 0.9% 아가로스 겔 전기영동으로 분리하였다.

[0258] PCR을 상기 설명된 최종 두 개의 PCR 생성물 둘 다를 융합하는 단일 증폭 생성물을 생성하기 위해 사용하였다. PCR 반응물 (50 µL)은 125 ng의 the TDH3 프로모터 함유 PCR 생성물, 114 ng의 PDC 종결 영역 함유 PCR 생성물, 1X Phusion HF 버퍼 (New England Biolabs), 100 pmol 각각의 프라이머 0611176 및 0611179, 200 µM 각각의 dATP, dCTP, dGTP, 및 dTTP, 및 1 유닛 Phusion™ High-Fidelity DNA Polymerase (New England Biolabs)를 함유하였다. PCR을 98 °C에서 30초 동안 1 사이클 후 이어서 각각 98 °C에서 10초 동안, 56 °C에서 20초 동안,

72 °C에서 2분 30초 동안 34 사이클, 72 °C에서 10분 동안 최종 연장으로 프로그램된 EPPENDORF® MASTERCYCLER® (Eppendorf Scientific)로 수행하였다. 열 순환 후, 약 1966개의 염기쌍 PCR 생성물을 겔에서 잘라냈고 제조사의 지시에 따라 NUCLEOSPIN® Extract II Kit (Macherey-Nagel, Bethlehem, PA, USA)를 사용하여 정제한 경우, PCR 반응 생성물을 TAE 버퍼 중의 0.9% 아가로스 겔 전기영동으로 분리하였다.

[0259] 상기 정제된 1,210 bp PCR 생성물 및 1,966 bp PCR 융합 생성물을 상기 설명된 바와 같이 125 ng pMhCt017 HindIII 내지 EcoRI 벡터 단편, 54 ng의 URA3 ORF의 3' 말단에 이어서 URA3 종결자를 함유하는 PCR 생성물, 200 ng의 TDH3 프로모터 및 PDC 종결자 융합 PCR 생성물, 1X In-Fusion 반응 버퍼 (Clontech) 및 1 µL의 IN-FUSION™ 효소 (Clontech)로 구성된 10 µL의 총 반응량에서 IN-FUSION™ Advantage PCR Cloning Kit (Clontech) HindIII 및 EcoRI 분해된 pMhCt017 단편에 삽입하였다. 반응물을 37 °C에서 15분 동안, 50 °C에서 15분 동안 배양하였고, 얼음 위에 두었다. 반응물을 40 µL의 TE 버퍼로 희석하였고 2.5 µL을 제조사의 지시에 따라 SoloPack Gold SuperCompetent 세포 (Agilent Technologies)를 형질전환하기 위해 사용하였다. 형질전환체를 2X YT + amp 플레이트에 도말하였고 37 °C에서 하룻밤 동안 배양하였다. 여러 결과의 형질전환체를 ApaLI 분해에 의해 원하는 PCR 생성물의 적절한 삽입을 위해 스크리닝하였다. 원하는 밴드 크기를 수득하는 클론을 DNA 시퀀싱에 의해 올바른 것으로 확인하였고 pMhCt069로 지정하였다.

[0260] 플라스미드 pMhCt069는 아이. 오리엔탈리스 URA3 마커의 3' 말단, 해당 URA3 종결자, URA3 프로모터 (이후 URA3 마커 밖으로 루프-아웃), TDH3 프로모터, 이소성 발현을 위해 원하는 유전자의 서브클로닝을 위한 XbaI 및 PacI 제한 부위, 및 C 위치의 3' 플랭킹 영역을 함유한다.

[0261] 실시예 2F: adh9091 위치에서 다수의 외인성 유전자를 발현하는 삽입 벡터 단편의 구조물

[0262] 내인성 아이. 오리엔탈리스 adh9091 유전자를 여기에 설명된 원하는 다수의 유전자를 발현하는 카세트로 대체하기 위해 다음 삽입 벡터 단편을 실시예 2E에 설명된 것과 유사한 접근법을 사용하여 설계하였다.

[0263] 왼쪽 단편의 구조물

[0264] 빈 벡터 왼쪽 구조물, pGREr125를 하기 설명된 바와 같이 다수의 단계로 클로닝하였다.

[0265] 아이. 오리엔탈리스 adh9091 위치에서 상동 재조합에 필요한 5' 플랭크를 포함하는 구조물 및 빈 발현 카세트 PDC 프로모터/TAL 종결자를 벡터 플라스미드 pCR2.1-TOPO (Invitrogen)으로 PCR 클로닝하였다. PDC 프로모터 단편을 프라이머 0611250 및 0611251을 사용하여 플라스미드 pACN5로부터 PCR 증폭하였다 (도 19). PCR 반응물 (50 µL)은 15 ng의 플라스미드 pACN5 DNA, 1X Phusion HF 버퍼 (New England Biolabs), 50 pmol 각각의 프라이머 0611250 및 0611251, 200 µM 각각의 dATP, dCTP, dGTP, 및 dTTP, 1.5 µL의 100% DMSO (New England Biolabs) 및 1 유닛 Phusion High-Fidelity DNA Polymerase (New England Biolabs)를 함유하였다. PCR을 98 °C에서 2분 동안 1 사이클 후 이어서 각각 98 °C에서 30초 동안, 55 °C에서 30초 동안, 72 °C에서 1분 30초 동안 35 사이클; 72 °C에서 7분 동안 최종 연장으로 프로그램된 EPPENDORF® MASTERCYCLER® (Eppendorf Scientific)로 수행하였다. 열 순환 후, 약 900 bp PCR 생성물을 겔에서 잘라냈고 제조사의 지시에 따라 QIAQUICK® Gel Extraction Kit (Qiagen)를 사용하여 정제한 경우, PCR 반응 생성물을 TBE 버퍼 중의 0.8% 아가로스 겔 전기영동으로 분리하였다. 단편의 5' 말단에 NotI 제한 부위 및 단편의 3' 말단에 PacI 및 XbaI 제한 부위가 있는 결과의 PCR 단편의 총 길이는 약 753 bp였다.

[0266] XbaI 및 PacI 제한 부위를 포함하는, 상기 PCR 생성물에 대한 5' 상동성을 함유하는 두 번째 PCR 단편을 프라이머 0611252 및 0611253을 사용하여 플라스미드 pACN5로부터 TAL 종결 영역을 증폭시키기 위해 생성하였다. PCR 반응물 (50 µL)은 15 ng의 플라스미드 pACN5 DNA (상기), 1X Phusion HF 버퍼 (New England Biolabs), 50 pmol 각각의 프라이머 0611252 및 0611253, 200 µM 각각의 dATP, dCTP, dGTP, 및 dTTP, 1.5 µL의 100% DMSO (New England Biolabs) 및 1 유닛 Phusion High-Fidelity DNA Polymerase (New England Biolabs)를 함유하였다. PCR을 98 °C에서 2분 동안 1 사이클 후 이어서 각각 98 °C에서 30초 동안, 55 °C에서 30초 동안, 72 °C에서 1분 30초 동안 35 사이클; 72 °C에서 7분 동안 최종 연장으로 프로그램된 EPPENDORF® MASTERCYCLER® (Eppendorf Scientific)로 수행하였다. 열 순환 후, 약 900 bp PCR 생성물을 겔에서 잘라냈고 제조사의 지시에 따라 QIAQUICK® Gel Extraction Kit (Qiagen)를 사용하여 정제한 경우, PCR 반응 생성물을 TBE 버퍼 중의 0.8% 아가로스 겔 전기영동으로 분리하였다. 단편의 5' 말단에 XbaI 및 PacI 제한 부위 및 3' 말단에 PmeI 제한 부위가 있는 결과의 PCR 단편의 총 길이는 약 435 bp였다.

[0267] 753 bp 단편 및 435 bp 단편을 프라이머 0611250 및 0611253을 사용하여 PCR로 함께 융합하였으며, 결과의 PDC 프로모터가 TAL 종결자의 업스트림인 1,149 bp 단편으로 이어진다. PCR 반응물 (50 µL)은 125 ng의 753 bp 단

편, 75 ng의 435 bp 단편, 1X Phusion HF 버퍼 (New England Biolabs), 50 pmol 각각의 프라이머 0611250 및 0611253, 200 μ M 각각의 dATP, dCTP, dGTP, 및 dTTP, 1.5 μ L의 100% DMSO (New England Biolabs) 및 1 유닛 Phusion High-Fidelity DNA Polymerase (New England Biolabs)를 함유하였다. PCR을 98 $^{\circ}$ C에서 2분 동안 1 사이클 후 이어서 각각 98 $^{\circ}$ C에서 30초 동안, 55 $^{\circ}$ C에서 30초 동안, 72 $^{\circ}$ C에서 1분 동안 35 사이클; 72 $^{\circ}$ C에서 7분 동안 최종 연장으로 프로그램된 EPPENDORF® MASTERCYCLER® (Eppendorf Scientific)로 수행하였다. 열 순환 후, 약 1149 bp PCR 생성물을 겔에서 잘라냈고 제조사의 지시에 따라 QIAQUICK® Gel Extraction Kit (Qiagen)를 사용하여 정제한 경우, PCR 반응 생성물을 TBE 버퍼 중의 0.8% 아가로스 겔 전기영동으로 분리하였다.

[0268] NotI 제한 부위를 포함하는 상기 1149 bp PCR 생성물에 대한 3' 상동성을 함유하는 PCR 단편을 아이. 오리엔탈 리스 adh9091 로커스에 대한 5' 플랭크를 증폭하기 위해 프라이머 0611254 및 0611255를 사용하여 생성하였다. PCR 반응물 (50 μ L)은 주형 DNA로 플라스미드 pHJJ27 (도 21) 15 ng, 1X Phusion HF 버퍼 (New England Biolabs), 50 pmol 각각의 프라이머 0611254 및 0611255, 200 μ M 각각의 dATP, dCTP, dGTP, 및 dTTP, 1.5 μ L의 100% DMSO (New England Biolabs) 및 1 유닛 Phusion High-Fidelity DNA Polymerase (New England Biolabs)를 함유하였다. PCR을 98 $^{\circ}$ C에서 2분 동안 1 사이클 후 이어서 각각 98 $^{\circ}$ C에서 30초 동안, 55 $^{\circ}$ C에서 30초 동안, 72 $^{\circ}$ C에서 1분 30초 동안 35 사이클; 72 $^{\circ}$ C에서 7분 동안 최종 연장으로 프로그램된 EPPENDORF® MASTERCYCLER® (Eppendorf Scientific)로 수행하였다. 열 순환 후, 약 900 bp PCR 생성물을 겔에서 잘라냈고 제조사의 지시에 따라 QIAQUICK® Gel Extraction Kit (Qiagen)를 사용하여 정제한 경우, PCR 반응 생성물을 TBE 버퍼 중의 0.8% 아가로스 겔 전기영동으로 분리하였다. 단편의 5' 말단에 HpaI 제한 부위 및 3' 말단에 NotI 제한 부위가 있는 결과의 PCR 단편의 총 길이는 약 891 bp였다.

[0269] 891 bp 단편을 약 2005 bp 단편을 발생시키는 프라이머 0611254 및 0611253을 사용하여 PCR로 1149 bp PDC 프로모터/TAL 종결자 단편의 업스트림과 융합하였다. PCR 반응물 (50 μ L)은 125 ng의 1149 bp 단편, 95 ng의 891 bp 단편, 1X Phusion HF 버퍼 (New England Biolabs), 50 pmol 각각의 프라이머 0611254 및 0611253, 200 μ M 각각의 dATP, dCTP, dGTP, 및 dTTP, 1.5 μ L의 100% DMSO (New England Biolabs) 및 1 유닛 Phusion High-Fidelity DNA Polymerase (New England Biolabs)를 함유하였다. PCR을 98 $^{\circ}$ C에서 2분 동안 1 사이클 후 이어서 각각 98 $^{\circ}$ C에서 30초 동안, 55 $^{\circ}$ C에서 30초 동안, 및 72 $^{\circ}$ C에서 2분 동안 35 사이클; 72 $^{\circ}$ C에서 7분 동안 최종 연장으로 프로그램된 EPPENDORF® MASTERCYCLER® (Eppendorf Scientific)로 수행하였다. 열 순환 후, 약 2005 bp PCR 생성물을 겔에서 잘라냈고 제조사의 지시에 따라 QIAQUICK® Gel Extraction Kit (Qiagen)를 사용하여 정제한 경우, PCR 반응 생성물을 TBE 버퍼 중의 0.8% 아가로스 겔 전기영동으로 분리하였다.

[0270] adh9091 위치에서 통합을 위한 5' 플랭크, PDC 프로모터 및 TAL 종결자를 포함하는, 결과의 2005 bp 단편을 pCR2.1-TOPO 벡터로 클로닝하였고 제조사의 지시에 따라 TOPO TA Cloning Kit (Invitrogen)를 사용하여 One-Shot TOP10 이. 콜리 세포에 형질전환하였다. 형질전환체를 2X YT + amp 플레이트에 도말하였고 37 $^{\circ}$ C에서 하룻밤 동안 배양하였다. 여러 결과의 형질전환체를 AvaI 분해에 의해 원하는 단편의 적절한 삽입을 위해 스크리닝 하였다. 원하는 밴드 크기를 수득하는 클론을 확인하였고 pGMEr112로 지정하였다. 플라스미드 pGMEr112는 adh9091 위치에서 상동 재조합에 대한 5' 플랭크에 이어서 빈 발현 카세트 PDC 프로모터/TAL 종결자를 포함한다.

[0271] URA3 프로모터 단편에 의해 구동되는 절단된 5' URA3 마커 유전자를 프라이머 0611283 및 0611263을 사용하여 플라스미드 pHJJ27로부터 PCR 증폭하였다 (도 21). PCR 반응물 (50 μ L)은 15 ng의 플라스미드 pHJJ27, 1X Phusion HF 버퍼 (New England Biolabs), 50 pmol 각각의 프라이머 0611263 및 0611283, 200 μ M 각각의 dATP, dCTP, dGTP, 및 dTTP, 1.5 μ L의 100% DMSO (New England Biolabs) 및 1 유닛 Phusion High-Fidelity DNA Polymerase (New England Biolabs)를 함유하였다. PCR을 98 $^{\circ}$ C에서 2분 동안 1 사이클 후 이어서 각각 98 $^{\circ}$ C에서 30초 동안, 55 $^{\circ}$ C에서 30초 동안, 및 72 $^{\circ}$ C에서 1분 30초 동안 35 사이클; 72 $^{\circ}$ C에서 7분 동안 최종 연장으로 프로그램된 EPPENDORF® MASTERCYCLER® (Eppendorf Scientific)로 수행하였다. 열 순환 후, 약 2005 bp PCR 생성물을 겔에서 잘라냈고 제조사의 지시에 따라 QIAQUICK® Gel Extraction Kit (Qiagen)를 사용하여 정제한 경우, PCR 반응 생성물을 TBE 버퍼 중의 0.8% 아가로스 겔 전기영동으로 분리하였다. 단편의 5' 말단에 HpaI 및 PmeI 제한 부위 및 그것의 3' 말단에 NheI 제한 부위가 있는 결과의 PCR 단편의 총 길이는 약 885 bp였다.

[0272] 결과의 885 bp 단편을 pCR2.1-TOPO 벡터로 클로닝하였고 제조사의 지시에 따라 TOPO TA Cloning Kit (Invitrogen)를 사용하여 One-Shot TOP10 이. 콜리 세포로 클로닝하였다. 형질전환체를 2X YT + amp 플레이트에 도말하였고 37 $^{\circ}$ C에서 하룻밤 동안 배양하였다. 여러 결과의 형질전환체를 BglI 분해에 의해 원하는 인서트

의 적절한 삽입을 위해 스크리닝하였다. 원하는 밴드 크기를 수득하는 클론을 확인하였고 pGMEr108로 지정하였다. 플라스미드 pGMEr108은 URA3 프로모터에 이어서 URA3 유전자의 절단된 5' 세그먼트를 포함하며, 이러한 단편은 5' 말단에 HpaI 및 PmeI 제한 부위에 의해 및 3' 말단에 NotI 제한 부위에 의해 플랭크된다.

[0273] 5' adh9091 플랭크에 이어서 구조물 PDC 프로모터/TAL 종결자를 포함하는, 플라스미드 pGMEr112 (상기)의 1998 bp HpaI 및 PmeI 제한 단편을 HpaI 및 PmeI에 의해 선형화 pGMEr108 (상기)의 4806 bp 벡터에 결합하였다. 1998 bp 인서트 단편 및 4806 bp 벡터 단편을 겔에서 잘라냈고 제조사의 지시에 따라 QIAQUICK® Gel Extraction Kit (Qiagen)를 사용하여 정제한 경우, 이중 제한 반응 생성물을 TBE 버퍼 중의 0.8% 아가로스 겔 전기영동으로 분리하였다. 결합 반응을 1:3 벡터:인서트 비율을 사용하여 수행하였다; 특히 반응을 2 μ L의 4806 bp 선형화 벡터, 6 μ L의 1998 bp 인서트 단편, 9 μ L의 2X QUICK 결합 반응 버퍼 및 1 μ L QUICK T4 DNA 리가제 (New England Biolabs)로 설정하였고, 제조사의 지시에 따라 수행하였다.

[0274] 5 μ L의 결합 생성물을 제조사의 지시에 따라 이. 콜리 XL10-Gold® Ultracompetent 세포 (Agilent Technologies)에 형질전환하였다. 형질전환체를 2X YT + amp 플레이트에 도말하였고 37 °C에서 하룻밤 동안 배양하였다. 여러 결과의 형질전환체를 HindIII 분해에 의해 원하는 인서트의 적절한 삽입을 위해 스크리닝하였다. 원하는 밴드 크기를 수득하는 클론을 확인하였고 pGMEr117로 지정하였다.

[0275] 플라스미드 pGMEr117은 5' adh9091 플랭크에 이어서, 빈 발현 카세트 PDC 프로모터/TAL 종결자 및 URA3 프로모터에 의해 구동되는 절단된 5' URA3 유전자를 포함한다. 추가적으로, 플라스미드 pGMEr117은 두 개의 다른 XbaI 제한 부위를 낳는다: 원하는 유전자를 삽입하기 위해 사용될 수 있는, PDC 프로모터 및 TAL 종결자 사이의 (및 제한 부위 PacI에 인접한) 첫 번째 제한 부위, 및 원래의 pCR2.1-TOPO 백본에서 유전되는 (inherited) 두 번째 XbaI 제한 부위. 이 두 번째 XbaI 제한 부위를 제한하기 위해, 플라스미드 pGMEr117을 제한 효소 ApaI로 분해하였고, 선형화 플라스미드를 제조사의 지시에 따라 효소 DNA Polymerase I, 큰 (클레노브 (Klenow)) 단편 (New England Biolabs)과 함께 처리하였다. 결과의 선형 벡터 (블런트 (blunt) 말단 함유)를 제한 효소 EcoRV로 분해하였으며, 이것은 XbaI 제한 부위를 포함하는 벡터로부터 43 bp 단편을 잘라낸다. 제한 반응 생성물을 TBE 버퍼 중의 0.8% 아가로스 겔 전기영동으로 분리하였고 6761 bp 벡터 단편을 겔에서 잘라냈고 제조사의 지시에 따라 QIAQUICK® Gel Extraction Kit (Qiagen)를 사용하여 정제하였다. 자가 결합 반응을 3 μ L의 선형화 벡터, 6 μ L의 멸균 2차 증류수, 10 μ L의 2X QUICK 결합 반응 버퍼 및 1 μ L QUICK T4 DNA 리가제 (New England Biolabs)로 설정하였고 제조사의 지시에 따라 수행하였다.

[0276] 5 μ L의 결합 생성물을 제조사의 지시에 따라 이. 콜리 XL10-Gold® Ultracompetent 세포 (Agilent Technologies)에 형질전환하였다. 형질전환체를 2X YT + amp 플레이트에 도말하였고 37 °C에서 하룻밤 동안 배양하였다. 여러 결과의 형질전환체를 XbaI 분해에 의해 원하는 인서트의 적절한 삽입을 위해 스크리닝하였다. 원하는 밴드 크기를 수득하는 클론을 확인하였고 pGMEr122로 지정하였다.

[0277] EN01 프로모터 단편을 프라이머 0611295 및 0611296을 사용하여 플라스미드 pACN43 (도 22)으로부터 PCR 증폭하였다. PCR 반응물 (50 μ L)은 15 ng의 플라스미드 pACN43, 1X Phusion HF 버퍼 (New England Biolabs), 50 pmol 각각의 프라이머 0611295 및 0611296, 200 μ M 각각의 dATP, dCTP, dGTP, 및 dTTP, 1.5 μ L의 100% DMSO (New England Biolabs) 및 1 유닛 Phusion High-Fidelity DNA Polymerase (New England Biolabs)를 함유하였다. PCR을 98 °C에서 2분 동안 1 사이클 후 이어서 각각 98 °C에서 30초 동안, 55 °C에서 30초 동안, 및 72 °C에서 1분 동안 35 사이클; 72 °C에서 7분 동안 최종 연장으로 프로그램된 EPPENDORF® MASTERCYCLER® (Eppendorf Scientific)로 수행하였다. 열 순환 후, 약 1009 bp PCR 생성물을 겔에서 잘라냈고 제조사의 지시에 따라 QIAQUICK® Gel Extraction Kit (Qiagen)를 사용하여 정제한 경우, PCR 반응 생성물을 TBE 버퍼 중의 0.8% 아가로스 겔 전기영동으로 분리하였다. 단편의 5' 말단에 PmeI 제한 부위 및 3' 말단에 ApaI 및 NruI 제한 부위가 있는 결과의 PCR 단편의 총 길이는 약 1009 bp였다.

[0278] NruI 및 ApaI 제한 부위를 포함하는, 상기 PCR 생성물에 대하여 5' 상동성을 함유하는 두 번째 PCR 단편을 플라스미드 pACN43 (도 22)로부터 RKI 종결 영역을 증폭하기 위해 프라이머 0611297 및 0611298을 사용하여 생성하였다. 반응물 (50 μ L)은 15 ng의 플라스미드 pACN43 DNA, 1X Phusion HF 버퍼 (New England Biolabs), 50 pmol 각각의 프라이머 0611297 및 0611298, 200 μ M 각각의 dATP, dCTP, dGTP, 및 dTTP, 1.5 μ L의 100% DMSO (New England Biolabs) 및 1 유닛 Phusion High-Fidelity DNA Polymerase (New England Biolabs)를 함유하였다. PCR을 98 °C에서 2분 동안 1 사이클 후 이어서 각각 98 °C에서 30초 동안, 55 °C에서 30초 동안, 및 72 °C에서 1분 동안 35 사이클; 72 °C에서 7분 동안 최종 연장으로 프로그램된 EPPENDORF® MASTERCYCLER® (Eppendorf Scientific)로 수행하였다. 열 순환 후, 약 438 bp PCR 생성물을 겔에서 잘라냈고 제조사의 지시에

따라 QIAQUICK® Gel Extraction Kit (Qiagen)를 사용하여 정제한 경우, PCR 반응 생성물을 TBE 버퍼 중의 0.8% 아가로스 겔 전기영동으로 분리하였다. 단편의 5' 말단에 NruI 및 ApaI 제한 부위 및 단편의 3' 말단에 PmeI 제한 부위가 있는 결과의 PCR 단편의 총 길이는 약 438 bp였다.

- [0279] 1009 bp 프로모터 단편 및 438 bp 종결자 단편을 프라이머 0611295 및 0611298을 사용하여 PCR로 함께 융합하였으며, EN01 프로모터가 RKI 종결자의 업스트림인 약 1447 bp 단편으로 이어진다. PCR 반응물 (50 μ L)은 125 ng의 1009 bp 단편, 65 ng의 438 bp 단편, 1X Phusion HF 버퍼 (New England Biolabs), 50 pmol 각각의 프라이머 0611295 및 0611298, 200 μ M 각각의 dATP, dCTP, dGTP, 및 dTTP, 1.5 μ L DMSO (New England Biolabs) 및 1 유닛 Phusion High-Fidelity DNA Polymerase (New England Biolabs)를 함유하였다. PCR을 98 °C에서 2분 동안 1 사이클 후 이어서 각각 98 °C에서 30초 동안, 55 °C에서 30초 동안, 및 72 °C에서 2분 동안 35 사이클; 72 °C에서 7분 동안 최종 연장으로 프로그램된 EPPENDORF® MASTERCYCLER® (Eppendorf Scientific)로 수행하였다. 열 순환 후, 약 1447개의 bp PCR 생성물을 겔에서 잘라냈고 제조사의 지시에 따라 QIAQUICK® Gel Extraction Kit (Qiagen)를 사용하여 정제한 경우, PCR 반응 생성물을 TBE 버퍼 중의 0.8% 아가로스 겔 전기영동으로 분리하였다.
- [0280] EN01 프로모터/RKI 종결자 구조물을 포함하는, 결과의 1447 bp 단편을 pCR2.1-TOPO 벡터로 클로닝하였고 제조사의 지시에 따라 TOPO TA Cloning Kit (Invitrogen)를 사용하여 One-Shot TOP10 이. 콜리 세포를 형질전환하였다. 형질전환체를 2X YT + amp 플레이트에 도말하였고 37 °C에서 하룻밤 동안 배양하였다. 여러 결과의 형질전환체를 BamHI 분해에 의해 원하는 단편의 적절한 삽입을 위해 스크리닝하였다. 원하는 밴드 크기를 수득하는 클론을 확인하였고 pGMEr114로 지정하였으며, 빈 발현 카세트 EN01 프로모터/RKI 종결자를 포함한다.
- [0281] 플라스미드 pGMEr122 및 pGMEr114를 제한 효소 PmeI로 37 °C에서 3시간 동안 분해하였다. 각각의 분해 반응을 중단하기 약 1시간 전에, 1 μ L의 소 위장 알칼린 포스파타제 (New England Biolabs)를 말단을 탈-인산화하고 자가 결찰을 방지하기 위해 각각의 분해 튜브에 추가하였다. 플라스미드 pGMEr122의 결과의 6761 bp 벡터 단편, 및 플라스미드 pGMEr114의 구조물 EN01 프로모터/TAL 종결자 (1439 bp)를 포함하는, 결과의 인서트 단편을 1X TBE 버퍼 중의 0.8% 아가로스 겔 전기영동으로 분리하였고, 겔에서 잘라냈고, 제조사의 지시에 따라 QIAQUICK® Gel Extraction Kit (Qiagen)를 사용하여 정제하였다.
- [0282] 플라스미드 pGMEr122의 3 μ L의 벡터 단편, 플라스미드 pGMEr114의 4 μ L의 인서트 단편, 2 μ L의 멸균 2차 증류수, 10 μ L의 2X QUICK 리가제 버퍼 및 1 μ L의 QUICK T4 리가제 (New England Biolabs)를 포함하는 다음 결찰 반응물을 제조하였고 제조사의 지시에 따라 수행하였다. 상기 결찰 반응물의 5 μ L 앨리퀀트를 제조사의 지시에 따라 XL10-Gold® Ultracompetent 이. 콜리 세포 (Agilent Technologies)에 형질전환하였다. 형질전환체를 2X YT + amp 플레이트에 도말하였고 37 °C에서 하룻밤 동안 배양하였다. 여러 결과의 형질전환체를 XbaI 및 BstBI를 사용하는 분해에 의해 원하는 인서트의 적절한 삽입을 위해 스크리닝하였다. 원하는 밴드 크기를 수득하는 클론을 시퀀싱으로 확인하였고 pGMEr125로 지정하였다. EN01 프로모터/RKI 종결자 구조물을 반대 방향으로 삽입하였으며, (a) 및 (b)로 지정된 두 개의 버전의 플라스미드 pGMEr125 (도 4 및 5)를 생성하였다.
- [0283] 플라스미드 pGMEr125a 및 pGMEr125b는 PDC 프로모터 영역, TAL 종결자, EN01 프로모터 영역, RKI 종결자, 아이. 오리엔탈리스 URA3 프로모터, 해당 URA3 마커의 5' 말단 및 아이. 오리엔탈리스 adh9091 위치의 5' 플랭킹 영역을 함유한다.
- [0284] 오른쪽 단편의 구조물
- [0285] 빈 벡터 오른쪽 구조물, pGREr121을 하기 설명된 바와 같이 다수의 단계로 클로닝하였다.
- [0286] TDH3 프로모터 단편을 프라이머 0611256 및 0611257을 사용하여 플라스미드 pACN23 (도 20)로부터 PCR 증폭하였다. PCR 반응물 (50 μ L)은 15 ng의 플라스미드 pACN23 DNA, 1X Phusion HF 버퍼 (New England Biolabs), 50 pmol 각각의 프라이머 0611256 및 0611257, 200 μ M 각각의 dATP, dCTP, dGTP, 및 dTTP, 1.5 μ L의 100% DMSO (New England Biolabs) 및 1 유닛 Phusion High-Fidelity DNA Polymerase (New England Biolabs)를 함유하였다. PCR을 98 °C에서 2분 동안 1 사이클 후 이어서 각각 98 °C에서 30초 동안, 55 °C에서 30초 동안, 및 72 °C에서 1분 30초 동안 35 사이클; 72 °C에서 7분 동안 최종 연장으로 프로그램된 EPPENDORF® MASTERCYCLER® (Eppendorf Scientific)로 수행하였다. 열 순환 후, 약 994 bp PCR 생성물을 겔에서 잘라냈고 제조사의 지시에 따라 QIAQUICK® Gel Extraction Kit (Qiagen)를 사용하여 정제한 경우, PCR 반응 생성물을 TBE 버퍼 중의 0.8% 아가로스 겔 전기영동으로 분리하였다. 단편의 5' 말단에 SfoI 제한 부위 및 단편의 3' 말단에 PacI 및 NruI 제한 부위가 있는 결과의 PCR 단편의 총 길이는 약 994 bp였다.

- [0287] NruI 및 PacI 제한 부위를 포함하는, 상기 PCR 생성물에 대한 5' 상동성을 함유하는 두 번째 PCR 단편을 0611258 및 0611259 플라스미드를 사용하여 프라이머 pACN23 (도 20)의 TKL 종결 영역을 증폭하기 위해 생성하였다. PCR 반응물 (50 μ L)은 15 ng의 플라스미드 pACN23 DNA, 1X Phusion HF 버퍼 (New England Biolabs), 50 pmol 각각의 프라이머 0611258 및 0611259, 200 μ M 각각의 dATP, dCTP, dGTP, 및 dTTP, 1.5 μ L의 100% DMSO (New England Biolabs) 및 1 유닛 Phusion High-Fidelity DNA Polymerase (New England Biolabs)를 함유하였다. PCR을 98 $^{\circ}$ C에서 2분 동안 1 사이클 후 이어서 각각 98 $^{\circ}$ C에서 30초 동안, 55 $^{\circ}$ C에서 30초 동안, 및 72 $^{\circ}$ C에서 1분 30초 동안 35 사이클; 72 $^{\circ}$ C에서 7분 동안 최종 연장으로 프로그램된 EPPENDORF® MASTERCYCLER® (Eppendorf Scientific)로 수행하였다. 열 순환 후, 약 469 bp PCR 생성물을 겔에서 잘라냈고 제조사의 지시에 따라 QIAQUICK® Gel Extraction Kit (Qiagen)를 사용하여 정제한 경우, PCR 반응 생성물을 TBE 버퍼 중의 0.8% 아가로스 겔 전기영동으로 분리하였다. 단편의 5' 말단에 PacI 및 NruI 제한 부위 및 단편의 3' 말단에 NotI 제한 부위가 있는 결과의 PCR 단편의 총 길이는 약 469 bp였다.
- [0288] 상기 994 bp 및 469 bp 단편을 프라이머 0611256 및 0611259를 사용하여 PCR로 함께 융합하였으며, TDH3 프로모터가 TKL 종결자의 업스트림인 약 1433 bp 단편으로 이어진다. PCR 반응물 (50 μ L)은 125 ng의 994 bp 단편, 60 ng의 469 bp 단편, 1X Phusion HF 버퍼 (New England Biolabs), 50 pmol 각각의 프라이머 061159 및 0611256, 200 μ M 각각의 dATP, dCTP, dGTP, 및 dTTP, 1.5 μ L의 100% DMSO (New England Biolabs) 및 1 유닛 Phusion High-Fidelity DNA Polymerase (New England Biolabs)를 함유하였다. PCR을 98 $^{\circ}$ C에서 2분 동안 1 사이클 후 이어서 각각 98 $^{\circ}$ C에서 30초 동안, 55 $^{\circ}$ C에서 30초 동안, 및 72 $^{\circ}$ C에서 1분 동안 35 사이클; 72 $^{\circ}$ C에서 7분 동안 최종 연장으로 프로그램된 EPPENDORF® MASTERCYCLER® (Eppendorf Scientific)로 수행하였다. 열 순환 후, 약 1433 bp PCR 생성물을 겔에서 잘라냈고 제조사의 지시에 따라 QIAQUICK® Gel Extraction Kit (Qiagen)를 사용하여 정제한 경우, PCR 반응 생성물을 TBE 버퍼 중의 0.8% 아가로스 겔 전기영동으로 분리하였다.
- [0289] NotI 제한 부위를 포함하는, 상기 1433 bp PCR 생성물의 3' 말단에 대한 5' 상동성을 함유하는 PCR 단편을 프라이머 0611260 및 0611261을 사용하여 adh9091 위치에 대한 3' 플랭크를 증폭시키기 위해 생성하였다. PCR 반응물 (50 μ L)은 주형 DNA로서 플라스미드 pHJJ27 DNA (도 21) 15 ng, 1X Phusion HF 버퍼 (New England Biolabs), 50 pmol 각각의 프라이머 0611260 및 0611261, 200 μ M 각각의 dATP, dCTP, dGTP, 및 dTTP, 1.5 μ L의 100% DMSO (New England Biolabs) 및 1 유닛 Phusion High-Fidelity DNA Polymerase (New England Biolabs)를 함유하였다. PCR을 98 $^{\circ}$ C에서 2분 동안 1 사이클 후 이어서 각각 98 $^{\circ}$ C에서 30초 동안, 55 $^{\circ}$ C에서 30초 동안, 및 72 $^{\circ}$ C에서 1분 30초 동안 35 사이클; 72 $^{\circ}$ C에서 7분 동안 최종 연장으로 프로그램된 EPPENDORF® MASTERCYCLER® (Eppendorf Scientific)로 수행하였다. 열 순환 후, 약 1019 bp PCR 생성물을 겔에서 잘라냈고 제조사의 지시에 따라 QIAQUICK® Gel Extraction Kit (Qiagen)를 사용하여 정제한 경우, PCR 반응 생성물을 TBE 버퍼 중의 0.8% 아가로스 겔 전기영동으로 분리하였다. 단편의 5' 말단에 NotI 제한 부위 및 단편의 3' 말단에 ApaI 제한 부위가 있는 결과의 PCR 단편의 총 길이는 약 1019 bp였다.
- [0290] 1019 bp 단편을 약 2405 bp 단편을 생성하는 프라이머 0611256 및 0611261을 사용하여 PCR로 1433 bp TDH3 프로모터/TKL 종결자 단편의 다운스트림과 융합하였다. PCR 반응물 (50 μ L)은 125 ng의 1433 bp 단편, 90 ng의 1019 bp 단편, 1X Phusion HF 버퍼 (New England Biolabs), 50 pmol 각각의 프라이머 0611256 및 0611261, 200 μ M 각각의 dATP, dCTP, dGTP, 및 dTTP, 1.5 μ L의 100% DMSO (New England Biolabs) 및 1 유닛 Phusion High-Fidelity DNA Polymerase (New England Biolabs)를 함유하였다. PCR을 98 $^{\circ}$ C에서 2분 동안 1 사이클 후 이어서 각각 98 $^{\circ}$ C에서 30초 동안, 55 $^{\circ}$ C에서 30초 동안, 및 72 $^{\circ}$ C에서 2분 동안 35 사이클; 72 $^{\circ}$ C에서 7분 동안 최종 연장으로 프로그램된 EPPENDORF® MASTERCYCLER® (Eppendorf Scientific)로 수행하였다. 열 순환 후, 약 2405 bp PCR 생성물을 겔에서 잘라냈고 제조사의 지시에 따라 QIAQUICK® Gel Extraction Kit (Qiagen)를 사용하여 정제한 경우, PCR 반응 생성물을 TBE 버퍼 중의 0.8% 아가로스 겔 전기영동으로 분리하였다.
- [0291] TDH3 프로모터/TAL 종결자 구조물의 adh9091 위치 다운스트림에서 3' 플랭크를 포함하는 결과의 2405 bp 단편을 pCR2.1-TOPO 벡터로 클로닝하였고 제조사의 지시에 따라 TOPO TA Cloning Kit (Invitrogen)를 사용하여 One-Shot TOP10 이. 콜리 세포에 형질전환하였다. 형질전환체를 2X YT + amp 플레이트에 도말하였고 37 $^{\circ}$ C에서 하룻밤 동안 배양하였다. 여러 결과의 형질전환체를 AvaI 분해에 의해 원하는 단편의 적절한 삽입을 위해 스크리닝하였다. 원하는 밴드 크기를 수득하는 클론을 확인하였고 pGMEr113로 지정하였다. 플라스미드 pGMEr113는 빈 발현 카세트 TDH3 프로모터/TKL 종결자가 선행하는 아이. 오리엔탈리스 adh9091 위치에서 상동 재조합을 위한 3' 플랭크를 포함한다.
- [0292] PCR을 프라이머 0611264 및 0611284를 사용하여 플라스미드 pHJJ27 (도 21)로부터 URA3 ORF의 절단된 3' 단편,

URA3 종결자, 및 URA3 프로모터 (상기 설명된 바와 같이 효모 숙주로 통합 후 마커 밖으로 루프-아웃되는 반복 영역의 역할을 함)을 증폭하기 위해 사용하였다. PCR 반응물 (50 μ L)은 15 ng의 플라스미드 pHJJ27 DNA, 1X Phusion HF 버퍼 (New England Biolabs), 50 pmol 각각의 프라이머 0611264 및 0611284, 200 μ M 각각의 dATP, dCTP, dGTP, 및 dTTP, 1.5 μ L의 100% DMSO (New England Biolabs) 및 1 유닛 Phusion High-Fidelity DNA Polymerase (New England Biolabs)를 함유하였다. PCR을 98 $^{\circ}$ C에서 2분 동안 1 사이클 후 이어서 각각 98 $^{\circ}$ C에서 30초 동안, 55 $^{\circ}$ C에서 30초 동안, 및 72 $^{\circ}$ C에서 1분 30초 동안 35 사이클; 72 $^{\circ}$ C에서 7분 동안 최종 연장으로 프로그램된 EPPENDORF® MASTERCYCLER® (Eppendorf Scientific)로 수행하였다. 열 순환 후, 약 1324 bp PCR 생성물을 겔에서 잘라냈고 제조사의 지시에 따라 QIAQUICK® Gel Extraction Kit (Qiagen)를 사용하여 정제한 경우, PCR 반응 생성물을 TBE 버퍼 중의 0.8% 아가로스 겔 전기영동으로 분리하였다. 단편의 5' 말단에 NheI 제한 부위 및 단편의 3' 말단에 ApaI 및 SfoI 제한 부위가 있는 결과의 PCR 단편의 총 길이는 약 994 bp였다.

[0293] 상기 겔-정제된 1324 bp 단편을 pCR2.1-TOPO 벡터로 클로닝하였고 제조사의 지시에 따라 TOPO TA Cloning Kit (Invitrogen)를 사용하여 One-Shot TOP10 이. 콜리 세포에 형질전환하였다. 형질전환체를 2X YT + amp 플레이트에 도말하였고 37 $^{\circ}$ C에서 하룻밤 동안 배양하였다. 여러 결과의 형질전환체를 HindIII 분해에 의해 원하는 인서트의 적절한 삽입을 위해 스크리닝하였다. 원하는 밴드 크기를 수득하는 클론을 확인하였고 pGMEr109로 지정하였다. 플라스미드 pGMEr109는 URA3 ORF의 3' 단편 및 URA3 종결자, 이어서 URA3 프로모터를 포함한다. 플라스미드 pGMEr109의 URA3 유전자의 3' 단편의 업스트림 부분은 플라스미드 pGMEr108로 클로닝된 절단된 5' URA3 단편의 극단을 갖는 460 bp 상동성을 낳는다. 상동성의 영역은 세그먼트 둘 다를 함유하는 구조물을 갖는 숙주 유기체의 동시-형질 전환에서 기능적 선택 마커를 생성하는 유전자의 두 개의 부분 사이의 재조합을 허용한다.

[0294] 플라스미드 pGMEr109를 KpnI로 분해하였고, 제조사의 지시에 따라 DNA Polymerase I, 큰 (클레노브) 단편 (New England Biolabs)을 처리하였다. 선형화 pGMEr109 플라스미드 (블런트 말단을 함유)를 BamHI로 분해하였다 생성물을 TBE 버퍼 중의 8% 아가로스 겔 전기영동으로 분리하였고 5247 bp 벡터 단편을 겔에서 잘라냈고 제조사의 지시에 따라 QIAQUICK® Gel Extraction Kit (Qiagen)를 사용하여 정제하였다.

[0295] 플라스미드 pGMEr113를 BamHI 및 EcoRV로 분해하였으며, URA3 종결자에 이어서 URA3 프로모터를 갖는 URA3 ORF의 구조물 TDH3 프로모터/TKL 종결자에 이어서 절단된 3' 단편을 낳는 2466 bp 단편을 발생시킨다. 이중 제한 반응 생성물을 TBE 버퍼 중의 0.8% 아가로스 겔 전기영동으로 분리하였고 약 2466 bp 벡터 단편을 겔에서 잘라냈고 제조사의 지시에 따라 QIAQUICK® Gel Extraction Kit (Qiagen)를 사용하여 정제하였다. 2466 bp BamHI/EcoRV 분해된 단편을 플라스미드 pGMEr109의 5247 bp 벡터 단편에 결합하였다. 결합 반응을 3 μ L의 5247 bp 선형화 벡터, 3 μ L의 2466 bp 인서트 단편, 3 μ L의 멸균 2차 증류수, 10 μ L의 2X QUICK 결합 반응 버퍼 및 1 μ L QUICK T4 DNA 리가제 (New England Biolabs)로 설정하였고, 제조사의 지시에 따라 수행하였다.

[0296] 5 μ L의 결합 생성물을 제조사의 지시에 따라 이. 콜리 XL10-Gold® Ultracompetent 세포 (Agilent Technologies)에 형질전환하였다. 형질전환체를 2X YT + amp 플레이트에 도말하였고 37 $^{\circ}$ C에서 하룻밤 동안 배양하였다. 여러 결과의 형질전환체를 XbaI 및 PacI로 분해에 의해 원하는 인서트의 적절한 삽입을 위해 스크리닝하였다. 원하는 밴드 크기를 수득하는 클론을 확인하였고 pGMEr121 (도 6)로 지정하였다.

[0297] 플라스미드 pGMEr121은 아이. 오리엔탈리스 URA3 마커의 3' 말단에 이어서 해당하는 URA3 프로모터, TDH3 프로모터, TKL 종결자 및 adh9091 위치의 3' 플랭킹 영역을 함유한다.

[0298] 실시예 2G: 아이. 오리엔탈리스 CNB1의 구성

[0299] 아이. 오리엔탈리스 CNB1는 하기 설명된 바와 같이 아이. 오리엔탈리스 CD1822로부터 구성된다 (아이. 오리엔탈리스 ATCC PTA-6658로부터 아이. 오리엔탈리스 CD1822의 생성에 대하여 실시예 1A 참조). 균주 CD1822에 함유된 URA3 유전자의 카피 둘 다를 유전적 설계를 위한 선택 마커로서 이 유전자의 사용을 허용하기 위해 삭제하였다. URA3은 유전자의 존재 (우라실 결핍 배지에서 성장에 의해) 및 부재 (5-플루오로오로트산의 존재시 성장에 의해) 둘 다에 이용 가능한 마커이기 때문에 효모 유전학을 위한 다용도 마커이다. URA3 유전자 중 하나의 붕괴를 반복 DNA 서열에 의해 플랭크된 MEL5 선택 마커를 함유하는 선택 카세트와 대체함으로써 수행하였다. MEL5 선택 카세트를 양성 균주 테스트를 MEL 마커 유전자의 손실에 대하여 스크리닝하였다. 두 번째 URA3 유전자의 손실은 5-플루오로오로트산의 존재시 성장에 의해 선택되었다.

[0300] CD1822를 2.8 μ g의 벡터 pMI458 (도 27)의 SacI/PspOMI 분해된 DNA. 플라스미드에 형질전환하였다. pMI458은 아이. 오리엔탈리스 URA3 유전자의 서열 업스트림 (P-1oURA3, SEQ ID NO: 253) 및 다운스트림 (T-1oURA3, SEQ

ID NO: 254)에 상동성인 DNA 단편에 의해 플랭크된, 아이. 오리엔탈리스 PGK 프로모터 (P-loPGK, SEQ ID NO: 247)에 의해 조절되는, 에스. 세레비시에 MEL5 유전자 (SEQ ID NO: 255)를 함유한다. P-loURA3 및 T-loURA3 단편은 아이. 오리엔탈리스 게놈과 상대적으로 같은 방향에 있다. 대략 500 MeI + 콜로니를 30 °C에서 5일 후에 얻었다. 열 개의 콜로니는 10 mL BFP (버터필드 포스페이트 버퍼) 튜브를 접종하고 25 µL을 DM1 X-α-gal 플레이트 상에 도말함으로써 분리된 단일 콜로니였다. 각각의 초기 분리체의 단일 파란 콜로니를 추가의 분석을 위해 YPD 상에서 골랐다.

[0301] PCR을 원하는 유전적 이벤트를 위한 형질전환체를 스크리닝하기 위해 사용하였다. PCR 스크리닝시 주형으로서 사용되는 게놈 DNA를 얻기 위해, 1.5 mL 하룻밤 동안 배양된 세포를 스크류-캡 미세원심분리 튜브에서 스핀 다운시켰고 0.2 mL의 2% Triton X-100, 1% SDS, 100 mM NaCl, 10 mM 트리스-HCl pH 8.0, 1 mM Na₂EDTA pH 8.0 용액에 재현탁하였다. 10 mM 트리스 pH 8.0/1 mM EDTA (Sigma) 및 0.3 g 유리 비드로 평형화된 0.2 mL의 페놀:클로로포름:이소아밀 알콜 (25:24:1) 혼합물을 추가하였다. 튜브를 Mini-BeadBeater-8 (BioSpec)에서 2분 동안 최고 속도에서 흔들었다. 0.2 mL의 TE를 추가하였고 튜브를 간단히 보텍스하였다. 수층을 16,100 x g로 5분 동안 원심분리에 의해 분리하였다. 상층액을 새 튜브로 제거하였고 1 mL의 100% 에탄올을 추가하였다. 튜브를 20 °C에서 30분 동안 배치하였고, 16,100 x g로 5분 동안 원심분리하였고, 액체를 부었다. DNA를 공기건조하였고 500 µL TE에 재현탁하였다.

[0302] 원하는 5' 교차 혼합을 위한 PCR 스크리닝을 프라이머 oCA405 및 oCA406을 사용하여 수행하였으며, 이것은 1.5 kbp 생성물을 생산한다. 원하는 3' 교차 혼합을 위한 PCR 스크리닝을 프라이머 WG26 및 CM647을 사용하여 수행하였으며, 이것은 1.6 kbp 생성물을 생산한다. pMI458을 생성하기 위해 사용된 URA3 영역 외부의 프라이머 (더 먼 업스트림 또는 다운스트림)는 oCA405 및 CM647이며, 이것은 야생형에 대해 3.2 kbp 생성물, pMI458 붕괴 대립유전자에 대해 5.0 kbp 생성물, 및 선택 마커가 밖으로 루프-아웃될 때 2.2 kbp 생성물을 생산한다. 이들 PCR 반응을 55 °C 어닐링 (annealing) 온도를 사용하여 수행하였다. PCR을 또한 네 개의 프라이머 접근법을 사용하여 URA3 오픈 리딩 프레임의 손실에 대해 스크리닝하기 위해 사용하였다. 프라이머 pJLJ28 및 pJLJ29는 액틴 유전자의 800 bp 단편을 증폭시켰고 프라이머 pJLJ30 및 pJLJ31은 URA3 유전자의 600 bp 단편을 증폭시켰다. 한 반응에서 모든 프라이머의 사용은 양성 내부 조절 (액틴 단편)을 제공한다. Roche의 Taq DNA 폴리머라제를 제조사의 프로토콜에 따라, 61 °C의 어닐링 온도로 사용하였다. 균주 1822ura het MEL-1 및 1822ura het MEL-2를 URA3 위치에서 MEL5 선택 카세트를 통합하는 것으로 확인하였다.

[0303] KtSEQ1a (SEQ ID NO: 256) 및 KtSEQ1b (SEQ ID NO: 257) 서열 사이의 재조합을 허용함으로써 MEL5 마커를 1822ura het MEL-1 및 1822ura het MEL-2의 게놈으로부터 제거하였다. MEL+ 균주를 YPD 배지에서 하룻밤 동안 대략 0.5 내지 2.0의 OD₆₀₀으로 키웠다. 배양물을 YPD 배지에서 약 0.00001의 OD₆₀₀으로 다시 희석하였다. 200 µL의 배양 희석액을 96웰 미세역가 플레이트로 옮겼다. 플레이트를 접착 커버로 덮었고, 6-7 시간 동안 (균주의 성장 속도에 의존하여, 대략 6번의 세포 분화) 최대 교반으로, 30 °C 배양기에서 배양하였다. 각각의 웰에서 100 µL (플레이트 당 약 ~1000 cfu)를 DM1 + X-α-gal 배지에 도말하였다. 플레이트를 색 변화를 관찰하기 위해 30 °C에서 하룻밤 동안 또는 상온에서 2일 동안 배양하였다. 흰색 콜로니 (추정상 mel-)를 유사한 배지로 스트리킹하였고, 상기 설명된 바와 같이 PCR에 의해 스크리닝하였다. 두 개의 독립적 루프-아웃, 1822ura het mel-1로서 구성된, 1822ura het MEL-1로부터 하나 및 1822ura het mel-2로서 구성된, 1822ura het mel-2로부터 다른 것을 발견하였다. 이상하게, 얻어진 흰색 콜로니의 엄청난 중요성은 2.2 kbp의 예상된 밴드를 제공하지 않았다.

[0304] ura- 유도체를 얻기 위해, YP5D 배지 (YP + 100 g/L 텍스트로스)에서 하룻밤 동안 성장시켰고 하룻밤 동안의 배양물의 앨리쿼트 ((0.5, 5 및 50 µL)를 ScD-2X FOA 플레이트에 도말하였다. FOA-저항성 콜로니를 단일 콜로니에 스트리킹하였고 ura- 표현형에 대해 확인하였다. 1822ura het mel-2로부터 2 개의 콜로니 및 1822ura het mel-1로부터 6개의 콜로니를 추가의 분석을 위해 골랐다. 이들 콜로니를 YPD에서 하룻밤 동안 성장시켰고 게놈 DNA를 상기 페놀/클로로포름 방법을 사용하여 추출하였다.

[0305] URA3 오픈 리딩 프레임의 존재를 PCR로 스크리닝하였다; 8개의 균주 중 아무것도 URA3 유전자를 함유하지 않았다. 1822ura het mel-1의 두 개의 ura-디센던트 (descendent)를 yJLJ3 (CNB1) 및 yJLJ4로 이름지었다. 게놈 시퀀싱에 기초하여, yJLJ3 (CNB1) 및 yJLJ4를 URA3 유전자 및 인근의 퍼메아제 유전자 둘 다의 삭제를 함유하도록 결정하였다; 바람직하게 URA3 유전자만 삭제될 것이다. 이 퍼메아제 유전자의 붕괴 없이 이 CD1822의 ura3 영역 요구체를 생성하기 위해, CD1822를 1 µg의 SacI/ApaI 분해된 pCM208 (도 28)에 형질전환하였다. 플라스미드 pCM208는 아이. 오리엔탈리스 URA3 유전자의 업스트림 (5' URA 플랭크 (근처), SEQ ID NO: 258) 및 다운스트림

(3' URA3 플랭크 (근처), SEQ ID NO: 259) 플랭킹 영역에 상동성인 DNA 서열을 함유한다. 대략 200 Mel+ 콜로니를 30 °C에서 5일 후 얻었다. 8개의 파란 콜로니를 ScD X-α-gal 플레이트에 스트리킹함으로써 분리하였다. PCR을 원하는 유전적 이벤트를 위해 형질전환체를 스크리닝하기 위해 사용하였다. 5' 교차 혼합 스크리닝을 프라이머 oJY11 및 oJY12를 사용하여 수행하였으며, 이것은 원하는 형질전환체의 0.9 kbp 생성물을 생산한다. 3' 교차 혼합 스크리닝을 프라이머 oJY13 및 oJY14를 사용하여 수행하였으며, 이것은 원하는 형질전환체의 1.0 kbp 생성물을 생산한다. 8개의 콜로니 중 3개는 원하는 PCR 생성물을 나타냈다. 이 콜로니에 대한 MEL 마커를 루프-아웃될 수 있고 두 번째 URA3 유전자를 상기 설명된 바와 같이 삭제하였다. 대안으로, yJLJ3 또는 yJLJ4로부터 유도된 균주에서 URA3 및 퍼메아제 유전자 삭제는 실시예 2H에 설명된 바와 같이, 원-스텝 형질전환에 고정될 수 있다.

[0306] 실시예 2H: URA3 선택 마커를 함유하는 MBin500 조절 균주의 구성

[0307] 상기 설명된 바와 같이, 여기에 사용된 아이. 오리엔탈리스 균주 지정 CNB1은 URA3 유전자의 동형 접합체의 삭제로 인한 유리딘 영양 요구체였다. URA3 위치의 이형 수리는 691 bp의 5' 플랭킹 DNA, 및 1500 bp의 3' 플랭킹 DNA를 갖는 URA3 유전자로 구성된 ura 고정 카세트를 함유하는 ura 고정 벡터, pJLJ62 (도 26)를 사용함으로써 이루어진다. ura 고정 카세트는 5' NotI 제한 부위 및 3' ApaI 제한 부위에 의해 플랭크된다. NotI 및 ApaI를 사용하는 제한 분해를 벡터 백본으로부터 2796 bp ura 고정 카세트를 제거하기 위해 수행하였다. 분해물을 제조사에 의해 명시된 바와 같이 QIAQUICK PCR Purification Kit (Qiagen)를 사용하여 정제하였다. DNA를 유리 멸균된 물에서 용출하였고 1 μg을 아이. 오리엔탈리스 CNB1에 형질전환하기 위해 사용하였다. 형질전환체를 ura 선택 플레이트에서 선택하였고, 유리딘 보충이 필요하지 않은 단일 콜로니를 MBin500로 지정하였다.

[0308] 실시예 2I: URA3 선택 마커의 제거

[0309] 통합 카세트에 존재하는 2개의 URA3 프로모터 영역의 재조합을 통해 URA3 선택 마커 유전자를 제거한 균주를 분리하기 위해, 원하는 ura+ 균주를 3 mL의 YP + 10% 글루코스 배지에 접종하였고 37 °C에서 적어도 4시간 동안 및 최대 하룻밤 동안 250 rpm으로 흔들면서 키웠다. 50-100 μL의 배양물을 ScD FOA 플레이트에 도말하였고 콜로니가 나타날 때까지 37 °C에서 48-60시간 동안 배양하였다.

[0310] FOA에서 성장은 FOA가 URA3 단백질에 의해 독성 화합물로 전환되기 때문에 URA3 마커의 제거를 위해 선택하였으며, ura+ 세포의 죽음을 일으킨다. 여러 FOA-저항성 콜로니를 37 °C의 YPD 플레이트 상에서 키움으로써 두 번 정제하였다. 이들 정제된 분리체를 여기에 설명된 바와 같이 PCR을 통해 적절한 URA3 루프-아웃에 대하여 스크리닝하였다.

[0311] 실시예 2J: 나트륨 도데실 술페이트 폴리아크릴아미드 겔 전기영동 (SDS-PAGE) 분석 및 효소 검정을 위해 변형된 효소 균주의 셰이크 플라스크 성장 과정

[0312] 4 mL의 ura 선택 배지를 14 mL 팔콘 튜브에 추가하였고 원하는 균주를 멸균 루프를 사용하여 이 배지에 접종하였다. 배양물을 37 °C에서 하룻밤 동안 (~16 시간) 250 rpm으로 흔들면서 키웠다. 아이. 오리엔탈리스 위치의 적어도 하나의 야생형 카피를 갖는 균주에 대하여, 하룻밤 동안의 배양물의 500 μL를 25 mL의 YP + 10% 글루코스 배지를 함유하는 125 mL 배플 플라스크에 추가하였다. pdcΔ/pdcΔ 균주에 대하여, 하룻밤 동안의 배양물 1 mL을 25 mL의 액체 YP + 100 g/L 텍스트로스 배지를 함유하는 125 mL 배플 플라스크에 추가하였다. 플라스크를 37 °C에서 250 rpm으로 흔들면서 키웠다. 배양물의 작은 앨리쿼트를 약 한 시간 간격으로 회수하였고 OD₆₀₀를 측정하였다. 배양물을 OD₆₀₀가 4 내지 6일 때까지 키웠다.

[0313] SDS-PAGE 분석을 위해 세포의 작은 샘플을 제조하기 위해, 2.5 OD 유닛에 해당하는 배양물의 양을 배양물로부터 취하였고 1.5 mL 튜브로 옮겼다. 세포를 16,100 x g로 펠렛화하였고, 상층액을 제거하였고, 세포 펠렛을 사용 전까지 -20 °C에 저장하였다.

[0314] 성장 플라스크에 남은 세포를 상온에서 2279 x g로 원심분리하여 수확하였고, 펠렛을 12.5 mL 0.85 M NaCl에 재현탁하였고, 상온에서 2279 x g로 원심분리하였다. 펠렛을 1 mL 0.85 M NaCl에 재현탁하였고, 재현탁된 세포를 2.0 mL 튜브로 옮겼고, 16,100 x g로 펠렛화하였다. 상층액을 제거하였고 펠렛을 일주일 이내에 효소 검정에 사용되면 -20 °C, 또는 더 장기간 저장을 위해 -80 °C에 저장하였다.

[0315] 2.5 OD 유닛에 해당하는 세포 펠렛의 SDS-PAGE 분석을 위해, 세포를 100 dH₂O에 재현탁하였고, 100 μL 0.2 M NaOH를 추가하였다. 샘플을 상온에서 5분 동안 배양하였고, 세포를 16,100 x g로 펠렛화하였고 100 μL SDS 샘플로 버퍼 (Bio-Rad Laboratories)에 재현탁하였다. 샘플을 95 °C에서 5분 동안 가열하였고 세포를 간결한 원심

분리로 펠렛화하였다. 1 내지 5 μL 의 상층액을 제조사의 지시에 따라 Criterion 8-16% Pre-Cast Gel (Bio-Rad Laboratories) 상에서 분석하였다. 밴드를 InstantBlue™ Coomassie-Based Staining Solution (Expedeon Protein Solutions, San Diego, CA, USA)을 사용하여 시각화하였다.

[0316] 실시에 2K: 생성물 분석을 위한 변형된 효모 균주의 셰이크 플라스크 성장 과정

[0317] 균주를 Ura 선택 플레이트 상에 단일 콜로니에 대하여 스트리킹하였고 30 °C에서 1-2일 동안 배양하였다. 시드 배양물을 Ura 선택 플레이트로부터 1-2개의 콜로니로 접종된 50 mL CNB1 셰이크 플라스크 배지를 함유하는 250 ml 배플 플라스크에서 제조하였다. 시드 배양물을 30 °C에서 약 18시간 동안 200 rpm으로 흔들면서 키웠다. 배양물의 작은 앨리퀀트를 4-6의 OD₆₀₀에 도달할 때까지 OD₆₀₀를 측정하기 위해 회수하였다. 존재하는 잔여 글루코스를 Uristix® Reagent Strip (Bayer, Elkhart, IN, USA)를 사용하여 측정하였다. OD₆₀₀=0.2로 50 mL CNB1 셰이크 플라스크 배지를 함유하는 125 ml 배플 플라스크를 접종하기 위해 시드 플라스크 배양을 사용하였다. 배양물을 30 °C에서 140 rpm으로 20시간 동안 흔들면서 배양하였다. 브로스의 샘플을 하기 설명된 바와 같이 분석하기 위해 제거하였다. 배양물의 최적의 밀도 (OD)를 측정하기 위해 샘플의 앨리퀀트를 사용하였고 존재하는 잔여 글루코스를 Uristix® Reagent Strip을 사용하여 측정하였다. 샘플의 나머지를 원심분리하였고 상층액을 생성물 분석에 사용하였다.

[0318] 실시에 2L: 생성물 분석을 위해 변형된 효모 균주의 발효 방법

[0319] 여기에 설명된 균주를 시드 증식 단계를 사용한 후 이어서 2 L 바이오리액터 (Applikon, Foster City, CA, USA)에서 단일 단계 발효를 사용하여 배양하였다.

[0320] 시드 단계 제조를 위해 25 mL의 1X DM2 배지 (KOH로 원하는 pH로 조정됨)를 125 mL 배플 플라스크에 추가한 후, 이어서 멸균 루프를 사용하여 원하는 균주로 접종하였다. 배양물을 원하는 온도에서 하룻밤 동안, 예를 들어, 약 16시간 250rpm으로 흔들면서 키웠다. 배양물의 작은 앨리퀀트를 4-6의 OD₆₀₀에 도달할 때까지 OD₆₀₀를 측정하기 위해 약 한 시간 간격으로 회수하였다.

[0321] 존재하는 잔여 글루코스를 Uristix® Reagent Strip (Bayer, Elkhart, IN, USA)을 사용하여 측정하였다. 12 mL의 배양물을 4 mL의 멸균 냉각된 75% 글리세롤에 추가하였고, 완전히 혼합하였고, 10분 동안 얼음 위에서 배양하였다. 배양물 및 글리세롤 혼합물을 재혼합하였고 1.0 mL을 10개의 멸균 1.8 mL 냉동 바이알 (Thermo Scientific, Rochester, NY, USA) 각각에 앨리퀀트하였고 -80 °C에서 저장하였다.

[0322] 1.5 L의 DM2 배지를 함유하는 2 L 바이오리액터를 접종하기 위해 25 mL의 시드 플라스크 배양을 사용하였다. 바이오리액터에서 발효를 약 30 °C - 40 °C의 온도에서 수행하였으며, 약 2.0-7.0의 범위에서 및 2-45 mmol/L/hr의 범위의 산소 흡수율 (OUR)로 이어지는 교반 및 에어레이션 (aeration) 조건 하에 pH를 조절하였다. 여기에 제공된 실시예에서, 바이오리액터에서 배양을 위한 온도, pH 및 OUR은 각각 30 °C, 4.0 및 25-30이었다.

[0323] 발효 브로스의 샘플을 분석을 위해 정기적으로 제거하였다. 간단히 말하면, 샘플의 앨리퀀트를 배양물의 최적의 밀도 (OD), 글루코스 농도 및 pH를 측정하기 위해 사용하였다. 샘플의 나머지를 원심분리하였다. 펠렛을 효소 검정을 위해 -80 °C에서 저장하였고, 상층액을 3-HP 및 다른 세포 외 화합물의 분석을 위해 사용하였다. 여기에 보고된 모든 3-HP 생성값은 달리 명시되지 않으면 발효의 48시간 시점의 것이다. 바이오리액터로부터 배출된 가스 중 이산화탄소 함량 및 산소 함량을 결정함으로써 발효 중 이산화탄소 생산 및 산소 소비를 측정하였다.

[0324] 실시에 2M: 변형된 효모 균주에 의해 생산된 3-HP 및 β -알라닌의 분석 방법

[0325] 배양물 샘플을 1% 포름산으로 10X 희석에 의해 산성화하였고 0.46 μm 96-웰 여과 플레이트를 통해 여과하였다. 샘플의 분해 물질 농도에 기초하여 물로 추가의 희석을 수행하였다. 추가의 10X 희석을 1 mM NH₄Ac, 0.1% NH₃ 및 5 mg/L의 ¹³C 일괄적으로 표지된 3-HP (3-HP에 대한 내부 표준으로서), 또는 1% 포름산 및 3 mg/L의 ¹³C 일괄적으로 표지된 β -알라닌 (β -알라닌에 대한 내부 표준으로서)의 샘플 버퍼로 수행하였다. 총 희석 인자는 약 100 내지 1000이었고, β -알라닌 또는 3-HP의 농도에 기초하여 사용하였다.

[0326] 2 μL 샘플을 표 6에 나열된 장치 설정 및 컬럼을 사용하여 Agilent 6410 Triple Quad MS/MS 검출기를 갖는 MassHunter 프로그램에 의해 조절된 Agilent 1200 HPLC (Agilent)로 주사하였다. 이온 단편 피크 영역의 정량의 그것의 안정한 동위원소 대응물에 대한 비율을 이온 억제 효과를 제거하기 위한 정량 및 장치 드리프팅 (drifting)에 사용하였다. 표준 편차는 매일 검정에서 5% 이하였다.

[0327] 표 6: β-알라닌 및 3-HP 분석을 위한 LC/MS/MS 설정

| | 3-HP (¹³ C 3-HP) | β-알라닌 (¹³ C β-알라닌) |
|-----------------------------|---------------------------------------|------------------------------------|
| 컬럼 | X 브리지 HILIC 실리카 3.5 μm, 2.1 x 150 mm | 아틀란티스 HILIC 실리카 3 μm, 2.1 x 150 mm |
| 용출 버퍼 | 62% 아세트니트릴, 0.35mM NH ₄ Ac | 38% 아세트니트릴, 0.6% 포름산 |
| 유속 (mL/분) | 0.30 | 0.30 |
| 컬럼 온도 | 50°C | 45°C |
| 유지 시간 (분) | 1.07 | 1.64 |
| 실행 시간 (분) | 3 | 3 |
| 전구체 이온 | 89 (92) | 90 (93) |
| 정량자로서 생성 이온 | 59 (61) | 72 (75) |
| 검증자로서 생성 이온 | 41 (43) | 30 (31) |
| 단편화 전압 | 50 | 70 |
| 충돌 에너지 | 정량자에 대해 5, 검증자에 대해 21 | 정량자에 대해 3, 검증자에 대해 7 |
| 극성 | 음성 | 양성 |
| 분무기 N ₂ 압력 (psi) | 10 | 11 |
| N ₂ 속도 (L/분) | 32 | 35 |
| N ₂ 온도 | 300°C | 340°C |
| 모세혈관 (V) | 4000 | 4000 |
| 델타 EMV | 450 | 400 |

[0328]

[0329] 실시예 3: 3-HP 발효 경로 유전자를 발현하는 변형된 효모 균주

[0330] 효모 숙주 세포에서, 다양한 3-HP 발효 경로에 수반되는 효소를, 단독으로 또는 조합으로, 암호화하는 하나 이상의 유전자가 발현될 수 있다. 3-HP 경로 효소는 외인성 유전자, 내인성 유전자, 또는 이들의 일부 조합으로부터 발현될 수도 있다. 발현되는 외인성 유전자는 유전자 발현 구조물, 예를 들어, 실시예 2에 설명된 발현 구조물을 사용하여 효모 세포로 도입될 수도 있다. 외인성 유전자는 단일 부위에 또는 다수의 위치에서 숙주 효모 게놈으로 통합될 수도 있고, 외인성 유전자의 통합은 하기 설명된 바와 같이 삽입 부위의 표적 유전자의 삭제 또는 붕괴와 결합될 수도 있다.

[0331] 실시예 3A: 아스파르테이트/말로네이트 세미알데히드 경로 유전자를 발현하는 변형된 효모 균주

[0332] PEP 및/또는 피루베이트, OAA, 아스파르테이트, β-알라닌, 및 말로네이트 세미알데히드 중간물을 이용하는 경로를 통해 3-HP를 생산하는 효모 세포는 경로에 수반된 하나 이상의 효소를 발현함으로써 설계될 수 있다. 발현된 유전자는 PPC, PYC, AAT, ADC, BAAT, gabT, 3-HPDH (말로네이트 세미알데히드 리덕타제 포함), HIBADH, 또는 4-히드록시부티레이트 데히드로게나제 유전자 중 하나 이상을 포함할 수도 있다.

[0333] 발현된 유전자는 숙주 세포에 네이티브한 유전자로부터 유도될 수도 있다. 예를 들어, 효모 숙주 세포가 아이. 오리엔탈리스인 경우, 발현된 유전자는 아이. 오리엔탈리스 PYC (예를 들어, SEQ ID NO: 2에 아미노산 서열을 암호화하는 및/또는 SEQ ID NO: 1에 설명된 뉴클레오티드 서열의 암호화 영역을 포함하는 아이. 오리엔탈리스 PYC 유전자), AAT (예를 들어, SEQ ID NO: 14에 설명된 아미노산 서열을 암호화하는 및/또는 SEQ ID NO: 13에 설명된 뉴클레오티드 서열의 암호화 영역을 포함하는 아이. 오리엔탈리스 AAT 유전자), BAAT (예를 들어, SEQ ID NO: 20에 설명된 아미노산 서열을 암호화하는 및/또는 SEQ ID NO: 19에 설명된 뉴클레오티드의 암호화 영역을 포함하는 아이. 오리엔탈리스 pyd4 상동체 유전자), 또는 3-HPDH (예를 들어, SEQ ID NO: 26에 설명된 아미노산 서열을 암호화하는 및/또는 SEQ ID NO: 25에 설명된 뉴클레오티드 서열의 암호화 영역을 포함하는 YMR226C 유전자에 대한 아이. 오리엔탈리스 상동체) 유전자로부터 유도될 수도 있다. 효모 숙주 세포가 또 다른 3-HP 내성 효모 균주인 경우, 유전자 서열은 업계에 알려진 기술을 사용하여 얻어질 수 있고 경로 유전자에 대하여 네이티브 상동체는 외인성으로 또는 외인성 조절 요소와 결합하여 발현될 수 있다. 네이티브 경로 유전자는 하나 이상의 PPC, PYC, AAT, BAAT, 및/또는 3-HPDH 유전자를 포함할 수도 있다.

[0334] 대안으로, 발현된 3-HP 유전자 중 하나 이상은 숙주 세포에 비-네이티브인 근원 유전자로부터 유도될 수도

있다. 예를 들어, 효모 숙주 세포가 아이. 오리엔탈리스인 경우, 세포는 하나 이상의 비-네이티브 PYC 유전자, 예를 들어, SEQ ID NO: 3의 아미노산 서열을 암호화하는 알. 스페로이데스 PYC 유전자, SEQ ID NO: 4의 아미노산 서열을 암호화하는 알. 에틀리 PYC 유전자, SEQ ID NO: 5 또는 6의 아미노산 서열을 암호화하는 피. 플루오레센스 PYC 유전자, SEQ ID NO: 7의 아미노산 서열을 암호화하는 씨. 글루타미쿰 PYC 유전자, SEQ ID NO: 7의 아미노산 서열을 암호화하는 씨. 글루타미쿰 PYC 유전자, 또는 SEQ ID NO: 8의 아미노산 서열을 암호화하는 에스. 펠리로티 PYC 유전자; 하나 이상의 비-네이티브 PPC 유전자, 예를 들어, SEQ ID NO: 10의 아미노산 서열을 암호화하는 이. 콜리 PPC 유전자, SEQ ID NO: 11의 아미노산 서열을 암호화하는 엠. 씨모아우토토로피쿰 PPC 유전자, 또는 SEQ ID NO: 12의 아미노산 서열을 암호화하는 씨. 페르프링겐스 PPC 유전자; 하나 이상의 비-네이티브 AAT 유전자, 예를 들어, SEQ ID NO: 16의 아미노산 서열을 암호화하는 이. 콜리 aspC 유전자 또는 SEQ ID NO: 15의 아미노산 서열을 암호화하는 에스. 세레비시에 AAT2 유전자; 하나 이상의 비-네이티브 ADC 유전자, 예를 들어, SEQ ID NO: 17의 아미노산 서열을 암호화하는 (및/또는 SEQ ID NO: 130, 145, 146, 또는 147에 설명된 뉴클레오티드 서열의 암호화 영역을 포함하는) 에스. 아베르미틸리스 panD 유전자, SEQ ID NO: 18의 아미노산 서열을 암호화하는 (및/또는 SEQ ID NO: 131에 설명된 뉴클레오티드 서열의 암호화 영역을 포함하는) 씨. 아세토부틸리쿰 panD 유전자, SEQ ID NO: 133의 아미노산 서열을 암호화하는 (및/또는 SEQ ID NO: 132에 설명된 뉴클레오티드 서열의 암호화 영역을 포함하는) 에이치. 필로리 ADC 유전자, SEQ ID NO: 135의 아미노산 서열을 암호화하는 (및/또는 SEQ ID NO: 134에 설명된 뉴클레오티드 서열의 암호화 영역을 포함하는) 바실루스 중 TS25 ADC 유전자, SEQ ID NO: 137의 아미노산 서열을 암호화하는 (및/또는 SEQ ID NO: 136에 설명된 뉴클레오티드 서열의 암호화 영역을 포함하는) 씨. 글루타미쿰 ADC 유전자, 또는 SEQ ID NO: 139의 아미노산 서열을 암호화하는 (및/또는 SEQ ID NO: 138, 148, 149, 150, 또는 151 중 어느 하나에 설명된 뉴클레오티드 서열의 암호화 영역을 포함하는) 비. 리체니포르미스 ADC 유전자; 하나 이상의 비-네이티브 BAAT 또는 gabT 유전자, 예를 들어, SEQ ID NO: 21의 아미노산 서열을 암호화하는 (및/또는 SEQ ID NO: 142에 설명된 뉴클레오티드 서열의 암호화 영역을 포함하는) 에스. 클루이베리 pyd4 유전자, SEQ ID NO: 22의 아미노산 서열을 암호화하는 (및/또는 SEQ ID NO: 140에 설명된 뉴클레오티드 서열의 암호화 영역을 포함하는) 에스. 아베르미틸리스 BAAT 유전자, SEQ ID NO: 23에 설명된 아미노산 서열을 암호화하는 에스. 아베르미틸리스 gabT 유전자, 또는 SEQ ID NO: 24에 설명된 아미노산 서열을 암호화하는 (및/또는 SEQ ID NO: 141에 설명된 뉴클레오티드 서열의 암호화 영역을 포함하는) 에스. 세레비시에 UGA1 유전자; 하나 이상의 비-네이티브 3-HPDH 유전자, 예를 들어, SEQ ID NO: 27의 아미노산 서열을 암호화하는 (및/또는 SEQ ID NO: 143에 설명된 뉴클레오티드 서열의 암호화 영역을 포함하는) 이. 콜리 ydfG 유전자 또는 SEQ ID NO: 129의 아미노산 서열을 암호화하는 (및/또는 SEQ ID NO: 144에 설명된 뉴클레오티드 서열의 암호화 영역을 포함하는) 에스. 세레비시에 YMR226C 유전자; 하나 이상의 비-네이티브 말로네이트 세미알데히드 리덕타제 유전자, 예를 들어, SEQ ID NO: 29에 설명된 아미노산 서열을 암호화하는 (및/또는 SEQ ID NO: 343에 설명된 뉴클레오티드 서열의 암호화 영역을 포함하는) 엠. 세둘라 말로네이트 세미알데히드 리덕타제 유전자; 하나 이상의 비-네이티브 HIBADH 유전자, 예를 들어, SEQ ID NO: 28에 설명된 아미노산 서열을 암호화하는 에이. 패칼리스 M3A 유전자, 각각 SEQ ID NO: 30 또는 SEQ ID NO: 31에 설명된 아미노산 서열을 암호화하는 피. 푸티다 KT2440 또는 E23440 mmsB 유전자, 또는 SEQ ID NO: 32에 설명된 아미노산 서열을 암호화하는 피. 애루기노사 PA01 mmsB 유전자; 및/또는 하나 이상의 비-네이티브 4-히드록시부티레이트 데히드로게나제 유전자, 예를 들어, SEQ ID NO: 33에 설명된 아미노산 서열을 암호화하는 알. 유티로파 H16 4hbd 유전자 또는 SEQ ID NO: 34에 설명된 아미노산 서열을 암호화하는 씨. 클루이베리 DSM 555 hbd 유전자를 발현하도록 설계될 수도 있다.

[0335] 실시예 3A-0: 아스파르트레이트/말로네이트 세미알데히드 경로 유전자를 발현하는 변형된 효모 균주에 대한 효모 활성 검정

[0336] 효소 검정을 위한 크루드 (crude) 무세포 추출물의 제조:

[0337] 여기에서 셰이크 플라스크 또는 바이오리액터 배양물의 지시된 세포를 상기 설명된 바와 같이, 원심분리, 제거된 상층액, 및 -80 °C에 저장된 세포 펠렛에 의해 수거하였다. CFE의 제조를 위해, 세포 펠렛을 녹였고, 포스페이트-완충된 식염수 (PBS)로 세척하였고 다시 원심분리에 의해 수거하였다. 상층액을 제거하였고 세포 펠렛을 2.0 mL 미세원심분리 튜브에서 1% 프로테아제 억제제 카테일 (Sigma의 P8215)을 함유하는, 대략 동등한 양의 용해 버퍼에 재현탁하였다. 약 300 µL의 0.5 mM 지르코니아 비드 (BioSpec)를 추가하였고, 세포 용해를 FastPrep ®-24 붕괴기 (MP Biomedicals)에서 3라운드 동안 6/20 초 설정으로 수행하였다. 샘플 튜브를 각 라운드 사이에 5분 동안 얼음 위에서 냉각하였다. 용해 후, 샘플을 4 °C에서 15분 동안 미세 원심분리기의 최대 속도로 원심분리하였다. 상층액을 새로운 1.5 mL 튜브로 옮겼고 얼음 위에 보관하였다. 제조사에 의해 제공된 지시에 따라 Bio-Rad 단백질 검정 시약 (Brad d Assay) 및 표준으로서 소 혈청 알부민을 사용하여 용해물의 총 단백질 농도

를 결정하였다.

[0338] *피루베이트 카르복실라제 (PYC) 활성:*

[0339] 여기에서 지시된 세포의 CFE에서 피루베이트 카르복실라제 활성을 다음과 같이 결정하였다. 검정 반응 혼합물에서 CFE와 결합될 때, 다음 최종 농도의 성분을 제공하는 스톱 반응 혼합 용액을 제조하였다: 트리스 (pH 8.0), 100 mM; NaHCO_3 , 10 mM; MgCl_2 , 5 mM; NADH, 0.2 mM; ATP, 1 mM; 아세틸 CoA, 1 mM; 피루베이트, 1 mM; 비오틴 (검정되는 PYC 효소에 필요하면), 5 μM ; 소 심장 말레이트 데히드로게나제, 0.02 유닛/mL. 이 용액의 270 μL 를 96-웰 미세역가 플레이트의 웰에 추가하였고 30 μL 의 적절히 희석된 CFE를 반응을 시작하기 위해 추가하였다. NADH의 소모를 SpectraMax 340 PC 플레이트 리더를 사용하여 340 nm에서 관찰하였다. 피루베이트 카르복실라제 활성은 나노몰 소모된 NADH/초/mg 단백질로서 표현된다.

[0340] *포스포에놀피루베이트 카르복실라제 (PPC) 활성:*

[0341] CFE에서 포스포에놀피루베이트 카르복실라제 (PPC) 활성을 다음과 같이 결정할 수도 있다. 검정 반응 혼합물에서 CFE와 결합될 때, 다음 최종 농도의 성분을 제공하는 스톱 반응 혼합 용액을 제조하였다: 트리스 (pH 8.0), 100 mM; NaHCO_3 , 10 mM; MgCl_2 , 5 mM; NADH, 0.1 mM; 아세틸 CoA, 0.5 mM; 포스포에놀피루베이트, 3.3 mM; 소 심장 (또는 돼지 심장) 말레이트 데히드로게나제, 0.02 유닛/mL. 이 혼합물의 270 μL 를 96-웰 미세역가 플레이트의 웰에 추가하였고 30 μL 의 적절히 희석된 CFE를 반응을 시작하기 위해 추가하였다. NADH의 소모는 SpectraMax 340 PC 플레이트 리더를 사용하여 340 nm에서 관찰된다.

[0342] *아스파르테이트 아미노트랜스퍼라제 (AAT) 활성:*

[0343] 여기에서 지시된 세포의 CFE에서 아스파르테이트 아미노트랜스퍼라제 활성을 다음과 같이 결정하였다. 검정 반응 혼합물에서 CFE와 결합될 때, 다음 최종 농도의 성분을 제공하는 스톱 반응 혼합 용액을 제조하였다: 100 mM 트리스 (pH 8.0), 100 mM; NaHCO_3 , 10 mM; MgCl_2 , 5 mM; NADH, 0.1 mM; 아스파르테이트, 1mM; α -케토글루타레이트, 1 mM; 및 말레이트 데히드로게나제, 0.02 유닛/mL. 일부 검정에서, 스톱 반응 혼합물은 또한 피리독살 5'-포스페이트 (0.1 mM)를 함유하였다. 이 용액의 270 μL 를 96-웰 미세역가 플레이트의 웰에 추가하였고 30 μL 의 적절히 희석된 CFE를 반응을 시작하기 위해 추가하였다. NADH의 소모를 SpectraMax 340 PC 플레이트 리더를 사용하여 340 nm에서 관찰하였다. 아스파르테이트 아미노트랜스퍼라제 활성은 nmol 소모된 NADH/초/mg 단백질로서 표현된다.

[0344] *아스파르테이트 데카르복실라제 (ADC) 활성:*

[0345] 여기에서 지시된 세포의 CFE에서 아스파르테이트 데카르복실라제 활성을 다음과 같이 결정하였다. 165 μL 의 100 mM NH_4Ac 버퍼 (pH 6.8), 및 25 μL 의 80 mM 아스파르테이트를 37 $^\circ\text{C}$ 로 온도 조절되는 96-웰 미세역가 플레이트의 각각의 웰에 추가하였다. 반응을 10 μL 의 CFE를 추가함으로써 개시하였다. 다른 시간 간격 (5, 10, 15, 20, 25, 30, 40, 60분)으로, 샘플의 20 μL 를 반응 혼합물로부터 회수하였고 퀀칭 (quenching) 버퍼 (2% 포름산 플러스 내부 표준으로서 ^{13}C 표지된 β -알라닌 2mg/L)의 180 μL 에 추가하였다. 여과 후, 샘플에서 β -알라닌을 LC/MS/MS로 분석하였다. 기울기를 β -알라닌 대 시간 플롯으로부터 얻었다. 기울기를 반응물의 총 세포 단백질 농도로 나눔으로써 활성을 계산하였다. ADC 활성은 μmol 형성된 β -알라닌/초/g 단백질로서 표현된다.

[0346] 변형된 ADC 검정을 일부 실험에서 사용하였다. 이들 경우에, 여기에서 지시된 세포의 CFE에서 ADC 활성을 다음과 같이 결정하였다. 110 μL 의 100 mM NH_4Ac 버퍼 (pH 7.6), 및 80 μL 의 25 mM 아스파르테이트 (NaOH로 중화한 후)를 40 $^\circ\text{C}$ 로 온도 조절된 96-웰 미세역가 플레이트의 각 웰에 추가하였다. 10 μL 의 CFE를 추가함으로써 반응을 개시하였다. 다른 시간 간격 (2, 4, 6, 8, 10분)으로, 샘플의 20 μL 를 반응 혼합물로부터 회수하였고 퀀칭 버퍼 (2% 포름산 플러스 내부 표준으로서 ^{13}C 로 표지되거나 2% 포름산에서 퀀칭되고 20% 메탄올/80% 물과 내부 표준으로서 2mg/L ^{13}C 표지된 β -알라닌을 1:10으로 옮긴다)의 180 μL 에 추가하였다. 여과 후, 샘플에서 β -알라닌을 LC/MS/MS로 분석하였다. 기울기를 β -알라닌 대 시간 플롯으로부터 얻었다. 기울기를 반응물의 총 세포 단백질 농도로 나눔으로써 활성을 계산하였다. ADC 활성은 μmol 형성된 β -알라닌/초/g 단백질로서 표현된다.

[0347] *β -알라닌 아미노트랜스퍼라제 (BAAT) 활성:*

- [0348] CFE에서 β -알라닌 아미노트랜스퍼라제 (BAAT) 활성을 다음과 같이 결정하였다. 100 mM의 NH_4HCO_3 (pH 7.6), 8 mM α -케토글루타레이트, 0.5 mM 아세틸-CoA, 0.1 mM 피리독살-5'-포스페이트, 및 200 mM β -알라닌을 함유하는 190 μL 의 반응 혼합물을 상온에서 96 웰 미세역가 플레이트에 추가하였다. 10 μL 의 CFE를 추가함으로써 반응을 개시하였다. 샘플 20 μL 각각 2, 4, 6, 8, 10, 12, 15, 및 20분에 채취하였고 쿨링 버퍼 (2.5% 포름산) 75 μL 에 추가하였다. 샘플을 중화하였고 5 μL 10 M NaOH 및 50 μL 100 mM NaCO_3 (pH 10)를 추가함으로써 pH를 조절하였다. 여과된 샘플을 주사시, OPA (o-프탈디알데히드) 시약, 10 mg/mL (Agilent Technologies 5061-3335)과 혼합함으로써 유도하였다. OPA로 유도된 글루타메이트를 형광 발광 검출에 의해 HPLC 분리 후 정량하였다 (340 nm에서 들뜸; 460 nm에서 방출). 15 μL 의 샘플을 5 μL 패킹 (Phenomenex 150 x 4.6 mm)과 함께 분석적 역상 제미니 C18 컬럼 (Gemini C18 column)으로 주사하였다. 컬럼을 62.5% 20 mM 포스페이트 버퍼 (pH 7.8) (A) 및 37.5% 메탄올 (B)로 평형화하였다. 선형 구배는 다음과 같다: 40% B로 램프, 0-0.3분; 40% B, 0.3-1분; 85% B로 램프, 1-1.75분; 85% B, 1.75-2.25분; 37.5%로 램프, 2.25-3분; 37.5% B, 3-4분. 유속은 2 mL/분이었다. 반응 버퍼에서 글루타메이트의 표준 곡선을 샘플의 농도를 결정하기 위해 사용하였다. 기울기를 [글루타메이트] 대 시간 플롯으로부터 얻었다. 기울기를 반응물의 총 세포 단백질 농도로 나눔으로써 활성을 계산하였다.
- [0349] *3-HP 데히드로게나제 (3-HPDH) 활성:*
- [0350] 여기에서 지시된 세포의 CFE에서 3-HP 데히드로게나제 활성을 다음과 같이 결정하였다. 100 mM NH_4HCO_3 (pH 7.6) 및 NADPH에서 희석 (전형적으로 100X 희석)된 CFE의 190 μL 를 37 $^\circ\text{C}$ 로 온도 조절된 6-웰 미세역가 플레이트의 각 웰에 추가하였다. 60mM 말로네이트 세미알데히드 (MSA, 10 mM H_2SO_4 중의 200 mM MSA 스톡 용액으로부터 10 mM H_2SO_4 로 새롭게 제조됨) 10 μL 를 추가함으로써 반응을 개시하였다. 20 μL 의 샘플을 각각 1, 2, 4, 6, 8, 10, 및 12분에 채취하였고, 끓는 물 80 μL 에서 쿨링하였다. 냉각 후, 쿨링된 혼합물 75 μL 를 2 mM NH_4Ac (pH 6.8) 및 3 mg/L의 ^{13}C 표지된 3-HP를 함유하는 버퍼 75 μL 와 혼합하였다. 여과 후, 샘플의 3-HP를 LC/MS/MS로 정량하였다. 기울기를 3-HP 대 시간 플롯으로부터 얻었다. 기울기를 반응물의 총 세포 단백질 농도로 나눔으로써 활성을 계산하였다. 3-HP 데히드로게나제 활성은 nmol 형성된 NADPH/초/mg 단백질로서 표현된다.
- [0351] 변형된 3-HPDH 검정을 일부 실험에 사용하였다. 이들 경우에서, 여기에서 지시된 세포의 CFE에서 3-HPDH 활성을 다음과 같이 결정하였다. 340 nm에서 시간의 흐름에 따른 말로네이트 세미-알데히드 감소를 다음 NADPH의 소실로 측정하였다. 말로네이트 세미-알데히드를 Yamada 및 Jacoby에 의해 개발된 프로토콜에 따라 내부에서 합성하였다 (Yamada, E.W., Jacoby, W.B., 1960, Direct conversion of malonic semialdehyde to acetyl-coenzyme A, Journal of Biological Chemistry, Volume 235, Number 3, pp. 589-594). 검정을 96 웰 마이크로플레이트에서 수행하였고, 최종부피는 200 μL 였다. 반응을 CCE 30 μL 를 검정 버퍼 (2 mM 말로네이트 세미-알데히드, 100 mM 트리스 pH 8.0 및 0.5 mM NADPH) 170 μL 에 추가함으로써 시작하였다. 340 nm에서 흡수율은 상온 ($\sim 25^\circ\text{C}$)에서 10분 동안 마이크로 플레이트 리더 (Spectra Max 340PC, Molecular Devices LLC, Sunnyvale, CA)에서 선행되었다. 한 유닛의 3-HPDH 활성은 말로네이트 세미-알데히드의 존재시 1분에 NADPH의 1 μmol 을 산화하기 위해 필요한 효소의 양으로 정의된다.
- [0352] 실시예 3A-1: adh1202 위치에서 아스파르테이트 데카르복실라제 (ADC)를 발현하는 삽입 벡터
- [0353] 여러 아스파르테이트 데카르복실라제 유전자를 아이. 오리엔탈리스에서 발현에 코돈-최적화하였고 GeneArt® (Burlingame, CA, USA)로 합성하였으며, 표 7에 나열된 플라스미드를 발생시킨다. 합성 유전자는 벡터 pMA-T에 도착하고 XbaI 및 PacI 제한 분해를 통해 벡터로부터 유도될 수 있다. 제한 단편을 PDC 프로모터 및 종결자에 의해 조절되는 유전자를 배치하고, 통합이 아이. 오리엔탈리스 adh1202 위치에서 발생하는 것을 허용하는 pMBa107의 같은 부위에 클로닝할 수 있다.

[0354] 표 7: 형질전환체 구성

| 구조 플라스미드 | 유전자 근원 | 유전자 번호 | 유전자 SEQ ID NO | 통합 구조 | 형질전환체 |
|----------|---|---------------|---------------|---------------|----------------------|
| 1051387 | 헬리코박터 필로리 (<i>Helicobacter pylori</i>) | P56065 | 132 | pWTY10-0033-1 | yWTY1-1 yWTY1-2 |
| 1051391 | 바실루스 종 TS25 | ZY440006.유전자3 | 134 | pWTY10-0033-2 | yWTY1-5 yWTY1-6 |
| 1051389 | 코리네박테리움 글루타미쿰 (<i>Corynebacterium glutamicum</i>) | Q9X4N0 | 136 | pWTY10-0033-3 | yWTY1-9 yWTY1-10 |
| 1051388 | 클로스트리듐 아세토부틸리쿰 (<i>Clostridium acetobutylicum</i>) | P58285 | 131 | pWTY10-0033-4 | yWTY1-13 yWTY1-14 |
| 1051390 | 바실루스 리체니포르미스 (<i>Bacillus licheniformis</i>) | Q65158 | 138 | pWTY10-0033-5 | yWTY1-17 yWTY1-18 |
| 1045172 | 스트렙토미세스 아베르미틸리스 (<i>Streptomyces avermitilis</i>) | | 130 | pWTY10-0033-7 | yWTY1-25 yWTY1-26 |

[0355]

[0356] 플라스미드 1045172, 105387, 105388, 105389, 105390, 및 105391을 XbaI 및 PacI로 분해하였고 TBE 버퍼를 사용하여 1.3% 아가로스 겔 상에서 실행하였다. ADC (panD) 유전자에 해당하는 각각의 분해로부터 400-500 bp의 단편을 겔에서 잘라냈고 QIAQUICK® Gel Extraction Kit (Qiagen)를 사용하여 아가로스로부터 추출하였다.

[0357] 플라스미드 pMBa107를 XbaI 및 PacI로 분해하였고, 소 위장 포스포타제 (New England Biolabs)를 처리하였고 TBE 버퍼 중의 아가로스 겔 전기영동 후 벡터 밴드를 정제하였다. XbaI 및 PacI 분해된 panD 단편을 제조사의 지시에 따라 T4 DNA 리가제 및 QUICK Ligation Kit (New England Biolabs)를 사용하여 이 정제된 pMBa107 벡터에 결합하였다. 결합 생성물을 제조사의 지시에 따라 XL10-GOLD ULTRA 세포 (Agilent Technologies)에 형질 전환하였다. 형질전환체를 2X YT+ amp 플레이트에 도말하였고 37 °C에서 하룻밤 동안 배양하였다. 각각의 반응으로부터 24개의 형질전환체를 2X YT + amp 플레이트에서 골랐다. 결과의 형질전환체 중 네 개 각각의 미니-프랩 DNA를 ApaI, NcoI 및 SacI 분해로 스크리닝하였다. 원하는 밴드 크기를 획득하는 클론을 DNA 시퀀싱에 의해 올바른 것으로 확인하였고 표 7에 나타난 바와 같이 지정하였다. 결과의 플라스미드는 adh1202 위치에서 각 ADC 유전자의 정방향으로 향하는 발현 카세트와 통합을 허용한다.

[0358] 각 통합 구조물 중 약 10 µg씩을 ApaI 및 KpnI로 분해하였고 89 mM 염기성 트리스-89 mM 붕산-2 mM 이나트륨 EDTA (TBE) 버퍼를 사용하여 1% 아가로스 겔 상에서 실행하였다. 각 플라스미드에 대하여 약 4450 bp의 단편을 겔에서 잘라냈고 제조사의 지시에 따라 QIAquick gel extraction kit (Qiagen)을 사용하여 추출하였다. 정제된 생성물의 농도는 39-138 ng/µl인 것으로 발견되었다. 통합 구조물 (ApaI 및 Kpn I로 분해됨)의 단편의 0.39-1.4 µg을 상기 설명된 바와 같이 아이. 오리엔탈리스 CNB1에 형질전환하였다. 형질전환체를 ura 선택 배지에 도말하였고 37 °C에서 배양하였고, ura 선택 배지에서 재-스트리킹하였고, 37 °C에서 하룻밤 동안 배양하였다. 게놈 DNA를 URA3+ 콜로니로부터 제조하였고 통합을 확인하기 위해 PCR로 체크하였다. 프라이머 0611718 및 0611632를 통합을 확인하기 위해 2.5 kbp 단편을 증폭하는데 사용하였다. 각각의 PCR 반응물은 25 µL의 최종 부피 중에 2.5 µL의 게놈 DNA, 각각 12.5 pM의 프라이머 0611718 및 0611632, 1X Crimson Taq™ 반응 버퍼 (New England Biolabs), 0.2 mM dNTP 혼합물, 0.625 유닛 Crimson Taq™ DNA 폴리머라제 (New England Biolabs)를 함유하였다. 증폭 반응물을 95 °C에서 30초 동안 1 사이클; 각각 95 °C에서 30초 동안, 50 °C에서 30초 동안, 및 68 °C에서 3분 동안 30 사이클; 및 68 °C에서 10분 동안 1 사이클로 프로그램된 EPPENDORF® MASTERCYCLER® (Eppendorf Scientific)에서 배양하였다.

[0359] 각각의 구조물에 대하여 2개의 URA3+ 확인된 형질전환체를 표 7에 나타난 바와 같이 지정하였다. 이들 균주는 adh1202에서 지시된 ADC 유전자에 대한 이형 접합체 (heterozygous)이며, 발현은 아이. 오리엔탈리스의 PDC 프로모터 및 종결자에 의해 구동된다.

[0360] 표 7에 나열된 균주 (및 음성 대조군으로서 균주 MBin500, 상기)의 PanD 발현 및 효소 활성을 테스트하였다. 각 균주의 하룻밤 동안의 배양물을 YPD에서 37 °C에서 하룻밤 동안 배양하였고, 37 °C에서 125 mL 배플 플라스크 중의 신선한 YPD 25 mL에 1:50으로 희석하였고, OD₆₀₀ ~2-8로 성장시켰다. 세포 펠렛을 CFE를 제조하기 위해 사

용하였으며, 이것을 상기 설명된 바와 같이 ADC 활성에 대하여 검정하였다. 대표적인 결과는 표 8에 나타난다.

표 8: 형질전환체 효소 활성 데이터

| 균주 | ADC 유전자의 근원 | 유전자 SEQ ID NO | ADC 활성 |
|----------------------|--------------------|------------------|--------------|
| MBin500 (대조군) | N/A | N/A | 0 |
| yWTY1-1 yWTY1-2 | 헬리코박터 필로리 | 132 | 0 |
| yWTY2-5 yWTY2-6 | 바실루스 종 TS25 | 134 | 0.31 0.15 |
| yWTY3-9 yWTY3-10 | 코리네박테리움 글루타미쿰 | 136 | 0.31 0.19 |
| yWTY4-13 yWTY4-14 | 클로스트리듐 아세토부틸리쿰 | 131 | 0.37 0.23 |
| yWTY5-17 yWTY5-18 | 바실루스 리체니포르미스 | 138 | 0.56 0.61 |
| yWTY7-25 yWTY7-26 | 스트렙토미세스 아베르미틸리스 | 130 | 0.23 0.30 |

다음으로, yWTY5-17 및 yWTY7-25의 동형 접합체 버전을 생성하였다. 먼저, ura-유도체 yWTY5-17 및 yWTY7-25를 상기 설명된 바와 같이 분리하였다. 게놈 DNA를 FOA-저항성 콜로니로부터 제조하였고 URA3 선택 가능한 마커의 손실을 확인하기 위해 상기 설명된 바와 같이 PCR로 체크하였다. 프라이머 0611718 및 0611632를 존재하는 ura 마커 및 ura 마커의 부재시 1100 bp 단편과 통합하는 2.4 kbp 단편을 증폭하기 위해 사용하였다. 프라이머 0611718 및 0611632를 갖는 1100 bp의 PCR 단편을 수득하는 yWTY5-17 및 yWTY7-25의 Ura-균주를 각각 M1Ba331 및 M1Ba332로 지정하였다.

pWTY10-0033-5 및 pWTY10-0033-7의 각각 10 µg을 ApaI, KpnI, 및 NcoI로 분해하였고 89 mM 염기성 트리스-89 mM 붕산-2 mM 이나트륨 EDTA (TBE) 버퍼를 사용하여 1% 아가로스 겔 상에서 수행하였다. 각 플라스미드에 대한 약 4450 bp의 단편을 겔에서 잘라냈고 QIAQUICK® Gel Extraction Kit (Qiagen)를 사용하여 추출하였다. pWTY10-0033-5 및 pWTY10-0033-7로부터 정제된 단편을 상기 설명된 바와 같이 각각 M1Ba331 및 M1Ba332에 형질 전환하였다. 형질전환체를 ura 선택 배지에 도말하였고 37 °C에서 하룻밤 동안 배양하였고, ura 선택 배지에 재-스트리킹하였고 37 °C에서 하룻밤 동안 배양하였다. 게놈 DNA를 제조하였고 Crimson Taq™ PCR을 상기 설명된 바와 같이 통합을 확인하기 위해 실행하였다. 프라이머 0611718 및 0611632는 존재하는 ura 마커와 통합하는 2.4 kbp 단편을 증폭하고, ura 마커의 부재시 1100 bp 단편을 증폭한다. 프라이머 0611718 및 0611632를 갖는 1100 bp 및 2.4 kbp의 단편을 수득하는 M1Ba331 및 M1Ba332의 형질전환체를 저장하였고 각각 M1Ba338 및 M1Ba337로 지정하였다. M1Ba337은 adh1202 위치에서 에스. 아베르미틸리스의 ADC 유전자에 대한 동형 접합체이고 M1Ba338은 adh1202 위치에서 비. 리체니포르미스의 ADC 유전자에 대한 동형 접합체이다. 균주 둘 다는 아이. 오리엔탈리스의 PDC 프로모터 및 종결자에 의한, 각각의 ADC 유전자의 조절을 갖는다.

panD 동형 접합체 균주 M1Ba337 및 M1Ba338의 ADC 발현 및 효소 활성을 이형 접합체 panD 균주 yWTY5-17 및 yWTY7-25와 비교하였다. 배양물을 YPD에서 37 °C에서 하룻밤 동안 성장시켰고, 37 °C에서 125 mL 배플 플라스크 중의 신선한 YPD 25 mL에 1:50으로 희석하였고, OD₆₀₀ ~2-8로 성장시켰다. 세포 펠렛을 CFE를 제조하기 위해 사용하였으며, 이것을 상기 설명된 바와 같이 ADC 활성에 대하여 검정하였다. 대표적인 결과는 표 9A에 나타난다.

표 9A: 형질전환체 효소 활성 데이터

| 균주 | panD 유전자의 근원 | 대립유전자 타입 | 유전자 SEQ ID NO | ADC 활성 (Exp. 1, Exp. 2) |
|------------------|--------------------|----------|------------------|----------------------------|
| MBin500 (대조군) | N/A | N/A | N/A | 0, 0 |
| yWTY5-17 | 바실루스 리체니포르미스 | 이형 접합체 | 138 | 0.6, 0.29 |
| M1Ba338 | 바실루스 리체니포르미스 | 동형 접합체 | 138 | 0, 0.22 |
| yWTY7-25 | 스트렙토미세스 아베르미틸리스 | 이형 접합체 | 130 | 0.19, 0.13 |
| M1Ba337 | 스트렙토미세스 아베르미틸리스 | 동형 접합체 | 130 | 0.28, 0.19 |

균주 M1Ba337 및 M1Ba338로부터 제조된 CFE에서 ADC 활성과 비교할 세 번째 독립적인 실험의 결과는 표 9B에 나

타난다.

[0369] 표 9B: 형질전환체 효소 활성 데이터

| 균주 | panD 유전자의 근원 | 대립유전자 타입 | 유전자 SEQ ID NO | ADC 활성 |
|---------------|-----------------|----------|---------------|--------|
| MBin500 (대조군) | N/A | N/A | N/A | 0 |
| yWTY5-17 | 바실루스 리체니포르미스 | 이형 접합체 | 138 | 0.218 |
| MBa338 | 바실루스 리체니포르미스 | 동형 접합체 | 138 | 0.453 |
| yWTY7-25 | 스트렙토미세스 아베르미틸리스 | 이형 접합체 | 130 | 0.087 |
| MBa337 | 스트렙토미세스 아베르미틸리스 | 동형 접합체 | 130 | 0.188 |

[0370]

[0371] 상기 샘플의 SDS-PAGE 분석은 MBa338의 panD 발현이 이들 균주 중에 가장 높다는 것을 나타냈다.

[0372] 여기에 설명된 방법을 사용하여, 균주 MBin500, MBa337 및 MBa338을 3-HP 생산에 대하여 바이오리액터로 평가하였다. 대조군 균주 MBin500은 검출 가능한 3-HP를 생산하지 않았다 (2개의 독립적 발효의 평균). 균주 MBa337은 1.33 g/L 3-HP를 생산하였고 (하나의 발효가 수행됨) 균주 MBa338은 3.15 g/L 3-HP를 생산하였다 (3개의 독립적 발효의 평균). 균주 MBa337 및 MBa338의 개개의 발효를 추가로 그것들의 3-HP 생산 능력 및 ADC 활성에 관하여 비교하였다 (표 10). 이들 발효에서 세포 질량의 차이를 설명하기 위해, 3-HP 생산 능력은 표 10에서 세포 질량의 유닛 당 3-HP 농도 ([g/L 3-HP]/[g/L 건조 세포 중량])으로 표현됨)로 보고된다. 결과는 바실루스 리체니포르미스 panD 유전자 (균주 MBa338) 대 스트렙토미세스 아베르미틸리스 panD 유전자 (균주 MBa337)를 사용할 때, 개선된 ADC 활성 및 3-HP 생산 능력을 나타낸다.

[0373] 표 10: 균주 MBa337 및 MBa338에서 3-HP 생산 능력 및 ADC 활성으로 통합

| 발효 시간 (시간) | MBa337 | | MBa338 | |
|------------|------------------------|---------|--------------------|---------|
| | ADC 활성 (mmol/분/g prot) | 3HP/DCW | 활성 (mmol/분/g prot) | 3HP/DCW |
| 11 | 0.005 | 0.00 | 0.021 | 0.024 |
| 22 | 0.011 | 0.05 | 0.055 | 0.131 |
| 31 | 0.003 | 0.04 | 0.029 | 0.159 |
| 48 | 0.001 | 0.05 | 0.018 | 0.142 |

[0374]

[0375] 실시예 3A-2: adh1202 위치에서 β-알라닌 아미노트랜스퍼라제 (BAAT) 또는 3-히드록시프로피온산 데히드로게나제 (3-HPDH)를 발현하는 삽입 벡터

[0376] 에스. 아베르미틸리스의 BAAT, 에스. 세레비시아의 UGA1, 에스. 클루이베리의 PYD4, 에스. 세레비시아의 YMR226c, 및 이. 콜리의 ydfG의 아이. 오리엔탈리스 코돈 최적화된 버전을 GeneArt로 합성하였으며, 하기 나열된 플라스미드를 생성한다. 합성 유전자는 벡터 pMA-T에 도달하였고 XbaI 및 PacI를 사용하는 분해를 통해 벡터로부터 유도될 수 있다. 분해된 단편은 pMBa107의 같은 부위로 클로닝될 수 있으며, PDC 프로모터 및 종결자에 의해 조절되는 유전자를 두고 통합이 adh1202 위치에서 발생하는 것을 허용한다.

[0377] 표 11: 형질전환체 구성

| 구조 플라스미드 | 유전자 | 유전자 근원 | SEQ ID NO | 통합 구조 | 형질전환체 |
|----------|------------------|-------------|-----------|---------|--------|
| 1045169 | gabT (UGA1) | 에스. 세레비시아 | 141 | pMBa122 | MBa310 |
| 1045170 | BAAT | 에스. 아베르미틸리스 | 140 | pMBa121 | MBa309 |
| 1045171 | BAAT (PYD4) | 에스. 클루이베리 | 142 | pMBa124 | MBa312 |
| 1045173 | 3-HPDH (YMR226c) | 에스. 세레비시아 | 144 | pMBa123 | MBa311 |
| 1045168 | 3-HPDH (ydfG) | 이. 콜리 | 143 | pMBa120 | MBa308 |

[0378]

[0379] 플라스미드 1054168, 1054169, 1054170, 1054171, 1054172, 및 1054173을 XbaI 및 PacI로 분해하였고 89 mM 염기성 트리스-89 mM 붕산-2 mM 이나트륨 EDTA (TBE) 버퍼를 사용하여 1% 아가로스 겔 상에서 실행하였다.

1045168의 761 (ydfG) bp, 1045169의 1430 (UGA1) bp, 1045170의 1370 (BAAT) bp, 1045171의 1442 (PYD4) bp, 또는 1045173의 814 (YMR226c) bp의 단편을 겔에서 잘라냈고 QIAQUICK® Gel Extraction Kit (Qiagen)를 사용하여 아가로스로부터 추출하였다. 플라스미드 pM1Ba107을 XbaI 및 PacI로 분해한 후 이어서 CIP를 처리하였으며 7.9 kbp 선형 단편을 생성하였다. 분해물을 QIAQUICK® PCR Purification Kit (Qiagen)를 사용하여 정제하였다. ydfG, UGA1, BAAT, PYD4, 또는 YMR226c의 분해된 단편을 여기에 설명된 바와 같이 T4 DNA 리가제를 사용하여 pM1Ba107 (XbaI 및 PacI로 분해되고 CIP가 처리됨)로 결합하였다. 결합 생성물을 제조사의 지시에 따라 One Shot® TOP10 화학적 컴피턴트 이. 콜리 세포 (Invitrogen)에 형질전환하였다. 형질전환체를 2X YT+amp 플레이트에 도말하였고 37 °C에서 하룻밤 동안 배양하였다. 다수의 결과의 형질전환체를 XbaI 및 PacI로 분해에 의해 스크리닝하였다. 원하는 밴드 크기를 수득하는 클론을 DNA 시퀀싱에 의해 올바른 것으로 확인하였고 ydfG, BAAT, UGA1, YMR226c, 또는 PYD4에 대하여 각각 pM1Ba120, pM1Ba121, pM1Ba122, pM1Ba123, 및 pM1Ba124로 지정하였다. 결과의 플라스미드는 adh1202 위치에서 원하는 유전자의 정방향을 향하는 발현 카세트와 통합을 허용한다.

[0380] 표 11의 통합 구조물을 아이. 오리엔탈리스에서 발현에 코돈-최적화된 원하는 유전자를 PDC 프로모터 및 종결자에 의해 조절되는 adh1202 위치로 통합하기 위해 사용하였다. 발현 카세트는 또한 여기에 설명된 바와 같이 ura- 숙주 내에서 형질전환체의 선택을 허용하는, URA3 선택 가능한 마커를 함유한다. 발현 카세트 및 adh1202 상동 영역은 플라스미드의 단편의 방출을 허용하는, ApaI 및 KpnI 제한 부위에 의해 플랭크된다.

[0381] 표 11의 통합 구조물의 각각 15 µg을 ApaI, KpnI, 및 NcoI로 분해하였고 89 mM 염기성 트리스-89 mM 붕산-2 mM 이나트륨 EDTA (TBE) 버퍼를 사용하여 1% 아가로스 겔 상에서 실행하였다. NcoI로 분해는 벡터 백본을 분해하고 아가로스 겔의 원하는 단편을 추출하는 것을 더 쉽게 만든다. pM1Ba120, pM1Ba121, pM1Ba122, pM1Ba123, 및 pM1Ba124의 각각 4884 bp, 5493 bp, 5553 bp, 4937 bp, 및 5565 bp의 단편을 겔에서 잘라냈고 QIAQUICK® Gel Extraction Kit (Qiagen)를 사용하여 아가로스로부터 추출하였다. 정제된 생성물의 농도는 80-120 ng/µL 인 것으로 발견되었다. pM1Ba120-4의 제한 단편의 0.8-1.2 µg을 여기에 사용된 바와 같이 아이. 오리엔탈리스 CNB1 (ura-)에 형질전환하였다. 형질전환체를 ura 선택 배지에 도말하였고 상온에서 60시간 동안 키웠다. 형질전환체를 ura 선택 배지 상에 재-스트리킹하였고 37 °C에서 하룻밤 동안 배양하였다.

[0382] 다수의 형질전환체 각각의 통합을 확인하기 위해 콜로니 PCR로 체크하였다. 올바른 통합을 통합의 5' 및 3' 말단을 체크하는 프라이머 쌍을 사용함으로써 확인하였고 하기 나열된다. 프라이머 0611717은 역방향으로 PDC 프로모터에서 어닐링하는 한편, 프라이머 0611225는 정방향으로 URA3 선택 가능한 마커에서 어닐링한다. 프라이머 0611631 및 0611632는 각각 정방향 및 역방향으로 통합의 부위의 외부에서 어닐링한다; 프라이머 0611717 및 0611631은 올바른 구성요소의 976 bp 단편을 증폭시킨다; 프라이머 0611225 및 0611632는 올바른 구성요소의 1.4 kbp 단편을 증폭시키고; 프라이머 0611631 및 0611632는 야생형 염색체를 나타내는 2.7 kbp 단편을 증폭시키고 통합을 위해 단편 ~5kbp를 증폭시킬 것이다. 게놈 DNA를 생성하기 위해, 각각의 형질전환체 중 하나의 콜로니를 37 °C에서 30분 동안 TE 중의 0.05 U/µL 리터카제 (Sigma, St. Louis, MO, USA)의 50 µL에서 배양한 후 이어서 95 °C에서 10분 동안 배양하였다. 통합을 확인하기 위해 여기에 설명된 바와 같이 PCR을 실행하였다. 0611717 및 0611631을 갖는 976 bp, 0611225 및 0611632를 갖는 1.4 kbp, 및 0611631 및 0611632를 갖는 2.7 kbp의 PCR 단편을 수득하는 각각의 이형 접합체 중 하나의 형질전환체를 저장하였고 표 11에 나타난 바와 같이 M1Ba308, M1Ba309, M1Ba310, M1Ba311, 및 M1Ba312로 지정하였다.

[0383] 형질전환체 M1Ba308, M1Ba309, M1Ba310, M1Ba311, 및 M1Ba312의 배양물을 YPD에서 37 °C에서 하룻밤 동안 키웠다. 배양물을 37 °C에서 125 mL 배플 플라스크 중의 신선한 YPD 25 mL에 1:50으로 희석하였고, OD₆₀₀ ~4-10으로 키웠다. 세포의 샘플을 여기에 설명된 방법을 사용하여 SDS-PAGE로 단백질 발현에 대하여 분석하였다. CFE를 또한 배양물의 세포 펠렛으로부터 제조하였고, 3P 데히드로게나제 활성을 CFE에서 여기에 설명된 방법을 사용하여 측정하였다. 균주 M1Ba310 및 M1Ba312로부터 각각 UGA1 및 PYD4의 발현을 유전자에 대하여 통합되지 않는 균주에 없는 ~53 KDa 밴드의 등장에 의해 SDS-PAGE로 검출하였다. M1Ba309에서 BAAT의 발현은 이들 조건 하에 SDS-PAGE로 검출되지 않았다. 표 12A는 균주의 CFE에서 3-HP 데히드로게나제 (3-HPDH) 활성을 나타낸다.

[0384] 표 12A: 형질전환체 효소 활성 데이터

| 균주 | 과발현된 유전자 | 유전자 근원 | 유전자 SEQ ID NO | 3-HPDH 활성 |
|---------|-----------------------|----------------|---------------|------------|
| MBin500 | N/A | N/A | N/A | 0.28, 0.24 |
| MBa310 | <i>gabT</i> (UGA1) | 에스. 세레비시아 | 141 | 0.39 |
| MBa309 | BAAT | 에스. 아베르미탈리스 | 140 | 0.39 |
| MBa312 | BAAT (PYD4) | 에스. 클루이베리 | 142 | 0.45 |
| MBa311 | 3-HPDH (YMR226c) | 에스. 세레비시아 | 144 | 1.1 |
| MBa308 | 3-HPDH (ydfG) | 이. 콜리 | 143 | 0.67 |

[0385]

[0386] 개선된 검정 조건을 사용하는 독립적인 실험에서, 균주 MBin500 (대조군), MBa310, MBa309 및 MBa312로부터 제조된 CFE에서 BAAT 활성을 비교하였다. 이 실험의 결과는 표 12B에 나타난다.

[0387] 표 12B: 형질전환체 효소 활성 데이터

| 균주 | 과발현된 유전자 | 유전자 근원 | 유전자 SEQ ID NO | BAAT 활성 |
|---------|-----------------------|----------------|---------------|---------|
| MBin500 | N/A | N/A | N/A | 0.67 |
| MBa310 | <i>gabT</i> (UGA1) | 에스. 세레비시아 | 141 | 9.05 |
| MBa309 | BAAT | 에스. 아베르미탈리스 | 140 | 0.42 |
| MBa312 | BAAT (PYD4) | 에스. 클루이베리 | 142 | 105.85 |

[0388]

[0389] 플라스미드 pMBa120-4 (상기)는 다음과 같은 발현 카세트를 플랭크하는 NotI 제한 부위를 함유한다: PDC 프로 모터, 원하는 유전자 (BAAT 또는 3-HPDH), PDC 종결자, 및 URA3 선택 마커. *adh1202* 위치에서 통합에 대한 상 동성은 NotI 제한 부위의 밖에 있다. 이들 플라스미드 모두는 정방향으로 발현 카세트를 갖는다.

[0390] 동형 접합체 통합 균주의 스크리닝의 용이함을 허용하기 위해 새로운 플라스미드를 역방향을 향하는 발현 카세 트로 구성하였다. 플라스미드 pMBa120-4를 NotI로 분해하였고 89 mM 염기성 트리스-89 mM 붕산-2 mM 이나트 른 EDTA (TBE) 버퍼를 사용하여 1% 아가로스 겔 상에서 실행하였다. 3.5 kbp (pMBa120), 4.2 kbp (pMBa121), 4.2 kbp (pMBa122), 3.5 kbp (pMBa123), 및 4.2 kbp (pMBa124)의 단편을 겔에서 잘라냈고 QIAQUICK® Gel Extraction Kit (Qiagen)를 사용하여 아가로스로부터 추출하였다. 이들 각각의 단편을 여기에 설명된 바와 같이 T4 DNA 리가제를 사용하여 5.2 kbp 선형 NotI/ura 없는 CIP 처리된 pHJJ76에 결합하였다. 결합 생성물을 제조사 의 지시에 따라 One Shot® TOP10 화학적 컴피턴트 이. 콜리 세포 (Invitrogen)에 형질전환하였다. 형질전환체 를 2X YT+ amp 플레이트에 도말하였고 37 °C에서 하룻밤 동안 배양하였다. 결과의 형질전환체 중 다수를 XbaI 및 KpnI 분해에 의해 스크리닝하였다. 원하는 밴드 크기를 수득하는 클론을 UGA1, YMR226c, PYD4, ydfG, 및 BAAT에 대하여 각각 pMBa131, pMBa132, pMBa133, pMBa134, 및 pMBa135로 지정하였다. 결과의 플라스미드 는 *adh1202* 위치에서 원하는 유전자의 통합을 허용하며, 발현 카세트는 역방향을 향한다.

[0391] MBa308-12의 Ura-유도체를 여기에 설명된 바와 같이 분리하였다. MBa308-12에 대한 다수의 FOA-저항성 콜로니 는 YPD 플레이트에서 37 °C에서 성장시킴으로써 두 번 정제된 콜로니였다. 게놈 DNA를 FOA-저항성 콜로니로부터 제조하였고 여기에 설명된 바와 같이 URA3 선택 가능한 마커의 손실을 확인하기 위해 PCR로 체크하였다. 프라이 머 0611631 및 0611632는 각각 정방향 및 역방향을 향하는 통합의 부위의 외부에서 어닐링한다. 프라이머 0611718은 *ura* 선택 가능한 마커의 PDC 종결자 업스트림에서 어닐링한다; 프라이머 0611632 및 0611631은 야생 형 염색체에 대한 2.7 kbp 단편을 증폭시키고; 프라이머 0611718 및 0611632는 존재하는 *ura* 마커 및 *ura* 마커 의 부재시 1100 bp 단편 통합에 대한 2.4 kbp 단편을 증폭시킨다. 0611718 및 0611632를 갖는 1100 bp, 및 0611631 및 0611632를 갖는 2.7 kbp의 PCR 단편을 수득하는 MBa308-12의 하나의 *ura*-균주를 저장하였고 MBa314 (MBa310의 *ura*-균주), MBa315 (MBa312의 *ura*-균주), MBa316 (MBa311의 *ura*-균주), MBa326 (MBa308의 *ura*-균주), 및 MBa328 (MBa309의 *ura*-균주)로 지정하였다.

[0392] pMBa131, pMBa132, pMBa133, 및 pMBa135의 각각 10-15 µg을 ApaI, KpnI, 및 NcoI으로 분해하였고 89 mM

염기성 트리스-89 mM 붕산-2 mM 이나트륨 EDTA (TBE) 버퍼를 사용하여 1% 아가로스 겔 상에서 실행하였다. NcoI 으로 분해는 아가로스 겔로부터 원하는 단편의 추출을 가능하게 한다. pMIBa131-3, pMIBa135 각각 5553 bp, 4937 bp, 5565 bp, 5493 bp의 단편을 겔에서 잘라냈고 QIAQUICK® Gel Extraction Kit (Qiagen)를 사용하여 아가로스로부터 추출하였다. 정제된 생성물의 농도는 67-80 ng/μL인 것으로 발견되었다. pMIBa131-3, 및 pMIBa135의 제한된 단편 중 0.67-0.8 μg을 여기에 설명된 바와 같이 MIBa314, MIBa316, MIBa315, 또는 MIBa328에 형질전환하였다. 형질전환체를 ura 선택 배지에 도말하였고 37 °C에서 하룻밤 동안 배양하였고, ura 선택 배지에 재-스트리킹하였고 37 °C에서 하룻밤 동안 배양하였다. 게놈 DNA를 URA3+ 콜로니로부터 제조하였고 여기에 설명된 바와 같이 두 번째 발현 카세트의 통합을 확인하기 위해 PCR로 체크하였으며, 원하는 유전자에 대한 균주 동형 접합체를 만든다. 프라이머 0611718 및 0611632는 상기 설명된 바와 같이 첫 번째 통합을 위한 1100 bp 단편을 증폭시키고, 프라이머 0611632 및 0611717은 역방향의 두 번째 통합을 위한 814 bp 단편을 증폭시킨다. 0611718 및 0611632를 갖는 1100 bp 단편 및 0611717 및 0611632를 갖는 814 bp 단편을 증폭시키는, 각각의 계통의 URA3+ 형질전환체를 MIBa317, MIBa318, MIBa319, 및 MIBa329로 지정하였다 (표 13 참조).

표 13: 형질전환체 유전자형

| 균주 | 유전자형 |
|---------|--|
| MIBa317 | adh1202Δ:: (PDC _{프로모터} -Opt.ScYMR226c, URA3-스카)/ adh1202Δ:: (PDC _{프로모터} -Opt.ScYMR226c, URA3) ura3-/ura3- |
| MIBa318 | adh1202Δ:: (PDC _{프로모터} -Opt.ScUGA1, URA3-스카)/ adh1202Δ:: (PDC _{프로모터} -Opt.ScUGA1, URA3) ura3-/ura3- |
| MIBa319 | adh1202Δ:: (PDC _{프로모터} -Opt.SkPYD4, URA3-스카)/ adh1202Δ:: (PDC _{프로모터} -Opt.SkPYD4, URA3) ura3-/ura3- |
| MIBa329 | adh1202Δ:: (PDC _{프로모터} -Opt.SaBAAT, URA3-스카)/ adh1202Δ:: (PDC _{프로모터} -Opt.SaBAAT, URA3) ura3-/ura3- |

YMR226c, UGA1, PYD4, 및 BAAT에 대한 균주 동형 접합체 또는 이형 접합체의 발현 및 효소 활성을 결정하였다. MIBa309-12, MIBa317-9 및 MIBa329의 하룻밤 동안의 배양물을 YPD에서 37 °C에서 키웠고, 37 °C에서 125 mL 배플 플라스크 중의 신선한 YPD 25 mL에 1:50으로 희석하였고, OD₆₀₀ ~4-10으로 키웠다. 세포의 샘플을 여기에 설명된 방법을 사용하여 SDS-PAGE로 단백질 발현에 대하여 분석하였다. CFE를 또한 배양물의 세포 펠렛으로부터 제조하였고, 3P 데히드로게나제 활성을 CFE에서 여기에 설명된 방법을 사용하여 측정하였다. SDS-PAGE 결과에 기초하여, 균주 MIBa310, MIBa318, MIBa312 및 MIBa319는 ~53 KDa의 질량을 갖는 단백질 (UGA1 또는 PYD4 유전자에 의해 암호화된 단백질의 예상된 크기)을 함유하였다. 이 단백질에 해당하는 밴드는 균주 MBin500의 SDS-PAGE 분석에서 관찰되지 않았다. 게다가, 동형 접합체 균주 MIBa318 및 MIBa319 각각 UGA1 및 PYD4의 발현은 해당하는 이형 접합체 균주 MIBa310 또는 MIBa312보다 더 컸다 (SDS-PAGE 분석에 의해 판단된 바와 같음). BAAT 발현은 이들 조건 하에 균주 MIBa309 또는 MIBa329에서 (SDS-PAGE에 의해) 검출되지 않았다. 표 14A는 균주 MBin500, MIBa311 및 MIBa317의 CFE에서 3-HP 데히드로게나제 ("3-HPDH") 활성을 나타낸다.

표 14A: 형질전환체 효소 활성 데이터

| 균주 | 과발현된 유전자 | 유전자 SEQ ID NO | 근원 | 대립유전자 타입 | 3-HPDH 활성 |
|------------------|---------------------|---------------------|--------------|-------------|--------------|
| MBin500 (대조군) | N/A | N/A | N/A | N/A | 0.13 |
| MIBa311 | 3-HPDH (YMR226c) | 144 | 에스. 세레비시아 | 이형 접합체 | 1.49 |
| MIBa317 | 3-HPDH (YMR226c) | 144 | 에스. 세레비시아 | 동형 접합체 | 2.85 |

개선된 검정 조건을 사용하는 독립적인 실험에서, 균주 MBin500 (대조군), MIBa319 및 MIBa329로부터 제조된 CFE에서 BAAT 활성을 비교하였다. 이 실험의 결과는 표 14B에 나타난다.

[0399] 표 14B: 형질전환체 효소 활성 데이터

| 균주 | 과발현된 유전자 | 유전자 SEQ ID NO | 근원 | BAAT 활성 |
|------------------|----------------|---------------------|----------------|---------|
| MBin500 (대조군) | N/A | N/A | N/A | 0.67 |
| MBa319 | BAAT (PYD4) | 142 | 에스. 클루이베리 | 228.01 |
| MBa329 | BAAT | 140 | 에스. 아베르미틸리스 | 0.38 |

[0400]

[0401]

균주 MBa317, MBa318, 및 MBa319의 Ura-유도체를 여기에 설명된 바와 같이 분리하였다. MBa317, MBa18, 및 MBa19에 대한 다수의 FOA-저항성 콜로니는 MBa317, MBa18, 및 MBa19 YPD 플레이트에서 37 °C에서 키움으로써 정제된 콜로니였다. 게놈 DNA를 FOA-저항성 콜로니로부터 제조하였고 URA3 선택 가능한 마커의 손실을 확인하기 위해 여기에 설명된 바와 같이 PCR로 체크하였다. 프라이머 0611718 및 0611632는 상기 설명된 바와 같이 첫 번째 통합을 나타내는 1100 bp 단편을 증폭시키고, 프라이머 0611632 및 0611717은 역방향으로 두 번째 통합의 존재를 나타내는 814 bp 단편을 증폭시킨다. 프라이머 0611718 및 0611631은 ura 마커 및 ura 마커 없이 1200 bp 단편을 갖는 두 번째 통합을 나타내는 2.6 kbp 단편을 증폭시켰다. 0611718 및 0611632를 갖는 1100 bp, 0611632 및 0611717을 갖는 814 bp, 또는 0611718 및 0611631을 갖는 1200 bp의 PCR 단편을 수득하는 MBa317 및 MBa318의 ura 균주를 저장하였고 각각 MBa320 및 MBa321로 지정하였다. ura 마커가 MBa319로부터 제거될 때 가능한 유전자 전환 이벤트는 발현 카세트 둘 다가 정방향을 향하도록 PCR에 의해 지시된 바와 같이 (0611632 및 0611631 프라이머를 갖는 2.7 kbp 단편 또는 0611632 및 0611717을 갖는 814 bp 단편은 아니지만, 0611718 및 0611632를 갖는 1100bp 단편은 증폭됨) MBa322의 결과로 발생하였다.

[0402]

실시에 3A-3: pdc 위치에서 아스파르테이트 1-데카르복실라제 (ADC), β -알라닌 아미노트랜스퍼라제 (BAAT), 및 3-히드록시프로피온산 데히드로게나제 (3-HPDH)를 발현하는 삽입 벡터의 왼쪽 단편의 구조를

[0403]

에스. 아베르미틸리스 ADC (SEQ ID NO: 130) 및 에스. 아베르미틸리스 BAAT (SEQ ID NO: 140)를 함유하는 왼쪽 단편

[0404]

ENO1 프로모터 및 PDC 종결 영역 사이에서 발현하는 유전자의 삽입을 허용하기 위해, pMhCt068 벡터를 XbaI 및 PacI로 분해하였고, 37 °C에서 30분 동안 10 유닛 소 위장 포스파타제 (New England Biolabs)를 처리하였고, TAE 버퍼 중의 0.9% 아가로스 겔 전기영동으로 정제하였고, 약 6.1 kbp 밴드를 겔에서 잘라냈고 제조사의 지시에 따라 NUCLEOSPIN® Extract II Kit (Macherey-Nagel)를 사용하여 정제하였다.

[0405]

사카로미세스 클루이베리 BAAT (pyd4) 유전자 (SEQ ID NO: 142)를 이후의 서브클로닝을 위한 제한 부위를 함유하는 프라이머 0611196 및 0611186으로 증폭하였다. PCR 반응물 (50 μ L)은 에스. 클루이베리 pyd4 유전자, 1X Pfx 증폭 버퍼 (Invitrogen), 100 pmol 프라이머 0611196 및 0611186 각각, 200 μ M dATP, dCTP, dGTP, 및 dTTP 각각, 1.5 μ L DMSO 및 2.5 유닛의 Platinum® Pfx DNA Polymerase (Invitrogen)를 함유하는 플라스미드의 미니-프렙의 1 내지 50배 희석액 중 1 μ L를 함유한다. PCR을 95 °C에서 2분 동안 1 사이클 후 이어서 각각 95 °C에서 30초 동안, 40.8 °C에서 30초 동안, 및 72 °C에서 1분 30초 동안 34 사이클, 72 °C에서 5분 동안 최종 연장으로 프로그램된 EPPENDORF® MASTERCYCLER® (Eppendorf Scientific)에서 수행하였다. 열 순환 후, 약 1428 bp PCR 생성물을 겔에서 잘라냈고 제조사의 지시에 따라 NUCLEOSPIN® Extract II Kit (Macherey-Nagel)를 사용하여 정제하는 경우, PCR 반응 생성물을 TAE 버퍼 중의 0.9% 아가로스 겔 전기영동으로 분리하였다.

[0406]

상기 생성된 pyd4 PCR 반응물을 XbaI 및 PacI로 분해하였고 TAE 버퍼 중의 0.9% 아가로스 겔 전기영동으로 정제하였고, 약 1.4 kbp 밴드를 겔에서 잘라냈고 제조사의 지시에 따라 NUCLEOSPIN® Extract II Kit (Macherey-Nagel)를 사용하여 정제하였다. 이 정제된 DNA를 1X QUICK 결찰 버퍼 (New England Biolabs), 100 ng XbaI/PacI pMhCt068 벡터, 70.5 ng Xba I/PacI pyd4 인서트, 및 1 μ L QUICK T4 DNA 리가제 (New England Biolabs)를 함유하는 결찰 반응물 (20 μ L)에서 상기 설명된 XbaI 및 PacI 제한된 pMhCt068 벡터에 클로닝하였다. 결찰 반응물을 상온에서 5분 동안 배양하였고, 얼음 위에서 냉각하였다. 제조사의 지시에 따라 SoloPack Gold SuperCompetent 세포 (Agilent Technologies)를 형질전환하기 위해 이 반응물 5 μ L를 사용하였다. 형질전환체를 2X YT+amp 플레이트에 도말하였고 37 °C에서 하룻밤 동안 배양하였다. 결과의 형질전환체 중 다수를 여기에 설명된 바와 같이 XbaI 및 PacI로 분해에 의한 원하는 PCR 생성물의 적절한 인서트에 대하여 스크리닝하였으며, 한 확인된 분리체는 "왼쪽+pyd4#1"로 지정되었다.

- [0407] SEQ ID NO: 17의 에스. 아베르미틸리스 ADC를 암호화하고 이. 콜리에서 발현에 코돈-최적화된 폴리뉴클레오티드를 프라이머 0611376 및 0611377로 증폭시켰다 (프라이머 0611376이 In-Fusion을 통해 pMhCt068에 삽입 후 NheI 제한 부위의 5' 말단에서 "T" 염기의 제거를 일으키며, 이것은 초기 pMhCt068 클론에 존재하는 원하지 않는 ATG 시작 코돈을 제거한다는 것을 나타냄). PCR 반응물 (50 μ L)은 1 내지 50배 희석된 pACN5의 미니-프랩 중 1 μ L, 1X ThermoPol 반응 버퍼 (New England Biolabs), 100 pmol 각각의 프라이머 0611376 및 0611377, 200 μ M 각각의 dATP, dCTP, dGTP, 및 dTTP, 2 μ L 100 mM $MgSO_4$, 및 2 유닛 VentR® (엑소-) DNA Polymerase (New England Biolabs)를 함유하였다. PCR을 94 $^{\circ}C$ 에서 2분 동안 1 사이클 후 이어서 각각 94 $^{\circ}C$ 에서 30초 동안, 54 $^{\circ}C$ 에서 30초 동안, 72 $^{\circ}C$ 에서 1분 동안 34 사이클, 72 $^{\circ}C$ 에서 10분 동안 최종 연장으로 프로그램된 EPPENDORF® MASTERCYCLER® (Eppendorf Scientific)로 수행하였다. 열 순환 후, 약 420 bp PCR 생성물을 겔에서 잘라냈고 제조사의 지시에 따라 NUCLEOSPIN® Extract II Kit (Macherey-Nagel)를 사용하여 정제한 경우, PCR 반응 생성물을 TAE 버퍼 중의 0.9% 아가로스 겔 전기영동으로 분리하였다.
- [0408] "왼쪽+pyd4#1" 플라스미드를 NheI 및 AscI로 분해하였고, 37 $^{\circ}C$ 에서 30분 동안 10 유닛 소 위장 포스파타제 (New England Biolabs)를 처리하였고, TAE 버퍼 중의 0.9% 아가로스 겔 전기영동으로 정제하였고, 약 7.5 kbp 밴드를 겔에서 잘라냈고 제조사의 지시에 따라 NUCLEOSPIN® Extract II Kit (Macherey-Nagel)를 사용하여 정제하였다. 이. 콜리에 최적화된 에스. 아베르미틸리스 ADC 유전자를 함유하는, 상기 정제된 PCR 단편을 NheI 및 AscI로 분해하였고 TAE 버퍼 중의 0.9% 아가로스 겔 전기영동으로 정제하였고, 약 420 bp 밴드를 겔에서 잘라냈고 제조사의 지시에 따라 NUCLEOSPIN® Extract II Kit (Macherey-Nagel)를 사용하여 정제하였다. 결과의 단편을 1X QUICK 결찰 버퍼 (New England Biolabs), 100 ng NheI/AscI "왼쪽+pyd4#1" 벡터, 31 ng의 NheI 및 AscI 분해된 panD 인서트, 및 1 μ L QUICK T4 DNA 리가제 (New England Biolabs)를 함유하는 결찰 반응물 (20 μ L)에서 선행화 "왼쪽+pyd4#1" 벡터에 결찰하였다. 결찰 반응물을 상온에서 5분 동안 배양하였고, 얼음 위에 두었다. 제조사의 지시에 따라 SoloPack Gold SuperCompetent 세포 (Agilent Technologies)를 형질전환하기 위해 이 반응물 5 μ L을 사용하였다. 형질전환체를 2X YT+amp 플레이트에 도말하였고 37 $^{\circ}C$ 에서 하룻밤 동안 배양하였다. 결과의 형질전환체 중 다수를 스크리닝하였다. 여기에 설명된 바와 같이 AscI 및 PvuII로 분해에 의해 이. 콜리에 최적화된 에스. 아베르미틸리스 ADC 유전자의 적절한 삽입을 위해 스크리닝하였다. 원하는 밴드 크기를 수득하는 클론을 DNA 시퀀싱에 의해 올바른 것으로 확인하였고 pMhCt070으로 지정하였다.
- [0409] 플라스미드 pMhCt070은 효모 숙주에서 발현에 코돈-최적화된 ADC 및 BAAT 상동체의 삽입을 위한 염기성 벡터의 역할을 한다. pMhCt070을 XbaI 및 PacI로 분해하였고, 37 $^{\circ}C$ 에서 30분 동안 10 유닛 소 위장 포스파타제 (New England Biolabs)를 처리하였고, TAE 버퍼 중의 0.9% 아가로스 겔 전기영동으로 정제하였고, 약 6.5 kbp 밴드를 겔에서 잘라냈고 제조사의 지시에 따라 NUCLEOSPIN® Extract II Kit (Macherey-Nagel)를 사용하여 정제하였다.
- [0410] 에스. 아베르미틸리스 BAAT (SEQ ID NO: 140)을 암호화하고 아이. 오리엔탈리스에서 발현에 최적화된 폴리뉴클레오티드를 함유하는, 상기 설명된 XbaI 및 PacI 분해된 단편을 다음과 같이 pMhCt070 컷 벡터에 결찰하였다: 1X QUICK 결찰 버퍼 (New England Biolabs), 42 ng의 XbaI 및 PacI 분해된 pMhCt070 벡터, 4 μ L의 코돈-최적화된 에스. 아베르미틸리스 BAAT XbaI 및 PacI 분해된 인서트, 및 1 μ L QUICK T4 DNA 리가제 (New England Biolabs)를 함유하는 결찰 반응물 (20 μ L). 결찰 반응물을 상온에서 5분 동안 배양하였고, 얼음 위에 두었다. 이 반응물 중 5 μ L를 제조사의 지시에 따라 SoloPack Gold SuperCompetent 세포 (Agilent Technologies)를 형질전환하기 위해 사용하였다. 형질전환체를 2X YT+amp 플레이트에 도말하였고 37 $^{\circ}C$ 에서 하룻밤 동안 배양하였다. 결과의 형질전환체 중 다수를 여기에 설명된 바와 같이 XbaI, PacI, 및 EcoRV 분해에 의해 원하는 BAAT ORF에 대하여 스크리닝하였다. 원하는 밴드 크기를 수득하는 클론을 DNA 시퀀싱에 의해 올바른 것으로 확인하였고 pMhCt072로 지정하였다.
- [0411] pMhCt072의 이. 콜리에서 발현에 코돈-최적화된 에스. 아베르미틸리스 ADC 유전자를 아이. 오리엔탈리스 (SEQ ID NO: 130)에서 발현에 코돈-최적화된 버전으로 대체하였다. 아이. 오리엔탈리스 (SEQ ID NO: 130) 코돈-최적화된 ADC 유전자 및 원하는 추가의 제한 부위 및 플랭킹 DNA를 프라이머 0611378 및 0611379로 증폭하였다. PCR 반응물 (50 μ L)은 코돈-최적화된 에스. 아베르미틸리스 panD (GeneArt®)를 함유하는 플라스미드의 1 내지 50배 희석된 미니-프랩 중 1 μ L, 1X ThermoPol 반응 버퍼 (New England Biolabs), 100 pmol 각각의 프라이머 0611378 및 0611379, 200 μ M 각각의 dATP, dCTP, dGTP, 및 dTTP, 2 μ L 100 mM $MgSO_4$, 및 2 유닛 VentR® (엑소-) DNA Polymerase (New England Biolabs)를 함유하였다. PCR을 94 $^{\circ}C$ 에서 2분 동안 1 사이클 후 이어서 각각 94 $^{\circ}C$ 에서 30초 동안, 54 $^{\circ}C$ 에서 30초 동안, 72 $^{\circ}C$ 에서 45초 동안 34 사이클, 72 $^{\circ}C$ 에서 10분 동안 최종 연

장으로 프로그램된 EPPENDORF® MASTERCYCLER® (Eppendorf Scientific)로 수행하였다. 열 순환 후, 약 420개의 염기쌍 PCR 생성물을 겔에서 잘라냈고 제조사의 지시에 따라 NUCLEOSPIN® Extract II Kit (Macherey-Nagel)를 사용하여 정제한 경우, PCR 반응 생성물을 TAE 버퍼 중의 1% 아가로스 겔 전기영동으로 분리하였다.

- [0412] pMhCt072의 미니-프랩 중 5 μ L를 XbaI 및 PacI로 분해하였고, 37 °C에서 30분 동안 10 유닛 소 위장 포스파타제 (New England Biolabs)를 처리하였고, TAE 버퍼 중의 1% 아가로스 겔 전기영동으로 정제하였고, 약 7.5 kbp 밴드를 겔에서 잘라냈고 제조사의 지시에 따라 NUCLEOSPIN® Extract II Kit (Macherey-Nagel)를 사용하여 정제하였다. 상기로부터 코돈-최적화된 에스. 아베르미틸리스 ADC 유전자를 함유하는 분리된 PCR 생성물을 다음 IN-FUSION™ Advantage PCR Cloning Kit (Clontech) 반응에서 벡터에 추가하였다: 10 μ L 반응 부피는 6 μ L의 pMhCt072 분해된 및 정제된 벡터, 1 μ L의 정제된 코돈-최적화된 panD PCR 생성물, 1X In-Fusion 반응 버퍼 (Clontech) 및 1 μ L의 IN-FUSION™ 효소 (Clontech)로 구성되었다. 반응물을 37 °C에서 15분 동안, 50 °C에서 15분 동안 배양하였고, 얼음 위에 두었다. 반응물을 40 μ L의 TE 버퍼로 희석하였고 제조사의 지시에 따라 SoloPack Gold SuperCompetent 세포 (Agilent Technologies)를 형질전환하기 위해 2.5 μ L를 사용하였다. 형질 전환체를 2X YT+amp 플레이트에 도말하였고 37 °C에서 하룻밤 동안 배양하였다. 결과의 형질전환체 중 다수를 NheI, AscI, 및 ClaI 분해에 의해 원하는 PCR 생성물의 적절한 삽입에 대하여 스크리닝하였다. 원하는 밴드 크기를 수득하는 클론을 DNA 시퀀싱에 의해 올바른 것으로 확인하였고 pMhCt074 (도 7)로 지정하였다.
- [0413] pMhCt074은 코돈-최적화된 에스. 아베르미틸리스 ADC의 발현을 구동하는 PDC 프로모터 (panD, SEQ ID NO: 130), TAL 종결자, 코돈-최적화된 에스. 아베르미틸리스 BAAT (SEQ ID NO: 140)의 발현을 구동하는 ENO1 프로모터, RKI 종결자, 아이. 오리엔탈리스 URA3 프로모터 및 아이. 오리엔탈리스 URA3 ORF의 5' 단편을 함유하는 왼쪽 PDC 표적 구조물이다.
- [0414] 에스. 아베르미틸리스 ADC (SEQ ID NO: 130) 및 에스. 클루이베리 BAAT (SEQ ID NO: 142)를 함유하는 왼쪽 단편
- [0415] 에스. 클루이베리 BAAT (pyd4)를 발현하는 왼쪽 DNA 구조물을 생성하기 위해, 아이. 오리엔탈리스 (SEQ ID NO: 142, 상기)에서 발현되는, 코돈-최적화된 에스. 클루이베리 BAAT 서열을 함유하는, XbaI 및 PacI 분해된 단편을 다음과 같이 상기 pMhCt070 분해된 벡터에 결합하였다: 1X QUICK 결합 버퍼 (New England Biolabs), 42 ng의 XbaI 및 PacI 분해된 pMhCt070 벡터, 4 μ L의 XbaI 및 PacI로 분해된, 코돈-최적화된 에스. 클루이베리 pyd4 인서트, 및 1 μ L QUICK T4 DNA 리가제 (New England Biolabs)를 함유하는 결합 반응물 (20 μ L). 결합 반응물을 상온에서 5분 동안 배양하였고, 얼음 위에 두었다. 제조사의 지시에 따라 SoloPack Gold SuperCompetent 세포 (Agilent Technologies)를 형질전환하기 위해 이 반응물 중 5 μ L를 사용하였다. 형질전환체를 2X YT+amp 플레이트에 도말하였고 37 °C에서 하룻밤 동안 배양하였다. 결과의 형질전환체 중 다수를 XbaI, PacI, 및 EcoRV 분해에 의해 원하는 pyd4 ORF의 적절한 삽입에 대하여 스크리닝하였다. 원하는 밴드 크기를 수득하는 클론을 DNA 시퀀싱에 의해 올바른 것으로 확인하였고 pMhCt073으로 지정하였다.
- [0416] 플라스미드 pMhCt073은 아이. 오리엔탈리스에서 발현에 코돈-최적화된, 원하는 에스. 클루이베리 BAAT (pyd4) 서열을 함유하지만 아이. 오리엔탈리스에서 발현에 코돈-최적화된 원하는 에스. 아베르미틸리스 ADC (panD) 서열을 함유하지 않는다. 이 ORF에서 이동하기 위해, pMhCt073의 미니-프랩 중 5 μ L를 XbaI 및 PacI로 분해하였고, 37 °C에서 30분 동안 10 유닛 소 위장 포스파타제 (New England Biolabs)를 처리하였고, TAE 버퍼 중의 1% 아가로스 겔 전기영동으로 정제하였다. 약 7.5 kbp 밴드를 겔에서 잘라냈고 제조사의 지시에 따라 NUCLEOSPIN® Extract II Kit (Macherey-Nagel)를 사용하여 정제하였다. 코돈-최적화된 에스. 아베르미틸리스 panD (상기)를 함유하는 분리된 PCR 생성물을 다음 IN-FUSION™ Advantage PCR Cloning Kit (Clontech) 반응물에서 벡터에 추가하였다: 10 μ L 반응 부피는 6 μ L의 pMhCt073 분해된 및 정제된 벡터, 1 μ L의 정제된 코돈-최적화된 panD PCR 생성물, 1X In-Fusion 반응 버퍼 (Clontech) 및 1 μ L의 IN-FUSION™ 효소 (Clontech)로 구성되었다. 반응물을 37 °C에서 15분 동안, 50 °C에서 15분 동안 배양하였고, 얼음 위에 두었다. 반응물을 40 μ L의 TE 버퍼로 희석하였고 제조사의 지시에 따라 SoloPack Gold SuperCompetent 세포 (Agilent Technologies)를 형질전환하기 위해 2.5 μ L를 사용하였다. 형질전환체를 2X YT+amp 플레이트에 도말하였고 37 °C에서 하룻밤 동안 배양하였다. 결과의 형질전환체 중 다수를 NheI, AscI, 및 ClaI 분해에 의해 원하는 PCR 생성물의 적절한 삽입을 위해 스크리닝하였다.
- [0417] 플라스미드 pMhCt076은 아이. 오리엔탈리스에서 발현에 코돈-최적화된 에스. 아베르미틸리스 ADC의 발현을 구동하는 PDC 프로모터 (panD, SEQ ID NO: 130), TAL 종결자, 아이. 오리엔탈리스에서 발현에 코돈-최적화된 에스. 클루이베리 BAAT의 발현을 구동하는 ENO1 프로모터 (pyd4, SEQ ID NO: 142), RKI 종결자, 아이. 오리엔탈리스

URA3 프로모터 및 아이. 오리엔탈리스 URA3 ORF의 5' 단편을 함유하는 왼쪽 PDC 표적 구조물이다.

- [0418] 플라스미드 pMhCt076는 PDC 프로모터 내로 약 200 bp에서 A에서 T로 뉴클레오티드 변화, pMhCt068 모체 벡터 (상기)에 존재하는 PDC 프로모터를 통해 방법의 2/3에서 G에서 T로 뉴클레오티드 변화를 함유하는 것을 결정하였다. 유전자 발현의 잠재적 변화에 대한 어떤 우려도 해결하기 위해, pMhCt076에 유사하지만 올바른 PDC 프로모터를 함유하는 구조물을 하기 설명된 바와 같이 클로닝하였다.
- [0419] pMhCt082의 미니-프랩 중 5 μ L를 NheI 및 PacI로 분해하였고, 37 $^{\circ}$ C에서 30분 동안 10 유닛 소 위장 포스파타제 (New England Biolabs)를 처리하였고, TAE 버퍼 중의 0.9% 아가로스 겔 전기영동으로 정제하였고, 약 4.7 kbp 밴드를 겔에서 잘라냈고 제조사의 지시에 따라 NUCLEOSPIN® Extract II Kit (Macherey-Nagel)를 사용하여 정제하였다. pMhCt076의 미니-프랩 중 4 μ L를 pMhCt076 NheI 및 PacI로 분해하였고 TAE 버퍼 중의 0.9% 아가로스 겔 전기영동으로 정제하였고, 약 3.3 kbp 밴드를 겔에서 잘라냈고 제조사의 지시에 따라 NUCLEOSPIN® Extract II Kit (Macherey-Nagel)를 사용하여 정제하였다. 정제된 4.7 kbp 벡터 및 3.3 kbp 인서를 1X QUICK 결찰 버퍼 (New England Biolabs), 3 μ L pMhCt082 벡터, 6 μ L pMhCt076 인서트, 및 1 μ L QUICK T4 DNA 리가제 (New England Biolabs)를 함유하는 결찰 반응물 (20 μ L)에서 함께 결찰하였다. 결찰 반응물을 상온에서 5분 동안 배양하였고, 얼음 위에 두었다.
- [0420] 이 반응물 중 5 μ L을 제조사의 지시에 따라 ONE SHOT® TOP10 화학적 컴피턴트 이. 콜리 세포 (Invitrogen)를 형질전환하기 위해 사용하였다. 형질전환체를 2X YT+amp 플레이트에 도말하였고 상온에서 3일 동안 배양하였다. 결과의 형질전환체 중 다수를 StuI 및 NotI로 분해에 의해 원하는 PCR 생성물에 대해 스크리닝하였다. 원하는 밴드 크기를 수득하는 클론을 DNA 시퀀싱에 의해 올바른 것으로 확인하였고 pMhCt083 (도 8)으로 지정하였다.
- [0421] 플라스미드 pMhCt083와 pMhCt076는 전자가 올바른 PDC 프로모터 서열을 함유하는 한편, 후자는 상기 설명된 하나의 A에서 T로 뉴클레오티드 변화 및 하나의 G에서 T로 뉴클레오티드 변화를 갖는 것을 제외하면 동일하다. 테스트는 pMhCt083 및 pMhCt077와 비교하여 pMhCt076 및 pMhCt077의 통합으로부터 에스. 아베르미틸리스 panD를 발현하는 균주의 panD 효소 활성에서 차이를 나타내지 않았다.
- [0422] 에스. 아베르미틸리스 ADC (SEQ ID NO: 130) 및 사카로미세스 세레비시애 사카로미세스 세레비시애 gabT (SEQ ID NO: 141)를 함유하는 왼쪽 단편
- [0423] pMhCt083의 미니-프랩 중 4 μ L을 XbaI 및 PacI로 분해하였고, 37 $^{\circ}$ C에서 30분 동안 10 유닛 소 위장 포스파타제 (New England Biolabs)를 처리하였고, TAE 버퍼 중의 0.9% 아가로스 겔 전기영동으로 정제하였고, 약 6.5 kbp 밴드를 겔에서 잘라냈고 제조사의 지시에 따라 NUCLEOSPIN® Extract II Kit (Macherey-Nagel)를 사용하여 정제하였다. 아이. 오리엔탈리스에서 발현에 코돈-최적화된 사카로미세스 세레비시애 gabT를 함유하는, XbaI 및 PacI로 분해된 단편 (UGA1, SEQ ID NO: 141)을 다음과 같이 pMhCt083 컷 벡터에 결찰하였다: 1X QUICK 결찰 버퍼 (New England Biolabs), XbaI 및 PacI로 분해된, 정제된 pMhCt083 벡터 1 μ L, 3 μ L 코돈-최적화된 에스. 세레비시애 UGA1 XbaI 및 PacI 분해된 인서트, 및 1 μ L QUICK T4 DNA 리가제 (New England Biolabs)를 함유하는 결찰 반응물 (20 μ L). 결찰 반응물을 상온에서 5분 동안 배양하였고, 튜브를 얼음 위에 두었다. 이 반응물 중 5 μ L을 제조사의 지시에 따라 SoloPack Gold SuperCompetent 세포 (Agilent Technologies)를 형질전환하기 위해 사용하였다. 형질전환체를 2X YT+amp 플레이트에 도말하였고 37 $^{\circ}$ C에서 하룻밤 동안 배양하였다. 결과의 형질전환체 중 다수를 XbaI 및 BglII 분해에 의해 원하는 BAAT ORF의 적절한 삽입에 대해 스크리닝하였다.
- [0424] 플라스미드 pMhCt087는 아이. 오리엔탈리스에서 발현에 코돈-최적화된 에스. 아베르미틸리스 ADC의 발현을 구동하는 PDC 프로모터 (panD, SEQ ID NO: 130), TAL 종결자, 아이. 오리엔탈리스에서 발현에 코돈 최적화된 에스. 세레비시애 gabT의 발현을 구동하는 ENO1 프로모터 (UGA1, SEQ ID NO: 141), RKI 종결자, 아이. 오리엔탈리스 URA3 프로모터 및 아이. 오리엔탈리스 URA3 ORF의 5' 단편을 함유하는 왼쪽 PDC 표적 구조물이다.
- [0425] 실시예 3A-4: pdc 위치에서 아스파르테이트 1-데카르복실라제 (ADC), β -알라닌 아미노트랜스퍼라제 (BAAT), 및 3-히드록시프로피온산 데히드로게나제 (3-HPDH)를 발현하는 삽입 벡터의 오른쪽 단편의 구조물
- [0426] 이. 콜리 3-HPDH (SEQ ID NO: 143)를 함유하는 오른쪽 단편
- [0427] pMhCt069 (상기)의 미니-프랩 중 2 μ L를 XbaI 및 PacI로 분해하였고, 37 $^{\circ}$ C에서 30분 동안 10 유닛 소 위장 포스파타제 (New England Biolabs)를 처리하였고, TAE 버퍼 중의 0.9% 아가로스 겔 전기영동으로 정제하였고, 약 2.2 kbp 밴드를 겔에서 잘라냈고 제조사의 지시에 따라 NUCLEOSPIN® Extract II Kit (Macherey-Nagel)를 사용하여 정제하였다. 아이. 오리엔탈리스에서 발현에 코돈-최적화된 이. 콜리 3-HPDH 유전자를 함유하는, XbaI 및 PacI로 분해된 단편 (ydfG, SEQ ID NO: 143)을 다음과 같이 pMhCt069 컷 벡터에 결찰하였다: 1X QUICK 결찰

버퍼 (New England Biolabs), XbaI 및 PacI로 분해된, 정제된 pMhCt069 벡터 2 μ L, XbaI 및 PacI로 분해된, 코돈-최적화된 이. 콜리 ydfG 인서트 4 μ L, 및 1 μ L QUICK T4 DNA 리가제 (New England Biolabs)를 함유하는 결찰 반응물 (20 μ L). 결찰 반응물을 상온에서 5분 동안 배양하였고, 얼음 위에 두었다. 이 반응물의 5 μ L를 제조사의 지시에 따라 SoloPack Gold SuperCompetent 세포 (Agilent Technologies)를 형질전환하기 위해 사용하였다. 형질전환체를 2X YT+amp 플레이트에 도말하였고 상온에서 3일 동안 배양하였다. 결과의 형질전환체 중 다수를 XbaI, PacI, 및 EcoRV 분해에 의해 원하는 ydfG ORF의 적절한 삽입에 대해 스크리닝하였다. 원하는 밴드 크기를 수득하는 클론을 여기에 설명된 바와 같이 DNA 시퀀싱에 의해 올바른 것으로 확인하였고 pMhCt075 (도 10)로 지정하였다.

[0428] 플라스미드 pMhCt075는 아이. 오리엔탈리스 URA3 ORF의 3' 단편, 아이. 오리엔탈리스의 URA3 종결자에 이어서 URA3 프로모터 (이후에 URA3 마커의 밖으로 루프-아웃), 아이. 오리엔탈리스 TDH3 프로모터에 의해 구동되는, 아이. 오리엔탈리스에서 발현에 코돈-최적화된 이. 콜리 3-HPDH 유전자 (ydfG, SEQ ID NO: 143), 및 아이. 오리엔탈리스 PDC 종결 영역을 함유하였다.

[0429] 사카로미세스 세레비시에 3-HPDH (SEQ ID NO: 144)를 함유하는 오른쪽 단편

[0430] 아이. 오리엔탈리스 (상기)에서 발현에 코돈-최적화된 에스. 세레비시에 YMR226C 유전자를 함유하는, XbaI 및 PacI로 분해된 단편을 다음과 같이 pMhCt069 컷 벡터에 결찰하였다: 1X QUICK 결찰 버퍼 (New England Biolabs), XbaI 및 PacI로 분해된, 정제된 pMhCt069 벡터 2 μ L, XbaI 및 PacI로 분해된, 코돈-최적화된 에스. 세레비시에 3-HPDH (YMR226C, SEQ ID NO: 144) 인서트 4 μ L 및, 1 μ L QUICK T4 DNA 리가제 (New England Biolabs)를 함유하는 결찰 반응물 (20 μ L). 결찰 반응물을 상온에서 5분 동안 배양하였고, 얼음 위에 두었다. 이 반응물의 5 μ L를 제조사의 지시에 따라 SoloPack Gold SuperCompetent 세포 (Agilent Technologies)를 형질전환하기 위해 사용하였다. 형질전환체를 2X YT+amp 플레이트에 도말하였고 상온에서 3일 동안 배양하였다. 결과의 형질전환체 중 다수를 XbaI, PacI, 및 EcoRV 분해에 의해 원하는 YMR226C ORF의 적절한 삽입에 대하여 스크리닝하였다. 원하는 밴드 크기를 수득하는 클론을 DNA 시퀀싱에 의해 올바른 것으로 확인하였고 pMhCt077 (도 11)로 지정하였다.

[0431] 플라스미드 pMhCt077은 아이. 오리엔탈리스 URA3 ORF의 3' 단편, 아이. 오리엔탈리스의 URA3 종결자에 이어서 URA3 프로모터 (이후에 URA3 마커의 밖으로 루프-아웃), 아이. 오리엔탈리스 TDH3 프로모터 아이. 오리엔탈리스에서 발현에 코돈-최적화된 에스. 세레비시에 3-HPDH 유전자 (YMR226C, SEQ ID NO: 144), 및 아이. 오리엔탈리스 PDC 종결 영역을 함유한다.

[0432] 실시예 3A-5: PDC 위치에서 아스파르트레이트 1-데카르복실라제 (ADC), β -알라닌 아미노트랜스퍼라제 (BAAT), 및 3-히드록시프로피온산 데히드로게나제 (3-HPDH)를 발현하는 이형 접합체 및 동형 접합체 효모 균주

[0433] 상기 실시예 3A-3 및 3A-4를 아이. 오리엔탈리스 PDC 위치에서 동시에 3개의 이소성 유전자의 발현을 표적화하는, 다양한 왼쪽 또는 오른쪽 구조물의 구조물을 설명한다. 형질전환 전에, 각 구조물 (하나의 원하는 왼쪽 구조물 및 하나의 원하는 오른쪽 구조물)의 약 10 μ L를 NotI로 pUC18 백본 벡터의 원하는 형질전환 DNA를 방출하기 위해 분해하였다. 대부분의 분해에 대해, 제한 효소 PvuI를 또한 NotI 분해와 함께 포함되었다. 제한 효소 PvuI는 pUC18 벡터 단편을 약 절반 정도 분해하며, 더 큰 조각으로 나누어지고, 원하는 DNA 단편은 겔 전기영동에 의해 더 쉬울 수 있다. 밴드를 함유하는 더 큰, 발현 카세트는 겔 전기영동에 의해 pUC18 백본 벡터의 원하는 형질전환 DNA의 단편을 pUC18 백본 벡터로부터 겔 전기영동으로 분리하였고, 겔에서 잘라냈고 제조사의 지시에 따라 NUCLEOSPIN® Extract II Kit (Macherey-Nagel)를 사용하여 정제하였다. 30 μ L 용출 버퍼를 용출 단계에 사용하였다. 총 10 μ L인 하나의 왼쪽 및 하나의 오른쪽 구조물의 등물의 비율을 아이. 오리엔탈리스 균주 CNB1 또는 적절한 유도체를 형질전환하기 위해 사용하였다. 형질전환체를 ura 선택 플레이트에서 선택하였고 성장을 위해 37 °C에 두었다. 다음날, 약 12개의 형질전환체를 골랐고 ura 선택 플레이트에 단일 콜로니에 대해 재스트리킹하였고 37 °C에서 키웠다. 다음날, 단일 콜로니를 각각의 초기 형질전환체에 의해 생성된 스트릭의 각각 골랐고 단일 콜로니에 대해 ura 선택 플레이트에 재스트리킹하였다. 37 °C에서 성장의 또 다른 밤 이후, 최종 단일 콜로니를 각각의 스트릭로부터 골랐고 ura 선택 플레이트에 재스트리킹하였고 37 °C에서 하룻밤 동안 키웠다. 단일 콜로니 정제 및 성장의 두 라운드 후에, 상기 설명된 바와 같이 발생된 원하는 표적화 통합을 확인하기 위해 PCR에 사용되는 게놈 DNA를 제조하였다. 왼쪽 및 오른쪽 구조물을 사용하여 PDC를 표적화 하기 위해, 프라이머 0611814, 0611554, 및 0611555를 사용하여 원하는 통합 이벤트의 확인을 결정하였다. 프라이머 0611554는 오른쪽 PDC 표적화 구조물에 존재하는 PDC 종결 영역의 3'의 아이. 오리엔탈리스 게놈 DNA에서 결합하고 중단을 향해 증폭하고; 프라이머 0611814는 오른쪽 구조물에 존재하는 TDH3 프로모터의 3' 말단 근처에 결

합한다. 프라이머 0611814 및 0611554를 함유하는, PCR의 1.9 kbp 밴드의 발생은 PDC 위치에서 원하는 통합 이벤트의 발생을 나타냈다. 프라이머 0611555 및 0611554를 함유하는, PCR의 1.4 kbp 밴드의 발생은 야생형 PDC 위치의 존재를 나타냈다. 이 통합 이벤트가 이배체 아이. 오리엔탈리스 CNB1의 첫 번째 표적화 이벤트이기 때문에 원하는 구성요소는 프라이머 0611814 및 0611554에 대한 1.9 kbp 밴드 및 프라이머 0611555 및 0611554의 1.4 kbp 밴드 둘 다를 나타낼 것이다. 각각의 플라스미드에 대하여 원하는 밴드 패턴을 제공하는 2개의 독립적인 형질전환체를 표 15에 나타난 바와 같이 지정하였다.

표 15: 형질전환체 유전자형

| 균주 | 플라스미드 w/ 왼쪽 단편 | 플라스미드 w/ 오른쪽 단편 | 유전자형 |
|----------|----------------------|-----------------------|---|
| yMhCt002 | pMhCt074 | pMhCt075 | pdcΔ:: (PDC 프로모터-Opt.SaPanD, ENO1 프로모터- Opt.SaBAAT, URA3, TDH3 프로모터- Opt.EcYdfG)/PDC ura3-/ura3- |
| 74/75 #1 | | | |
| yMhCt004 | pMhCt083 | pMhCt077 | pdcΔ:: (PDC 프로모터-Opt.SaPanD, ENO1 프로모터- Opt.SkPyd4, URA3, TDH3 프로모터- Opt.ScYMR226C)/PDC ura3-/ura3- |
| 83/77 #2 | | | |
| yMhCt005 | pMhCt087 | pMhCt077 | pdcΔ:: (PDC 프로모터-Opt.SaPanD, ENO1 프로모터- Opt.ScUGA1, URA3, TDH3 프로모터- Opt.ScYMR226C)/PDC ura3-/ura3- |
| 87/77 #2 | | | |

그 다음, yMhCt004 또는 yMhCt005의 ura-유도체를 상기 설명된 바와 같이 분리하였다. 각각의 모체 균주의 다수의 FOA 저항성 콜로니의 게놈 DNA를 프라이머 0611815 및 0611817로 원하는 루프-아웃 이벤트에 대하여 PCR로 스크리닝하였다. 프라이머 0611815는 왼쪽 구조물의 RKI 종결자에서 어닐링하고 URA3 프로모터로 증폭시킨다. 프라이머 0611817은 TDH3 프로모터에서 어닐링하고 URA3 카세트를 향해 증폭시킨다. 838 bp 밴드의 존재는 원하는 바와 같이 ura3 스카 (scar) 부위 (모체 균주에서 2개의 URA3 프로모터 사이에서 상동 재조합 후 단일 URA3 프로모터의 왼쪽 뒤)만의 존재를 나타내는 한편, 약 2.2 kbp의 밴드는 온전한 URA3 프로모터-URA3 ORF-URA3 종결자-URA3 프로모터 카세트의 존재를 나타내며, 원하는 재조합 이벤트가 발생하지 않았다는 것을 나타낸다. 상기 설명된 바와 같이 Crimson Taq™ DNA Polymerase로 함께 PCR 반응을 수행하였다. 모체 균주 yMhCt004의 하나의 FOA 저항성 콜로니를 yMhCt012로 지정하였고, 모체 균주 yMhCt005의 하나의 FOA 저항성 콜로니를 yMhCt007로 지정하였고, 원하는 828 bp 밴드를 제공하였다.

균주 yMhCt012를 pMhCt083 및 pMhCt077의 선형 DNA에 형질전환하는 한편, yMhCt007을 pMhCt087 및 pMhCt077의 선형 DNA에 형질전환하였다. 두 번의 라운드의 단일 콜로니 정제 후, 다수의 형질전환체의 게놈 DNA를 pMhCt087 및 pMhCt077의 선형 DNA에 형질전환하였다. 두 번의 라운드의 단일 콜로니 정제에서, 각각의 모체 균주의 다수의 형질전환체의 게놈 DNA를 PCR로 스크리닝하였다.

828 bp 밴드 (원래 표적화된 PDC 위치의 ura3 스카 영역의 증폭의) 및 2.2 kbp 밴드 (다른 염색체상에서 두 번째 통합 이벤트의 URA3 프로모터-URA3 ORF-URA3 종결자-URA3 프로모터 카세트의 통합의)를 둘 다 갖는 각각의 모체 균주의 두 번의 독립적으로 분리된 구성요소를 표 16에 나타난 바와 같이 지정하였다.

표 16: 형질전환체 유전자형

| 균주 설계 | 모체 균주 | 유전자형 |
|----------|----------|--|
| yMhCt013 | yMhCt012 | pdcΔ:: (PDC 프로모터-Opt.SaPanD, ENO1 프로모터-Opt.SkPyd4, URA3- 스카, TDH3 프로모터-Opt.ScYMR226C) /pdcΔ:: (PDC 프로모터- Opt.SaPanD, ENO1 프로모터-Opt.SkPyd4-URA3, TDH3 프로모터- Opt.ScYMR226C) ura3-/ura3- |
| yMhCt014 | | |
| yMhCt008 | yMhCt007 | pdcΔ:: (PDC 프로모터-Opt.SaPanD, ENO1 프로모터-Opt.ScUGA1, URA3-스카, TDH3 프로모터-Opt.ScYMR226C) /pdcΔ:: (PDC 프로모터- Opt.SaPanD, ENO1 프로모터-Opt.ScUGA1, URA3, TDH3 프로모터- Opt.ScYMR226C) ura3-/ura3- |
| yMhCt009 | | |

yMhCt008의 ura-유도체를 상기 설명된 바와 같이 분리하였다. yMhCt008 모체 균주의 다수의 FOA 저항성 콜로니의 게놈 DNA를 여기에 설명된 바와 같이 프라이머 0611815 및 0611817을 갖는 원하는 루프-아웃 이벤트에 대하여 PCR로 스크리닝하였다. 828 bp 밴드의 존재는 원하는 바와 같이 ura3 스카 부위 (모체 균주에서 2개의 URA3

프로모터 사이에서 상동 제조합 후 단일 URA3 프로모터 왼쪽 뒤)만의 존재를 나타냈다. 828 bp 밴드만을 나타내는 분리체를 여기에 설명된 바와 같이 프라이머 0611555 및 0611554를 사용하여 추가로 스크리닝하였다. 프라이머 0611555 및 0611554를 함유하는 PCR에서 약 1.4 kbp 밴드의 발생은 야생형 PDC 위치의 존재를 나타냈다. 두 개의 염색체상의 PDC 위치가 손실되었다는 것을 나타내는, 이 밴드가 없는 분리체를 yMhCt010으로 지정하였다.

[0442] 균주를 셰이크 플라스크에서 키웠고 CFE를 제조하였고 여기에 설명된 바와 같이 아스파르테이트 데카르복실라제 (ADC) 활성 및 3-HP 데히드로게나제 (3-HPDH) 활성에 대하여 검정하였다. 다수의 검정 세트 (시험 1-4로 나타남)에 대한 실험 결과는 표 17에 나타난다. 표 17의 시험 1의 균주를 여기에 설명된 바와 같이 SDS-PAGE로 분석하였다. 시험 1의 균주 74/75 #1 및 균주 yMhCt002는 시험 1의 대조군 균주 (Mbin500)에 존재하지 않는 27 kD에서 SDS-PAGE 분석의 밴드를 제공한다. 이 단백질 밴드의 크기는 그것의 ydfG 유전자에 의해 암호화된 단백질과 그것의 정체성이 일치한다.

[0443] 표 17: 형질전환체 효소 활성 데이터

| 시험 | 균주 | 과발현된 유전자 | 대립유전자 타입 | ADC 활성 | 3-HPDH 활성 |
|----|---------------|--|----------|------------|-----------|
| 1 | Mbin500 (대조군) | N/A | N/A | Not tested | 0.25 |
| 2 | | | | 0.00 | 0.12 |
| 3 | | | | 0.00 | 0.23 |
| 4 | | | | 0.00 | 0.45 |
| 4 | | | | 0.00 | 0.54 |
| 1 | 74/75 #1 | ADC (SEQ ID NO: 130) 3-HPDH (SEQ ID NO: 143) BAAT (SEQ ID NO: 140) | 이형 접합체 | Not tested | 0.68 |
| 1 | yMhCt002 | ADC (SEQ ID NO: 130) 3-HPDH (SEQ ID NO: 143) BAAT (SEQ ID NO: 140) | 이형 접합체 | Not tested | 0.93 |
| 2 | yMhCt004 | ADC (SEQ ID NO: 130) 3-HPDH (SEQ ID NO: 144) BAAT (SEQ ID NO: 142) | 이형 접합체 | 0.26 | 1.81 |
| 4 | | | | 0.30 | 2.20 |
| 4 | | | | 0.24 | 1.81 |
| 2 | 83/77 #2 | ADC (SEQ ID NO: 130) 3-HPDH (SEQ ID NO: 144) BAAT (SEQ ID NO: 142) | 이형 접합체 | 0.25 | 1.80 |
| 2 | yMhCt005 | ADC (SEQ ID NO: 130) 3-HPDH (SEQ ID NO: 144) gabT (SEQ ID NO: 141) | 이형 접합체 | 0.17 | 1.51 |
| 4 | | | | 0.17 | 1.26 |
| 4 | | | | 0.14 | 1.24 |
| 2 | 87/77 #2 | ADC (SEQ ID NO: 130) 3-HPDH (SEQ ID NO: 144) gabT (SEQ ID NO: 141) | 이형 접합체 | 0.20 | 1.68 |
| 3 | yMhCt005 | ADC (SEQ ID NO: 130) 3-HPDH (SEQ ID NO: 144) gabT (SEQ ID NO: 141) | 이형 접합체 | 0.16 | 1.27 |
| 3 | yMhCt008 | ADC (SEQ ID NO: 130) 3-HPDH (SEQ ID NO: 144) gabT (SEQ ID NO: 141) | 동형 접합체 | 0.80 | 2.29 |
| 4 | | | | 0.25 | 0.85 |
| 4 | | | | 0.42 | 1.37 |
| 3 | yMhCt009 | ADC (SEQ ID NO: 130) 3-HPDH (SEQ ID NO: 144) gabT (SEQ ID NO: 141) | 동형 접합체 | 0.45 | 1.33 |
| 4 | | | | 0.22 | 0.55 |
| 4 | | | | 0.43 | 1.15 |
| 4 | yMhCt013 | ADC (SEQ ID NO: 130) 3-HPDH (SEQ ID NO: 144) BAAT (SEQ ID NO: 142) | 동형 접합체 | 0.47 | 0.65 |
| 4 | | | | 0.50 | 0.82 |
| 4 | | | | 0.34 | 0.51 |
| 4 | yMhCt014 | ADC (SEQ ID NO: 130) 3-HPDH (SEQ ID NO: 144) BAAT (SEQ ID NO: 142) | 동형 접합체 | 0.54 | 1.06 |

[0444]

[0445] 또 다른 검정 세트에 대한 실험 결과는 표 17에 나타난다 (시험 2). 표 17의 시험 2의 균주를 또한 여기에 설명된 바와 같이 SDS-PAGE로 분석하였다. Mbin500을 제외한 시험 2의 모든 균주는 SDS-PAGE 분석에서 29 kD에서 밴드를 제공하였다. 이 단백질 밴드의 크기는 YMR226c 유전자에 의해 암호화된 단백질과 정체성이 일치한다. 시험 2에 대하여 균주 yMhCt005 및 87/77 #2는 이 시험에 대한 세 개의 다른 샘플에 존재하지 않는 53 kD에서 제공하였다. 이 단백질 밴드의 크기는 UGA1 유전자에 의해 암호화된 단백질과 그것의 정체성이 일치한다.

[0446] 또 다른 검정 세트에 대하여 실험 결과는 표 17에 나타난다 (시험 3). 표 17의 시험 3의 균주를 또한 여기에 설명된 바와 같이 SDS-PAGE로 분석하였다. Mbin500을 제외한 시험 3의 모든 균주는 SDS-PAGE 분석에서 53 kD 및 29 kD에서 밴드를 제공하였다. 이 단백질 밴드의 크기는 UGA1 및 YMR226c 유전자 각각에 의해 암호화된 단백질과 정체성이 일치한다. 시험 3의 균주 Mbin500 및 yMhCt005는 이 시험을 위해 yMhCt008 및 yMhCt009에 존재하지 않는 SDS-PAGE에서 64 kD에서 밴드를 나타냈다. 이 단백질 밴드의 크기는 아이. 오리엔탈리스 CNB1의 네이티브 피루베이트 데카르복실라제 (PDC) 유전자에 의해 암호화된 단백질과 정체성이 일치한다.

[0447] 또 다른 검정 세트에 대하여 실험 결과는 표 17에 나타난다 (시험 4). 표 17의 시험 4의 균주를 또한 여기에 설명

된 바와 같이 SDS-PAGE로 분석하였다. MBin500을 제외한 시험 4의 모든 균주는 29 kD에서 밴드를 제공하였다. 이 단백질 밴드의 크기는 YMR226c 유전자에 의해 암호화된 단백질과 정체성이 일치한다. 시험 4의 균주 yMhCt005, yMhCt008, 및 yMhCt009는 53 kD에서 밴드를 나타냈다. 이 밴드의 크기는 UGA1 유전자에 의해 암호화된 단백질과 일치한다. 시험 4의 균주 yMhCt013, 및 yMhCt014는 53 kD에서 희미한 밴드를 나타냈다. 이 밴드의 크기는 PYD4 유전자에 의해 암호화된 단백질과 일치한다. 시험 4의 균주 MBin500, yMhCt004, 및 yMhCt005는 균주 yMhCt008, yMhCt009, yMhCt013, 및 yMhCt014에 없는 64 kD에서 밴드를 나타냈다. 이 단백질 밴드의 크기는 아이. 오리엔탈리스 CNB1의 네이티브 피루베이트 데카르복실라제 (PDC) 유전자에 의해 암호화된 단백질과 정체성이 일치한다.

[0448] 여기에 설명된 방법을 사용하여, 3-HP 생산에 대하여 바이오리액터에서 평가된 균주 MBin500 및 yMhCt008을 테스트하였다. 대조군 균주 MBin500은 검출 가능한 3-HP를 생산하지 않았다 (두 개의 독립적인 발효의 평균). 균주 yMhCt008은 2.45 g/L 3-HP를 생산하였다 (12개의 독립적인 발효의 평균).

[0449] 실시예 3A-6: adh1202 위치에서 β -알라닌 아미노트랜스퍼라제 (BAAT) 또는 3-히드록시프로피온산 데히드로게나제 (3-HPDH)를 발현하고, pdc 위치에서 아스파르테이트 1-데카르복실라제 (ADC), β -알라닌 아미노트랜스퍼라제 (BAAT), 및 3-히드록시프로피온산 데히드로게나제 (3-HPDH)를 발현하는 효모 균주

[0450] pMhCt077, pMhCt083, 및 pMhCt087 (상기) 20 μ g을 NotI 및 PvuII로 분해하였고 89 mM 염기성 트리스-89 mM 붕산-2 mM 이나트륨 EDTA (TBE) 버퍼를 사용하여 1% 아가로스 겔 상에서 실행하였다. 각각 pMhCt077, pMhCt083, 및 pMhCt087로부터 3815 bp, 5332 bp, 또는 5320 bp의 NotI 분해된 단편을 겔에서 잘라냈고 QIAQUICK® Gel Extraction Kit (Qiagen)를 사용하여 아가로스로부터 추출하였다. NotI 분해된 pMhCt077 560 ng 및 NotI 분해된 pMhCt083 또는 pMhCt087 560 ng을 균주 MIBa320, MIBa321, 및 MIBa322에 형질전환하였다. MIBa320을 pMhCt077/83 및 pMhCt077/87 조합과 형질전환하였다. MIBa321을 pMhCt077/87과 형질전환하였고 MIBa322를 여기에 설명된 바와 같이 pMhCt077/83과 형질전환하였다. 형질전환체를 ura 선택 배지에 도말하였고 상온에서 약 60 시간 동안 배양하였다. 형질전환체를 ura 선택 배지에 재-스트리킹하였고 37 °C에서 하룻밤 동안 배양하였다. 게놈 DNA를 URA3+ 콜로니로부터 제조하였고 발현 카세트의 통합을 확인하기 위해 여기에 설명된 바와 같이 PCR로 체크하였다. 프라이머 쌍 611814 및 611554는 통합을 나타내는 1.9 kbp 단편을 증폭시켰다. 프라이머 쌍 611555 및 611554는 야생형 위치를 나타내는 1.4 kbp 단편을 증폭시켰다. 611554 및 611814를 갖는 1.9 kbp 및 611555 및 611554를 갖는 1.4 kbp의 PCR 단편을 증폭시키는 각각의 계통의 하나의 URA3+ 형질전환체를 저장하였다; 이들은 MIBa323, MIBa324, MIBa325, 및 MIBa327로 지정된다 (유전자형에 대하여 표 18 참조). 프로모터 및 종결자를 아이. 오리엔탈리스 유전자로부터 유도하였다.

[0451] 표 18: 형질전환체 유전자형

| 균주 설계 | 유전자형 |
|---------|---|
| MIBa323 | adh1202 Δ ::(PDC _{프로모터} -Opt.ScYMR226c, URA3-스카)/ adh1202 Δ ::(PDC _{프로모터} -Opt.ScYMR226c, URA3-스카) pdc Δ ::(PDC _{프로모터} -Opt.SaPanD, ENO1 _{프로모터} -Opt.ScUGA1, URA3, TDH3 _{프로모터} -Opt.ScYMR226c)/PDC ura3-/ura3- |
| MIBa324 | adh1202 Δ ::(PDC _{프로모터} -Opt.ScYMR226c, URA3-스카)/ adh1202 Δ ::(PDC _{프로모터} -Opt.ScYMR226c, URA3-스카) pdc Δ ::(PDC _{프로모터} -Opt.SaPanD, ENO1 _{프로모터} -Opt.SkPyd4, URA3, TDH3 _{프로모터} -Opt.ScYMR226c)/PDC ura3-/ura3- |
| MIBa325 | adh1202 Δ ::(PDC _{프로모터} -Opt.ScUGA1, URA3-스카)/ adh1202 Δ ::(PDC _{프로모터} -Opt.ScUGA1, URA3-스카) pdc Δ ::(PDC _{프로모터} -Opt.SaPanD, ENO1 _{프로모터} -Opt.ScUGA1, URA3, TDH3 _{프로모터} -Opt.ScYMR226c)/PDC ura3-/ura3- |
| MIBa327 | adh1202 Δ ::(PDC _{프로모터} -Opt.SkPyd4, URA3-스카)/ adh1202 Δ ::(PDC _{프로모터} -Opt.SkPyd4, URA3-스카) pdc Δ ::(PDC _{프로모터} -Opt.SaPanD, ENO1 _{프로모터} -Opt.SkPyd4, URA3, TDH3 _{프로모터} -Opt.ScYMR226c)/PDC ura3-/ura3- |

[0452] 균주를 셰이크 플라스크에서 키웠고 CFE를 제조하였고 3-HP 데히드로게나제 (3-HPDH) 활성에 대하여 여기에 설명된 바와 같이 검정하였다. 결과는 표 19에 나타난다.

[0454] 표 19: 형질전환체 효소 활성 데이터

| 균주 | 과발현된 유전자 | 3-HPDH 활성 |
|------------------|--|--------------|
| MBin500 (대조군) | N/A | 0.14 |
| MIba314 | gabT (SEQ ID NO: 141) | 0.09 |
| MIba318 | gabT (SEQ ID NO: 141) | 0.41 |
| MIba321 | gabT (SEQ ID NO: 141) | 0.08 |
| MIba325 | gabT (SEQ ID NO: 141) ADC (SEQ ID NO: 130) 3-HPDH (SEQ ID NO: 144) | 1.15 |
| MIba315 | BAAT (SEQ ID NO: 142) | 0.12 |
| MIba319 | BAAT (SEQ ID NO: 142) | 0.17 |
| MIba322 | BAAT (SEQ ID NO: 142) | 0.09 |
| MIba327 | BAAT (SEQ ID NO: 142) ADC (SEQ ID NO: 130) 3-HPDH (SEQ ID NO: 144) | 0.98 |
| MIba316 | 3-HPDH (SEQ ID NO: 144) | 0.48 |
| MIba317 | 3-HPDH (SEQ ID NO: 144) | 2.15 |
| MIba320 | 3-HPDH (SEQ ID NO: 144) | 1.07 |
| MIba323 | 3-HPDH (SEQ ID NO: 144) ADC (SEQ ID NO: 130) gabT (SEQ ID NO: 141) | 2.66 |
| MIba324 | 3-HPDH (SEQ ID NO: 144) ADC (SEQ ID NO: 130) BAAT (SEQ ID NO: 142) | 2.12 |

[0455]

[0456] MIba323, MIba324, MIba325 및 MIba327의 ura-유도체를 여기에 설명된 바와 같이 분리하였다. 게놈 DNA를 FOA-저항성 콜로니로부터 제조하였고 URA3 선택 가능한 마커의 손실을 확인하기 위해 여기에 설명된 바와 같이 PCR로 체크하였다. 프라이머 0611815는 왼쪽 구조물의 RKI 종결자에서 어닐링하고 URA3 프로모터를 향해 증폭시켰고 프라이머 0611817은 TDH3 프로모터에서 어닐링하고 URA3은 다시 URA3 카세트를 향해 증폭시켰다. 828 bp 밴드의 존재는 원하는 바와 같이 URA3 스카 부위 (모체 균주에서 2개의 URA3 프로모터 사이에서 상동 재조합 후 단일 URA3 프로모터 왼쪽 뒤)만의 존재를 나타내는 한편, 약 2.2 kbp의 밴드는 온전한 URA3 프로모터-URA3 ORF-URA3 종결자-URA3 프로모터 카세트의 존재를 나타내고, 원하는 재조합 이벤트가 일어나지 않았다는 것을 나타낸다. 프라이머 0611815 및 0611817을 갖는 828 bp의 PCR 단편을 수득하는, MIba323, MIba324, MIba325, 및 MIba327의 ura-균주를 저장하였고 각각 MIba335, MIba333, MIba334, 및 MIba336으로 지정하였다.

[0457] 균주 MIba333 및 MIba334를 pMhCt077 및 pMhCt087의 단편과 형질전환하였고, 균주 MIba335 및 MIba336을 MIba320-2 형질전환체에 대한 섹션에서 상기 설명된 바와 같이 pMhCt077 및 pMhCt083의 단편과 형질전환하였다. 형질전환체를 여기에 설명된 바와 같이 ura 선택 배지 상에서 키움에 의해 선택하였다. 게놈 DNA를 URA3+ 콜로니로부터 제조하였고 발현 카세트의 통합을 확인하기 위해 여기에 설명된 바와 같이 PCR로 체크하였다. 프라이머 0611815를 왼쪽 구조물의 RKI 종결자에서 어닐링하였고 URA3을 향해 증폭시켰다. 프라이머 0611817은 TDH3 프로모터에서 어닐링하였고 URA3 카세트를 향해 다시 증폭시켰다. 828 bp 밴드의 존재는 첫 번째 통합을 위해 원하는 바와 같이 URA3 스카 부위 (모체 균주에서 2개의 URA3 프로모터 사이에서 상동 재조합 후 단일 URA3 프로모터 왼쪽 뒤)만의 존재를 나타내고, 약 2.2 kbp의 밴드는 두 번째 통합을 위해 온전한 URA3 프로모터-URA3 ORF-URA3 종결자-URA3 프로모터 카세트의 존재를 나타낸다. 프라이머 쌍 0611815 및 0611816은 ura 마커가 존재할 때 625 bp 단편을 증폭시킨다. 프라이머 0611555 및 0611554는 PDC 위치가 존재할 때 1.4 kbp 단편을 증폭시킨다. 동형 접합체 통합체는 이들 프라이머를 갖는 단편을 증폭시키지 않아야 한다. 프라이머 0611815 및 0611817을 갖는 828 bp 및 2.2 kbp, 프라이머 0611815 및 0611816을 갖는 625 bp의 PCR 단편을 증폭시키고 프라이머 0611555 및 0611554를 갖는 단편을 증폭시키지 않는, 각각의 계통의 하나의 URA3+ 형질전환체를 저장하였다; 이들을 MIba340, MIba341, MIba345, 및 MIba348로 지정하였다 (표 20A 참조). 프로모터 및 종결자를 아 이. 오리엔탈리스 유전자로부터 유도하였다.

[0458] 표 20A: 형질전환체 유전자형

| 균주 설계 | 유전자형 |
|---------|---|
| MIBa345 | adh1202Δ::(PDC 프로모터-Opt.ScYMR226c URA3-스카)/adh1202Δ::(PDC 프로모터-Opt.ScYMR226c URA3-스카) pdcΔ::(PDC 프로모터-Opt.SaPanD ENO1 프로모터-Opt.ScUGA1, URA3-스카, TDH3 프로모터-Opt.ScYMR226c)/ pdcΔ::(PDC 프로모터-Opt.SaPanD, ENO1 프로모터-Opt.ScUGA1, URA3, TDH3 프로모터-Opt.ScYMR226c) ura3-/ura3- |
| MIBa348 | adh1202Δ::(PDC 프로모터-Opt.ScYMR226c URA3-스카)/adh1202Δ::(PDC 프로모터-Opt.ScYMR226c URA3-스카) pdcΔ::(PDC 프로모터-Opt.SaPanD, ENO1 프로모터-Opt.SkPyd4, URA3-스카, TDH3 프로모터-Opt.ScYMR226c) / pdcΔ::(PDC 프로모터-Opt.SaPanD, ENO1 프로모터-Opt.SkPyd4, URA3, TDH3 프로모터-Opt.ScYMR226c) ura3-/ura3- |
| MIBa340 | adh1202Δ::(PDC 프로모터-Opt.ScUGA1 URA3-스카)/adh1202Δ::(PDC 프로모터-Opt.ScUGA1 URA3-스카) pdcΔ::(PDC 프로모터-Opt.SaPanD, ENO1 프로모터-Opt.ScUGA1, URA3-스카, TDH3 프로모터-Opt.ScYMR226c)/ pdcΔ::(PDC 프로모터-Opt.SaPanD, ENO1 프로모터-Opt.ScUGA1, URA3, TDH3 프로모터-Opt.ScYMR226c) ura3-/ura3- |
| MIBa341 | adh1202Δ::(PDC 프로모터-Opt.SkPyd4 URA3-스카)/adh1202Δ::(PDC 프로모터-Opt.SkPyd4 URA3-스카) pdcΔ::(PDC 프로모터-Opt.SaPanD, ENO1 프로모터-Opt.SkPyd4, URA3-스카, TDH3 프로모터-Opt.ScYMR226c) / pdcΔ::(PDC 프로모터-Opt.SaPanD, ENO1 프로모터-Opt.SkPyd4, URA3, TDH3 프로모터-Opt.ScYMR226c) ura3-/ura3- |

[0459]

[0460] 균주 MBin500 (대조균), MIBa345, MIBa348, MIBa340 및 MIBa341로부터 제조된 CFE에서 아스파르테이트 1-데카르복실라제 (ADC), 베타-알라닌 아미노트랜스퍼라제 (BAAT) 및 3-HP 데히드로게나제 (3-HPDH) 활성을 비교하였다. 이 실험의 결과는 표 20B에 나타난다.

[0461] 표 20B: 형질전환체 효소 활성 데이터

| 균주 | 과발현된 유전자 | 유전자 근원 | ADC 활성 | BAAT 활성 | 3-HPDH 활성 |
|---------------|--------------------------|-------------|--------|---------|-----------|
| MBin500 (대조균) | N/A | N/A | 0.002 | 0.61 | 0.4 |
| MIBa345 | YMR226c (SEQ ID NO: 144) | एस. 세레비시아 | 0.194 | 14.37 | 79.5 |
| | ADC (SEQ ID NO: 130) | एस. 아베르미틸리스 | | | |
| | UGA1 (SEQ ID NO: 141) | एस. 세레비시아 | | | |
| MIBa348 | YMR226c (SEQ ID NO: 144) | एस. 세레비시아 | 0.169 | 173.67 | 76.7 |
| | ADC (SEQ ID NO: 130) | एस. 아베르미틸리스 | | | |
| | PYD4 (SEQ ID NO: 142) | एस. 클루비베리 | | | |
| MIBa340 | YMR226c (SEQ ID NO: 144) | एस. 세레비시아 | 0.239 | 19.51 | 64.8 |
| | ADC (SEQ ID NO: 130) | एस. 아베르미틸리스 | | | |
| | UGA1 (SEQ ID NO: 141) | एस. 세레비시아 | | | |
| MIBa341 | YMR226c (SEQ ID NO: 144) | एस. 세레비시아 | 0.22 | 386.92 | 65.5 |
| | ADC (SEQ ID NO: 130) | एस. 아베르미틸리스 | | | |
| | PYD4 (SEQ ID NO: 142) | एस. 클루비베리 | | | |

[0462]

[0463] 실시예 3A-7: adh1202 위치에서 아스파르테이트 1-데카르복실라제 (ADC)를 발현하는 다수의 뉴클레오타이드 서열을 갖는 삽입 벡터의 왼쪽 단편

- [0464] 구조물을 아이. 오리엔탈리스 adh1202 위치에서 ADC (SEQ ID NO: 17)을 암호화하는 뉴클레오타이드의 4개의 카피를 포함하도록 설계하였다. PDC 위치에 대하여 여기에 설명된 것들과 유사한 접근법으로, 왼쪽 및 오른쪽 구조물을 상동 재조합을 허용하도록 설계하였다. 통합 벡터 일반적인 설계 및 원하는 재조합 이벤트는 도 3에 나타난다. 이 접근법을 또한 하기 실시예에 설명된 바와 같이 대체 ADC (SEQ ID NO: 139)의 발현에 사용하였다.
- [0465] 재조합을 같은 ADC 서열을 암호화하는 뉴클레오타이드 서열의 다수의 카피 사이에서 발생하는 것으로부터 방지하기 위해, 아이. 오리엔탈리스 (SEQ ID NO: 130, 145, 146, 및 147)에서 발현에 코돈-최적화된, 4개의 별개의 뉴클레오타이드 서열을 SEQ ID NO: 17의 같은 ADC 서열을 암호화하도록 설계하였다. 추가적으로, 구조물의 초기 설계를 아이. 오리엔탈리스의 ald5680 위치를 표적화하도록 설계하였기 때문에, adh1202 표적화 서열은 클로닝의 후반 단계에서 이들 벡터로 통합된다. ald5680 구조물은 ald5680에서 panD의 4개의 추가적 카피를 갖는, adh120에서 panD의 이소성 4개의 카피에 대하여 이미 동형인 아이. 오리엔탈리스 CNB1 균주에서 두 번째 위치를 표적화하기 위해 사용될 수 있다.
- [0466] 왼쪽 ald5680 표적화 벡터를 다음과 같이 제조하였다. 원하는 추가적 제한 부위 및 플랜킹 DNA와 함께, ald5680 ORF의 5'인 서열을 함유하는 PCR 생성물을 프라이머 0612271 및 0612272로 증폭시켰다. PCR 반응물 (50 μ L)은 1 내지 50배 희석된 pHJJ75 미니-프랩 플라스미드 DNA (도 23) 1 μ L, 1X iProof™ HF 버퍼 (Bio-Rad Laboratories), 100 pmol 각각의 프라이머 0612271 및 0612272, 200 μ M 각각의 dATP, dCTP, dGTP, 및 dTTP, 및 1 유닛의 iProof™ High-Fidelity DNA Polymerase (Bio-Rad Laboratories)를 함유하였다. PCR을 98 °C에서 30초 동안 1 사이클 후 이어서 각각 98 °C에서 10초 동안, 59 °C에서 20초 동안, 72 °C에서 45초 동안 32 사이클, 72 °C에서 10분 동안 최종 연장으로 프로그램된 EPPENDORF® MASTERCYCLER® (Eppendorf Scientific)로 수행하였다. 열 순환 후, 약 930개의 염기쌍 PCR 생성물을 겔에서 잘라냈고 제조사의 지시에 따라 NUCLEOSPIN® Extract II Kit (Macherey-Nagel)를 사용하여 정제한 경우, PCR 반응 생성물을 TAE 버퍼 중의 1% 아가로스 겔 전기영동으로 분리하였다.
- [0467] 대량의 DNA의 정제를 허용하기 위해, 상기 설명된 PCR 반응을 주형 DNA로 정제된 930 bp PCR 생성물을 사용하여 반복하였다. 5개의 50 μ L 반응물을 설정하였고 1 μ L의 정제된 930 bp PCR 생성물이 주형 DNA로 pHJJ75 플라스미드 (상기)를 대체하는 것을 제외한 상기 설명된 조건으로 증폭시켰다. 열 순환 후, 증폭된 930 bp 생성물을 상기와 같이 정제하였다.
- [0468] PDC promo-optPanD-ENO1-UGA1를 함유하는 단편 (이것은 SEQ ID NO: 130의 서열을 암호화하는, 아이. 오리엔탈리스 코돈-최적화된 에스. 아베르미틸리스 ADC를 함유한다)을 NotI 및 EcoRI 분해를 통해 pMhCt087 (상기)에서 잘라냈다. pMhCt087의 미니-프랩 (midi-prep)의 10 μ g을 NotI 및 EcoRI으로 분해하였고 TAE 버퍼 중의 0.9% 아가로스 겔 전기영동으로 정제하였다. 약 4.4 kbp 밴드를 겔에서 잘라냈고 제조사의 지시에 따라 NUCLEOSPIN® Extract II Kit (Macherey-Nagel)를 사용하여 정제하였다.
- [0469] 클로닝을 위해 원하는 추가적 제한 부위 및 플랜킹 DNA를 갖는 URA3 분열 마커의 5' 절반을 함유하는 PCR 생성물을 프라이머 0612273 및 0612274를 사용하여 증폭하였다. PCR 반응물 (50 μ L)은 1 내지 50배 희석된 pMhCt082 미니-프랩 플라스미드 DNA (상기), 1X iProof™ HF 버퍼 (Bio-Rad Laboratories), 100 pmol 각각의 프라이머 0612273 및 0612274, 200 μ M 각각의 dATP, dCTP, dGTP, 및 dTTP, 및 1 유닛의 iProof™ High-Fidelity DNA Polymerase (Bio-Rad Laboratories)를 함유하였다. PCR을 98 °C에서 30초 동안 1 사이클 후 이어서 각각 98 °C에서 10초 동안, 59 °C에서 20초 동안, 72 °C에서 45초 동안 32 사이클, 72 °C에서 10분 동안 최종 연장으로 프로그램된 EPPENDORF® MASTERCYCLER® (Eppendorf Scientific)로 수행하였다. 열 순환 후, 약 960개의 염기쌍 PCR 생성물을 겔에서 잘라냈고 제조사의 지시에 따라 NUCLEOSPIN® Extract II Kit (Macherey-Nagel)를 사용하여 정제한 경우, PCR 반응 생성물을 TAE 버퍼 중의 1% 아가로스 겔 전기영동으로 분리하였다.
- [0470] 상기 DNA 단편에 대한 수령자 벡터를 생성하기 위해, 플라스미드 pUC19 (Yanisch-Perron, C., Vieira, J. and Messing, J. (1985) Gene, 33, 103-119)를 HindIII 및 EcoRI으로 분해하였고, 37 °C에서 30분 동안 10 유닛 소위장 포스포타제 (New England Biolabs)를 처리하였고, TAE 버퍼 중의 0.9% 아가로스 겔 전기영동으로 정제하였고, 약 2.6 kbp 밴드를 겔에서 잘라냈고 제조사의 지시에 따라 NUCLEOSPIN® Extract II Kit (Macherey-Nagel)를 사용하여 정제하였다.
- [0471] 상기 정제된 930 bp, 4.4 kbp, 및 960 bp DNA 단편을 HindIII 및 EcoRI로 분해된 pUC19 벡터 150 ng, ald5680 플랜킹 DNA를 함유하는 930 bp DNA 56 ng, NotI 및 EcoRI으로 분해된 pMhCt087의 PDC promo-optPanD-ENO1-UGA1 단편 250 ng, URA3 분열 마커의 5' 절반을 함유하는 960 bp PCR 생성물 5' 55 ng, 1X In-Fusion 반응 버퍼 (Clontech) 및 IN-FUSION™ 효소 (Clontech) 1 μ L로 구성된, 10 μ L의 총 반응량에서 IN-FUSION™ Advantage

PCR Cloning Kit (Clontech)를 사용하여 분해된 pUC19 단편에 삽입하였다. 반응물을 37 °C에서 15분 동안, 50 °C에서 15분 동안 배양하였고, 얼음 위에 두었다. 반응물을 40 µL의 TE 버퍼로 희석하였고 2.5 µL를 제조사의 지시에 따라 SoloPack Gold SuperCompetent 세포 (Agilent Technologies)를 형질전환하기 위해 사용하였다. 형질전환체를 2X YT+amp 플레이트에 도말하였고 상온에서 3일 동안 배양하였다. 결과의 형질전환체 중 다수를 SalI 및 HpaI 분해에 의해 원하는 PCR 생성물의 적절한 삽입에 대해 스크리닝하였다. 원하는 밴드 크기를 수득하는 클론을 DNA 시퀀싱에 의해 올바른 것으로 확인하였고 pMhCt089로 지정하였다.

[0472] 그 다음, pMhCt089에서 UGA1 ORF를 아이. 오리엔탈리스에서 발현에 코돈-최적화된 에스. 아베르미틸리스 panD 유전자로 대체하였으며 이것은 SEQ ID NO: 17의 ADC를 암호화한다. 에스. 아베르미틸리스 panD 버전 r1 (SEQ ID NO: 145)을 벡터 pMA-T에서 GeneArt®로 합성하였다. 플라스미드 pMA-T를 XbaI 및 PacI로 분해하였고, 결과의 단편을 TAE 버퍼 중의 0.9% 아가로스 겔 전기영동으로 분리하였고, 약 434 bp 밴드를 겔에서 잘라냈고 제조사의 지시에 따라 NUCLEOSPIN® Extract II Kit (Macherey-Nagel)를 사용하여 정제하였다. 플라스미드 pMhCt089를 XbaI 및 PacI로 분해하였고, 37 °C에서 30분 동안 10 유닛 소 위장 포스파타제 (New England Biolabs)를 처리하였고, 결과의 단편을 TAE 버퍼 중의 0.9% 아가로스 겔 전기영동으로 분리하였고, 약 4.5 kbp 밴드를 겔에서 잘라냈고 제조사의 지시에 따라 NUCLEOSPIN® Extract II Kit (Macherey-Nagel)를 사용하여 정제하였다. pMhCt089 벡터 및 에스. 아베르미틸리스 panD 버전 r1을 1X QUICK 결찰 버퍼 (New England Biolabs), 2 µL XbaI/PacI pMhCt089 벡터, 2 µL XbaI/PacI 에스. 아베르미틸리스 panD 버전 r1 인서트, 및 1 µL QUICK T4 DNA 리가제 (New England Biolabs)를 함유하는 결찰 반응물 (20 µL)과 합쳤다. 결찰 반응물을 상온에서 5분 동안 배양하였고, 얼음 위에 두었다. 이 반응물 중 5 µL를 제조사의 지시에 따라 ONE SHOT® TOP10 화학적 컴피턴트 이. 콜리 세포 (Invitrogen)를 형질전환하기 위해 사용하였다. 형질전환체를 2X YT+amp 플레이트에 도말하였고 상온에서 3일 동안 배양하였다. 결과의 형질전환체 중 다수를 XbaI 및 PacI 분해에 의해 원하는 PCR 생성물의 적절한 삽입에 대해 스크리닝하였다. 원하는 밴드 크기를 수득하는 클론을 DNA 시퀀싱에 의해 올바른 것으로 확인하였고 pMhCt092로 지정하였다.

[0473] 왼쪽 구조물에 대한 최종 클로닝 단계는 pMhCt092에 존재하는 ald5680 5' 상동 영역을 adh1202 5' 상동 영역으로 대체하는 것이었다. 클로닝을 위해 원하는 추가의 제한 부위 및 플랜킹 DNA와 함께 adh1202 ORF의 서열 5'을 함유하는 PCR 생성물을 프라이머 0612470 및 0612471로 증폭하였다. PCR 반응물 (50 µL)은 1 µL의 1 내지 50 배 희석된 pGMEr140 미니-프랩 플라스미드 DNA (PCR 증폭된 영역이 동일한 여기에 설명된 pMBa107의 유도체), 1X iProof™ HF 버퍼 (Bio-Rad Laboratories), 100 pmol 각각의 프라이머 0612470 및 0612471, 200 µM 각각의 dATP, dCTP, dGTP, 및 dTTP, 및 1 유닛의 iProof™ High-Fidelity DNA Polymerase (Bio-Rad Laboratories)를 함유하였다. PCR을 98 °C에서 30초 동안 1 사이클 후 이어서 각각 98 °C에서 10초 동안, 59 °C에서 20초 동안, 72 °C에서 30초 동안 34 사이클, 72 °C에서 10분 동안 최종 연장으로 프로그램된 EPPENDORF® MASTERCYCLER® (Eppendorf Scientific)로 수행하였다. 열 순환 후, 약 790 bp PCR 생성물을 겔에서 잘라냈고 제조사의 지시에 따라 NUCLEOSPIN® Extract II Kit (Macherey-Nagel)를 사용하여 정제한 경우, PCR 반응 생성물을 TAE 버퍼 중의 1% 아가로스 겔 전기영동으로 분리하였다.

[0474] 상기 PCR 생성물에 대한 수령자 벡터를 생성하기 위해, 플라스미드 pMhCt092를 HpaI 및 NotI로 분해하였고, 37 °C에서 30분 동안 10 유닛 소 위장 포스파타제 (New England Biolabs)를 처리하였고, TAE 버퍼 중의 0.9% 아가로스 겔 전기영동으로 정제하였고, 약 7.0 kbp 밴드를 겔에서 잘라냈고 제조사의 지시에 따라 NUCLEOSPIN® Extract II Kit (Macherey-Nagel)를 사용하여 정제하였다. PCR 생성물 및 선형 벡터를 HpaI 및 NotI로 분해된 pMhCt092 벡터 120 ng, PCR 생성물을 함유하는 adh1202 5' 상동체 30 ng, 1X In-Fusion 반응 버퍼 (Clontech) 및 IN-FUSION™ 효소 (Clontech) 1 µL로 구성된, 10 µL의 총 반응량에서 IN-FUSION™ Advantage PCR Cloning Kit (Clontech)를 사용하여 합쳤다. 반응물을 37 °C에서 15분 동안, 50 °C에서 15분 동안 배양하였고, 얼음 위에 두었다. 반응물을 40 µL의 TE 버퍼로 희석하였고 2.5 µL를 제조사의 지시에 따라 SoloPack Gold SuperCompetent 세포 (Agilent Technologies)를 형질전환하기 위해 사용하였다. 형질전환체를 2X YT+amp 플레이트에 도말하였고 37 °C에서 하룻밤 동안 배양하였다. 결과의 형질전환체 중 다수를 BamHI 및 PstI로 분해에 의해 원하는 PCR 생성물의 적절한 삽입에 대해 스크리닝하였다. 원하는 밴드 크기를 수득하는 클론을 DNA 시퀀싱에 의해 올바른 것으로 확인하였고 pMhCt095 (도 12)로 지정하였다.

[0475] 플라스미드 pMhCt095는 아이. 오리엔탈리스에서 발현에 코돈-최적화된 에스. 아베르미틸리스 ADC 유전자 (SEQ ID NO: 130)의 발현을 구동하는 PDC 프로모터, 아이. 오리엔탈리스 RKI 종결자, 아이. 오리엔탈리스 URA3 프로모터 및 아이. 오리엔탈리스 URA3 ORF의 5' 단편을 함유하는, 왼쪽 아이. 오리엔탈리스 adh1202 표적화 구조물이다.

- [0476] 실시예 3A-8: adh1202 위치에서 아스파르테이트 1-데카복실라제 (ADC)를 발현하는 다수의 뉴클레오타이드 서열을 갖는 삽입 벡터의 오른쪽 단편
- [0477] 클로닝을 위해 원하는 추가의 제한 부위 및 플랭킹 DNA와 함께 아이. 오리엔탈리스 URA3 ORF의 3' 단편을 함유하는 PCR 생성물을 프라이머 0612275 및 0612276으로 증폭하였다. PCR 반응물 (50 μ L)은 1 내지 50배 희석된 pMhCt069 미니-프랩 플라스미드 DNA (상기) 1 μ L, 1X iProof™ HF 버퍼 (Bio-Rad Laboratories), 100 pmol 각각의 프라이머 0612275 및 0612276, 200 μ M 각각의 dATP, dCTP, dGTP, 및 dTTP, 및 1 유닛의 iProof™ High-Fidelity DNA Polymerase (Bio-Rad Laboratories)를 함유하였다. PCR을 98 °C에서 30초 동안 1 사이클 후 이어서 각각 98 °C에서 10초 동안, 59 °C에서 20초 동안, 72 °C에서 45초 동안 32 사이클, 72 °C에서 10분 동안 최종 연장으로 프로그램된 EPPENDORF® MASTERCYCLER® (Eppendorf Scientific)로 수행하였다. 열 순환 후, 약 1155 bp PCR 생성물을 겔에서 잘라냈고 제조사의 지시에 따라 NUCLEOSPIN® Extract II Kit (Macherey-Nagel)를 사용하여 정제한 경우, PCR 반응 생성물을 TAE 버퍼 중의 1% 아가로스 겔 전기영동으로 분리하였다.
- [0478] TDH3 프로모터, 이소성 유전자의 삽입을 위한 XbaI 및 PacI 제한 부위, 및 PDC 종결자를 함유하는 단편을 NotI 및 PmeI로 분해를 통해 pMhCt069에서 잘라냈다. pMhCt069의 미니-프랩의 10 μ g을 NotI 및 PmeI로 분해하였고 TAE 버퍼 중의 0.9% 아가로스 겔 전기영동으로 정제하였고, 약 1.85 kbp 밴드를 겔에서 잘라냈고 제조사의 지시에 따라 NUCLEOSPIN® Extract II Kit (Macherey-Nagel)를 사용하여 정제하였다.
- [0479] 클로닝을 위해 원하는 추가의 제한 부위 및 플랭킹 DNA와 함께, ald5680 ORF의 서열 3'을 함유하는 PCR 생성물을 프라이머 0612277 및 0612278로 증폭시켰다. PCR 반응물 (50 μ L)은 1 내지 50배 희석된 pHJJ75 미니-프랩 플라스미드 DNA (도 23) 1 μ L, 1X iProof™ HF 버퍼 (Bio-Rad Laboratories), 100 pmol 각각의 프라이머 0612277 및 0612278, 200 μ M 각각의 dATP, dCTP, dGTP, 및 dTTP, 및 1 유닛의 iProof™ High-Fidelity DNA Polymerase (Bio-Rad Laboratories)를 함유하였다. PCR을 98 °C에서 30초 동안 1 사이클 후 이어서 각각 98 °C에서 10초 동안, 59 °C에서 20초 동안, 72 °C에서 45초 동안 32 사이클, 72 °C에서 10분 동안 최종 연장으로 프로그램된 EPPENDORF® MASTERCYCLER® (Eppendorf Scientific)로 수행하였다. 열 순환 후, 약 1155 bp PCR 생성물을 겔에서 잘라냈고 제조사의 지시에 따라 NUCLEOSPIN® Extract II Kit (Macherey-Nagel)를 사용하여 정제한 경우, PCR 반응 생성물을 TAE 버퍼 중의 1% 아가로스 겔 전기영동으로 분리하였다.
- [0480] 상기 정제된 1155 bp, 1.85 kbp, 및 844 bp DNA 단편을 HindIII 및 EcoRI로 분해된 pUC19 벡터의 단편 150 ng, URA3 분열 마커의 3' 부분을 함유하는 1155 bp DNA 66 ng, PmeI 및 NotI으로 분해되고 TDH3 프로모터, 이소성 유전자의 삽입을 위한 XbaI 및 PacI 제한 부위, 및 pMhCt069의 PDC 종결자를 함유하는 1.85 kbp 단편 106 ng, 3' ald5680 플랭킹 DNA를 함유하는 844 bp PCR 생성물 48 ng, 1X In-Fusion 반응 버퍼 (Clontech) 및 IN-FUSION™ 효소 (Clontech) 1 μ L로 구성된, 10 μ L의 총 반응량에서 IN-FUSION™ Advantage PCR Cloning Kit (Clontech)를 사용하여 EcoRI 및 HindIII로 분해된 pUC19에 삽입하였다. 반응물을 37 °C에서 15분 동안, 50 °C에서 15분 동안 배양하였고, 얼음 위에 두었다. 반응물을 40 μ L의 TE 버퍼로 희석하였고 2.5 μ L를 제조사의 지시에 따라 SoloPack Gold SuperCompetent 세포 (Agilent Technologies)를 형질전환하기 위해 사용하였다. 형질전환체를 2X YT+amp 플레이트에 도말하였고 상온에서 3일 동안 배양하였다. 결과의 형질전환체 중 다수를 SalI 및 HpaI로 분해에 의해 원하는 PCR 생성물의 적절한 삽입에 대해 스크리닝하였다. 원하는 밴드 크기를 수득하는 클론을 DNA 시퀀싱에 의해 올바른 것으로 확인하였고 ald5680 right #20으로 지정하였다.
- [0481] TKL 종결자, PGK1 프로모터, XbaI 및 PacI 제한 부위, 및 더 짧은 버전의 PDC 종결 영역을 다음과 같이 ald5680 right #20의 TDH3 프로모터 및 3' ald5680 플랭킹 DNA 사이에 추가하였다. 클로닝을 위해 원하는 추가의 제한 부위 및 플랭킹 DNA와 함께, TKL 종결자를 프라이머 0612356 및 0612357로 증폭시켰다. 원하는 PCR 생성물을 어닐링 온도를 위한 온도 구매 및 일부 반응에서 DMSO를 사용하여 증폭시켰다. 4개의 동일한 PCR 반응을 제조하였으며, 각각의 PCR 반응물 (50 μ L)은 1 내지 50배의 pACN23 미니-프랩 플라스미드 DNA (도 20) 1 μ L, 1X iProof™ HF 버퍼 (Bio-Rad Laboratories), 100 pmol 각각의 프라이머 0612356 및 0612357, 200 μ M 각각의 dATP, dCTP, dGTP, 및 dTTP, 및 1 유닛의 iProof™ High-Fidelity DNA Polymerase (Bio-Rad Laboratories)를 함유하였다. 4개의 튜브의 두 번째 세트를 반응 각각이 DMSO의 1.5 μ L의 추가를 포함했다는 것을 제외하고 상기와 같이 설정하였다. PCR을 98 °C에서 30초 동안 1 사이클 후 이어서 각각 98 °C에서 10초 동안, X °C에서 20초 동안, 72 °C에서 45초 동안 32 사이클, 72 °C에서 10분 동안 최종 연장으로 프로그램된 EPPENDORF® MASTERCYCLER® (Eppendorf Scientific)로 수행하였으며, X = 47.6 °C, 51.8 °C, 57.1 °C, 또는 62.1 °C이다. DMSO가 있고 없는 PCR을 나타난 각각의 어닐링 온도에 대하여 실행하였다. 열 순환 후, TAE 버퍼 중의 1% 아가로스 겔 전기영동으로 분리하였다. 이 겔의 시각화는 가장 높은 어닐링 온도에서 DMSO가 있고 가장 낮은 어닐링 온도에서 DMSO가 없이 수행된 PCR 반응이 원하는 844 bp 생성물의 가장 높은 수득율을 제공한다는 것을 나타냈

다. 이들 4개의 PCR을 결합하였고, 약 844개의 염기쌍 PCR 생성물을 겔에서 잘라냈고 제조사의 지시에 따라 NUCLEOSPIN® Extract II Kit (Macherey-Nagel)를 사용하여 정제한 경우, TAE 버퍼 중의 1% 아가로스 겔 전기영동으로 분리하였다.

[0482] 원하는 PGK1 프로모터 영역의 PCR 증폭을 두 단계 공정으로 수행하였다. 먼저, PGK1 프로모터 DNA를 함유하는 PCR 생성물을 다음 프라이머 0612150 및 0612151로 증폭 후 클로닝하였다. PCR 반응물 (50 μ L)은 3 μ L의 pJLJ49 미니-프래그 DNA (도 25), 1X Pfx 증폭 버퍼 (Invitrogen), 2 mM MgSO₄, 100 pmol 각각의 프라이머 0612150 및 0612151, 200 μ M 각각의 dATP, dCTP, dGTP, 및 dTTP, 및 1.25 유닛 Platinum® Pfx DNA Polymerase (Invitrogen)를 함유하였다. PCR을 95 °C에서 2 분 동안 1 사이클 후 이어서 각각 95 °C에서 1분 동안, 55 °C에서 1분 동안, 및 72 °C에서 3분 동안 25 사이클, 72 °C에서 10분 동안 최종 연장으로 프로그램된 EPPENDORF® MASTERCYCLER® (Eppendorf Scientific)로 수행하였다. 열 순환 후, 약 630 bp PCR 생성물을 겔에서 잘라냈고 제조사의 지시에 따라 NUCLEOSPIN® Extract II Kit (Macherey-Nagel)를 사용하여 정제한 경우, PCR 반응 생성물을 TAE 버퍼 중의 1% 아가로스 겔 전기영동으로 분리하였다.

[0483] 약 630 bp PCR 생성물을 제조사의 지시에 따라 시퀀싱용 Zero Blunt® TOPO® PCR Cloning Kit (Invitrogen)를 사용하여 pCR4™BLUNT TOPO® (Invitrogen) 벡터에 클로닝하였다. 6 μ L의 총 반응량에서 0.5 또는 4 μ L 630 bp PCR 생성물, 1 μ L 염용액 (Invitrogen) 및 1 μ L pCR4™BLUNT TOPO® (Invitrogen)를 상온에서 15분 동안 함께 배양하였다. 각 클로닝 반응물 2 μ L를 제조사의 지시에 따라 One Shot® TOP10 화학적 컴피턴트 이. 콜리 세포 (Invitrogen)에 형질전환하였다. 형질전환체를 LB+kan 플레이트에 도말하였고 37 °C에서 하룻밤 동안 배양하였다. 결과의 형질전환체 중 다수를 EcoRI 분해에 의해 원하는 PCR 생성물의 적절한 삽입을 위해 스크리닝하였다. 원하는 밴드 크기를 수득하는 클론을 DNA 시퀀싱에 의해 올바른 것으로 확인하였고 PGK1_in_TOPO로 지정하였다.

[0484] PGK1_in_TOPO의 PGK1 프로모터를 다음과 같이 PCR 주형으로 사용 전에 분리하였고 정제하였다. PGK1_in_TOPO의 미디-프래그 25 μ L를 XbaI 및 PacI로 분해하였고 TAE 버퍼 중의 0.9% 아가로스 겔 전기영동으로 정제하였고, 약 640 bp 밴드를 겔에서 잘라냈고 제조사의 지시에 따라 NUCLEOSPIN® Extract II Kit (Macherey-Nagel)를 사용하여 정제하였다.

[0485] 클로닝을 위해 원하는 추가의 제한 부위 및 플랭킹 DNA와 함께, PGK1 프로모터를 온도 구배를 사용하여 프라이머 0612358 및 0612359로 증폭하였다. 8개의 동일한 PCR 반응물을 설정하였고, 각각의 PCR 반응물 (50 μ L)은 PGK1_in_TOPO의 XbaI 및 PacI 분해를 통해 정제된 PGK1 프로모터 DNA 20 ng, 1X Herculase 반응 버퍼 (Agilent Technologies), 100 pmol 각각의 프라이머 0612358 및 0612359, 200 μ M 각각의 dATP, dCTP, dGTP, 및 dTTP, 및 2.5 유닛의 Herculase HotStart DNA Polymerase (Agilent Technologies)를 함유하였다. PCR을 98 °C에서 30 초 동안 1 사이클 후 이어서 각각 98 °C에서 10초 동안, X °C에서 20초 동안, 72 °C에서 45초 동안 32 사이클, 72 °C에서 10분 동안 최종 연장으로 프로그램된 EPPENDORF® MASTERCYCLER® (Eppendorf Scientific)로 수행하였으며, X = 53.7 °C, 55.4 °C, 57.6 °C, 60.0 °C, 62.4 °C, 64.8 °C, 66.9 °C, 68.6 °C이다. 열 순환 후, 각각의 PCR 반응물 10 μ L를 TAE 버퍼 중의 1% 아가로스 겔 전기영동으로 분리하였다. 이 겔의 시각화는 가장 높은 어닐링 온도로 수행된 4개의 PCR 반응물이 원하는 약 700 bp 생성물의 가장 높은 수득율을 제공한다는 것을 나타냈다. 이들 4개의 PCR을 합쳤고, 약 700개의 염기쌍 PCR 생성물을 겔에서 잘라냈고 제조사의 지시에 따라 NUCLEOSPIN® Extract II Kit (Macherey-Nagel)를 사용하여 정제한 경우, TAE 버퍼 중의 1% 아가로스 겔 전기영동으로 분리하였다.

[0486] 플라스미드 ald5680_right #20은 PDC 종결 영역으로서 아이. 오리엔탈리스 PDC ORF의 약 870 bp의 영역 다운스트림을 함유한다. 하지만, 이 영역은 종결자로서 적절한 기능에 대하여 필요 이상으로 클 가능성이 있고 그것의 현재 크기로 유지되면 PDC 위치에 대하여 원하지 않는 상동 재조합을 위한 촉매의 역할을 할 수도 있다. 그러므로, 더 작은 버전으로 ald5680_right #20에서 PDC 종결자를 대체하는 PCR 생성물을 프라이머 0612360 및 0612361로 증폭하였다. 원하는 PCR 생성물을 어닐링 온도에 대한 온도 구배 및 일부 반응에서 DMSO를 사용하여 증폭하였다. 4개의 동일한 PCR 반응물을 설정하였고, 각각의 PCR 반응물 (50 μ L)은 1 내지 50배 희석된 pJLJ49 미니-프래그 플라스미드 DNA (도 25) 1 μ L, 1X iProof™ HF 버퍼 (Bio-Rad Laboratories), 100 pmol 각각의 프라이머 0612360 및 0612361, 200 μ M 각각의 dATP, dCTP, dGTP, 및 dTTP, 및 1 유닛의 iProof™ High-Fidelity DNA Polymerase (Bio-Rad Laboratories)를 함유하였다. 4개의 튜브의 두 번째 세트를 반응물 각각이 1.5 μ L의 DMSO의 추가를 포함한다는 것을 제외하고 상기와 같이 설정하였다. PCR을 98 °C에서 30초 동안 1 사이클 후 이어서 각각 98 °C에서 10초 동안, X °C에서 20초 동안, 72 °C에서 45초 동안 32 사이클, 72 °C에서 10분 동안 최

중 연장으로 프로그램된 EPPENDORF® MASTERCYCLER® (Eppendorf Scientific)로 수행하였으며, X = 47.6 °C, 51.8 °C, 57.1 °C, 또는 62.1 °C이다. DMSO가 있고 없는 PCR을 나타난 각각의 어닐링 온도에 대하여 실행하였다. 열 순환 후, 각각의 PCR 반응물 10 µL를 TAE 버퍼 중의 1% 아가로스 겔 전기영동으로 분리하였다. 이 겔의 시각화는 DMSO를 가지고 수행된 4개의 PCR 반응물이, 어닐링 온도에 관계없이, 원하는 약 338 bp 생성물의 가장 높은 수득율을 제공한다는 것을 나타냈다. 이들 4개의 PCR을 합쳤고, 약 338개의 염기쌍 PCR 생성물을 겔에서 잘라냈고 제조사의 지시에 따라 NUCLEOSPIN® Extract II Kit (Macherey-Nagel)를 사용하여 정제한 경우, TAE 버퍼 중의 1% 아가로스 겔 전기영동으로 분리하였다.

[0487] PCR을 약 700 bp PGK1 함유 PCR 생성물을 338 bp PDC 종결자 생성물에 융합하는 단일 증폭 생성물을 생성하기 위해 사용하였다. PCR 반응물 (50 µL)은 107 ng의 PGK1 함유 PCR 생성물, 56 ng의 PDC 종결자 함유 PCR 생성물, 1X Phusion HF 버퍼 (New England Biolabs), 100 pmol 각각의 프라이머 0612358 및 0612361, 200 µM 각각의 dATP, dCTP, dGTP, 및 dTTP, 및 1 유닛의 Phusion™ High-Fidelity DNA Polymerase (New England Biolabs)를 함유하였다. PCR을 98 °C에서 30초 동안 1 사이클 후 이어서 각각 98 °C에서 10초 동안, 56 °C에서 20초 동안, 72 °C에서 2분 45초 동안 32 사이클, 72 °C에서 10분 동안 최종 연장으로 프로그램된 EPPENDORF® MASTERCYCLER® (Eppendorf Scientific)로 수행하였다. 열 순환 후, 약 1020개의 염기쌍 PCR 생성물을 겔에서 잘라냈고 제조사의 지시에 따라 NUCLEOSPIN® Extract II Kit (Macherey-Nagel)를 사용하여 정제한 경우, PCR 반응 생성물을 TAE 버퍼 중의 0.9% 아가로스 겔 전기영동으로 분리하였다.

[0488] 플라스미드 ald5680_right #20을 XbaI 및 NotI로 분해하였고, 37 °C에서 30분 동안 10 유닛 소 위장 포스파타제 (New England Biolabs)를 처리하였고, TAE 버퍼 중의 0.9% 아가로스 겔 전기영동으로 정제하였다. 약 5.6 kbp의 밴드를 겔에서 잘라냈고 제조사의 지시에 따라 NUCLEOSPIN® Extract II Kit (Macherey-Nagel)를 사용하여 정제하였다.

[0489] 상기 정제된 487 bp 및 1020 bp PCR 생성물을 150 ng의 XbaI 및 NotI로 분해된 ald5680_right #20 벡터, 13 ng의 TKL 종결자 PCR 생성물, 28 ng의 PGK1 프로모터-PDC 종결자 PCR 생성물, 1X In-Fusion 반응 버퍼 (Clontech) 및 1 µL의 IN-FUSION™ 효소 (Clontech)로 구성된, 10 µL의 총 반응량에서 IN-FUSION™ Advantage PCR Cloning Kit (Clontech)를 사용하여 분해된 ald5680_right#20 단편에 삽입하였다. 반응물을 37 °C에서 15분 동안, 50 °C에서 15분 동안 배양하였고, 얼음 위에 두었다. 반응물을 40 µL의 TE 버퍼로 희석하였고 2.5 µL를 제조사의 지시에 따라 SoloPack Gold SuperCompetent 세포 (Agilent Technologies)를 형질전환하기 위해 사용하였다. 형질전환체를 2X YT+amp 플레이트에 도말하였고 37 °C에서 하룻밤 동안 배양하였다. 결과의 형질전환체 중 다수를 AccI 분해에 의해 원하는 PCR 생성물의 적절한 삽입을 위해 스크리닝하였다. 원하는 밴드 크기를 수득하는 클론을 DNA 시퀀싱에 의해 올바른 것으로 확인하였고 pMhCt091로 지정하였다.

[0490] 플라스미드 pMhCt091은 아이. 오리엔탈리스 URA3 ORF의 3' 단편, 아이. 오리엔탈리스 TDH3 프로모터에 이어서 원하는 유전자의 추가를 위한 NheI 및 AscI 제한 부위, 아이. 오리엔탈리스 TKL 종결자, 아이. 오리엔탈리스 PGK1 프로모터에 이어서 두 번째 원하는 유전자의 추가를 위한 XbaI 및 PacI 제한 부위, 및 3' ald5680 위치에 상동 재조합을 표적화하는 플랭킹 DNA를 함유하는 빈 오른쪽 아이. 오리엔탈리스 ald5680 표적화 구조물이다.

[0491] 에스. 아베르미틸리스 panD 버전 r5 (SEQ ID NO: 146)을 벡터 1075328_SaPanD_r5에서 GeneArt®로 합성하였다. XbaI 및 PacI 제한 부위를 PCR로 pMhCt091의 5' 클로닝 부위로 클로닝을 위해 원하는 NheI 및 AscI 부위로 변화시켰다. PCR 반응물 (50 µL)은 1 내지 50배 희석된 1075328_SaPanD_r5 미니-프랩 플라스미드 DNA (GeneArt®) 1 µL, 1X iProof™ HF 버퍼 (Bio-Rad Laboratories), 100 pmol 각각의 프라이머 0612378 및 0612379, 200 µM 각각의 dATP, dCTP, dGTP, 및 dTTP, 1.5 µL DMSO 및 1 유닛의 iProof™ High-Fidelity DNA Polymerase (Bio-Rad Laboratories)를 함유하였다. PCR을 98 °C에서 30초 동안 1 사이클 후 이어서 각각 98 °C에서 10초 동안, 59 °C에서 20초 동안, 72 °C에서 30초 동안 34 사이클, 72 °C에서 10분 동안 최종 연장으로 프로그램된 EPPENDORF® MASTERCYCLER® (Eppendorf Scientific)로 수행하였다. 열 순환 후, 약 471 bp PCR 생성물을 겔에서 잘라냈고 제조사의 지시에 따라 NUCLEOSPIN® Extract II Kit (Macherey-Nagel)를 사용하여 정제한 경우, PCR 반응 생성물을 TAE 버퍼 중의 0.9% 아가로스 겔 전기영동으로 분리하였다.

[0492] 플라스미드 pMhCt091 (상기)을 NheI 및 AscI로 분해하였고, 37 °C에서 30분 동안 10 유닛 소 위장 포스파타제 (New England Biolabs)를 처리하였고, TAE 버퍼 중의 0.9% 아가로스 겔 전기영동으로 정제하였다. 약 6.9 kbp의 밴드를 겔에서 잘라냈고 제조사의 지시에 따라 NUCLEOSPIN® Extract II Kit (Macherey-Nagel)를 사용하여 정제하였다. 상기 정제된 panD r5 함유 PCR 생성물을 NheI 및 AscI로 분해된 pMhCt091 벡터 150 ng, panD r5

PCR 생성물 19 ng, 1X In-Fusion 반응 버퍼 (Clontech) 및 1 μ L IN-FUSION™ 효소 (Clontech)로 구성된, 10 μ L의 총 반응량에서 IN-FUSION™ Advantage PCR Cloning Kit (Clontech)를 사용하여 분해된 pMhCt091 단편에 삽입하였다. 반응물을 37 °C에서 15분 동안, 50 °C에서 15분 동안 배양하였고, 얼음 위에 두었다. 반응물을 40 μ L의 TE 버퍼로 희석하였고 2.5 μ L를 제조사의 지시에 따라 SoloPack Gold SuperCompetent 세포 (Agilent Technologies)를 형질전환하기 위해 사용하였다. 형질전환체를 2X YT+amp 플레이트에 도말하였고 37 °C에서 하룻밤 동안 배양하였다. 결과의 형질전환체 중 다수를 SmaI 분해에 의해 원하는 PCR 생성물의 적절한 삽입을 위해 스크리닝하였다. 원하는 밴드 크기를 수득하는 클론을 DNA 시퀀싱에 의해 올바른 것으로 확인하였고 pMhCt093으로 지정하였다.

[0493] 플라스미드 pMhCt093를 XbaI 및 PacI로 분해하였고, 37 °C에서 30분 동안 10 유닛 소 위장 포스파타제 (New England Biolabs)를 처리하였고, TAE 버퍼 중의 0.9% 아가로스 겔 전기영동으로 정제하였고, 약 7.4 kbp 밴드를 겔에서 잘라냈고 제조사의 지시에 따라 NUCLEOSPIN® Extract II Kit (Macherey-Nagel)를 사용하여 정제하였다. 에스. 아베르미틸리스 panD 버전 r2 (SEQ ID NO: 147)을 아이. 오리엔탈리스에서 발현에 코돈-최적화하였고 벡터 pMA-T에서 GeneArt®로 합성하였다. 플라스미드를 XbaI 및 PacI로 분해하였고 결과의 단편을 TAE 버퍼 중의 0.9% 아가로스 겔 전기영동으로 분리하였다. 약 434 bp의 밴드를 겔에서 잘라냈고 제조사의 지시에 따라 NUCLEOSPIN® Extract II Kit (Macherey-Nagel)를 사용하여 정제하였다.

[0494] 상기 정제된 ~434 bp 단편을 1X QUICK 결찰 버퍼 (New England Biolabs), XbaI 및 PacI로 분해된 pMhCt093 벡터 2 μ L, 상기 ~434 bp 단편 2 μ L, 및 1 μ L QUICK T4 DNA 리가제 (New England Biolabs)를 함유하는 결찰 반응물 (20 μ L)에서 XbaI 및 PacI로 분해된 pMhCt093 벡터에 클로닝하였다. 결찰 반응물을 상온에서 5분 동안 배양하였고, 튜브를 얼음 위에 두었다. 이 반응물의 5 μ L를 제조사의 지시에 따라 SoloPack Gold SuperCompetent 세포 (Agilent Technologies)를 형질전환하기 위해 사용하였다. 형질전환체를 2X YT+amp 플레이트에 도말하였고 상온에서 3일 동안 배양하였다. 결과의 형질전환체 중 다수를 XbaI 및 PacI로 분해에 의해 원하는 PCR 생성물의 적절한 삽입을 위해 스크리닝하였다. 분리체 pMhCt094를 이후의 작업을 위해 선택하였다. 오른쪽 구조물에 대한 최종 클로닝 단계는 adh1202 3' 상동 영역을 갖는 pMhCt094에 존재하는 ald5680 3' 상동 영역을 대체하는 것이다. 클로닝을 위해 원하는 추가의 제한 부위 및 플랭킹 DNA와 함께, adh1202 ORF의 서열 3'을 함유하는 PCR 생성물을 프라이머 612472 및 612473으로 증폭하였다. PCR 반응물 (50 μ L)은 1 내지 50배 희석된 pGMEr140 (상기) 미니-프랩 플라스미드 DNA 1 μ L, 1X iProof™ HF 버퍼 (Bio-Rad Laboratories), 100 pmol 각각의 프라이머 0612472 및 0612473, 200 μ M 각각의 dATP, dCTP, dGTP, 및 dTTP, 및 1 유닛의 iProof™ High-Fidelity DNA Polymerase (Bio-Rad Laboratories)를 함유하였다. PCR을 98 °C에서 30초 동안 1 사이클 후 이어서 각각 98 °C에서 10초 동안, 59 °C에서 20초 동안, 72 °C에서 30초 동안 34 사이클, 72 °C에서 10분 동안 최종 연장으로 프로그램된 EPPENDORF® MASTERCYCLER® (Eppendorf Scientific)로 수행하였다. 열 순환 후, 약 620개의 염기쌍 PCR 생성물을 겔에서 잘라냈고 제조사의 지시에 따라 NUCLEOSPIN® Extract II Kit (Macherey-Nagel)를 사용하여 정제한 경우, PCR 반응 생성물을 TAE 버퍼 중의 1% 아가로스 겔 전기영동으로 분리하였다.

[0495] 상기 PCR 생성물에 대한 수령자 벡터를 생성하기 위해, 플라스미드 pMhCt094를 SacII 및 NotI로 분해하였고, 37 °C에서 30분 동안 10 유닛 소 위장 포스파타제 (New England Biolabs)를 처리하였고, TAE 버퍼 중의 0.9% 아가로스 겔 전기영동으로 정제하였고, 약 7.0 kbp 밴드를 겔에서 잘라냈고 제조사의 지시에 따라 NUCLEOSPIN® Extract II Kit (Macherey-Nagel)를 사용하여 정제하였다. PCR 생성물 및 선형 벡터를 SacII 및 NotI로 분해된 pMhCt094 벡터 191 ng, adh1202 3' 상동체 함유 PCR 생성물 36 ng, 1X In-Fusion 반응 버퍼 (Clontech) 1 μ L의 IN-FUSION™ 효소 (Clontech)로 구성된, 10 μ L의 총 반응량에서 IN-FUSION™ Advantage PCR Cloning Kit (Clontech)를 사용하여 합쳤다. 반응물을 37 °C에서 15분 동안, 50 °C에서 15분 동안 배양하였고, 얼음 위에 두었다. 반응물을 40 μ L의 TE 버퍼로 희석하였고 2.5 μ L를 제조사의 지시에 따라 SoloPack Gold SuperCompetent 세포 (Agilent Technologies)를 형질전환하기 위해 사용하였다. 형질전환체를 2X YT+amp 플레이트에 도말하였고 37 °C에서 하룻밤 동안 배양하였다. 결과의 형질전환체 중 다수를 NsiI 및 PvuI로 분해에 의해 원하는 PCR 생성물의 적절한 삽입을 위해 스크리닝하였다. NsiI 및 PvuI. 원하는 밴드 크기를 수득하는 클론을 DNA 시퀀싱에 의해 올바른 것으로 확인하였고 pMhCt096 (도 13)로 지정하였다.

[0496] 플라스미드 pMhCt096은 아이. 오리엔탈리스 URA3 ORF의 3' 단편, 아이. 오리엔탈리스에서 발현에 코돈-최적화된 세 번째 에스. 아베르미틸리스 ADC 유전자의 발현을 구동하는 아이. 오리엔탈리스 TDH3 프로모터 (SEQ ID NO: 146), 아이. 오리엔탈리스 TKL 종결자, 아이. 오리엔탈리스에서 발현에 코돈-최적화된 네 번째 에스. 아베르미틸리스 ADC 유전자의 발현을 구동하는 아이. 오리엔탈리스 PGK1 프로모터 (SEQ ID NO: 147), 아이. 오리엔탈리스 PDC 종결자, 및 3' adh1202 위치에 대한 상동 재조합을 표적화하는 플랭킹 DNA를 함유하는 오른쪽 아이. 오

리엔탈리스 adh1202 표적화 구조물이다.

[0497] 실시예 3A-9: pdc 위치에서 아스파르테이트 1-데카르복실라제 (ADC), β -알라닌 아미노트랜스퍼라제 (BAAT), 및 3-히드록시프로피온산 데히드로게나제 (3-HPDH); 및 adh1202 위치에서 4개의 뉴클레오티드 서열로부터 아스파르테이트 1-데카르복실라제 (ADC)를 발현하는 효모 균주

[0498] 상기 실시예 3A-7 및 3A-8은 adh1202 위치에서 아이. 오리엔탈리스에서 발현에 코돈-최적화된 에스. 아베르미틸리스 ADC 유전자의 4개의 뉴클레오티드 변이의 발현을 표적화하는 왼쪽 및 오른쪽 구조물의 생성을 설명한다. 아이. 오리엔탈리스 CNB1에 형질전환 전에, 10 μ g의 pMhCt095를 pUC19 백본 벡터로부터 원하는 형질전환 DNA를 방출하기 위해 HpaI 및 SacII로 분해하였다. ~5 kbp 발현 카세트를 함유하는 밴드를 겔 전기영동으로 pUC19 백본 DNA로부터 분리하였고, 겔에서 잘라냈고, 제조사의 지시에 따라 NUCLEOSPIN® Extract II Kit (Macherey-Nagel)를 사용하여 정제하였다. 30 μ L의 용출 버퍼를 용출 단계에 사용하였다. 총 10 μ L인 pMhCt095 및 pMhCt096 선형 형질전환 DNA의 등몰의 비율을 균주 yMhCt010 (상기)을 형질전환하기 위해 사용하였다. 형질전환체를 ura 선택 플레이트에서 선택하였고 성장을 위해 37 °C에 두었다. 약 12개의 형질전환체를 다음날 골랐고 단일 콜로니에 대하여 ura 선택 플레이트에 재스트리킹하였고 37 °C에서 하룻밤 동안 키웠고, 단일 콜로니를 골랐다.

[0499] 각각의 초기 형질전환체에 의해 발생한 스트리크를 각각 골랐고 ura 선택 플레이트에 재스트리킹하였다. 37 °C에서 성장의 또 다른 밤 이후, 최종 단일 콜로니를 각각의 스트리크로부터 골랐고 ura 선택 플레이트에 재스트리킹하였고 37 °C에서 하룻밤 동안 키웠다. 단일 콜로니 정제 및 성장의 두 라운드 후에, 여기에 설명된 바와 같이 발생한 원하는 표적화 통합을 확인하기 위해 PCR에 사용되는 게놈 DNA를 제조하였다. 프라이머 0611718은 pMhCt096에 존재하는 PDC 종결 영역에서 결합하는 한편, 프라이머 0611632는 adh1202 위치에서 표적화된 영역의 DNA 3'을 결합하고 안티-센스 방향으로 증폭한다. 이들 프라이머를 가지고 PCR로부터 약 727 bp 밴드의 발생은 adh1202 위치에서 원하는 통합 이벤트의 발생을 나타냈다.

[0500] PCR 반응물 (25 μ L)은 스크리닝되는 균주에 대한 0.5 μ L 게놈 DNA, 1X Crimson Taq™ 반응 버퍼 (New England Biolabs), 25 pmol의 센스 프라이머, 25 pmol의 안티-센스 프라이머, 200 μ M 각각의 dATP, dCTP, dGTP, 및 dTTP, 및 0.625 유닛의 Crimson Taq™ DNA Polymerase (New England Biolabs)를 함유하였다. PCR을 95 °C에서 30초 동안 1 사이클 후 이어서 각각 95 °C에서 30초 동안, 50 °C에서 30초 동안, 및 68 °C에서 2.5분 동안 32 사이클, 68 °C에서 10분 동안 최종 연장으로 프로그램된 EPPENDORF® MASTERCYCLER® (Eppendorf Scientific)로 수행하였다. 열 순환 후, PCR 반응 생성물을 TAE 버퍼 중의 1% 아가로스 겔 전기영동으로 분리하였고 밴드의 크기를 시각화하였고 상기 설명된 바와 같이 설명하였다. 원하는 727 bp 밴드를 제공하는 2개의 독립적으로 분리된 형질전환체를 yMhCt019 또는 95/96 2로 지정하였다 (표 21의 유전자형 참조).

[0501] 표 21: 형질전환체 유전자형

| 균주 | 모체 균주 | 유전자형 |
|----------|----------|--|
| yMhCt019 | yMhCt010 | adh1202 Δ ::(PDC 프로모터-Opt.SaPanD r10, ENO1 프로모터-Opt.SaPanD r1, URA3, TDH3 프로모터-Opt.SaPanD r5, PGK1 프로모터-Opt.SaPanD r2)/ADH1202 pdc Δ ::(PDC 프로모터-Opt.SaPanD ENO1 프로모터-Opt.ScUGA1, URA3-스카, TDH3 프로모터-Opt.ScYMR226C)/pdc Δ ::(PDC 프로모터-Opt.SaPanD, ENO1 프로모터-Opt.ScUGA1, URA3-스카, TDH3 프로모터-Opt.ScYMR226C) ura3-/ura3- |
| 95/96 2 | | |

[0502]

[0503] yMhCt019의 ura-유도체를 여기에 설명된 바와 같이 분리하였다. yMhCt019의 다수의 FOA 저항성 콜로니의 게놈 DNA를 프라이머 0611815 및 0612795로 원하는 루프-아웃 이벤트를 위해 PCR로 스크리닝하였다. 프라이머 0611815는 왼쪽 구조물의 RKI 종결자에서 어닐링하고 URA3 프로모터를 향해 증폭시키는 한편, 프라이머 0612795는 (pMhCt096 또는 내인성 adh1202 위치의) 3' adh1202 상동체 내에서 5' 영역을 향해 어닐링한다. 상기 yMhCt019의 분리에 대해 연장 단계의 길이가 3.5분으로 변화한다는 것을 제외하고 설명된 바와 같이 PCR 반응을 수행하였다. 이들 프라이머로 3.7 kbp 밴드의 발생은 원하는 루프-아웃 이벤트가 발생하였고 URA3 프로모터 스카가 변형된 adh1202 위치에 남아있다는 것을 나타내는 한편, 약 5.1 kbp 밴드는 온전한 URA3 프로모터-URA3 ORF-URA3 종결자-URA3 프로모터 카세트의 존재를 나타낼 것이며, 원하는 재조합 이벤트가 발생하지 않았다는 것을 나타낸다. 원하는 3.7 kbp 밴드를 제공하는 균주를 유지하였고 yMhCt021로 지정하였다.

[0504] adh1202에서 다수의 panD 발현 카세트에 대한 yMhCt021 동형 접합체의 유도체를 분리하기 위해, yMhCt021을 상기 설명된 바와 같이 선형 pMhCt095 및 pMhCt096에 형질전환하였다. 단일 콜로니 정제 및 생장의 두 라운드 후에, 발생된 원하는 표적화 통합을 확인하기 위해 PCR에 사용되는 게놈 DNA를 제조하였다. 잔여 야생형 adh1202 위치에 대하여 pMhCt095 및 pMhCt096 단편의 올바른 표적화를 프라이머 0612891 및 0612893으로 확인하였다. 프라이머 0612891은 pMhCt095의 r1 중단 후에 SaPanD r1의 3' 영역 플러스 PacI 부위의 절반에서 어닐링한다. 프라이머 0612893은 SaPanD r5의 극단의 5' 영역에서 어닐링하고, NheI 부위 및 pMhCt096의 선도를 포함하며, 역상보성 방향으로 증폭시킨다.

[0505] 이들 프라이머로 3.2 kbp 밴드의 발생은 yMhCt021의 잔여 adh1202 야생형 위치에서 pMhCt095 및 pMhCt096를 통한 두 번째 통합 이벤트로부터 예상된 바와 같이 온전한 URA3 프로모터-URA3 ORF-URA3 종결자-URA3 프로모터 카세트의 존재를 나타내는 한편, 약 1.7 kbp 밴드는 다른 adh1202 위치에서 URA3 스카 부위의 존재 (초기 통합 이벤트 및 이후의 URA3 마커 루프-아웃으로부터)를 나타낼 것이다. 상기 yMhCt019의 분리에 대하여 연장 단계의 길이가 3.5분으로 변화한다는 것을 제외하고 설명된 바와 같이 PCR 반응을 수행하였다. 이들 밴드 크기 둘 다를 제공하는 균주를 yMhCt022로 지정하였다 (표 22의 유전자형 참조).

[0506] 표 22: 형질전환체 유전자형

| 균주 | 모체 균주 | 유전자형 |
|----------|----------|--|
| yMhCt022 | yMhCt021 | adh1202Δ:: (PDC 프로모터 -Opt.SaPanD r10, ENO1 프로모터 -Opt.SaPanD r1, URA3, TDH3 프로모터 -Opt.SaPanD r5, PGK1 프로모터 -Opt.SaPanD r2)/adh1202Δ:: (PDC 프로모터 -Opt.SaPanD r10, ENO1 프로모터 -Opt.SaPanD r1, URA3-스카, TDH3 프로모터 -Opt.SaPanD r5, PGK1 프로모터 -Opt.SaPanD r2) pdcΔ:: (PDC 프로모터 -Opt.SaPanD, ENO1 프로모터 -Opt.ScUGA1, URA3-스카, TDH3 프로모터 -Opt.ScYMR226C)/pdcΔ:: (PDC 프로모터 -Opt.SaPanD, ENO1 프로모터 -Opt.ScUGA1, URA3-스카, TDH3 프로모터 -Opt.ScYMR226C) ura3-/ura3- |

[0507]

[0508] 균주를 여기에서와 같이 셰이크 플라스크에서 키웠고 CFE를 제조하였고 아스파르테이트 데카르복실라제 (ADC) 활성에 대하여 검정하였다. 실험 결과는 표 23A에 나타난다.

[0509] 표 23A: 형질전환체 효소 활성 데이터

| 균주 | 과발현된 유전자 | ADC 활성 |
|---------------|--|--------|
| MBin500 (대조군) | N/A | 0.00 |
| yMhCt019 | ADC (SEQ ID NOs: 130, 145, 146, 및 147), gabT (SEQ ID NO: 141), 3-HPDH (SEQ ID NO: 144) | 2.18 |
| 95/96 2 | | 2.52 |

[0510]

[0511] 균주 MBin500 (대조군), yMhCt019, 95/96-2, yMhCt008 및 yMhCt022로부터 제조된 CFE에서 아스파르테이트 1-데카르복실라제 (ADC), 베타-알라닌 아미노트랜스퍼라제 (BAAT) 및 3HP 데히드로게나제 (3-HPDH) 활성을 비교하였다. 이 실험의 결과는 표 23B에 나타난다.

[0512] 표 23B: 형질전환체 효소 활성 데이터

| 균주 | 과발현된 유전자 | 유전자 근원 | ADC 활성 | BAAT 활성 | 3HP DH 활성 |
|------------------|--|--------------------|-----------|------------|--------------|
| MBin500 (대조군) | N/A | N/A | 0.002 | 0.61 | 0.4 |
| yMhCt019 | YMR226c (SEQ ID NO: 144) | 에스. 세레비시아 | 0.789 | 14.53 | 72.0 |
| | ADC (SEQ ID NOs: 130, 145, 146 및 147) | 에스. 아베르미틸리 스 | | | |
| | UGA1 (SEQ ID NO: 141) | 에스. 세레비시아 | | | |
| 95/96-2 | YMR226c (SEQ ID NO: 144) | 에스. 세레비시아 | 0.891 | 20.42 | 73.2 |
| | ADC (SEQ ID NOs: 130, 145, 146 및 147) | 에스. 아베르미틸리 스 | | | |
| | UGA1 (SEQ ID NO: 141) | 에스. 세레비시아 | | | |
| yMhCt008 | YMR226c (SEQ ID NO: 144) | 에스. 세레비시아 | 0.272 | 13.14 | 61.1 |
| | ADC (SEQ ID NO: 130) | 에스. 아베르미틸리 스 | | | |
| | UGA1 (SEQ ID NO: 141) | 에스. 세레비시아 | | | |
| yMhCt022 | YMR226c (SEQ ID NO: 144) | 에스. 세레비시아 | 1.233 | 15.44 | 66.7 |
| | ADC (SEQ ID NOs: 130, 145, 146 및 147) | 에스. 아베르미틸리 스 | | | |
| | UGA1 (SEQ ID NO: 141) | 에스. 세레비시아 | | | |

[0513]

[0514] 균주 yMhCt019 및 95/96 2를 또한 여기에 설명된 바와 같이 SDS-PAGE로 분석하였다. 균주 둘 다는 53 kD, 29 kD, ~14 kD, 및 ~3 kD에서 단백질 밴드를 나타냈다. 53 kD 및 29 kD 단백질 밴드의 크기는 UGA1 및 YMR226c 유전자에 의해 암호화된 단백질의 크기와 각각 일치한다. 14 및 3 kD 단백질 밴드의 결합된 크기는 panD 유전자에 의해 암호화된 번역 후 분할 단백질 (post-translationally cleaved protein)과 일치한다. 53 kD, 29 kD, 14 kD 및 3 kD 단백질은 대조군 균주 MBin500의 SDS-PAGE 분석에서 관찰되지 않았다.

[0515] 여기에 설명된 방법을 사용하여, 균주 MBin500 및 yMhCt019를 3-HP 생산에 대해 바이오리액터에서 평가하였다. 대조군 균주 MBin500는 검출 가능한 3-HP를 생산하지 않았다 (2개의 독립적 발효의 평균). 균주 yMhCt019는 5.23 g/L 3-HP를 생산하였다 (3개의 독립적 발효의 평균).

[0516] 실시예 3A-10: adh1202 위치에서 4개의 뉴클레오타이드 서열로부터 아스파르테이트 1-데카르복실라제 (ADC)를 발현하는 효모 균주

[0517] 추가의 구조물을 adh1202 위치에서 비. 리체니포르미스로부터 대안의 ADC를 암호화하는 뉴클레오타이드 (SEQ ID NO: 139)의 4개의 카피를 포함하도록 설계하였다. 상기 설명된 것들과 유사한 접근으로, 왼쪽 및 오른쪽 구조물을 아이. 오리엔탈리스 adh1202 위치에서 상동 재조합을 허용하도록 설계하였다.

[0518] 왼쪽 단편의 구조물

[0519] 플라스미드 pMhCt095 (상기)를 여기에 설명된 바와 같이 XbaI 및 PacI로 분해하였고 TBE 버퍼 중의 1% 아가로스 겔 전기영동으로 정제하였다. 약 7.3 kbp에서 밴드를 겔에서 잘라냈고 제조사의 지시에 따라 NUCLEOSPIN®

Extract II Kit (Macherey-Nagel)를 사용하여 정제하였다.

- [0520] 바실루스 리체니포르미스 아스파르테이트 데카르복실라제 (ADC) panD 유전자는 아이. 오리엔탈리스에서 발현에 코돈-최적화되었고 (버전 1; SEQ ID NO: 149) 합성으로 플라스미드 1110206으로 구성되었다 (GeneArt®). 플라스미드 1110206을 여기에 설명된 바와 같이 XbaI 및 PacI로 분해하였고 TBE 버퍼 중의 1% 아가로스 겔 전기영동으로 정제하였다. 약 380 bp에서 밴드를 겔에서 잘라냈고 제조사의 지시에 따라 NUCLEOSPIN® Extract II Kit (Macherey-Nagel)를 사용하여 정제하였다.
- [0521] ~380 bp 정제된 단편을 10mM ATP (New England Biolabs), 및 1 μ L T4 리가제 (New England Biolabs)와 함께 분해된 pMhCt095 60.5 ng, 1110206의 380 bp 단편 6.3 ng, 1 μ L 10X 결찰 버퍼로 구성된, 10 μ L의 총 반응량에서 T4 리가제 (New England Biolabs)를 사용하여 7.3 kbp pMhCt095 선형화 벡터에 결찰하였다. 반응물을 상온에서 1.5시간 동안 배양하였고 반응물의 엘리쿼트 3 μ L를 제조사의 지시에 따라 ONE SHOT® TOP10 화학적 컴피턴트 이. 콜리 세포 (Invitrogen)에 형질전환하였다. 형질전환체를 2X YT+amp 플레이트에 도말하였고 37 °C에서 하룻밤 동안 배양하였다. 결과의 형질전환체 중 다수를 XbaI 및 PacI를 사용하는 제한 분해에 의한 적절한 삽입에 대하여 스크리닝하였다. 올바른 분해된 밴드 크기를 수득하는 클론을 pMeJi309로 지정하였다.
- [0522] 플라스미드 pMeJi309를 NheI 및 AscI로 분해하였고 TBE 버퍼 중의 1% 아가로스 겔 전기영동으로 정제하였다. 약 7.3 kbp의 밴드를 겔에서 잘라냈고 제조사의 지시에 따라 NUCLEOSPIN® Extract II Kit (Macherey-Nagel)를 사용하여 정제하였다.
- [0523] 바실루스 리체니포르미스 아스파르테이트 데카르복실라제 panD 유전자는 다시 아이. 오리엔탈리스에서 발현에 코돈-최적화되었고 (버전 2; SEQ ID NO: 148) 합성으로 플라스미드 1110205로 구성되었다 (GeneArt®).
- [0524] PCR을 50 μ L의 최종 부피로 3 μ L 1110205, 25 pM 각각의 프라이머 0612695 및 0612724, 1X pfx 증폭 버퍼 (Invitrogen), 2 mM MgSO₄, 1.25 유닛 Platinum® pfx DNA Polymerase (Invitrogen)를 함유하는 혼합물에서 수행하였다. 증폭 반응물을 95 °C에서 2분 동안 1 사이클, 각각 95 °C에서 1분 동안, 55 °C에서 1분 동안, 및 72 °C에서 1분 동안 25 사이클로 프로그램된 EPPENDORF® MASTERCYCLER® (Eppendorf Scientific)에서 수행하였다.
- [0525] 증폭 반응의 PCR 생성물을 TBE 버퍼 중의 1% 아가로스 겔 전기영동으로 정제하였다. 약 400 bp 밴드를 겔에서 잘라냈고 제조사의 지시에 따라 NUCLEOSPIN® Extract II Kit (Macherey-Nagel)를 사용하여 정제하였다. 정제된 PCR 생성물을 93 ng pMeJi309 벡터 단편, 52 ng 상기 PCR 생성물, 2 μ L 1X IN-FUSION™ 반응 버퍼 (Clontech), 및 1 μ L의 IN-FUSION™ 효소 (Clontech)를 함유하는 반응물에서 IN-FUSION™ Advantage PCR Cloning Kit (Clontech)를 사용하여 상기 분해된 pMeJi309 벡터에 삽입하였다. 반응물을 37 °C에서 15분 동안, 50 °C에서 15분 동안 배양하였고, 얼음 위에 두었다. 반응물의 2.5 μ L 샘플을 제조사의 지시에 따라 ONE SHOT® TOP10 화학적 컴피턴트 이. 콜리 세포 (Invitrogen)에 형질전환하였다. 형질전환체를 2X YT+amp 플레이트에 도말하였고 37 °C에서 하룻밤 동안 배양하였다. 결과의 형질전환체 중 다수를 XbaI 및 PacI를 사용하는 제한 분해에 의해 적절한 삽입에 대해 스크리닝하였다. 올바른 분해된 밴드 크기 DNA를 수득하는 클론을 시퀀싱에 의해 올바른 것으로 확인하였고 pMeJi310-2 (도 14)로 지정하였다.
- [0526] 오른쪽 단편의 구조물
- [0527] 플라스미드 pMhCt096 (상기)를 XbaI 및 PacI로 분해하였고 TBE 버퍼 중의 1% 아가로스 겔 전기영동으로 정제하였다. 약 4.8 kbp에서 밴드를 겔에서 잘라냈고 제조사의 지시에 따라 NUCLEOSPIN® Extract II Kit (Macherey-Nagel)를 사용하여 정제하였다.
- [0528] 바실루스 리체니포르미스 아스파르테이트 데카르복실라제 panD 유전자는 다시 아이. 오리엔탈리스에서 발현에 코돈-최적화되었고 (버전 3; SEQ ID NO: 151) 합성으로 플라스미드 1110208로 구성되었다 (GeneArt®). 플라스미드 1110208을 XbaI 및 PacI로 분해하였고 TBE 버퍼 중의 1% 아가로스 겔 전기영동으로 정제하였고 약 380 bp 밴드를 겔에서 잘라냈고 제조사의 지시에 따라 NUCLEOSPIN® Extract II Kit (Macherey-Nagel)를 사용하여 정제하였다.
- [0529] 상기 380 bp 단편을 10mM ATP, 및 1 μ L T4 리가제로 72.2 ng 분해된 pMhCt096, 6.9 ng 1110208의 380 bp 단편, 1 μ L 10X 결찰 버퍼로 구성된, 10 μ L의 총 반응량에서 T4 리가제 (New England Biolabs)를 사용하여 7.3 kbp pMhCt096 선형화 벡터로 결찰하였다. 반응물을 상온에서 1.5시간 동안 배양하였고 반응물의 3 μ L 엘리쿼트를 제조사의 지시에 따라 ONE SHOT® TOP10 화학적 컴피턴트 이. 콜리 세포 (Invitrogen)에

형질전환하였다. 형질전환체를 2X YT+amp 플레이트에 도말하였고 37 °C에서 하룻밤 동안 배양하였다. 결과의 형질전환체 중 다수를 NheI 및 AscI를 사용하는 제한 분해에 의해 적절한 삽입에 대해 스크리닝하였다. 올바른 분해된 밴드 크기를 수득하는 클론을 pMeJi311로 지정하였다.

[0530] 플라스미드 pMeJi311를 효소 NheI 및 AscI로 분해하였고 TBE 버퍼 중의 1% 아가로스 겔 전기영동으로 정제하였고, 약 7.3 kbp 밴드를 겔에서 잘라냈고 제조사의 지시에 따라 NUCLEOSPIN® Extract II Kit (Macherey-Nagel)를 사용하여 정제하였다.

[0531] 바실루스 리체니포르미스 아스파르테이트 데카르복실라제 panD 유전자는 다시 아이. 오리엔탈리스에서 발현에 코돈-최적화되었고 (버전 4; SEQ ID NO: 150) 합성으로 플라스미드 1110207로 구성되었다 (GeneArt®). PCR을 50 µL의 최종 부피로 3 µL 1110207, 25 pM 각각의 0612698 및 0612725, 1X pfx 증폭 버퍼 (Invitrogen), 2 mM MgSO₄, 1.25 유닛 Platinum® pfx DNA Polymerase (Invitrogen)를 함유하는 혼합물에서 수행하였다. 증폭 반응물을 95 °C에서 2분 동안 1 사이클, 각각 95 °C에서 1분 동안, 55 °C에서 1분 동안, 및 72 °C에서 1분 동안 25 사이클, 및 72 °C에서 3분 동안 1 사이클로 프로그램된 EPPENDORF® MASTERCYCLER® (Eppendorf Scientific)에서 수행하였다.

[0532] 증폭 반응의 PCR 생성물을 TBE 버퍼 중의 1% 아가로스 겔 전기영동으로 정제하였다. 약 400 bp에서 밴드를 겔에서 잘라냈고 제조사의 지시에 따라 NUCLEOSPIN® Extract II Kit (Macherey-Nagel)를 사용하여 정제하였다. 정제된 PCR 생성물을 65.1 ng의 pMeJi311 NheI 내지 AscI 분해된 벡터 단편, 85 ng의 상기 PCR 생성물, 2 µL 1X IN-FUSION™ 반응 버퍼 (Clontech), 및 1 µL의 IN-FUSION™ 효소 (Clontech)를 함유하는 반응물에서 IN-FUSION™ Advantage PCR Cloning Kit (Clontech)를 사용하여 상기 분해된 pMeJi311 벡터에 삽입하였다. 반응물을 37 °C에서 15분 동안, 50 °C에서 15분 동안 배양하였고, 얼음 위에 두었다. 반응물의 2.5 µL 샘플을 제조사의 지시에 따라 ONE SHOT® TOP10 화학적 컴피턴트 이. 콜리 세포 (Invitrogen)에 형질전환하였다. 형질전환체를 2X YT+amp 플레이트에 도말하였고 상온에서 2일 동안 배양하였다. 결과의 형질전환체 중 다수를 NheI 및 AscI를 사용하는 제한 분해에 의해 적절한 삽입에 대해 스크리닝하였다. 올바른 분해된 밴드 크기 DNA를 수득하는 클론을 시퀀싱에 의해 올바른 것으로 확인하였고 pMeJi312-2 (도 15)로 지정하였다.

[0533] 왼쪽 및 오른쪽 단편의 통합

[0534] 여기에 설명된 바와 같이 플라스미드 pMeJi310-2를 HpaI 및 SacII로 분해하였고 플라스미드 pMeJi312-2를 EcoRI 및 SacII로 분해하였다. 이것들을 TBE 버퍼 중의 1% 아가로스 겔 전기영동으로 정제하였고, 2개의 약 5 kbp 밴드를 겔에서 잘라냈고 제조사의 지시에 따라 NUCLEOSPIN® Extract II Kit (Macherey-Nagel)를 사용하여 정제하였다.

[0535] 아이. 오리엔탈리스 CNB1을 분해된 pMeJi310-2 및 pMeJi312-2 DNA와 형질전환하였고 올바른 위치 표적화 및 형질전환을 여기에 설명된 바와 같이 Crimson Taq (New England Biolabs) PCR로 확인하였다. 프라이머 0612794 및 0611245는 약 3.17kbp 밴드를 수득하였다; 프라이머 612479 및 0611632는 약 1.48 kbp 밴드를 수득하였다; 및 프라이머 611248 및 612795는 약 2.3 kbp 밴드를 수득하였다. 발현 카세트의 적절한 통합에 대하여 예상된 밴드를 제공하는 균주를 MeJi409-2로 지정하였다. 균주 MeJi409-2의 ura-유도체를 상기 설명된 바와 같이 얻었다.

[0536] 여기에 설명된 방법을 사용하여, 균주 MBin500 및 MeJi409-2를 3-HP 생산에 대하여 발효 바이오리액터에서 평가하였다. 대조군 균주 MBin500는 검출 가능한 3-HP를 생산하지 않았다 (2개의 독립적 발효의 평균). 균주 MeJi409-2 (하나의 발효)는 4.62 g/L를 생산하였다. 다른 (예를 들어, 이후의) 발효와 비교하여 이들 발효에서 세포 질량의 양의 차이를 설명하기 위하여, 세포 질량의 유닛 당 3-HP 농도 ([g/L 3-HP]/[g/L 건조 세포 중량]로 표현됨)를 MeJi409-2에 대하여 0.20인 것으로 계산하였다.

[0537] 실시예 3A-11: adh9091 위치에서 아스파르테이트 1-데카르복실라제 (ADC) 및 아스파르테이트 아미노트랜스퍼라제 (AAT)를 발현하는 효모 균주

[0538] SEQ ID NO: 14의 아이. 오리엔탈리스 아스파르테이트 아미노트랜스퍼라제 (AAT)를 암호화하는 뉴클레오티드 서열 (SEQ ID NO: 13)을 프라이머 0611268 및 0611269를 사용하여 아이. 오리엔탈리스 게놈 DNA로부터 PCR 증폭시켰다. PCR 반응물 (50 µL)은 50 ng의 균주 아이. 오리엔탈리스 게놈 DNA, 1X Phusion HF 버퍼 (New England Biolabs), 50 pmol 각각의 프라이머 0611268 및 0611269, 200 µM 각각의 dATP, dCTP, dGTP, 및 dTTP, 1.5 µL의 100% DMSO (New England Biolabs) 및 1 유닛의 Phusion High-Fidelity DNA Polymerase (New England Biolabs)를 함유하였다. PCR을 98 °C에서 2분 동안 1 사이클 후 이어서 각각 98 °C에서 30초 동안, 55 °C에서

30초 동안, 및 72 °C에서 1분 30초 동안 32 사이클, 72 °C에서 7분 동안 최종 연장으로 프로그램된 EPPENDORF® MASTERCYCLER® (Eppendorf Scientific)로 수행하였다. 열 순환 후, 약 1278 bp PCR 생성물을 겔에서 잘라냈고 제조사의 지시에 따라 QIAQUICK® Gel Extraction Kit (Qiagen)를 사용하여 정제한 경우, PCR 반응 생성물을 TBE 버퍼 중의 0.8% 아가로스 겔 전기영동으로 분리하였다. 단편의 5' 말단에 NruI 제한 부위 및 그것의 3' 말단에 PacI 제한 부위를 갖는, 결과의 PCR 단편의 총 길이는 약 1278 bp이다.

[0539] AAT 유전자 CDS (SEQ ID NO: 13)를 포함하는 상기 결과의 1278 bp 단편을 pCR2.1-TOPO 벡터에 클로닝하였고 제조사의 지시에 따라 One-Shot TOP10 이. 콜리 세포 (Invitrogen)에 형질전환하였다. 형질전환체를 2X YT + amp 플레이트에 도말하였고 37 °C에서 하룻밤 동안 배양하였다. 결과의 형질전환체 중 다수를 Bam HI 분해에 의해 원하는 단편의 적절한 삽입에 대해 스크리닝하였다. 원하는 밴드 크기를 수득하는 클론을 시퀀싱으로 확인하였고 pGMEr111로 지정하였다.

[0540] 플라스미드 pGMEr121 및 pGMEr111을 제한 효소 PacI 및 NruI로 이중-분해하였다. 플라스미드 pGMEr121로부터, 결과의 7695 bp 벡터 단편 및 플라스미드 pGMEr111로부터, AAT 암호화 서열을 포함하는 1272 bp 인서트 단편을 1X TBE 버퍼 중의 0.8% 아가로스 겔 전기영동으로 분리하였고, 겔에서 잘라냈고, 제조사의 지시에 따라 QIAQUICK® Gel Extraction Kit (Qiagen)를 사용하여 정제하였다.

[0541] 결찰 반응을 3 µL의 벡터 단편, 4 µL의 인서트 단편, 2 µL의 멸균 2차 증류수, 10 µL의 2X QUICK 리가제 버퍼 및 1 µL의 QUICK T4 리가제 (QUICK Ligation Kit, New England Biolabs)로 설정하였고 제조사의 지시에 따라 수행하였다. 상기 결찰 반응물의 5 µL 앨리퀀트를 제조사의 지시에 따라 XL10-Gold® Ultracompetent 이. 콜리 세포 (Agilent Technologies)에 형질전환하였다. 형질전환체를 2X YT+amp 플레이트에 도말하였고 37 °C에서 하룻밤 동안 배양하였다. 결과의 형질전환체 중 다수를 SmaI/PpuMI 이중 분해에 의해 원하는 인서트의 적절한 삽입에 대해 스크리닝하였다. 원하는 밴드 크기를 수득하는 클론을 시퀀싱으로 확인하였고 pGMEr126 (도 16)으로 지정하였다.

[0542] 플라스미드 pGMEr126는 아이. 오리엔탈리스 AAT 발현 카세트를 포함하며, 여기에서 유전자 전사는 아이. 오리엔탈리스 TDH3 프로모터 및 TKL 종결자에 의해 조절되고, URA3 암호화 서열 및 URA3 프로모터의 절단된 3' 영역, 업스트림에 의해, 및 아이. 오리엔탈리스 adh9091 위치를 갖는 3' 상동 영역, 다운스트림에 의해 플랭크된다.

[0543] 아이. 오리엔탈리스에서 발현에 코돈-최적화된 에스. 아베르미틸리스 panD 유전자 (SEQ ID NO: 130)를 프라이머 061166 및 0611662를 사용하여 GeneArt®에서 받은 pMA-T 벡터로부터 PCR 증폭하였다. PCR 반응물 (50 µL)은 50 ng의 균주 플라스미드 DNA, 1X Phusion HF 버퍼 (New England Biolabs), 50 pmol 각각의 프라이머 0611661 및 0611662, 200 µM 각각의 dATP, dCTP, dGTP, 및 dTTP, 1.5 µL의 100% DMSO (New England Biolabs) 및 1 유닛의 Phusion High-Fidelity DNA Polymerase (New England Biolabs)를 함유하였다. PCR을 98 °C에서 2분 동안 1 사이클 후 이어서 각각 98 °C에서 30초 동안, 59 °C에서 30초 동안, 72 °C에서 30초 동안 35 사이클, 72 °C에서 7분 동안 최종 연장으로 프로그램된 EPPENDORF® MASTERCYCLER® (Eppendorf Scientific)로 수행하였다. 열 순환 후, 약 453 bp PCR 생성물을 겔에서 잘라냈고 제조사의 지시에 따라 QIAQUICK® Gel Extraction Kit (Qiagen)를 사용하여 정제한 경우, PCR 반응 생성물을 TBE 버퍼 중의 0.8% 아가로스 겔 전기영동으로 분리하였다. 단편의 5' 말단에 NruI 제한 부위 및 3' 말단에 ApaI 제한 부위를 갖는 결과의 PCR 단편의 총 길이는 약 453 bp였다.

[0544] 에스. 아베르미틸리스 panD 유전자의 코돈-최적화된 버전을 포함하는, 결과의 453 bp 단편을 pCR2.1-TOPO 벡터에 클로닝하였고 제조사의 지시에 따라 One-Shot TOP10 이. 콜리 세포 (Invitrogen)에 형질전환하였다. 형질전환체를 2X YT + amp 플레이트에 도말하였고 37 °C에서 하룻밤 동안 배양하였다. 결과의 형질전환체 중 다수를 EcoRI 분해에 의해 원하는 단편의 적절한 삽입에 대해 스크리닝하였다. 원하는 밴드 크기를 수득하는 플라스미드를 시퀀싱으로 확인하였고 pGMEr127로 지정하였다.

[0545] 플라스미드 pGMEr127 및 pGMEr125(a)를 제한 효소 NruI 및 Apa I로 분해하였다. 분해 반응을 중단하기 전에, 끝을 탈-인산화하고 자가-결찰을 방지하기 위해 1 µL의 소 위장 알칼린 포스파타제 (New England Biolabs)를 pGMEr125(a) 분해물에 추가하였다. 결과의 플라스미드 pGMEr125(a) (상기)의 8818 bp 벡터 단편 및 플라스미드 pGMEr127 (상기)의 에스. 아베르미틸리스 panD 유전자 (SEQ ID NO: 130)의 코돈-최적화된 버전을 포함하는, 440 bp 인서트 단편을 1X TBE 버퍼 중의 0.8% 아가로스 겔 전기영동으로 분리하였고, 겔에서 잘라냈고, 제조사의 지시에 따라 QIAQUICK® Gel Extraction Kit (Qiagen)를 사용하여 정제하였다.

[0546] 결찰 반응을 4 µL의 벡터 단편, 4 µL의 인서트 단편, 9 µL의 2X QUICK 리가제 버퍼 및 1 µL의 QUICK T4 리

가제 (New England Biolabs)로 설정하였고 제조사의 지시에 따라 수행하였다. 상기 결찰 반응물의 5 μ L 엘리퀴트를 제조사의 지시에 따라 XL10-Gold® Ultracompetent 이. 콜리 세포 (Stratagene)에 형질전환하였다. 형질전환체를 2X YT+amp 플레이트에 도말하였고 37 °C에서 하룻밤 동안 배양하였다. 결과의 형질전환체 중 다수를 BamHI 분해에 의해 원하는 인서트의 적절한 삽입을 위해 스크리닝하였다. 원하는 밴드 크기를 수득하는 클론을 확인하였고 pGMEr130 (도 17)으로 지정하였다.

[0547] 플라스미드 pGMEr130은 다음 단편으로 구성된 구조물을 포함한다: 아이. 오리엔탈리스 adh9091 위치의 5' 플랭크, 아이. 오리엔탈리스 PDC 프로모터/TAL 종결자를 갖는 빈 발현 카세트, 아이. 오리엔탈리스 ENO1 프로모터 및 RKI 종결자에 의해 조절되는 panD 발현 카세트 (SEQ ID NO: 130 함유), 및 아이. 오리엔탈리스 URA3 프로모터에 의해 조절되는 URA3 마커 유전자의 절단된 5' 단편.

[0548] 효모 균주 아이. 오리엔탈리스 CNB1가 에스. 아베르미틸리스 panD 유전자 (SEQ ID NO: 130) 및 아이. 오리엔탈리스 AAT 유전자 (SEQ ID NO: 13)의 아이. 오리엔탈리스 코돈-최적화된 버전을 발현할 수 있는지 결정하기 위해, 발현 플라스미드 pGMEr130 및 pGMEr126을 구성하였다. 플라스미드 pGMEr130은 (5'-3'에서) 아이. 오리엔탈리스 adh9091 위치에서 구조물의 게놈 통합을 위한 5' 플랭킹 영역, ENO1 프로모터 및 RKI 종결자에 의해 조절되는 panD 발현 카세트, 및 URA3 프로모터에 의해 구동되는 URA3 선택 마커의 절단된 5' 부분을 포함한다. 플라스미드 pGMEr126은 (5'-3'에서) URA3 선택 마커의 3' 부분, TDH3 프로모터 및 TKL 종결자에 의해 조절되는 AAT 유전자 발현 카세트, 및 아이. 오리엔탈리스 adh9091 위치에서 구조물의 게놈 통합을 위한 3' 플랭크를 포함한다. 모든 프로모터 및 종결자를 아이. 오리엔탈리스로부터 유도하였다.

[0549] 플라스미드 pGMEr126을 제한 효소 EcoRI로 분해하였으며, 이것은 원하는 4758 bp 단편을 자르는 한편, 플라스미드 pGMEr130을 형질전환에 필요한 5034 bp 단편을 생성하는 제한 효소 HindIII로 분해하였다. 4758 bp 및 5034 bp 단편을 1X TBE 버퍼 중의 0.8% 아가로스 겔 전기영동으로 분리하였고 겔에서 잘라냈고, 및 제조사의 지시에 따라 QIAQUICK® Gel Extraction Kit (Qiagen)를 사용하여 정제하였다.

[0550] 아이. 오리엔탈리스 CNB1을 배양하였고 여기에 설명된 바와 같이 4758 bp 및 5034 bp 선형 단편 둘 다의 약 500 ng로 동시-형질전환하였다. 8개의 형질전환체 균주를 얻었고 셰이크 플라스크에서 배양하였다. 결과의 브로스를 SDS-PAGE를 수행하기 위해 사용하였고,

[0551] 아이. 오리엔탈리스 (SEQ ID NO: 130)에서 발현에 코돈-최적화된 에스. 아베르미틸리스 panD 유전자 및 아이. 오리엔탈리스 AAT 유전자 (SEQ ID NO: 130)의 발현을 검출하기 위해 트리스-HCl (Bio-Rad Laboratories) 겔을 사용하였다. 양성 균주를 yGMEr008로 지정하였고 그것의 브로스를 또한 상기 설명된 바와 같이 ADC 및 AAT 활성 수준을 결정하기 위해 사용하였다.

[0552] 실시예 3A-12: adh9091 위치에서 피루베이트 카르복실라제 (PYC), 아스파르테이트 1-데카르복실라제 (ADC) 및 아스파르테이트 아미노트랜스퍼라제 (AAT)를 발현하는 효모 균주

[0553] SEQ ID NO: 2의 아이. 오리엔탈리스 피루베이트 카르복실라제 (PYC)를 암호화하는 뉴클레오티드 서열을 프라이머 0611266 및 0611267을 사용하여 아이. 오리엔탈리스 게놈 DNA로부터 PCR 증폭하였다. PCR 반응물 (50 μ L)은 50 ng의 아이. 오리엔탈리스 게놈 DNA, 1X Phusion HF 버퍼 (New England Biolabs), 50 pmol 각각의 프라이머 0611266 및 0611267, 200 μ M 각각의 dATP, dCTP, dGTP, 및 dTTP, 1.5 μ L의 100% DMSO (New England Biolabs) 및 1 유닛의 Phusion High-Fidelity DNA Polymerase (New England Biolabs)를 함유하였다. PCR을 98 °C에서 2분 동안 1 사이클 후 이어서 각각 98 °C에서 30초 동안, 59 °C에서 30초 동안, 72 °C에서 1분 동안 35 사이클, 72 °C에서 7분 동안 최종 연장으로 프로그램된 EPPENDORF® MASTERCYCLER® (Eppendorf Scientific)로 수행하였다. 열 순환 후, 약 3557 bp PCR 생성물을 겔에서 잘라냈고 제조사의 지시에 따라 QIAQUICK® Gel Extraction Kit (Qiagen)를 사용하여 정제한 경우, PCR 반응 생성물을 TBE 버퍼 중의 0.8% 아가로스 겔 전기영동으로 분리하였다. 결과의 PCR 단편은 단편의 5' 말단에 XbaI 제한 부위 및 그것의 3' 말단에 PacI 제한 부위를 갖는다.

[0554] 아이. 오리엔탈리스 PYC 유전자 CDS (SEQ ID NO: 1)를 포함하는, 결과의 3557 bp 단편을 pCR2.1-TOPO 벡터에 클로닝하였고 제조사의 지시에 따라 One-Shot TOP10 이. 콜리 세포 (Invitrogen)에 형질전환하였다. 형질전환체를 2X YT + amp 플레이트에 도말하였고 37 °C에서 하룻밤 동안 배양하였다. 결과의 형질전환체 중 다수를 EcoRI 분해에 의해 원하는 단편의 적절한 삽입에 대해 스크리닝하였다. 원하는 밴드 크기를 수득하는 6개의 클론을 확인하였고 pGMEr132.7, pGMEr132.14, pGMEr132.16, pGMEr132.25, pGMEr132.27 및 pGMEr132.30으로 지정하였다. 시퀀싱 분석은 플라스미드 pGMEr132.14가 적절한 PYC CDS를 갖지만 CDS 5' 말단에서 XbaI 제한 부위가 없어진다

는 것을 나타냈다. 이 제한 부위가 발현 플라스미드 pGMEr125 (변형된 XbaI 부위를 갖는 PYC CDS 상에서 5' 말단을 포함)의 PYC CDS를 삽입하기 위해 필요하기 때문에, 플라스미드 pGMEr132.14의 315 bp HindIII 단편을 플라스미드 pGMEr132.7의 315 bp HindIII 단편으로 대체하였으며, 이것은 올바른 XbaI 부위를 포함하는 PYC CDS의, 바뀌지 않은 5' 말단을 갖는다. 플라스미드 pGMEr132.14의, 결과의 7173 bp HindIII 벡터 단편, 및 플라스미드 pGMEr132.7의 315 bp HindIII 인서트 단편을 1X TBE 버퍼 중의 0.8% 아가로스 겔 전기영동으로 분리하였고, 겔에서 잘라냈고, 제조사의 지시에 따라 QIAQUICK® Gel Extraction Kit (Qiagen)를 사용하여 정제하였다.

[0555] 결찰 반응을 4 μ L의 벡터 단편, 5 μ L의 인서트 단편, 10 μ L의 2X QUICK 리가제 버퍼 및 1 μ L의 QUICK T4 리가제 (New England Biolabs)로 설정하였고 제조사의 지시에 따라 수행하였다. 상기 결찰 반응물의 5 μ L 엘리퀀트를 제조사의 지시에 따라 One-Shot TOP10 이. 콜리 세포 (Invitrogen)에 형질전환하였다. 형질전환체를 2X YT+amp 플레이트에 도말하였고 37 °C에서 하룻밤 동안 배양하였다. 결과의 형질전환체 중 다수를 BamHI 및 XbaI 이중 분해에 의해 원하는 인서트의 적절한 삽입에 대해 스크리닝하였다. 원하는 밴드 크기를 수득하는 클론을 확인하였고 pGMEr133으로 지정하였다.

[0556] PDC 프로모터의 아이. 오리엔탈리스 PYC CDS 다운스트림을 삽입하기 위해, 플라스미드 pGMEr125(b), 플라스미드 pGMEr125(b) 및 pGMEr133 (상기)를 PacI 및 XbaI로 분해하였다. 플라스미드 pGMEr125(b)의 결과의 8188 bp 벡터 단편, 및 플라스미드 pGMEr133의 아이. 오리엔탈리스 PYC CDS (SEQ ID NO: 1)를 포함하는, 3553 bp 인서트 단편을 1X TBE 버퍼 중의 0.8% 아가로스 겔 전기영동으로 분리하였고, 겔에서 잘라냈고, 제조사의 지시에 따라 QIAQUICK® Gel Extraction Kit (Qiagen)를 사용하여 정제하였다.

[0557] 결찰 반응을 3 μ L의 벡터 단편, 6 μ L의 인서트 단편, 10 μ L의 2X QUICK 리가제 버퍼 및 1 μ L의 QUICK T4 리가제 (New England Biolabs)로 설정하였고 제조사의 지시에 따라 수행하였다. 상기 결찰 반응물의 5 μ L 엘리퀀트를 제조사의 지시에 따라 One-Shot TOP10 이. 콜리 세포 (Invitrogen)에 형질전환하였다. 형질전환체를 2X YT+amp 플레이트에 도말하였고 37 °C에서 하룻밤 동안 배양하였다. 결과의 형질전환체 중 다수를 BamHI 분해에 의해 원하는 단편의 적절한 삽입에 대해 스크리닝하였다. 원하는 밴드 크기를 수득하는 클론을 확인하였고 pGMEr136으로 지정하였다.

[0558] 플라스미드 pGMEr136은 아이. 오리엔탈리스 adh9091 위치의 5' 플랭크, 아이. 오리엔탈리스 PDC 프로모터 및 TAL 종결자에 의해 조절되는 아이. 오리엔탈리스 PYC 유전자 발현 카세트 (SEQ ID NO: 1), 아이. 오리엔탈리스 ENO1 프로모터/RKI 종결자를 갖는 빈 발현 카세트, 및 URA3 프로모터에 의해 조절되는 아이. 오리엔탈리스 URA3 마커 유전자의 절단된 5' 단편을 포함한다.

[0559] 약 5 μ g의 플라스미드 pGMEr136 (상기) 및 4 μ g의 플라스미드 pGMEr127을 제한 효소 ApaI 및 NruI로 분해하였다. 플라스미드 pGMEr136의 결과의 11729 bp 벡터 단편, 및 pGMEr127의 아이. 오리엔탈리스 (SEQ ID NO: 130)에서 발현에 코돈-최적화된 에스. 아베르미틸리스 panD 유전자를 포함하는, 결과의 인서트 단편을 제조사의 지시에 따라 NucleoSpin® Extract II (Macherey-Nagel)를 사용하여 1X TBE 버퍼 중의 0.8% 아가로스 겔 전기영동으로 정제하였다.

[0560] 5 μ L의 벡터 단편, 4 μ L의 인서트 단편, 9 μ L의 2X QUICK 리가제 버퍼 및 1 μ L의 QUICK T4 리가제 (New England Biolabs)를 포함하는 결찰 반응을 설정하였다. 반응물을 상온에서 1시간 동안 배양하였다. 상기 결찰 반응물의 5 μ L을 제조사의 지시에 따라 ONE SHOT® TOP10 화학적 컴피턴트 이. 콜리 세포 (Invitrogen)에 형질전환하였다. 형질전환체를 2X YT+amp 플레이트에 도말하였고 37 °C에서 하룻밤 동안 배양하였다. 결과의 형질전환체 중 다수를 BamHI 분해에 의해 원하는 인서트에 대해 스크리닝하였다. 원하는 밴드 크기를 수득하는 클론을 선택하였고 pGMEr137 (도 18)로 지정하였다.

[0561] 플라스미드 pGMEr137은 아이. 오리엔탈리스 PDC 프로모터 및 TAL 종결자에 의해 전사 조절되는 아이. 오리엔탈리스 PYC 유전자 (SEQ ID NO: 1), 아이. 오리엔탈리스 ENO1 프로모터 및 RKI 종결자에 의해 전사 조절되는 아이. 오리엔탈리스 (SEQ ID NO: 130)에서 발현에 코돈-최적화된 에스. 아베르미틸리스 panD 유전자, URA3 프로모터에 이어서 URA3 마커의 5' 말단 및 아이. 오리엔탈리스 adh9091 위치의 5' 플랭크 영역을 포함한다.

[0562] 플라스미드 pGMEr126 (상기)를 제한 효소 EcoRI으로 분해하였으며, 이것은 원하는 4758 bp 단편을 잘랐다; 한편 플라스미드 pGMEr137 (상기)를 8400 bp 단편을 생성하는 제한 효소 HpaI 및 NheI로 분해하였다. 4758 bp 및 8400 bp 단편 둘 다를 1X TBE 버퍼 중의 0.8% 아가로스 겔 전기영동으로 분리하였다; 밴드를 겔에서 잘라냈고 제조사의 지시에 따라 NucleoSpin® Extract II (Macherey-Nagel)를 사용하여 정제하였다. 아이. 오리엔탈리스 CNB1을 배양하였고 여기에 설명된 바와 같이 4758 bp 및 8400 bp 선형 단편 약 500 ng에 형질전환하였으며, 형

질 전환체 yGMEr009를 발생시킨다.

[0563] 실시예 3A-13: pdc 위치에서 아스파르테이트 1-테카르복실라제 (ADC), β -알라닌 아미노트랜스퍼라제 (BAAT), 및 3-히드록시프로피온산 데히드로게나제 (3-HPDH)를 발현하고, adh9091 위치에서 피루베이트 카르복실라제 (PYC), 아스파르테이트 1-테카르복실라제 (ADC), 및 아스파르테이트 아미노트랜스퍼라제 (AAT)를 발현하는 효모 균주

[0564] 이미 아스파르테이트 아미노트랜스퍼라제 (AAT), β -알라닌 아미노트랜스퍼라제 (BAAT), 및 3-HP 데히드로게나제 (3-HPDH)를 과발현하는 균주에서 피루베이트 카르복실라제 (PYC) 및 아스파르테이트 아미노트랜스퍼라제 (AAT) 활성을 증가시키기 위해, 균주 yMhCt010 (상기)을 상기 설명된 바와 같이 pGMEr137 (상기) 및 pGMEr126 (상기)의 선형 단편에 형질전환하였다. 단일 콜로니 정제 및 생장의 두 라운드 후에, 상기 설명된 바와 같이 발생한 원하는 표적화 통합을 확인하기 위해 PCR에 사용되는 게놈 DNA를 제조하였다. pGMEr137 및 pGMEr126 단편의 adh9091 위치로 올바른 표적화를 프라이머 0611814 및 0612055를 사용하여 확인하였다. 프라이머 0611814는 pGMEr126의 TDH3 프로모터의 3' 말단에서 어닐링하고 3' 방향으로 증폭시킨다. 프라이머 0612055는 pGMEr126에 존재하는 adh9091 3' 플랭킹 상동체의 3'을 어닐링하고, 통합 DNA가 상동 재조합을 통해 올바른 위치를 표적화한 경우, 이 프라이머 쌍으로 PCR 생성물의 증폭만이 발생할 것이다. 프라이머 0611814 및 0612055를 함유하는 PCR의 약 3066 bp 밴드의 존재는 adh9091 위치에서 발생한 pGMEr126 및 pGMEr137 단편의 원하는 통합을 나타낸다.

[0565] pGMEr137 및 pGMEr126의 선형 단편을 갖는 yMhCt010의 다수의 독립적 형질전환체의 단일 콜로니 정제 및 생장의 두 라운드 후에, 상기 설명된 바와 같이 발생한 원하는 표적화 통합을 확인하기 위해 PCR에 사용되는 게놈 DNA를 제조하였다. 프라이머 0611814 및 0612055를 함유하는 PCR의 약 3066 bp 밴드를 제공하는 3개의 독립적으로 분리된 균주를 yMhCt020, GMErin010#2, 및 GMErin010#3으로 지정하였다. 이들 균주는 pdc 위치 둘 다 및 adh9091 위치 중 하나에서 해당하는 ADC (SEQ ID NO: 17)를 암호화하는 폴리뉴클레오티드 (SEQ ID NO: 130); pdc 위치 둘 다에서 해당하는 gabT (SEQ ID NO: 24)를 암호화하는 폴리뉴클레오티드 (SEQ ID NO: 141); pdc 위치 둘 다에서 해당하는 3-HPDH (SEQ ID NO: 129)를 암호화하는 폴리뉴클레오티드 (SEQ ID NO: 144); adh9091 위치 중 하나에서 해당하는 PYC (SEQ ID NO: 2)를 암호화하는 폴리뉴클레오티드 (SEQ ID NO: 1); 및 adh9091 위치 중 하나에서 AAT (SEQ ID NO: 14)를 암호화하는 폴리뉴클레오티드를 함유한다 (표 24 참조).

[0566] 표 24: 형질전환체 유전자형

| 균주 | 모체 균주 | 유전자형 |
|--------------|----------|---|
| yMhCt020 | yMhCt010 | adh9091 Δ ::(PDC 프로모터- <i>pycCNB1</i> , ENO1 프로모터-SaPanD(역방향), URA3, TDH3 프로모터- <i>aat</i>)/ADH9091 pdc Δ ::(PDC 프로모터-Opt.SaPanD, ENO1 프로모터-Opt.ScUGA1, URA3-스카, TDH3 프로모터-Opt.ScYMR226C)/pdc Δ ::(PDC 프로모터-Opt.SaPanD, ENO1 프로모터-Opt.ScUGA1, URA3-스카, TDH3 프로모터-Opt.ScYMR226C) <i>ura3-ura3</i> - |
| GMErin010 #2 | | |
| GMErin010 #3 | | |

[0567] 균주 yMhCt020, GMErin010#2, 및 GMErin010#3를 셰이크 플라스크에서 키웠고 CFE를 제조하였고 PYC, 여기에 사용된 바와 같이 AAT 및 3-HPDH 활성에 대해 검정하였다. 실험 결과는 표 25에 나타난다.

[0569] 표 25: 형질전환체 효소 활성 데이터

| 균주 | 과발현된 유전자 | PYC 활성 | AAT 활성 | ADC 활성 | 3-HPDH 활성 |
|------------------|---|-----------|-----------|-----------|--------------|
| MBin500 (대조군) | N/A | 0.14 | 0.03 | 0.00 | 0.14 |
| yMhCt008 | ADC (SEQ ID NO: 130), gabT (SEQ ID NO: 141), 3-HPDH (SEQ ID NO: 144) | 0.22 | 2.05 | 0.64 | 2.95 |
| GMEr009-2 | ADC (SEQ ID NO: 130), PYC (SEQ ID NO: 1), AAT (SEQ ID NO: 13) | 1.15 | 24.62 | 0.05 | 0.07 |
| yMhCt020 | ADC (SEQ ID NO: 130), gabT (SEQ ID NO: 141), 3-HPDH (SEQ ID NO: 144), PYC (SEQ ID NO: 1), AAT (SEQ ID NO: 13) | 2.74 | 35.02 | 0.56 | 1.68 |
| GMErin010 #2 | | 1.97 | 40.32 | 0.59 | 1.87 |
| GMErin010 #3 | | 1.50 | 23.79 | 0.42 | 3.30 |

- [0571] 표 25의 균주들을 또한 여기에 설명된 바와 같이 SDS-PAGE로 분석하였다. MBin500 및 GMEr009-2는 단백질 밴드가 4개의 다른 샘플에 없는 64 kD인 것을 나타냈다. 이들 단백질의 질량은 아이. 오리엔탈리스 CNB1의 네이티브 피루베이트 데카르복실라제로서 그것들의 정체성과 일치한다. 균주 yMhCt008, yMhCt020, GMErin010 #2, 및 GMErin010#3 53 kD 및 29 kD에서 밴드를 나타냈다. 이들 단백질의 질량은 UGA1 및 YMR226c 유전자에 의해 암호화된 단백질의 질량과 각각 일치한다. 균주 GMEr009-2, yMhCt020, GMErin010 #2, 및 GMErin010 #3 모두는 46.3 kD에서 밴드를 나타냈다. 이들 단백질의 질량은 ATT 유전자에 의해 암호화된 단백질의 질량과 일치한다.
- [0572] 실시예 3A-14: adh1202 위치 둘 다에서 4개의 뉴클레오티드 서열의 아스파르테이트 1-데카르복실라제 (ADC)를 발현하는 효모 균주
- [0573] SEQ ID NO: 139 (상기)의 비. 리체니포르미스 ADC를 암호화하는 뉴클레오티드의 4개의 카피를 함유하는 MeJi409-2의 ura3-유도체를 상기 설명된 FOA 카운터-선택 루프-아웃 프로토콜을 사용하여 분리하였다. 모체 균주 MeJi409-2의 다수의 FOA 저항성 콜로니의 게놈 DNA를 프라이머 0611815 및 0611817을 갖는 원하는 루프-아웃 이벤트에 대해 스크리닝하였다. 프라이머 0611815는 왼쪽 구조물의 RKI 종결자에서 어닐링하였고 ura3 프로모터를 향해 증폭시켰다. 프라이머 0611817은 TDH3 프로모터에서 어닐링하였고 ura3 카세트를 향해 증폭시켰다. 828 bp 밴드의 존재는 ura3 스카 부위 (모체 균주에서 2개의 URA3 프로모터 사이에서 상동 재조합 후 단일 URA3 프로모터 왼쪽 뒤)만의 존재를 나타내는 한편, 2.2 kbp의 밴드는 온전한 URA3 프로모터-URA3 ORF-URA3 종결자-URA3 프로모터 카세트의 존재를 나타내며, 원하는 재조합 이벤트가 발생하지 않았다는 것을 나타낸다. PCR 반응을 상기 설명된 바와 같이 Crimson Taq™ DNA Polymerase (New England Biolabs)로 수행하였다. 모체 균주 MeJi409-2의 하나의 FOA 저항성 콜로니를 MeJi411로 지정하였고, 원하는 828 bp 밴드를 제공하였다.
- [0574] 왼쪽 및 오른쪽 단편의 통합
- [0575] 여기에 설명된 바와 같이 플라스미드 pMeJi310-2를 HpaI 및 SacII로 분해하였고 및 플라스미드 pMeJi312-2를 EcoRI 및 SacII로 분해하였다. 이것들을 TBE 버퍼 중의 1% 아가로스 겔 전기영동으로 정제하였고, 2개의 약 5 kbp 밴드를 겔에서 잘라냈고 제조사의 지시에 따라 NUCLEOSPIN® Extract II Kit (Macherey-Nagel)를 사용하여 정제하였다.
- [0576] MeJi411를 분해된 pMeJi310-2 및 pMeJi312-2 DNA 및 올바른 위치 표적화에 형질전환하였고 형질전환체를 여기에 설명된 바와 같이 Crimson Taq (New England Biolabs)로 확인하였다. 프라이머 0611225 및 0611632는 약 5 kbp 밴드를 수득하였다; 프라이머 0611815 및 0611632는 ura 마커를 갖는 약 6 kbp 밴드, 및 ura 마커가 없는 4.5 kbp를 수득하였다. 프라이머 0611631 및 0612579는 야생형 adh1202 위치가 온전할 때 (이 밴드를 나타내지 않은 균주를 선택하였다) 약 936 bp 밴드를 수득하였다. 발현 카세트의 적절한 통합에 대하여 예측된 밴드를 제공한 균주를 MeJi412로 지정하였다.
- [0577] 실시예 3A-15: PDC 프로모터에 의해 조절되는 아스파르테이트 1-데카르복실라제 (ADC)를 암호화하는 뉴클레오티드의 2개의 카피 및 TDH3 프로모터에 의해 조절되는 2개의 카피와 함께, adh1202 위치에서 아스파르테이트 1-데카르복실라제 (ADC)를 암호화하는 뉴클레오티드의 4개의 카피를 발현하는 효모 균주
- [0578] 이 실시예는 아이. 오리엔탈리스 PDC 프로모터에 의해 조절되는 2개의 카피 및 아이. 오리엔탈리스 TDH3 프로모터에 의해 조절되는 2개의 카피와 함께, adh1202 위치에서 SEQ ID NO: 139의 비. 리체니포르미스를 암호화하는 뉴클레오티드의 4개의 카피를 포함하도록 설계된 구조물을 설명한다. 상기 설명된 것과 유사한 접근법으로, 왼쪽 및 오른쪽 구조물을 아이. 오리엔탈리스 CNB1 adh1202 위치에서 상동 재조합을 허용하도록 설계하였다.
- [0579] 왼쪽 단편의 구조물
- [0580] 플라스미드 pMeJi310-2 (상기; 도 14 참조)를 XbaI 및 StuI로 분해한 후 이어서 CIP를 처리하였고 여기에 설명된 바와 같이 TBE 버퍼 중의 1% 아가로스 겔 전기영동으로 정제하였다. 약 6.7 kbp에서 밴드를 겔에서 잘라냈고 제조사의 지시에 따라 QIAQUICK® Gel Extraction Kit (Qiagen)를 사용하여 정제하였다. PDC 프로모터를 NotI로 분해에 의해 pMeJi310-2에서 자른 후 이어서 클레노브로 필-인 (fill-in) 반응시켰고 그 다음 NheI로 분해하였다. 약 708 bp에서 밴드를 겔에서 잘라냈고 제조사의 지시에 따라 QIAQUICK® Gel Extraction Kit (Qiagen)를 사용하여 정제하였다.
- [0581] 정제된 단편의 708 bp를 pMeJi310-2의 6.7 kbp 단편 1 μ L, pMeJi310-2의 708 bp 단편 1 또는 5 μ L, 10mM ATP (New England Biolabs)가 들어있는 1 μ L 10X 결찰 버퍼, 및 1 μ L T4 리가제 (New England Biolabs)로 구성된, 10 μ L의 총 반응량에서 T4 리가제 (New England Biolabs)를 사용하여 6.7 kbp pMeJi310-2 선형화 벡

터에 결찰하였다.

- [0582] 반응물을 16 °C에서 하룻밤 동안 배양하였고 반응물의 4 µL 엘리퀼트를 제조사의 지시에 따라 ONE SHOT® TOP10 화학적 컴피턴트 이. 콜리 세포 (Invitrogen)를 사용하여 형질전환하였다. 형질전환체를 2X YT+amp 플레이트에 도말하였고 37 °C에서 하룻밤 동안 배양하였다. 결과의 형질전환체 중 다수를 ApaLI를 사용하는 제한 분해에 의해 적절한 삽입에 대해 스크리닝하였다. 올바른 분해된 밴드 크기를 수득하는 클론을 pM1Ba137로 지정하였다.
- [0583] *오른쪽 단편의 구조물*
- [0584] 플라스미드 pMeJi312-2 (상기)를 XbaI 및 StuI로 분해한 후 이어서 CIP를 처리하였고 여기에 설명된 바와 같이 TBE 버퍼 중의 1% 아가로스 겔 전기영동으로 정제하였다. 약 6.8 kbp에서 밴드를 겔에서 잘라냈고 제조사의 지시에 따라 QIAQUICK® Gel Extraction Kit (Qiagen)를 사용하여 정제하였다.
- [0585] TDH3 프로모터를 PmeI 및 NheI로 분해에 의해 pMeJi312-2에서 잘라냈다. 약 966 bp에서 밴드를 겔에서 잘라냈고 제조사의 지시에 따라 QIAQUICK® Gel Extraction Kit (Qiagen)를 사용하여 정제하였다.
- [0586] 정제된 단편의 966 bp를 pMeJi312-2의 6.8 kbp 1 µL, pMeJi312-2의 966 bp 단편 1 또는 5 µL, 10mM ATP (New England Biolabs)가 들어있는 10X 결찰 버퍼 1 µL, 및 1 µL T4 리가제 (New England Biolabs)로 구성된, 10 µL의 총 반응량에서 T4 리가제 (New England Biolabs)를 사용하여 6.8 kbp pMeJi312-2 선형화 벡터에 결찰하였다. 반응물을 16 °C에서 약 6시간 동안 배양하였고 반응물의 4 µL 엘리퀼트를 제조사의 지시에 따라 ONE SHOT® TOP10 화학적 컴피턴트 이. 콜리 세포 (Invitrogen)에 형질전환하였다. 형질전환체를 2X YT+amp 플레이트에 도말하였고 37 °C에서 하룻밤 동안 배양하였다. 결과의 형질전환체 중 다수를 SalI를 사용하는 제한 분해에 의해 적절한 삽입에 대해 스크리닝하였다. 올바른 분해된 밴드 크기를 수득하는 클론을 pM1Ba136으로 지정하였다.
- [0587] *왼쪽 및 오른쪽 단편의 통합*
- [0588] 플라스미드 pM1Ba137를 HpaI 및 SacII로 분해하였고 여기에 설명된 바와 같이 플라스미드 pM1Ba136을 EcoRI 및 SacII로 분해하였다. 이것들은 TBE 버퍼 중의 1% 아가로스 겔 전기영동으로 정제하였고, 2개의 약 5 kbp 밴드를 겔에서 잘라냈고 제조사의 지시에 따라 QIAQUICK® Gel Extraction Kit (Qiagen)를 사용하여 정제하였다.
- [0589] 아이. 오리엔탈리스 CNB1을 분해된 pM1Ba137 및 pM1Ba136 DNA에 형질전환하였고 올바른 위치 표적화 및 형질전환을 여기에 설명된 바와 같이 Crimson Taq (New England Biolabs) PCR로 확인하였다. 프라이머 0611717 및 0611631은 약 2.5 kbp 및 955 bp의 밴드를 수득하였다; 프라이머 0611718 및 0611632는 약 733 bp 밴드를 수득하였다; 프라이머 0612794 및 0611245는 약 2.7 kbp 밴드를 수득하였다; 프라이머 0611225 및 0612795는 약 4.2 kbp 밴드를 수득하였다. 발현 카세트의 적절한 통합에 대해 예측된 밴드를 제공하는 균주를 M1Ba351로 지정하였다.
- [0590] *M1Ba351로부터 ura 마커의 제거*
- [0591] M1Ba351의 ura-유도체를 상기 설명된 바와 같이 분리하였다. M1Ba351의 다수의 FOA 저항성 콜로니의 게놈 DNA를 프라이머 0611815 및 0611817로 원하는 루프-아웃에 대해 PCR로 스크리닝하였다. 프라이머 0611815는 왼쪽 구조물의 RKI 종결자에서 어닐링하고 ura3 프로모터를 향해 증폭시킨다. 프라이머 0611817은 TDH3 프로모터에서 어닐링하고 ura3 카세트를 향해 증폭시킨다. 828 bp 밴드의 존재는 원하는 바와 같이 ura3 스카 부위 (모체 균주의 2개의 URA3 프로모터 사이에서 상동 재조합 후 단일 URA3 프로모터 왼쪽 뒤)의 존재를 나타내는 한편, 약 2.2 kbp의 밴드는 온전한 URA3 프로모터-URA3 ORF-URA3 종결자-URA3 프로모터 카세트의 존재를 나타내며, 원하는 재조합 이벤트가 발생하지 않았다는 것을 나타낸다.
- [0592] 상기 설명된 바와 같이 Crimson Taq™ DNA Polymerase (New England Biolabs)로 PCR을 수행하였다. 상기 프라이머로 828 bp 단편을 수득하는 FOA 저항성 콜로니를 SEQ ID NO: 139의 비. 리체니포르미스 ADC를 암호화하는 뉴클레오티드 서열 SEQ ID NO: 138의 4개의 카피가 온전하게 남아있다는 것을 확인하기 위해, 2.7 kbp 생성물을 수득하는 프라이머 0612794 및 0611245 및 4 kbp 생성물을 수득하는 프라이머 0611815 및 0612795로 추가로 테스트하였다. 모체 균주 M1Ba351의 하나의 FOA 저항성 콜로니를 M1Ba353으로 지정하였고, 3개의 프라이머 세트 모두로 원하는 PCR 생성물을 제공하였다.
- [0593] *역 발현 카세트 오른쪽 단편의 구조물*

[0594] 플라스미드 pM1Ba136은 정방향으로 가는 2개의 발현 카세트를 함유한다. 동형 접합체 균주의 스크리닝을 용이하게 하기 위해, pM1Ba136의 panDb1 발현 카세트가 역방향에 위치하는 경우 새로운 플라스미드를 구성하였다. 플라스미드 pM1Ba136 (상기)을 NotI 및 PmeI로 분해한 후 이어서 여기에 설명된 바와 같이 클레노브로 필-인 반응시켰고 및 TBE 버퍼 중의 1% 아가로스 겔 전기영동으로 정제하였다. 약 3.4 및 4.4 kbp에서 밴드를 겔에서 잘라냈고 제조사의 지시에 따라 QIAQUICK® Gel Extraction Kit (Qiagen)를 사용하여 정제하였다. pM1Ba136의 4.4 kbp 단편에 CIP를 처리하였고 제조사의 지시에 따라 QIAQUICK® PCR Purification Kit (Qiagen)를 사용하여 정제하였다.

[0595] pM1Ba136의 3.4 kbp 정제된 단편을 pM1Ba136의 4.4 kbp 단편 1 µL, pM1Ba136의 3.4 kbp 단편 1 또는 5 µL, 10mM ATP (New England Biolabs)가 들어있는 10X 결찰 버퍼 1 µL, 및 1 µL T4 리가제 (New England Biolabs)로 구성된, 10 µL의 총 반응량에서 T4 리가제 (New England Biolabs)를 사용하여 4.4 kbp pM1Ba136 벡터에 결찰하였다. 반응물을 16 °C에서 하룻밤 동안 배양하였고 배양물의 4 µL 앨리쿼트를 제조사의 지시에 따라 ONE SHOT® TOP10 화학적 컴피턴트 이. 콜리 세포 (Invitrogen)에 형질전환하였다. 형질전환체를 2X YT+amp 플레이트에 도말하였고 37 °C에서 하룻밤 동안 배양하였다. 결과의 형질전환체 중 다수를 PacII 및 EcoRI를 사용하는 제한 분해에 의해 적절한 삽입에 대해 스크리닝하였다. 올바른 분해된 밴드 크기를 수득하는 클론을 pM1Ba138로 지정하였다.

[0596] 왼쪽 및 오른쪽 단편의 통합

[0597] 플라스미드 pM1Ba137을 HpaI 및 SacII로 분해하였고 여기에 설명된 바와 같이 플라스미드 pM1Ba138을 EcoRI 및 SacII로 분해하였다. 이들은 TBE 버퍼 중의 1% 아가로스 겔 전기영동으로 정제하였고, 2개의 약 5 kbp 밴드를 겔에서 잘라냈고 제조사의 지시에 따라 QIAQUICK® Gel Extraction Kit (Qiagen)를 사용하여 정제하였다.

[0598] M1Ba353를 분해된 pM1Ba137 및 pM1Ba138 DNA에 형질전환하였고 올바른 위치 표적화 및 형질전환을 제조사의 지시에 따라 Phire® Plant Direct PCR Kit (Finnzymes)를 사용하여 PCR로 확인하였다. 프라이머 0611718 및 0611632는 약 733 bp 밴드를 수득하였다 (첫 번째 통합이 존재하는지 확인하기 위해); 프라이머 0612367 및 0611632는 약 960 bp 밴드를 수득하였다 (두 번째 카피가 통합되었는지 확인하기 위해); 프라이머 0611631 및 0612579는 야생형 adh1202 위치가 존재하는 경우 약 936 bp 밴드를 수득하였다 (이 밴드의 결핍은 wt adh1202 위치의 손실을 확인한다). 발현 카세트의 적절한 통합에 대하여 예측된 밴드를 제공하는 2개의 균주를 저장하였고 M1Ba355 및 M1Ba356으로 지정하였다.

[0599] MeJi409-2, MeJi412, M1Ba351 및 M1Ba355에서 아스파르테이트 1-데카르복실라제 활성

[0600] 균주 MeJi409-2, MeJi412, M1Ba351 및 M1Ba355를 셰이크 플라스크에서 키웠고 CFE를 제조하였고 여기에 설명된 바와 같이 아스파르테이트 1-데카르복실라제 (ADC) 활성에 대하여 검정하였다. 결과는 표 26에 나타난다. 각 균주에 대한 활성은 2개의 독립적 셰이크 플라스크 배양물의 평균이다. 여기에 설명된 방법을 사용하여, 균주 MeJi409-2, MeJi412, M1Ba351 및 M1Ba355를 또한 3-HP 생산에 대하여 바이오리액터에서 테스트하였다. 이들 바이오리액터 실험의 결과는 또한 표 26에 나타난다. 이들 발효에서 세포 질량의 차이에 대해 설명하기 위해, 표 26에 나타난 3-HP 생산 능력은 세포 질량의 유닛 당 3-HP 농도로서 표현된다 (g/L 3-HP/g/L 건조 세포 중량으로 표현됨). 결과는 세포에서 ADC 활성의 수준이 증가함에 따라 3-HP 생산 능력이 증가한다는 것을 나타낸다.

[0601] 표 26: 형질전환체 ADC 활성 및 3-HP 생산 데이터

| 균주 | 과발현된 유전자 | ADC 활성 | 3-HP/DCW |
|-----------|----------------------|--------|----------|
| MBin500 | N/A | 0 | 0 |
| MeJi409-2 | ADC (SED ID NO: 138) | 0.629 | 0.19 |
| MeJi412 | | 1.151 | 0.43 |
| M1Ba351 | | 0.659 | 0.32 |
| M1Ba355 | | 1.173 | 0.52 |

[0602]

[0603] 실시예 3A-16: PDC 위치에서 피루베이트 카르복실라제 (PYC)를 발현하는 플라스미드 구조물

[0604] PDC 위치에서 통합에 대하여 SEQ ID NO: 1의 뉴클레오티드 서열을 함유하는 플라스미드 pANN28 (SEQ ID NO: 2의 아이. 오리엔탈리스 PYC를 암호화함)를 하기 설명된 바와 같이 구성하였다.

[0605] 아이. 오리엔탈리스 PDC의 업스트림 및 다운스트림 플랭킹 영역을 제조사의 지시에 따라 주형으로 게놈 DNA Pfu

폴리머라제 (Stratagene)를 사용하여 PCR로 증폭하였다. 프라이머 oANN7 및 oANN8은 업스트림 영역을 플랭킹하는 독특한 제한 부위의 통합을 허용하고 프라이머 oANN9 및 oANN10은 다운스트림 영역을 플랭킹하는 독특한 제한 부위의 통합을 허용하였다. PCR 생성물을 여기에 설명된 바와 같이 TBE 버퍼 중의 1% 아가로스 겔 전기영동으로 정제하였다. 각각의 PCR 생성물에 대한 약 800 bp의 밴드를 겔에서 잘라냈고 제조사의 지시에 따라 Gel Extraction Kit (Qiagen)를 사용하여 정제하였다. 정제된 PCR 생성물을 TOPO 벡터 (Invitrogen)에 클로닝하였고 제조사의 지시에 따라 일렉트로-컴피턴트 (electro-competent) 이. 콜리 DH10B 세포 (Invitrogen)에 형질전환하였다. 결과의 형질전환체 중 다수를 PCR 생성물을 생성하기 위해 사용된 같은 프라이머로 콜로니 PCR에 의해 적절한 삽입에 대하여 스크리닝하였다. 양성 클론을 추가로 시퀀싱에 의해 확인하였다. 올바른 PDC 다운스트림 플랭크를 수득하는 클론을 pANN04로 지정하였다. 올바른 PDC 업스트림 플랭크를 수득하는 클론을 pANN07로 지정하였다.

[0606] 플라스미드 pANN04를 ApaI 및 SacI로 분해하였다 (벡터/백본으로 사용됨); 플라스미드 pANN04를 NotI 및 SacI로 분해하였다; 플라스미드 pANN07을 NotI 및 ApaI로 분해하였다. 각각의 단편을 여기에 설명된 바와 같이 TBE 버퍼 중의 1% 아가로스 겔 전기영동으로 정제하였다. 벡터에 대한 약 3.5 kbp의 밴드, 및 각각의 인서트에 대한 약 1 kbp를 겔에서 잘라냈고 제조사의 지시에 따라 Gel Extraction Kit (Qiagen)를 사용하여 정제하였다. 정제된 생성물을 49 ng의 벡터, 120 ng의 다운스트림 인서트, 41 ng의 업스트림 인서트, 10mM ATP (New England Biolabs)가 들어있는 10X 결찰 버퍼 1 μ L, 및 1 μ L T4 리가제 (New England Biolabs)로 구성된, 10 μ L의 총 반응량에서 T4 리가제 (New England Biolabs)를 사용하여 결찰하였다. 반응물을 상온에서 30분 동안 배양하였고 반응물의 2 μ L 앨리퀴트를 제조사의 지시에 따라 일렉트로-컴피턴트 이. 콜리 OneShot TOP10 세포 (Invitrogen)에 형질전환하였다. 형질전환체를 LB + 카나마이신 플레이트에 도말하였고 37 $^{\circ}$ C에서 하룻밤 동안 배양하였다. 결과의 형질전환체 중 다수를 프라이머 oANN7 및 oANN10 (약 1.7 kbp의 밴드를 수득함)로 콜로니 PCR에 의해 적절한 삽입에 대해 스크리닝하였다. 올바른 삽입물을 수득하는 클론을 pANN12로 지정하였다.

[0607] pGMEr137 (상기)의 아이. 오리엔탈리스 PYC 암호화 서열 (SEQ ID NO: 1)을 암호화된 효소의 아미노산 서열을 변화시키지 않는 3개의 EcoRI 제한 부위를 제거하도록 돌연변이 유발을 지시하는 부위를 변형시켰다. 플라스미드 pGMEr137를 제조사의 지시에 따라 Multi change kit (Stratagene)를 사용하여 상기 언급된 제한 부위의 제거에 사용된 프라이머 oANN13, oANN14 및 oANN15를 갖는 주형으로서 사용하였다. 결과의 형질전환체 중 다수를 EcoRI를 사용하는 제한 분해에 대해 스크리닝하였다. 양성 클론을 시퀀싱으로 추가로 확인하였다. 올바른 pyc 암호화 서열을 수득하는 클론을 pANN14로 지정하였다.

[0608] 플라스미드 pJY39 (도 29)를 XhoI 및 PacI로 분해하였다; 플라스미드 pACN5 (상기; 도 19 참조)를 XhoI 및 XbaI로 분해하였다; 플라스미드 pANN14를 XbaI 및 PacI로 분해하였다. 각각의 단편을 여기에 설명된 바와 같이 TBE 버퍼 중의 1% 아가로스 겔 전기영동으로 정제하였다. 벡터에 대한 약 8 kbp, 첫 번째 인서트에 대한 약 700 bp, 및 PYC를 암호화하는 두 번째 인서트에 대한 약 3.6 kbp의 밴드를 겔에서 잘라냈고 제조사의 지시에 따라 Gel Extraction Kit (Qiagen)를 사용하여 정제하였다. 정제된 생성물을 51 ng의 벡터, 49 ng의 첫 번째 인서트, 210 ng의 두 번째 인서트, 10 mM ATP (New England Biolabs)가 들어있는 10X 결찰 버퍼 1 μ L, 및 1 μ L T4 리가제 (New England Biolabs)로 구성된, 10 μ L의 총 반응량에서 T4 리가제 (New England Biolabs)를 사용하여 결찰하였다. 반응물을 상온에서 30분 동안 배양하였고 반응물의 2 μ L 앨리퀴트를 제조사의 지시에 따라 일렉트로-컴피턴트 이. 콜리 OneShot TOP10 세포 (Invitrogen)에 형질전환하였다. 형질전환체를 LB + 카나마이신 플레이트에 도말하였고 37 $^{\circ}$ C에서 하룻밤 동안 배양하였다. 결과의 형질전환체 중 다수를 프라이머 oJLJ57 및 oJLJ43 (약 1 kbp의 밴드를 수득함), 프라이머 oJLJ45 및 oANN16 (약 730 bp의 밴드를 수득함), 및 프라이머 oANN20 및 oJY45 (약 1.2 kbp의 밴드를 수득함)로 콜로니 PCR에 의해 적절한 삽입에 대해 스크리닝하였다. 올바른 삽입물을 수득하는 클론을 pANN15로 지정하였다.

[0609] 플라스미드 pANN12 및 pANN15를 NotI로 분해하였다. 플라스미드 pANN15를 백본의 추가의 분류 및 원하는 단편의 개선된 분리를 위해 추가적으로 NcoI로 분해하였다. 분해된 pANN12를 제조사의 지시에 따라 Qiagen kit를 사용하여 정제하였다. NotI 단편을 여기에 설명된 바와 같이 TBE 버퍼 중의 아가로스 겔 전기영동으로 정제하였다. 약 5 kbp (pANN12) 및 약 6.3 kbp (pANN15)의 밴드를 제조사의 지시에 따라 Gel Extraction Kit (Qiagen)를 사용하여 정제하였다.

[0610] pANN15의 정제된 생성물을 50 ng의 벡터, 115 ng의 인서트, 10mM ATP (New England Biolabs)가 들어있는 10X 결찰 버퍼 1 μ L, 및 1 μ L T4 리가제 (New England Biolabs)로 구성된, 10 μ L의 총 반응량에서 T4리가제를 사용하여 pANN12 선형화 벡터에 결찰하였다. 반응물을 상온에서 1.5시간 동안 배양하였고 반응물의 2 μ L 앨리퀴트를 제조사의 지시에 따라 일렉트로-컴피턴트 이. 콜리 OneShot TOP10 세포 (Invitrogen)에 형질전환하였다.

형질전환체를 LB + 카나마이신 플레이트에 도말하였고 37 °C에서 하룻밤 동안 배양하였다. 결과의 형질전환체 중 다수를 프라이머 oANN20 및 oJY45 (약 1.2 kbp의 밴드를 수득함)로 콜로니 PCR에 의해 적절한 삽입을 위해 스크리닝하였다. 올바른 삽입물을 수득하는 클론을 인서트 방향을 분류하기 위해 SacI/EcoRI로 및 SacI/EcoRV로 제한 효소 분해로 추가로 스크리닝하였다. 업스트림 PDC 플랭크 근처의, ura3 마커를 수득하는 클론을 pANN27로 지정하였다. 다운스트림 PDC 플랭크 근처의, ura3 마커를 수득하는 클론을 pANN28로 지정하였다.

[0611] 실시예 3A-17: adh1202 위치에서 4개의 뉴클레오타이드 서열의 아스파르테이트 1-데카르복실라제 (ADC) 및 pdc 위치에서 피루베이트 카르복실라제 (PYC)를 발현하는 효모 균주

[0612] 본 실시예는 adh1202 위치에서 SEQ ID NO: 139의 비. 리체니포르미스 ADC를 암호화하는 뉴클레오타이드의 4개의 카피 및 pdc 위치에서 SEQ ID NO: 2의 아이. 오리엔탈리스 PYC를 암호화하는 뉴클레오타이드를 발현하는 효모 균주의 구조물을 설명한다.

[0613] MeJi412로부터 ura 마커의 제거

[0614] MeJi412의 ura-유도체를 상기 설명된 바와 같이 분리하였다. MeJi412의 다수의 FOA 저항성 콜로니를 프라이머 0611815 및 0611817로 원하는 루프-아웃 이벤트에 대해 콜로니 PCR로 스크리닝하였다. 프라이머 0611815는 왼쪽 구조물의 RKI 종결자에서 어닐링하고 ura3 프로모터를 향해 증폭시킨다. 프라이머 0611817은 TDH3 프로모터에서 어닐링하고 ura3 카세트를 향해 증폭시킨다. 869 bp의 존재는 원하는 바와 같이 ura3 스카 부위 (모체 균주에서 2개의 URA3 프로모터 사이에서 상동 재조합 후 단일 URA3 프로모터 왼쪽 뒤)만의 존재를 나타내는 한편, 약 2.6 kbp의 밴드는 온전한 URA3 프로모터-URA3 ORF-URA3 종결자-URA3 프로모터 카세트의 존재를 나타내며, 원하는 재조합 이벤트가 발생하지 않았다는 것을 나타낸다. 상기 설명된 바와 같이 Phire® Plant Direct PCR Kit (Finnzymes)로 PCR 반응을 수행하였다. 상기 프라이머로 869 bp 단편을 수득하는 FOA 저항성 콜로니를 SEQ ID NO: 139의 비. 리체니포르미스 ADC를 암호화하는 뉴클레오타이드 서열 SEQ ID NO: 138의 4개의 카피가 온전하게 남아있다는 것을 확인하기 위해, 3.8 kbp 생성물을 수득하는 프라이머 0612794 및 0611817, 및 3.7 kbp 생성물을 수득하는 프라이머 611815 및 612795로 추가로 테스트하였다. 모체 균주 MeJi412의 하나의 FOA 저항성 콜로니는 3개의 프라이머 세트 모두로 원하는 PCR 생성물을 제공하였다.

[0615] 단편의 통합

[0616] 플라스미드 pANN28 (상기)를 AscI 및 SacI로 분해하였고 TBE 버퍼 중의 아가로스 겔 전기영동으로 정제하였다. 약 7.1 kbp에서 밴드를 겔에서 잘라냈고 제조사의 지시에 따라 겔 Purification Kit (Qiagen)를 사용하여 정제하였다.

[0617] 균주 MeJi413을 pANN28의 분리되고 정제된 단편에 형질전환하였고 올바른 위치 표적화 및 형질전환을 제조사의 지시에 따라 콜로니 PCR (Failsafe (Failsafe), 믹스 E, Epicenter)로 확인하였다. 프라이머 oANN12 및 oJLJ44는 약 1 kbp 밴드를 수득하였다; 프라이머 oANN11 및 oANN16은 약 1.3 kbp 밴드를 수득하였다. 발현 카세트의 적절한 통합에 대해 예측된 밴드를 제공하는 균주를 yANN35로 지정하였다.

[0618] yANN35의 ura-유도체를 상기 설명된 바와 같이 분리하였다. 다수의 FOA 저항성 콜로니를 프라이머 oANN12 및 oJY44로 원하는 루프-아웃 이벤트에 대하여 콜로니 PCR로 스크리닝하였다. 프라이머 oANN12는 다운스트림 플랭킹 영역의 외부에서 어닐링한다. 프라이머 oJY44는 TAL 종결자로 어닐링한다. 1.5 kbp 밴드의 존재는 ura3 스카 부위 (모체 균주에서 2개의 URA3 프로모터 사이에서 상동 재조합 후 단일 URA3 프로모터 왼쪽 뒤)만의 존재를 나타내는 한편, 약 2.8 kbp의 밴드는 온전한 URA3 프로모터-URA3 ORF-URA3 종결자-URA3 프로모터 카세트의 존재를 나타내며, 원하는 재조합 이벤트가 발생하지 않았다는 것을 나타낸다. 상기 설명된 바와 같이 Failsafe DNA Polymerase (Epicenter)로 PCR 반응을 수행하였다. 이 이벤트에 양성인 분리체를 프라이머 oANN16 및 oANN11로 콜로니 PCR로 추가로 확인하였다. 하나의 FOA 저항성 콜로니를 yANN37로 지정하였다.

[0619] 균주 yANN37를 pANN28의 원하는 및 정제된 단편에 형질전환하였고 올바른 위치 표적화 및 형질전환을 제조사의 지시에 따라 콜로니 PCR (Failsafe, 믹스 E, Epicenter)로 확인하였다. 예비 스크리닝을 PDC 유전자에 특이적인 프라이머 oHJJ116 및 oHJJ117로 수행하였다. 약 500 bp의 밴드는 유전자의 존재 및 따라서 원하는 통합에 대한 음성 결과를 나타낸다. PDC의 삭제에 대해 양성인 분리체를 추가의 PCR 반응물로 추가로 확인하였다. 프라이머 oANN11 및 oANN16은 약 1.3 kbp 밴드를 수득하였다. 프라이머 oANN12 및 oJLJ44는 약 1 kbp 밴드를 수득하였다; 프라이머 oANN12 및 oJY44는 약 1.5 kbp 밴드 및 약 2.9 kbp 밴드 (첫 번째 및 두 번째 통합 이벤트 각각에 해당함)를 수득하였다.

[0620] 추가적으로, adh1202 위치에서 이전의 통합 이벤트를 상기 설명된 바와 같이 콜로니 PCR로 확인하였다. 프라이

며 0611631 및 0611245는 약 3.8 kbp 밴드를 수득하였다. 프라이머 0611245 및 oNovo3은 약 3 kbp 밴드를 수득하였다. 프라이머 0611815 및 0612795는 약 3.6 kbp 밴드를 수득하였다. 발현 카세트의 적절한 통합에 대한 예측된 밴드를 제공하는 균주를 yANN41로 지정하였다.

[0621] 실시예 3A-18: adh1202 위치에서 4개의 뉴클레오티드 서열의 아스파르트테이트 1-데카복실라제 (ADC) 및 pdc 위치에서 피루베이트 카복실라제 (PYC)를 발현하는 효모 균주

[0622] 본 실시예는 adh1202 위치에서 SEQ ID NO: 139의 비. 리체니포르미스 ADC를 암호화하는 뉴클레오티드의 4개의 카피 (PDC 프로모터에 의해 조절되는 2개의 카피 및 TDH3 프로모터에 의해 조절되는 2개의 카피와 함께) 및 pdc 위치에서 SEQ ID NO: 2의 아이. 오리엔탈리스 PYC를 암호화하는 뉴클레오티드를 발현하는 효모 균주의 구조물을 설명한다.

[0623] *MBa355의 ura 마커의 제거*

[0624] MBa355의 ura-유도체를 상기 설명된 바와 같이 분리하였다. MBa355의 다수의 FOA 저항성 콜로니의 게놈 DNA를 프라이머 0611815 및 0611718로 원하는 루프-아웃 이벤트에 대해 PCR로 스크리닝하였다. 약 500 bp 밴드의 존재는 원하는 바와 같이 ura3 스카 부위 (모체 균주에서 2개의 URA3 프로모터 사이에서 상동 재조합 후 단일 URA3 프로모터 왼쪽 뒤)만의 존재를 나타내는 한편, 약 1.9 kbp의 밴드는 온전한 URA3 프로모터-URA3 ORF-URA3 종결자-URA3 프로모터 카세트의 존재를 나타내며, 원하는 재조합 이벤트가 발생하지 않았다는 것을 나타낸다. PCR을 제조사의 지시에 따라 Phire® Plant Direct PCR kit (Finnzymes)를 사용하여 수행하였다. 상기 프라이머로 약 500 bp 단편을 수득하는 FOA 저항성 콜로니를 SEQ ID NO: 139의 비. 리체니포르미스를 암호화하는 뉴클레오티드 서열 SEQ ID NO: 138의 4개의 카피가 온전하게 남아있다는 것을 확인하기 위해, 3.5 kbp 생성물을 수득하는 프라이머 0611631 및 0611245, 및 4.5 kbp 생성물을 수득하는 프라이머 0611815 및 0611632로 추가로 테스트하였다. 그것들은 adh1202에서 첫 번째 변형이 존재한다는 것을 확인하기 위해 프라이머 0611815 및 0611817을 사용하는 PCR로 테스트하였다. 이들 PCR 프라이머는 828 bp 단편을 수득하였다. 모체 균주 MBa355의, 하나의 FOA 저항성 콜로니를 MBa357로 지정하였고, 4개의 프라이머 세트 모두로 원하는 PCR 생성물을 제공하였다

[0625] 플라스미드 pANN28 (상기)을 AscI 및 SacI로 분해하였고 TBE 버퍼 중의 1% 아가로스 겔 전기영동으로 정제하였다. 약 7.1 kbp 밴드를 겔에서 잘라냈고 제조사의 지시에 따라 NUCLEOSPIN® Extract II Kit (Macherey-Nagel)를 사용하여 정제하였다.

[0626] MBa357을 분해된 pANN28 DNA에 형질전환하였고 올바른 위치 표적화 및 형질전환을 Phire Plant Direct PCR Kit (Finnzymes)를 사용하는 PCR로 확인하였다. 프라이머 0611622 및 0611552는 약 850 bp 밴드를 수득하였다; 프라이머 0611245 및 0612794는 약 2.8 kbp 밴드를 수득하였다; 프라이머 0611815 및 0612795는 약 3.9 kbp 밴드를 수득하였다. 발현 카세트의 적절한 통합에 대해 예측된 밴드를 제공하는 균주를 McTs241로 지정하였다.

[0627] McTs241의 ura-유도체를 이전에 설명된 바와 같이 분리하였다. McTs241의 다수의 FOA 저항성 콜로니를 프라이머 0611423 및 0611554로 원하는 루프-아웃 이벤트 및 ura 마이너스 선택 플레이트에서 성장에 대해 PCR로 스크리닝하였다. 4.6 kbp 밴드의 존재는 원하는 바와 같이 ura3 스카 부위 (모체 균주에서 2개의 URA3 프로모터 사이에서 상동 재조합 후 단일 URA3 프로모터 왼쪽 뒤)만의 존재를 나타내는 한편, 약 5.9 kbp의 밴드는 온전한 URA3 프로모터-URA3 ORF-URA3 종결자-URA3 프로모터 카세트의 존재를 나타내며, 원하는 재조합 이벤트가 발생하지 않았다는 것을 나타낸다. PCR 반응을 상기 설명된 바와 같이 Phire Plant Direct PCR Kit (Finnzymes)를 사용하여 수행하였다. 원하는 루프-아웃 이벤트를 갖는 모체 균주 McTs241의 하나의 FOA 저항성 콜로니를 McTs247로 지정하였다.

[0628] 동형 접합체 통합을 생성하기 위해, McTs247을 분해된 pANN28 DNA에 형질전환하였고 올바른 위치 표적화 및 형질전환을 Phire Plant Direct PCR Kit (Finnzymes)를 사용하여 PCR로 확인하였다. 첫 번째 스크리닝으로서 형질전환체를 pdc 위치가 온전하고 따라서 PDC 위치에서 PYC의 동형 접합체 통합이 발생하지 않는 경우에만 약 850 bp 밴드를 수득해야하는 프라이머 0611552 및 0611553을 가지고 PCR로 스크리닝하였다. 본 PCR의 밴드에 음성인 것들을 프라이머 0611555 및 0611554로 추가적인 PCR로 스크리닝하였다. 이들 프라이머로 생성물을 PDC가 온전하고 따라서 PDC 위치에서 PYC의 동형 접합체 통합이 아닌 경우 1.4 kbp 밴드를 증폭시켜야만 한다. 형질전환체의 추가의 스크리닝을 약 850 bp 밴드를 수득하는 프라이머 0611622 및 0611552를 사용하는 PCR로 수행하였다; 프라이머 0611245 및 0612794는 약 2.8 kbp 밴드를 수득하였다; 프라이머 0611815 및 0612795는 약 3.9 kbp 밴드를 수득하였다. 발현 카세트의 적절한 통합에 대해 예측된 밴드를 제공하는 균주를 McTs253으로 지정하였다.

- [0629] 실시예 3A-19: 균주 MeJi412, yANN35, yANN41, M1Ba355, McTs241 및 McTs253의 PYC 활성, ADC 활성 및 3-HP 생산 능력
- [0630] 균주 MeJi412, yANN35, yANN41, M1Ba355, McTs241 및 McTs253을 셰이크 플라스크에서 키웠고 CFE를 제조하였고 여기에 설명된 바와 같이 피루베이트 카르복실라제 (PYC) 활성 및 아스파르테이트 1-데카르복실라제 (ADC) 활성에 대하여 검정하였다. 결과는 표 27에 나타난다. 균주 MeJi412, yANN35, yANN41, M1Ba355, McTs241 및 McTs253을 또한 여기에 설명된 방법을 사용하여 3-HP 생산에 대해 바이오리액터에서 테스트하였다. 이들 바이오리액터 실험의 결과는 또한 표 27에 나타난다. 이들 발효에서 세포 질량의 차이에 대해 설명하기 위해, 나타난 3-HP 생산 능력은 세포 질량의 유닛 당 3-HP 농도로 표현된다 (g/L 3-HP/g/L 건조 세포 중량으로 표현됨). 결과는 세포의 PYC 활성의 수준이 증가함에 따라, 3-HP 생산 능력이 증가한다는 것을 나타낸다.
- [0631] 표 27: 형질전환체 PYC 및 ADC 활성 및 3-HP 생산 능력
- | 균주 | 과발현된 유전자 | PYC 활성 | ADC 활성 | 3-HP/DCW |
|---------|--|--------|--------|----------|
| MeJi412 | ADC (SED ID NO: 138) | 6.8 | 1.151 | 0.43 |
| yANN35 | ADC (SED ID NO: 138) PYC (SEQ ID NO: 1) | 47.2 | 1.090 | 0.64 |
| yANN41 | ADC (SED ID NO: 138) PYC (SEQ ID NO: 1) | 49.0 | 1.263 | 1.30 |
| M1Ba355 | ADC (SED ID NO: 138) | 6.9 | 1.173 | 0.52 |
| McTs241 | ADC (SED ID NO: 138) PYC (SEQ ID NO: 1) | 24.8 | 1.119 | 0.76 |
| McTs253 | ADC (SED ID NO: 138) PYC (SEQ ID NO: 1) | 55.5 | 1.347 | 1.30 |
- [0632]
- [0633] 실시예 3A-20: adh1202 위치에서 4개의 뉴클레오티드 서열의 아스파르테이트 1-데카르복실라제 (ADC) 및 베타-알라닌 아미노트랜스퍼라제 (BAAT)의 삭제를 발현하는 효모 균주
- [0634] 본 실시예는 adh1202 위치에서 SEQ ID NO: 139의 비. 리체니포르미스 ADC를
- [0635] 암호화하는 뉴클레오티드의 4개의 카피 및 성능 및 SEQ ID NO: 20의 BAAT (PYD4)를 암호화하는 네이티브 아이. 오리엔탈리스 유전자의 삭제를 발현하는 효모 균주의 구조물을 설명한다.
- [0636] 아이. 오리엔탈리스 BAAT (PYD4) 삭제 플라스미드의 구조물
- [0637] 플라스미드 pM1Ba123 (상기)를 NotI, KpnI, ApaI로 분해하였고 여기에 설명된 바와 같이 TBE 버퍼 중의 1% 아가로스 겔 전기영동으로 정제하였다. 약 3.6 kbp 및 3.8 kbp의 2개의 밴드를 겔에서 잘라냈고 제조사의 지시에 따라 NUCLEOSPIN® Extract II Kit (Macherey-Nagel)를 사용하여 정제하였다. 이들 2개의 조각은 플라스미드 백본 및 SEQ ID NO: 144의 에스. 세레비시iae 3-HPDH 유전자 (YMR226c)를 갖는 ura 선택 카세트를 포함하였다.
- [0638] 업스트림 아이. 오리엔탈리스 PYD4 상동체에 대한 PCR 생성물을 이전에 프라이머 0613178 및 0613180를 사용하는 것으로 설명된 바와 같이 아이. 오리엔탈리스 MBin500 게놈 DNA를 사용하는 PCR 증폭으로 생성하였다. 각각의 프라이머 50 pmol을 50 µL의 최종 부피로 주형으로서 MBin500 게놈 DNA 0.5 µL, 0.2 mM 각각의 dATP, dGTP, dCTP, dTTP, 1X Expand High-Fidelity Buffer (Roche), 3.5 U Expand High-Fidelity 효소 믹스 (Roche)를 함유하는 PCR 반응물에 사용하였다. 증폭 반응을 95 °C에서 3분 동안 1 사이클 및 각각 95 °C에서 30초 동안, 55 °C에서 30초 동안, 및 72 °C에서 1분 동안 30 사이클로 프로그램된 EPPENDORF® MASTERCYCLER® 5333 (Eppendorf)에서 수행하였다. 사이클 후, 반응물을 72 °C에서 5분 동안 배양하였고 추가로 가공될 때까지 10 °C로 냉각하였다. 프라이머 0613178 및 0613180를 사용하는 PCR의 약 800 bp 밴드 및 프라이머 0613179 및 0613181을 사용하는 PCR의 약 900 bp 밴드는 TBE 버퍼 중의 1% 아가로스 겔 전기영동으로 정제하였다. 밴드를 겔에서 잘라냈고 제조사의 지시에 따라 NUCLEOSPIN® Extract II Kit (Macherey-Nagel)를 사용하여 정제하였다.
- [0639] PYD4 업스트림 PCR 생성물, PYD4 다운스트림 PCR 생성물, 및 pM1Ba123 NotI/KpnI/ApaI 분해된 플라스미드를 제조사의 지시에 따라 IN-FUSION HD™ (Clontech Laboratories, Inc.)으로 반응물에 모았다. In-FUSION 반응물 2 µL를 제조사의 지시에 따라 ONE SHOT® TOP10 화학적 컴피턴트 이. 콜리 세포 (Invitrogen)에 형질전환하였다. 회수 단계 후, 형질전환 반응물의 100 µL 앨리퀀트 2개를 mL 당 100 µg 앰피실린이 보충된 150 mM 2X YT 플레이트에 도말하였다. 플레이트를 37 °C에서 하룻밤 동안 배양하였다. 추정의 재조합 클론을 선택 플레이트에서 선택하였고 플라스미드 DNA를 BIOROBOT® 9600 (Qiagen)을 사용하여 각각 제조하였다. 클론을 제한 분해 및 시퀀싱으로 분석하였다. 올바른 서열을 갖는 플라스미드를 시퀀싱으로 확인하였고 pMcTs61로 명명하였다.

- [0640] 플라스미드 pMcTs61은 PDC 프로모터, 에스. 세레비시아의 YMR226c 유전자, 및 PDC 종결자를 함유한다. 이들 원하지 않는 세그먼트를 제거하기 위해, pMcTs61을 EcoRI 및 XhoI로 분해한 후 이어서 블린트 말단을 생성하기 위해 클레노브 단편을 추가하였다. 7.1 kbp 단편을 TBE 버퍼 중의 1% 아가로스 겔 전기영동으로 정제하였다. 밴드를 겔에서 잘라냈고 제조사의 지시에 따라 NUCLEOSPIN® Extract II Kit (Macherey-Nagel)를 사용하여 정제하였다. 분해된 블린트 플라스미드를 T4 DNA 리가제 (New England Biolabs, Ipswich, MA, USA)를 사용하여 함께 결합하였다. 반응 혼합물은 20 μ L의 총 부피에서 1X T4 DNA 리가제 버퍼, 1 μ L T4 DNA 리가제, 5 μ L pMcTs61 분해된 및 블린트화 정제된 DNA를 함유하였다. 반응물을 상온에서 2시간 동안 배양하였다. 결합 반응물의 10 μ L 샘플을 제조사의 지시에 따라 ONE SHOT® TOP10 화학적 컴피턴트 이. 콜리 세포 (Invitrogen)를 형질 전환하기 위해 사용하였다. 회수 단계 후, 형질전환 반응물의 100 μ L 엘리쿼트 2개를 mL 당 100 μ g 암피실린 이 보충된 150 mM 2X YT 플레이트에 도말하였다. 플레이트를 37 °C에서 하룻밤 동안 배양하였다. 추정된 재조합 클론 선택 플레이트로부터 선택하였다. 클론을 콜로니 PCR로 분석하였다. 각각의 콜로니로부터 주형 DNA를 50 μ L 멸균수에 1개의 콜로니를 용해시킴으로써 제조하였고, 95 °C에서 10분 동안 가열하였고, 이후 사용할 때까지 얼음 위에서 냉각하였다. 프라이머 0612911 및 0612909를 형질전환체를 스크리닝하기 위해 사용하였다. 이들 프라이머로 PCR 반응은 플라스미드가 올바른 경우 1 kbp 밴드를 증폭시킬 것이다. 각각의 프라이머 10 pmol을 20 μ L의 최종 부피에서 2 μ L 콜로니 DNA 주형, 0.1mM 각각의 dATP, dGTP, dCTP, dTTP, 1X Crimson Taq 반응 버퍼 (New England Biolabs), 1 U Crimson Taq DNA Polymerase (New England Biolabs)를 함유하는 PCR 반응물에 사용하였다. 증폭 반응을 95 °C에서 3분 동안 1 사이클; 및 각각 95 °C에서 30초 동안, 55 °C에서 30초 동안, 및 72 °C에서 3분 동안 30 사이클로 프로그램된 EPPENDORF® MASTERCYCLER® 5333 (Eppendorf)에서 수행하였다. 사이클 후, 반응물을 72 °C에서 5분 동안 배양하였고 추가로 가공될 때까지 10 °C에서 냉각하였다. 5 μ L의 PCR 반응물에서 1 kbp PCR 단편을 TAE 버퍼 중에 에티뮴 브로마이드가 들어있는 1% TAE-아가로스 겔 상에서 시각화하였다. 올바른 크기의 PCR 생성물을 갖는 하나의 형질전환체를 선택하였고 pMcTs64 (도 30)로 명명하였다. pMcTs64의 플라스미드 DNA를 BIOROBOT® 9600 (Qiagen)을 사용하여 제조하였다.
- [0641] pMcTs64 구조물을 사용하여 MeJi413으로부터 네이티브 아이. 오리엔탈리스 BAAT (PYD4)의 삭제
- [0642] 플라스미드 pMcTs64 (상기; 도 30 참조)를 ApaI, NcoI, KpnI로 분해하였고 TBE 버퍼 중의 1% 아가로스 겔 전기영동으로 정제하였다. 약 3.3 kbp 밴드를 겔에서 잘라냈고 제조사의 지시에 따라 NUCLEOSPIN® Extract II Kit (Macherey-Nagel)를 사용하여 정제하였다.
- [0643] 균주 MeJi413 (상기)을 분해된 pMcTs64 DNA에 형질전환하였고 올바른 위치 표적화 및 형질전환을 Phire Plant Direct PCR Kit (Finnzymes)를 사용하여 PCR로 확인하였다. 프라이머 0612908 및 0613242는 약 1.7 kbp 밴드를 수득하였다; 프라이머 0613241 및 0612909는 삭제 카세트의 통합을 확인하는 약 1.5 kbp 밴드를 수득하였다. 프라이머 0611815 및 0611632는 약 4.2 kbp 밴드를 수득하였다; 프라이머 0611817 및 0611631은 adh1202 위치에서 ADC 카세트가 온전하다는 것을 확인하기 위해 약 4.8 kbp 밴드를 수득한다. 삭제 카세트 및 ADC 카세트의 적절한 통합에 대해 예측된 밴드를 제공하는 균주를 McTs225로 지정하였다.
- [0644] McTs225의 ura-유도체를 이전에 설명된 바와 같이 분리하였다. McTs225의 다수의 FOA 저항성 콜로니를 프라이머 0612911 및 0612910으로 원하는 루프-아웃 이벤트에 대하여 PCR로 스크리닝하였다. 1.1 kbp 밴드의 존재는 원하는 바와 같이 ura3 스카 부위 (모체 균주에서 2개의 URA3 프로모터 사이에서 상동 재조합 후 단일 URA3 프로모터 왼쪽 뒤)만의 존재를 나타내는 한편, 약 2.5 kbp의 밴드는 온전한 URA3 프로모터-URA3 ORF-URA3 종결자-URA3 프로모터 카세트의 존재를 나타내며, 원하는 재조합 이벤트가 발생하지 않았다는 것을 나타낸다. 프라이머 0611815 및 0611632는 약 4.2 kbp 밴드를 수득하였다; 프라이머 0611817 및 0611631은 adh1202 위치에서 ADC 카세트가 온전하다는 것을 확인하기 위해 약 4.8 kbp 밴드를 수득한다. PCR 반응을 상기 설명된 바와 같이 Phire Plant Direct PCR Kit (Finnzymes)를 사용하여 수행하였다. 원하는 루프-아웃 이벤트를 갖는 모체 균주 McTs225의 하나의 FOA 저항성 콜로니를 McTs228로 지정하였다.
- [0645] SEQ ID NO: 20의 아이. 오리엔탈리스 BAAT (PYD4)를 암호화하는 네이티브 유전자의 동형 접합체 삭제를 생성하기 위해, McTs228을 분해된 pMcTs64에 형질전환하였고 올바른 위치 표적화 및 형질전환을 Phire Plant Direct PCR Kit (Finnzymes)를 사용하여 PCR로 확인하였다. 2개의 프라이머 세트를 PYD4 위치 삭제하도록 PCR로 스크리닝하기 위해 사용하였다. 프라이머 0613550 및 0612910은 PYD4의 동형 접합체 삭제가 발생하지 않았다는 것을 나타내는, PYD4 위치가 온전한 경우 약 700 bp 밴드를 수득한다. 추가적으로 형질전환체를 PYD4가 삭제되지 않은 경우 약 600 bp를 수득하는 프라이머 0612911 및 0613551로 스크리닝하였다. 아이. 오리엔탈리스 PYD4 위치에 음성인 형질전환체를 약 3.5 kbp 및 2.1 kbp 밴드를 수득하는 프라이머 0613242 및 0613243으로 추가로 스크리닝하였다; 프라이머 0612908 및 0613243은 약 1.7 kbp 밴드를 수득하였다; 프라이머 0612909 및 0612911은

약 950 bp 밴드를 수득하였다. 약 4.8 kbp를 수득하는 프라이머 0611817 및 0611631 및 약 4.2 kbp 밴드를 수득하는 프라이머 611815 및 612712를 갖는, adh1202 위치의 ADC 카세트는 온전한 것으로 확인되었다. 발현 카세트의 적절한 통합에 대해 예측된 밴드를 제공하는 균주를 McTs236으로 지정하였다.

[0646] McTs236의 ura-유도체를 이전에 설명된 바와 같이 분리하였다. McTs236의 다수의 FOA 저항성 콜로니를 약 2.1 kbp 밴드를 수득하는 프라이머 0613242 및 0613243으로 원하는 루프-아웃 이벤트에 대해 PCR로 스크리닝하였다. 약 3 kbp를 수득하는 프라이머 0611245 및 0612794 및 약 3.6 kbp 밴드를 수득하는 프라이머 0611815 및 0612795를 갖는, ADH1202 위치의 ADC 카세트는 온전한 것으로 확인되었다. 발현 카세트의 적절한 통합에 대해 예측된 밴드를 제공하는 균주를 McTs245로 지정하였다.

[0647] 여기에 설명된 방법을 사용하여, 균주 M1Ba372 및 McTs245를 3-HP 생산에 대해 바이오리액터에서 테스트하였다. 이들 발효에서 세포 질량의 차이를 설명하기 위해, 3-HP 생산 능력은 세포 질량의 유닛 당 3-HP 농도로서 표현된다 (g/L 3-HP/g/L 건조 세포 중량으로 표현됨). 균주 M1Ba372 및 McTs245에 대한 g/L 3-HP/g/L 건조 세포 중량은 각각 1.66 및 0.16이었다. 이들 결과는 이 유전자의 삭제가 3-HP 생산 능력의 10배 감소로 이어지기 때문에, 아이. 오리엔탈리스의 네이티브 PYD4 유전자가 베타-알라닌의 말로네이트 세미알데히드로 전환의 원인이 된다는 것을 제안한다.

[0648] 실시예 3A-21: adh1202 위치에서 4개의 뉴클레오티드 서열의 아스파르테이트 1-데카르복실라제 (ADC) 및 3-HP 데히드로게나제 (3-HPDH)의 삭제를 발현하는 효모 균주

[0649] 본 실시예는 adh1202 위치에서 SEQ ID NO: 139의 비. 리체니포르미스 ADC를 암호화하는 뉴클레오티드의 4개의 카피 및 SEQ ID NO: 26의 3-HPDH를 암호화하는 네이티브 아이. 오리엔탈리스 유전자의 삭제를 발현하는 효모 균주의 구조물 및 성능을 설명한다.

[0650] 아이. 오리엔탈리스 3-HPDH 삭제 플라스미드의 구조물

[0651] 플라스미드 pM1Ba123 (상기)를 NotI, KpnI, ApaI로 분해하였고 여기에 설명된 바와 같이 TBE 버퍼 중의 1% 아가로스 겔 전기영동으로 정제하였다. 약 3.6 kbp 및 3.8 kbp에서 2개의 밴드를 겔에서 잘라냈고 제조사의 지시에 따라 NUCLEOSPIN® Extract II Kit (Macherey-Nagel)를 사용하여 정제하였다. 이들 2개의 조각은 플라스미드 백본 및 SEQ ID NO: 144의 에스. 세레비시에 3-HPDH 유전자를 갖는 ura 선택 카세트를 포함했다.

[0652] 업스트림 아이. 오리엔탈리스 3-HPDH 상동체에 대한 PCR 생성물을 프라이머 0613183 및 0613184를 사용하여 이전에 설명된 바와 같이 제조된 아이. 오리엔탈리스 MBin500 게놈 DNA로부터 증폭하였다. 다운스트림 아이. 오리엔탈리스 3-HPDH 상동체 조각을 프라이머 0613185 및 0613186을 사용하여 아이. 오리엔탈리스 MBin500 게놈 DNA로부터 증폭하였다. 각각의 프라이머 50 pmol을 50 μ L의 최종 부피에 주형으로서 MBin500 게놈 DNA 0.5, 0.2 mM 각각의 dATP, dGTP, dCTP, dTTP, 2% DMSO, 1X Phusion HF 버퍼 (Finnzymes), 2U Phusion® Hot Start High-Fidelity DNA Polymerase (Finnzymes)를 함유하는 PCR 반응물에 사용하였다. 증폭 반응을 95 °C에서 3분 동안 1 사이클; 및 각각 95 °C에서 30초 동안, 55 °C에서 30초 동안, 및 72 °C에서 1분 동안 30 사이클로 프로그램된 EPPENDORF® MASTERCYCLER® 5333 (Eppendorf Scientific, Inc., Westbury, NY, USA)에서 수행하였다. 사이클 후, 반응물을 72 °C에서 5분 동안 배양하였고 추가로 가공될 때까지 10 °C에서 냉각하였다. 프라이머 0613183 및 0613184의 PCR의 약 640 bp 밴드 및 프라이머 0613185 및 0613186의 PCR의 약 670 bp 밴드를 TBE 버퍼 중의 1% 아가로스 겔 전기영동으로 정제하였다. 밴드를 겔에서 잘라냈고 제조사의 지시에 따라 NUCLEOSPIN® Extract II Kit (Macherey-Nagel)를 사용하여 정제하였다.

[0653] 아이. 오리엔탈리스 3-HPDH 업스트림 PCR 생성물, 아이. 오리엔탈리스 3-HPDH 다운스트림 PCR 생성물, 및 pM1Ba123 NotI/KpnI/ApaI 분해된 플라스미드를 제조사의 지시에 따라 IN-FUSION HD™ (Clontech Laboratories, Inc.)으로 반응물에 모았다. In-FUSION 반응물의 2 μ L을 제조사의 지시에 따라 ONE SHOT® TOP10 화학적 컴피턴트 이. 콜리 세포 (Invitrogen)에 형질전환하였다. 회수 기간 후에, 형질전환 반응물의 2개의 100 μ L 엘리쿼트를 mL 당 100 μ g 앰피실린이 보충된 150 mM 2X YT 플레이트에 도말하였다. 플레이트를 37 °C에서 하룻밤 동안 배양하였다. 추정된 재조합 클론을 선택 플레이트로부터 선택하였고 플라스미드 DNA를 BIOROBOT® 9600 (Qiagen)를 사용하여 각각 제조하였다. 클론을 제한 분해 및 시퀀싱으로 분석하였다. 올바른 서열을 갖는 플라스미드를 시퀀싱으로 확인하였고 pMcTs60으로 명명하였다.

[0654] 플라스미드 pMcTs60은 PDC 프로모터, 에스. 세레비시에의 YMR226c 유전자, 및 PDC 종결자를 함유한다. 이들 원하지 않는 세그먼트를 제거하기 위해, pMcTs60을 NotI 및 XbaI로 분해하였고 3-HPDH 상동 영역 및 플라스미드 백본을 함유하는 약 5 kbp 밴드를 TBE 버퍼 중의 1% 아가로스 겔 전기영동으로 정제하였다. 밴드를 겔에서 잘라

냈고 제조사의 지시에 따라 NUCLEOSPIN® Extract II Kit (Macherey-Nagel)를 사용하여 정제하였다. *ura3* 선택 카세트를 프라이머 0613416 및 0613417을 갖는 하나 및 프라이머 0613418 및 0613419를 갖는 다른 것의, 2개의 PCR 반응으로 증폭시켰다. 각각의 프라이머 50 pmol을 50 µL의 최종 부피에 주형으로서 pMcTs60 플라스미드 DNA 0.5 µL, 0.2 mM 각각의 dATP, dGTP, dCTP, dTTP, 1X Expand High-Fidelity Buffer (Roche), 3.5 U Expand High-Fidelity Enzyme Mix (Roche)를 함유하는 PCR 반응물에 사용하였다. 증폭 반응을 95 °C에서 3분 동안 1 사이클; 및 각각 95 °C에서 30초 동안, 55 °C에서 30초 동안, 및 72 °C에서 1분 동안 30 사이클로 프로그램된 EPPENDORF® MASTERCYCLER® 5333 (Eppendorf Scientific, Inc., Westbury, NY, USA)에서 수행하였다. 사이클 후, 반응물을 72 °C에서 5분 동안 배양하였고 추가로 가공될 때까지 10 °C에서 냉각하였다. 프라이머 0613416 및 0613417의 PCR의 약 700 bp 밴드 및 프라이머 0613418 및 0613419의 PCR의 약 1 kbp 밴드를 TBE 버퍼 중의 1% 아가로스 겔 전기영동으로 정제하였다. 밴드를 겔에서 잘라냈고 제조사의 지시에 따라 NUCLEOSPIN® Extract II Kit (Macherey-Nagel)를 사용하여 정제하였다. 아이. 오리엔탈리스 3-HPDH 상동 영역을 함유하는 플라스미드 백본 및 *ura3* 카세트 PCR 생성물을 제조사의 지시에 따라 IN-FUSION HD™ (Clontech Laboratories, Inc.)으로 반응물에 모았다. In-FUSION 반응의 2 µL를 제조사의 지시에 따라 SoloPack Gold SuperCompetent Cell (Stratagene)에 형질전환하였다. 회수 기간 후, 형질전환 반응물의 2개의 100 µL 엘리쿼트를 mL 당 100 µg 엡피실린이 보충된 150 mM 2X YT 플레이트에 도말하였다. 플레이트를 37 °C에서 하룻밤 동안 배양하였다. 추정된 재조합 클론을 선택 플레이트로부터 선택하였고 플라스미드 DNA를 BIOROBOT® 9600 (Qiagen)을 사용하여 각각 제조하였다. 클론을 제한 분해로 분석하였고 올바른 제한 분해 패턴을 갖는 플라스미드를 pMcTs65 (도 3 1)로 명명하였다.

- [0655] *pMcTs65* 구조물을 사용하여 *MeJi413*으로부터 네이티브 아이. 오리엔탈리스 3-HPDH의 삭제
- [0656] 플라스미드 pMcTs65 (상기; 도 31 참조)를 *Apa*I, *Sph*, *Kpn*I로 분해하였고 TBE 버퍼 중의 1% 아가로스 겔 전기영동으로 정제하였다. 약 2.9 kbp 밴드를 겔에서 잘라냈고 제조사의 지시에 따라 NUCLEOSPIN® Extract II Kit (Macherey-Nagel)를 사용하여 정제하였다.
- [0657] 균주 *MeJi413* (상기)를 분해된 pMcTs65 DNA에 형질전환하였고 올바른 위치 표적화 및 형질전환을 Phire Plant Direct PCR Kit (Finnzymes)를 사용하여 PCR로 확인하였다. 프라이머 0613034 및 0613035는 삭제 카세트의 통합을 확인하기 위해 약 2.7 kbp 밴드를 수득하였다. 프라이머 0611815 및 0611632는 약 4.2 kbp 밴드를 수득하였다; 프라이머 0611817 및 0611631은 *adh1202* 위치에서 ADC 카세트가 온전하다는 것을 확인하기 위해 약 4.8 kbp 밴드를 수득하였다. 삭제 카세트 및 ADC 카세트의 적절한 통합에 대해 예측된 밴드를 제공하는 균주를 McTs229로 지정하였다.
- [0658] McTs229의 *ura*-유도체를 이전에 설명된 바와 같이 분리하였다. McTs229의 다수의 FOA 저항성 콜로니를 프라이머 0613034 및 0613241로 원하는 루프-아웃 이벤트에 대해 PCR로 스크리닝하였다. 1.4 kbp 밴드의 존재는 원하는 바와 같이 *ura3* 스카 부위 (모체 균주에서 2개의 URA3 프로모터 사이에서 상동 재조합 후 단일 URA3 프로모터 왼쪽 뒤)만의 존재를 나타내는 한편, 약 2.8 kbp의 밴드는 온전한 URA3 프로모터-URA3 ORF-URA3 종결자-URA3 프로모터 카세트의 존재를 나타내며, 원하는 재조합 이벤트가 발생하지 않았다는 것을 나타낸다. 1.9 kbp 밴드의 존재는 이것들이 삭제에 대해 이형 접합체이기 때문에 이들 형질전환체에 존재하는 야생형 위치를 나타낸다. 프라이머 0611815 및 0611632는 약 4.2 kbp 밴드를 수득하였다; 프라이머 0611631 및 0612366은 *adh1202* 위치에서 ADC 카세트가 온전하다는 것을 확인하기 위해 약 4.5 kbp 밴드를 수득하였다. PCR 반응을 상기 설명된 바와 같이 Phire Plant Direct PCR Kit (Finnzymes)를 사용하여 수행하였다. 원하는 루프-아웃 이벤트를 갖는 모체 균주 McTs225의 하나의 FOA 저항성 콜로니를 McTs238로 지정하였다.
- [0659] SEQ ID NO: 26의 아이. 오리엔탈리스 3-HPDH를 암호화하는 네이티브 유전자의 동형 접합체 삭제를 생성하기 위해, McTs238을 분해된 pMcTs65에 형질전환하였고 올바른 위치 표적화 및 형질전환을 Phire Plant Direct PCR Kit (Finnzymes)를 사용하여 PCR로 확인하였다. 2개의 프라이머 세트를 YMR226c 위치 삭제에 대해 PCR로 스크리닝하기 위해 사용하였다. 프라이머 0613034 및 0613747은 3-HPDH의 동형 접합체 삭제가 발생하지 않는다는 것을 나타내는 3-HPDH 위치가 온전한 경우 약 500 bp 밴드를 수득할 것이다. 추가적으로 형질전환체를 3-HPDH가 삭제되지 않은 경우 약 660 bp 밴드를 수득하는 프라이머 0613746 및 0613241로 스크리닝하였다.
- [0660] 아이. 오리엔탈리스 3-HPDH 위치에 음성인 야생형인 형질전환체를 약 2.8 kbp 및 1.4 kbp 밴드를 수득하는 프라이머 0613034 및 0613241로 추가에 형질전환하였다; 프라이머 0612908 및 0613241은 약 1.5 kbp 밴드를 수득하였다; 프라이머 0613034 및 0612909는 약 1 kbp 밴드를 수득하였다. 약 3 kbp를 수득하는 프라이머 0611245 및 0612794 및 약 3.6 kbp 밴드를 수득하는 프라이머 0611815 및 0612795를 가지고 ADH1202 위치에서 ADC 카세트

가 온전한 것으로 확인하였다. 발현 카세트의 적절한 통합에 대해 예측된 밴드를 갖는 균주를 McTs244로 지정하였다.

[0661] McTs244의 ura-유도체를 이전에 설명된 바와 같이 분리하였다. McTs244의 다수의 FOA 저항성 콜로니를 약 1.4 kbp 밴드를 수득하는 프라이머 0613034 및 0613241로 원하는 루프-아웃 이벤트에 대해 PCR로 스크리닝하였다. ADH1202 위치에서 ADC 카세트를 약 3 kbp 밴드를 수득하는 프라이머 0611245 및 0612794 및 약 3.6 kbp 밴드를 수득하는 프라이머 0611815 및 0612795를 가지고 ADH1202 위치에서 ADC 카세트가 온전한 것으로 확인하였다. 루프-아웃 이벤트를 갖는 모체 균주 McTs244의 하나의 FOA 저항성 콜로니를 McTs259로 지정하였다.

[0662] 여기에 설명된 방법을 사용하여, 균주 M1Ba372 및 McTs244를 3-HP 생산에 대해 바이오리액터에서 테스트하였다. 이들 발효에서 세포 질량의 차이를 설명하기 위해, 3-HP 생산 능력은 세포 질량의 유닛 당 3-HP 농도로 표현된다 (g/L 3-HP/g/L 건조 세포 중량으로 표현됨). 균주 M1Ba372 및 McTs259에 대한 g/L 3-HP/g/L 건조 세포 중량은 각각 1.66 및 < 0.1이었다. 이들 결과는 이 유전자의 삭제가 3-HP 생산을 폐지하기 때문에, 아이. 오리엔탈리스의 네이티브 3-HPDH가 말로네이트 세미알데히드의 3-HP로 전환의 원인이 된다는 것을 나타낸다.

[0663] 실시예 3A-22: pdc 위치에서 피루베이트 카르복실라제 (PYC), 아스파르테이트 아미노트랜스퍼라제 (AAT), β -알라닌 아미노트랜스퍼라제 (BAAT), 및 3-히드록시프로피온산 데히드로게나제 (3-HPDH); 및 adh1202 위치에서 4개의 뉴클레오티드의 아스파르테이트 1-테카르복실라제 (ADC)를 발현하는 효모 균주

[0664] adh1202 위치에서 SEQ ID NO: 139의 비. 리체니포르미스 ADC를 암호화하는 뉴클레오티드의 4개의 카피를 또한 함유하는 균주에서 아이. 오리엔탈리스 pdc 위치에서 SEQ ID NO: 2의 아이. 오리엔탈리스 PYC를 암호화하는 뉴클레오티드 서열 SEQ ID NO: 1, SEQ ID NO: 14의 아이. 오리엔탈리스 AAT를 암호화하는 뉴클레오티드 서열 SEQ ID NO: 13, SEQ ID NO: 21의 에스. 클루이베리 BAAT를 암호화하는 뉴클레오티드 서열 SEQ ID NO: 142, 및 SEQ ID NO: 129의 에스. 세레비시에 3-HPDH를 암호화하는 뉴클레오티드 서열 SEQ ID NO: 144를 포함하도록 추가적인 구조물을 설계하였다. 상기 설명된 것과 유사한 접근법으로, 왼쪽 및 오른쪽 구조물을 아이. 오리엔탈리스 CNB1 pdc 위치에서 상동 재조합을 허용하도록 설계하였다. 이들 구조물을 제조하였고 하기 설명된 바와 같이 M1Ba357에 형질전환하였다.

[0665] 왼쪽 단편의 구조물

[0666] SEQ ID NO: 14의 아이. 오리엔탈리스 AAT를 암호화하는 뉴클레오티드 서열 SEQ ID NO: 13을 제조사의 지시에 따라 주형 (Pfu Polymerase, Stratagene)으로서 플라스미드 pGMEr126 (도 16)을 사용하여 PCR로 증폭하였다. 프라이머 oANN1 및 oANN2는 유전자 암호화 서열을 플랭킹하는 독특한 제한 부위의 통합을 허용하였다. PCR 생성물을 여기에 설명된 바와 같이 TBE 버퍼 중의 1% 아가로스 겔 전기영동으로 정제하였다. 약 1.3 kbp의 밴드를 겔에서 잘라냈고 제조사의 지시에 따라 Gel Extraction Kit (Qiagen)를 사용하여 정제하였다. 정제된 PCR 생성물을 ApaI 및 NruI로 분해하였고 여기에 설명된 바와 같이 겔 정제하였다.

[0667] 플라스미드 pGMEr135 (ENO1 프로모터/RKI 종결자 인서트가 반대 방향인 것을 제외하고, 상기 pGMEr136과 동일함)를 ApaI 및 NruI로 분해하였고 여기에 설명된 바와 같이 TBE 버퍼 중의 아가로스 겔 전기영동으로 정제하였다. 약 11.7 kbp의 밴드를 겔에서 잘라냈고 제조사의 지시에 따라 Gel Extraction Kit (Qiagen)를 사용하여 정제하였다.

[0668] 정제된 1.3 kbp PCR 생성물을 49.8 ng의 벡터, 354 ng의 1.3 kbp 인서트, 10mM ATP (New England Biolabs)가 들어있는 10X 결합 버퍼 1 μ L, 및 1 μ L T4 리가제 (New England Biolabs)로 구성된, 10 μ L의 총 반응량의 T4 리가제 (New England Biolabs)를 사용하여 11.7 kbp pGMEr135 선형화 벡터에 결합하였다. 반응물을 상온에서 30분 동안 배양하였고 반응물의 2 μ L 앨리퀀트를 제조사의 지시에 따라 일렉트로-컴피턴트 이. 콜리 DH10B 세포 (Invitrogen)에 형질전환하였다. 형질전환체를 LB + 카나마이신 플레이트에 도말하였고 37 °C에서 하룻밤 동안 배양하였다. 결과의 형질전환체 중 다수를 프라이머 oHJ2 및 oANN1 (약 2.3 kbp의 밴드를 수득) 및 프라이머 oANN5 및 oANN6 (약 877 bp의 밴드를 수득)로 콜로니 PCR에 의해 적절한 삽입에 대해 스크리닝하였다. 올바른 삽입물을 수득하는 클론 및 서열을 여기에 설명된 바와 같이 ApaI 및 NruI로 분해하였다. 약 1.3 kbp의 밴드를 겔에서 잘라냈고 Gel Extraction Kit (Qiagen)를 사용하여 정제하였다.

[0669] SEQ ID NO: 2의 아이. 오리엔탈리스 PYC를 암호화하는 원하는 뉴클레오티드 서열 SEQ ID NO: 1을 함유하는 플라스미드 pGMEr137 (상기; 도 18 참조)를 ApaI 및 NruI로 분해하였고 여기에 설명된 바와 같이 TBE 버퍼 중의 아가로스 겔 전기영동으로 정제하였다. 약 11.7 kbp의 밴드를 겔에서 잘라냈고 제조사의 지시에 따라 Gel Extraction Kit (Qiagen)를 사용하여 정제하였다.

- [0670] 정제된 1.3 kbp PCR 생성물을 49.4 ng의 벡터, 54 ng의 1.3 kbp 인서트, 10mM ATP (New England Biolabs)가 들어있는 10X 결찰 버퍼 1 μ L, 및 1 μ L T4 리가제 (New England Biolabs)로 구성된, 10 μ L의 총 반응량의 T4 리가제 (New England Biolabs)를 사용하여 11.7 kbp pGEM137 선형 벡터에 결찰하였다. 반응물을 상온에서 30 분 동안 배양하였고 반응물의 2 μ L 엘리퀴트를 제조사의 지시에 따라 일렉트로-컴피턴트 이. 콜리 DH10B 세포 (Invitrogen)에 형질전환하였다. 형질전환체를 LB + 카나마이신 플레이트에 도말하였고 37 $^{\circ}$ C에서 하룻밤 동안 배양하였다. 결과의 형질전환체 중 다수를 프라이머 oHJ2 및 oANN1 (약 2.3 kbp의 밴드를 수득함) 및 프라이머 oANN5 및 oANN6 (약 877 bp의 밴드를 수득함)로 콜로니 PCR에 의해 적절한 삽입에 대해 스크리닝하였다. 올바른 삽입물을 수득하는 클론 및 서열을 pANN02로 지정하였다.
- [0671] 반대 방향으로 서열을 암호화하는 ATT를 갖는 왼쪽 단편의 구조물
- [0672] 플라스미드 pANN02를 PmeI로 분해하였고 여기에 설명된 바와 같이 TBE 버퍼 중의 아가로스 겔 전기영동으로 정제하였다. 약 10.3 kbp의 밴드 및 약 2.7 kbp의 밴드를 겔에서 잘라냈고 제조사의 지시에 따라 Gel Extraction Kit (Qiagen)를 사용하여 정제하였다. pANN2의 10.3 kbp 벡터 단편을 CIP (New England Biolabs)로 탈인산화하였고 제조사의 지시에 따라 Purification Kit (Qiagen)로 정제하였다. pANN02의 2.7 kbp 단편을 36 ng의 벡터, 28 ng의 인서트, 10mM ATP (New England Biolabs)가 들어있는 10X 결찰 버퍼 1 μ L, 및 1 μ L T4 리가제 (New England Biolabs)로 구성된, 10 μ L의 총 반응량의 T4 리가제 (New England Biolabs)를 사용하여 pANN02의 탈인산화된 10.3 kbp 선형화 벡터에 결찰하였다. 반응물을 상온에서 30분 동안 배양하였고 반응물의 2 μ L 엘리퀴트를 제조사의 지시에 따라 일렉트로-컴피턴트 이. 콜리 TOP10 세포 (Invitrogen)에 형질전환하였다. 형질전환체를 LB + 카나마이신 플레이트에 도말하였고 37 $^{\circ}$ C에서 하룻밤 동안 배양하였다. 결과의 형질전환체 중 다수를 프라이머 oJY44 및 oHJ1 (약 1.3 kbp의 밴드를 수득함)로 콜로니 PCR로 적절한 삽입에 대해 스크리닝하였다. 올바른 삽입물을 수득하는 클론을 pANN5로 지정하였다.
- [0673] 오른쪽 단편의 구조물
- [0674] 2개의 비. 리체니포르미스 ADC 암호화 영역 및 아이. 오리엔탈리스 PDC 위치 3' 표적화 플랭킹 DNA를 함유하는 오른쪽 구조물을 다음과 같이 구성하였다. pMhCt071 플라스미드 (에스. 세레비시에 3-HPDH ORF가 아이. 오리엔탈리스에 최적화된 코돈이 아닌 것을 제외하고 상기 pMhCt077과 동일한 플라스미드)를 PmeI 및 PacI로 분해하였고, 10 유닛 소 위장 포스포타제 (New England Biolabs)를 처리하였고, TAE 버퍼 중의 0.9% 아가로스 겔 전기영동으로 정제하였고, 약 4.7 kbp 밴드를 겔에서 잘라냈고 제조사의 지시에 따라 NUCLEOSPIN® Extract II Kit를 사용하여 정제하였다.
- [0675] 플라스미드 pMeJi312-2 (상기; 도 15 참조)를 2개의 비. 리체니포르미스 ADC 발현 카세트를 추출하기 위해 PmeI 및 PacI로 분해하였고 TAE 버퍼 중의 0.9% 아가로스 겔 전기영동으로 정제하였다. 약 2.8 kbp 밴드를 겔에서 잘라냈고 제조사의 지시에 따라 NUCLEOSPIN® Extract II Kit를 사용하여 정제하였다.
- [0676] 1X QUICK 결찰 버퍼 (New England Biolabs), 1 μ L pMhCt071 벡터의 4.7 kbp 단편, 3 μ L pMeJi312-2의 2.8 kbp 인서트, 및 1 μ L QUICK T4 DNA 리가제 (New England Biolabs)를 함유하는 결찰 반응물 (20 μ L)에서 pMeJi312-2의 이중 비. 리체니포르미스 ADC 암호화 영역을 함유하는 단편을 선형화 pMhCt071 벡터 단편에 결찰하였다. 결찰 반응물을 상온에서 5분 동안 배양하였고, 튜브를 얼음 위에 두었다. 이 반응물 중 5 μ L를 제조사의 지시에 따라 SoloPack Gold SuperCompetent 세포 (Agilent Technologies)를 형질전환하기 위해 사용하였다. 형질전환체를 2X YT+ amp 플레이트에 도말하였고 37 $^{\circ}$ C에서 하룻밤 동안 배양하였다. 결과의 형질전환체 중 다수를 ApaLI 분해에 의해 원하는 단편의 적절한 결찰에 대해 스크리닝하였다. 원하는 밴드 크기를 수득하는 클론을 유지하였고 pMhCt110으로 지정하였다.
- [0677] 플라스미드 pMhCt110을 XbaI 및 PacI로 분해한 후 이어서 CIP를 처리하였고 여기에 설명된 바와 같이 TBE 버퍼 중의 1% 아가로스 겔 전기영동으로 정제하였다. 약 7.1 kbp에서 밴드를 겔에서 잘라냈고 제조사의 지시에 따라 QIAQUICK® Gel Extraction Kit (Qiagen)를 사용하여 정제하였다.
- [0678] SEQ ID NO: 144의 코돈-최적화된 에스. 세레비시에 3-HPDH 암호화 서열을 XbaI 및 PacI로 분해에 의해 pM1Ba123 (상기)로부터 잘라냈고 여기에 설명된 바와 같이 TBE 버퍼 중의 1% 아가로스 겔 전기영동으로 정제하였다. 약 814 bp에서 밴드를 겔에서 잘라냈고 제조사의 지시에 따라 QIAQUICK® Gel Extraction Kit (Qiagen)를 사용하여 정제하였다. 814 bp 정제된 단편을 1 μ L의 분해된 pMhCt110, 1 또는 pM1Ba123의 814 bp 단편 7 μ L, 10mM ATP (New England Biolabs)가 들어있는 10X 결찰 버퍼 1 μ L, 및 1 μ L T4 리가제 (New England Biolabs)로 구성된, 10 μ L의 총 반응량의 T4 리가제 (New England Biolabs)를 사용하여 pMhCt110의 7.1 kbp

단편에 결찰하였다. 반응물을 상온에서 30분 동안 배양하였고 반응물의 4 μ L 엘리쿼트를 제조사의 지시에 따라 ONE SHOT® TOP10 화학적 컴피턴트 이. 콜리 세포 (Invitrogen)에 형질전환하였다. 형질전환체를 2X YT+amp 플레이트에 도말하였고 상온에서 일주일 동안 배양하였다. 결과의 형질전환체 중 다수를 효소 XbaI 및 PacI 및 AscI 및 PacI의 2개의 조합을 사용하는 제한 분해에 의해 적절한 삽입에 대해 스크리닝하였다. 각 분해물에서 올바른 분해된 밴드 크기를 수득하는 클론을 pMIBa142로 지정하였다.

[0679] 플라스미드 pMIBa142를 AscI로 분해한 후 이어서 클레노브로 필-인 반응시켰고 그 다음 NheI로 분해하고 CIP를 처리하였다. 분해물을 여기에 설명된 바와 같이 TBE 버퍼 중의 1% 아가로스 겔 전기영동으로 정제하였다. 약 7.5 kbp에서 밴드를 겔에서 잘라냈고 제조사의 지시에 따라 QIAQUICK® Gel Extraction Kit (Qiagen)를 사용하여 정제하였다.

[0680] SEQ ID NO: 21의 에스. 클루이베리 BAAT를 암호화하는 뉴클레오티드 서열 SEQ ID NO: 142를 PacI로 분해에 의해 pMIBa124 (상기)로부터 잘라낸 후 이어서 클레노브로 필-인 반응시켰고 그 다음 XbaI로 분해하였다. 분해물을 여기에 설명된 바와 같이 TBE 버퍼 중의 1% 아가로스 겔 전기영동으로 정제하였다. 약 1.4 kbp에서 밴드를 겔에서 잘라냈고 제조사의 지시에 따라 QIAQUICK® Gel Extraction Kit (Qiagen)를 사용하여 정제하였다.

[0681] pMIBa124로부터 1.4 kbp 정제된 단편을 1 μ L의 분해된 pMIBa142, pMIBa124의 1.4 단편 7 μ L, 10mM ATP (New England Biolabs)가 들어있는 10X 결찰 버퍼 1 μ L, 및 1 μ L T4 리가제 (New England Biolabs)로 구성된, 10 μ L의 총 반응량의 T4 리가제 (New England Biolabs)를 사용하여 pMIBa142의 7.5 kbp 단편에 결찰하였다. 반응물을 16 °C에서 2.5시간 동안 배양하였고 전체 결찰물을 제조사의 지시에 따라 Sure Cell (Agilent)에 형질전환하였다. 형질전환체를 2X YT+amp 플레이트에 도말하였고 37 °C에서 하룻밤 동안 배양하였다. 결과의 형질전환체 중 다수를 StuI 및 PmeI로 제한 분해에 의해 적절한 삽입에 대해 스크리닝하였다. 올바른 분해된 밴드 크기를 수득하는 클론을 pMIBa144로 지정하였다.

[0682] 왼쪽 및 오른쪽 단편의 MIBa357로 통합

[0683] 여기에 설명된 바와 같이 플라스미드 pANN5를 NotI 및 NheI로 분해하였고 플라스미드 pMIBa144를 NotI로 분해하였다. 이것들을 TBE 버퍼 중의 1% 아가로스 겔 전기영동으로 정제하였고, pANN5의 8.2 kbp 단편 및 pMIBa144의 6 kbp 단편을 겔에서 잘라냈고 제조사의 지시에 따라 NUCLEOSPIN® Extract II Kit (Macherey-Nagel)를 사용하여 정제하였다.

[0684] MIBa357을 분해된 pANN5 및 pMIBa144 DNA에 형질전환하였고 올바른 위치 표적화 및 형질전환을 제조사의 지시에 따라 Phire® Plant Direct PCR Kit (Finnzymes)를 사용하여 PCR로 확인하였다. 프라이머 0611552 및 0613695는 약 4.1 kbp 밴드를 수득하였다 (pdc 위치에서 왼쪽 절반 통합을 확인하기 위해); 프라이머 0612358 및 0611554는 약 2.5 kbp 밴드를 수득하였다 (pdc 위치에서 오른쪽 절반 통합을 확인하기 위해); 0611245 및 0611631은 약 3.5 kbp 밴드를 수득하였다 (왼쪽 절반 4X ADC 통합이 adh1202 위치에 남아있다는 것을 확인하기 위해); 및 프라이머 0611815 및 0611632는 약 4.6 kbp 밴드를 수득하였다 (오른쪽 절반 4X ADC 통합이 adh1202 위치에 남아있다는 것을 확인하기 위해). pdc 위치에서 발현 카세트의 적절한 통합에 대해 예측된 밴드를 제공하고 adh1202 위치에서 발현 카세트를 유지하는 하나의 분리체를 저장하였고 MIBa360으로 지정하였다.

[0685] MIBa360의 ura 마커의 제거

[0686] MIBa360의 ura-유도체를 상기 설명된 바와 같이 분리하였다. MIBa360의 다수의 FOA 저항성 콜로니의 게놈 DNA를 프라이머 0611815 및 0613689로 원하는 루프-아웃 이벤트에 대해 PCR로 스크리닝하였다. 약 1.9 kbp 밴드의 존재는 원하는 바와 같이 ura3 스카 부위 (모체 균주에서 2개의 URA3 프로모터 사이에서 상동 재조합 후 단일 URA3 프로모터 왼쪽 뒤)를 갖는 ura 마커의 제거를 나타내는 한편, 약 3.3 kbp의 밴드는 온전한 URA3 프로모터-URA3 ORF-URA3 종결자-URA3 프로모터 카세트의 존재를 나타내며, 원하는 재조합 이벤트가 발생하지 않았다는 것을 나타낸다. PCR을 제조사의 지시에 따라 Phire® Plant Direct PCR kit (Finnzymes)를 사용하여 수행하였다. 상기 프라이머로 약 1.9 kbp 단편을 수득하는 2개의 FOA 저항성 콜로니를 저장하였고 MIBa363 및 MIBa364로 지정하였다.

[0687] 역 발현 카세트 오른쪽 단편의 구조물

[0688] 플라스미드 pMIBa144는 원하는 ADC 및 정방향으로 가는 3-HPDH 발현 카세트를 함유한다. 동형 접합체 균주의 스크리닝을 쉽게 하기 위해, pMIBa144의 ADC 발현 카세트가 역방향에 위치하는 경우 새로운 플라스미드를 구성하였다. 플라스미드 pMIBa144 (상기)를 StuI 및 PmeI로 분해하였고 TBE 버퍼 중의 1% 아가로스 겔 전기영동으로 정제하였다. 여기에 설명된 바와 같이. 약 6.1 kbp 및 2.8 kbp에서 밴드를 겔에서 잘라냈고 제조사의 지시에 따

라 QIAQUICK® Gel Extraction Kit (Qiagen)를 사용하여 정제하였다. pM1Ba144의 6.1 kbp 단편에 CIP를 처리하였고 제조사의 지시에 따라 QIAQUICK® PCR Purification Kit (Qiagen)를 사용하여 정제하였다.

[0689] pM1Ba144의 2.8 kbp 정제된 단편을 pM1Ba144의 6.1 kbp 단편 1 μ L, pM1Ba144의 2.8 kbp 단편 7 μ L, 10mM ATP (New England Biolabs)가 들어있는 10X 결찰 버퍼 1 μ L, 및 1 μ L T4 리가제 (New England Biolabs)로 구성된, 10 μ L의 총 반응량의 T4 리가제 (New England Biolabs)를 사용하여 6.1 kbp pM1Ba144 선형화 벡터에 결찰하였다. 반응물을 16 °C에서 4시간 동안 배양하였고 반응물의 4 μ L 엘리퀴트를 제조사의 지시에 따라 ONE SHOT® TOP10 화학적 컴피턴트 이. 콜리 세포 (Invitrogen)에 형질전환하였다. 형질전환체를 2X YT+amp 플레이트에 도말하였고 37 °C에서 하룻밤 동안 배양하였다. 결과의 형질전환체 중 다수를 SphI 및 XbaI를 사용하는 제한 분해에 의해 적절한 삽입에 대해 스크리닝하였다. 올바른 분해된 밴드 크기를 수득하는 클론을 pM1Ba146으로 지정하였다.

[0690] 왼쪽 및 오른쪽 단편의 M1Ba363으로 통합

[0691] 여기에 설명된 바와 같이 플라스미드 pANN5를 NotI 및 NheI로 분해하였고 플라스미드 pM1Ba146을 NotI로 분해하였다. 이것들을 TBE 버퍼 중의 1% 아가로스 겔 전기영동으로 정제하였고, pANN5의 8.2 kbp 단편 및 pM1Ba146의 6.2 kbp 단편을 겔에서 잘라냈고 제조사의 지시에 따라 QIAQUICK® PCR Purification Kit (Qiagen)를 사용하여 정제하였다.

[0692] M1Ba363을 분해된 pANN5 및 pM1Ba144 DNA에 형질전환하였고 올바른 위치 표적화 및 형질전환을 제조사의 지시에 따라 Phire® Plant Direct PCR kit (Finnzymes) 또는 Kapa Robust DNA Polymerase를 사용하여 PCR로 확인하였다. pdc 위치에서 통합을 확인하기 위해, 다음 프라이머 쌍을 사용하였다. 프라이머 0613689 및 0611815는 약 1.9 kbp 밴드를 수득하였다; 프라이머 0612366 및 0611554는 약 2.5 kbp 밴드를 수득하였다; 0613688 및 0611815는 약 3.2 kbp 밴드를 수득하였다; 0611622 및 0611552는 약 945 bp 밴드를 수득하였다. adh1202에서 통합을 체크하기 위해 다음 프라이머 쌍을 사용하였다. 프라이머 0611245 및 0612794는 약 2.8 kbp 밴드를 수득하였고 프라이머 0611815 및 0612795는 약 3.9 kbp 밴드를 수득하였다. pdc 위치에서 발현 카세트의 적절한 통합에 대해 예측된 밴드를 제공하고 adh1202 위치에서 발현 카세트를 유지하는 2개의 분리체를 저장하였고 M1Ba372 및 M1Ba373으로 지정하였다.

[0693] M1Ba372의 ura 마커의 제거

[0694] M1Ba372의 ura-유도체를 상기 설명된 바와 같이 분리하였다. M1Ba372의 다수의 FOA 저항성 콜로니의 게놈 DNA를 프라이머 0611815 및 0613688로 원하는 루프-아웃 이벤트에 대해 PCR로 스크리닝하였다. 약 2.1 kbp 밴드의 존재는 원하는 바와 같이 ura3 스카 부위 (모체 균주에서 2개의 URA3 프로모터 사이에서 상동 재조합 후 단일 URA3 프로모터 왼쪽 뒤)를 갖는 ura 마커의 제거를 나타내는 한편, 약 3.3 kbp의 밴드는 온전한 URA3 프로모터-URA3 ORF-URA3 종결자-URA3 프로모터 카세트의 존재를 나타내며, 원하는 재조합 이벤트가 발생하지 않았다는 것을 나타낸다. M1Ba372의 FOA 저항성 콜로니를 또한 첫 번째 염색체의 변형을 확인하기 위해 프라이머 0611815 및 0613689 (1.9 kbp 단편을 증폭시킨다) 및 pdc 위치의 손실을 확인하기 위해 0611552 및 0611553 (pdc 위치가 존재하는 경우 850 bp 단편을 증폭시킨다)을 가지고 PCR로 스크리닝하였다. PCR을 제조사의 지시에 따라 Phire® Plant Direct PCR kit (Finnzymes)을 사용하여 수행하였다. 0611815 및 0613689로 약 1.9 kbp 단편을 수득하지만, 0613688 및 0611815 또는 0611552 및 0611553으로 단편을 증폭시키지 않는 FOA 저항성 콜로니를 저장하였고 M1Ba375로 지정하였다. M1Ba375의 유전자형은 표 28에 나타난다.

[0695] 표 28: 형질전환체 유전자형

| 균주 | 유전자형 |
|---------|--|
| M1Ba375 | adh1202Δ:: (PDC 프로모터-Opt.Blpand PDC 프로모터-Opt.Blpand Ura3-스카 TDH3 프로모터 Opt.Blpand TDH3 프로모터 Opt.Blpand)/adh1202Δ:: (PDC 프로모터-Opt.Blpand PDC 프로모터-Opt.Blpand Ura3-스카 [TDH3 프로모터 Opt.Blpand TDH3 프로모터 Opt.Blpand]역방향) pdcΔ:: (PDC 프로모터-CNB1pyc ENO1 프로모터-CNB1aat Ura3-스카 TDH3 프로모터-OptSkPYD4 PGK1 프로모터-OptScYMR226c/ pdcΔ:: (PDC 프로모터-CNB1pyc ENO1 프로모터-CNB1aat Ura3-스카 TDH3 프로모터-OptSkPYD4 PGK1 프로모터-OptScYMR226c ura3-/ura3- |

[0696]

[0697] 실시예 3A-23: 글리세롤 3-포스페이트 데히드로게나제 (GPD) 유전자에 대하여 삭제된 효모 균주

- [0698] 균주 MIBa375의 GPD 유전자의 삭제
- [0699] 추가적 구조물을 숙주 아이. 오리엔탈리스 게놈의 글리세롤 3-포스페이트 데히드로게나제 유전자 (SEQ ID NO: 117, 이것은 SEQ ID NO: 118의 GPD를 암호화한다)의 카피 둘 다를 삭제하도록 설계하였다. 이들 구조물은 GPD 유전자의 서열 업스트림에 대하여 상동성인 약 1003 bp의 뉴클레오티드 서열 및 GPD 유전자의 서열 다운스트림에 대하여 상동성인 약 852 bp의 서열을, 사이에서 클로닝된 a T_{PDC}-URA3 마커 카세트 (PDC 종결자-URA3 프로모터-URA3 유전자-URA3 종결자-URA3 프로모터)와 함께, 함유하였다.
- [0700] GPD의 영역 업스트림 및 다운스트림을 제조사의 지시에 따라 Pfu DNA Polymerase를 사용하여 아이. 오리엔탈리스 CNB1 게놈 DNA로부터 증폭시켰다. 업스트림 영역은 ApaI 부위를 함유하였다; 이것을 ApaI 인식 서열에서 뉴클레오티드 중 하나에 대하여 mismatches로 설계된 중첩된 프라이머를 사용하는 PCR로 제거하였다. 프라이머 쌍 oACN48/oACN51 및 oACN49/oACN50을 사용하여 업스트림 영역에 대한 이들 2개의 중첩된 단편을 증폭시켰다. 이들 PCR 생성물을 1% 아가로스 겔 상에서 분리하였고 이로부터 잘라냈고 제조사의 지시에 따라 Qiaquick Gel Extraction Kit (Qiagen)를 사용하여 정제하였고, 프라이머 oACN48/oACN49 (정방향 ApaI/역방향 NotI 부위를 가짐)로 PCR의 두 번째 라운드에 대한 주형으로서 사용하였다. 다운스트림 영역을 제조사의 지시에 따라 Pfu DNA Polymerase를 사용하여 아이. 오리엔탈리스 게놈 DNA의 프라이머 oACN52/oACN53 (정방향 NotI/역방향 SacI 부위를 가짐)로 증폭시켰다. PCR 생성물 둘 다를 겔 정제하였고 벡터 pCR® BluntII-TOPO® (Invitrogen)에 별도로 클로닝하였다. 분리체를 플라스미드 DNA의 제한 분해에 의해 원하는 인서를 갖는 것으로 확인하였고, 시퀀싱으로 확인하였다. 벡터 pACN58은 클로닝된 업스트림 단편을 함유하였고 pACN59는 다운스트림 단편을 함유하였다. 플라스미드 pACN58을 업스트림 플랭크를 방출하기 위해 ApaI/NotI로 분해하였고, 플라스미드 pACN59를 다운스트림 플랭크를 방출하기 위해 SacI/NotI로 분해하였고 pACN59를 벡터 백본을 제공하기 위해 ApaI/SacI로 분해하였다. 3개의 원하는 단편을 1% 아가로스 겔 상에서 분리하였고, 잘라냈고 정제하였고, T4 리가제 (New England Biolabs)를 사용하는 3-조각 결찰 반응으로 결찰하였다. 결찰 반응물을 이. 콜리 TOP10 일렉트로컴피턴트 세포에 형질전환하였고 형질전환체를 플라스미드 DNA의 제한 분해로 확인하였다. 이 공정 중에, 다운스트림 영역의 추가적인 SacI 부위를 검출하였으며, 이것은 853 bp의 다운스트림 영역을 생성하였다 (1 kbp에서 반대). 원하는 인서를 갖는 2개의 분리체를 pACN62 및 pACN63으로 명명하였다.
- [0701] T_{PDC1}-URA3 카세트를 NotI 분해 및 겔 정제를 사용하여 벡터 pJLJ8 (도 32)로부터 분리하였다. 이 단편을 NotI로 분해되고 탈인산화된 pACN62 (상기)에 결찰하였고, 결찰물을 이. 콜리 DH10B 일렉트로컴피턴트 세포 (Invitrogen)에 형질전환하였다. URA3 인서를 갖는 콜로니를 프라이머 oACN48/oJLJ44 및 oACN48/oJLJ46을 사용하여 PCR로 확인하였다. 프라이머 oJLJ44는 다운스트림 URA3 프로모터의 말단에서 어닐링하고 T_{PDC1}-URA3 카세트로부터 바깥쪽으로 증폭시킨다. 프라이머 oJLJ46은 PDC 종결자의 5' 말단에서 어닐링하고 T_{PDC1}-URA3 카세트로부터 바깥쪽으로 증폭시킨다. 벡터 pHJJ56은 다운스트림 GPD 영역에 직면한 URA3을 함유하고 pHJJ57은 업스트림 GPD 영역에 직면한 URA3을 함유한다.
- [0702] 플라스미드 pHJJ56 및 pHJJ57을 KpnI 및 ApaI로 분해에 의해 선형화하였고 삭제 카세트를 함유하는 단편을 겔 추출로 정제하였다. 선형화된 pHJJ56을 ura-균주 MIBa375에 형질전환하였다. 단일 콜로니를 정제를 위해 제스트 리킹하였고 프라이머 oJLJ44, oJLJ46, oACN54 및 oACN55를 사용하여 원하는 GPD 삭제를 위해 PCR로 테스트하였다. 세포를 40 µL Y-용해 버퍼 및 2 µL 지물리아제 (ZymoResearch)에서 37 °C에서 30분 동안 용해시켰고 용해 반응물 중 1 µL를 25 µL PCR 반응물에서 사용하였다. PCR 반응물은 55 °C의 어닐링 온도 및 다음 사이클링 프로파일을 가지고, 제조사의 지시에 따라 Failsafe DNA 폴리머라제 및 버퍼 E를 사용하였다: 94 °C에서 2분 동안 1 사이클; 각각 94 °C에서 30초 동안, 55 °C에서 30초 동안, 및 72 °C에서 1.5분 동안 29 사이클; 및 72 °C에서 3분 동안 1 사이클. GPD 너아웃의 하나의 카피를 갖는 균주는 약 0.9 및 1.2 kbp의 밴드를 생산하였고 yHJN1 및 yHJN2로 명명하였다.
- [0703] 균주 yHJN1 및 yHJN2를 YPD 배지에서 하룻밤 동안 키웠고 URA3 마커의 손실로 선택하기 위해 ScD-2X FOA 배지에 도말하였다. 단일 콜로니를 YPD 상에서 정제하였고 ura-표현형을 확인하기 위해 ScD-ura 및 YPD 배지에 패치하였다. Ura-콜로니를 첫 번째 통합을 확인하기 위해 사용된 같은 PCR 반응물을 사용하여 너아웃을 유지하는지 확인하였다. yHJN1의 ura-유도체를 yHJN3으로 명명하였고 yHJN2의 ura-유도체를 yHJN4로 명명하였다.
- [0704] 선형화된 pHJJ57을 yHJN3 및 yHJN4에 형질전환하였고 단일 콜로니를 ScD-ura 배지 상에서 정제하였다. GPD 너아웃의 2개의 카피의 존재를 한 반응에서 프라이머 oJLJ44, oACN54, 및 oACN55를 사용하여, 및 두 번째 반응에서 프라이머 oJLJ46, oACN54, 및 oACN55를 사용하여 PCR로 확인하였다. 프라이머 oACN54는 GPD에 대한 업스트림

플랭킹 서열의 약 37 bp 업스트림 영역으로 어닐링하는 한편, oACN55는 다운스트림 플랭크의 약 24 bp 다운스트림 영역으로 어닐링한다. 전 반응은 GPD의 카피 둘 다가 삭제되는 경우 약 900 및 1050 bp의 밴드를 생산하고, 후 반응은 약 1025 및 1200 bp의 밴드를 생산한다. GPD Nick아웃의 2개의 카피를 갖는 콜로니는 삭제의 단일 카피를 갖는 것들보다 ScD-ura 플레이트에서 더 천천히 자랐다. 삭제된 GPD 유전자의 2개의 카피를 갖는 균주를 yHJN7 (yHJN3으로부터 유도됨) 및 yHJN8 (yHJN4로부터 유도됨)로 명명하였다.

[0705] 균주 M1Ba372, yHJN7 및 yHJN8을 여기에 설명된 방법을 사용하여, 글리세롤 생산에 대해 바이오리액터에서 테스트하였다. 균주 M1Ba372는 48시간에 29.5 g/L 글리세롤을 생산하였다. 검출 불가능한 글리세롤은 발효 중에 균주 yHJN7 또는 yHJN8에 의해 생산되었다. 최종 발효 브로스에서 글리세롤의 부재는 발효 브로스로부터 3-HP의 회수 및 정제시 이점을 제공할 수도 있다.

[0706] 3A-24: adh1202 위치에서 4개의 뉴클레오타이드 서열의 아스파르트레이트 1-데카르복실라제 (ADC), adh9091 위치에서 3-HP 데히드로게나제 (3-HPDH) 및 네이티브 3-HP 데히드로게나제 (3-HPDH)의 삭제를 발현하는 효모 균주

[0707] adh9091 위치에서 아이. 오리엔탈리스 3-HPDH의 통합을 위한 플라스미드 구조물

[0708] SEQ ID NO: 26의 아이. 오리엔탈리스 3-HPDH를 암호화하는 SEQ ID NO: 25 및 SEQ ID NO: 20의 아이. 오리엔탈리스 BAAT (PYD4)를 암호화하는 SEQ ID NO: 19의 뉴클레오타이드 서열을 프라이머 0611954 및 0611957 (3-HPDH에 대해) 또는 0611997 및 0611998 (PYD4)을 사용하여 이전에 설명된 바와 같이 제조된 MBin500 아이. 오리엔탈리스 게놈 DNA로부터 증폭시켰다. 증폭 중에 프라이머 0611954는 PYD4의 5' 말단에 kozak 서열 (TAAA) 및 NheI 부위를 추가하였고, 프라이머 0611957은 3-HPDH의 3' 말단에 PacI 부위를 추가하였다. 증폭 중에 프라이머 0611997은 PYD4의 5' 말단에 kozak 서열 (TAAA) 및 PacI 부위를 추가하였고, 프라이머 0611998은 3-HPDH의 3' 말단에 PacI 부위를 추가하였다. 각각의 프라이머의 50 pmol을 50 µL의 최종 부피에 주형으로서 MBin500 게놈 DNA 50 ng, 0.2 mM 각각의 dATP, dGTP, dCTP, dTTP, 1X Expand High-Fidelity Buffer (Roche), 및 2.6 유닛의 Expand High-Fidelity DNA Polymerase (Roche)를 함유하는 PCR 반응물에서 사용하였다. PCR을 95 °C에서 3분 동안 1 사이클 후 이어서 각각 95 °C에서 1분 동안, 55 °C에서 1분 동안, 및 72 °C에서 1분 (3-HPDH PCR) 또는 2분 (PYD4 PCR) 동안 30 사이클, 72 °C에서 5분 동안 최종 연장으로 프로그램된 EPPENDORF® MASTERCYCLER® (Eppendorf Scientific)로 수행하였다. 열 순환 후, 약 831 bp 3-HPDH 또는 1.4 kbp PYD4 PCR 생성물을 겔에서 잘라냈고 제조사의 지시에 따라 QIAQUICK® Gel Extraction Kit (Qiagen)를 사용하여 정제한 경우, PCR 반응 생성물을 TAE 버퍼 중의 1.0% 아가로스 겔 전기영동으로 분리하였다. 5 µL의 정제된 3-HPDH 또는 PYD4를 제조사의 지시에 따라 TOPO-TA Cloning Kit (Invitrogen)를 사용하여 pCR2.1 (Invitrogen)에 클로닝하였다.

[0709] 형질전환체를 2X YT + amp 플레이트에 도말하였고 37 °C에서 하룻밤 동안 배양하였다. 결과의 형질전환체 중 다수를 EcoRI으로 분해에 의해 스크리닝하였다. 원하는 인서트 크기를 수득하는 클론을 DNA 시퀀싱에 의해 올바른 것으로 확인하였다. 3-HPDH를 함유하는 하나의 클론을 pMBin190으로 지정하였고, PYD4를 함유하는 또 다른 하나를 pMBin193으로 지정하였다.

[0710] 1.4 kbp PYD4 밴드를 겔에서 잘라냈고 제조사의 지시에 따라 QIAQUICK® Gel Extraction Kit (Qiagen)를 사용하여 정제한 경우, 플라스미드 pMBin193을 XbaI 및 PacI로 분해하였고 TAE 버퍼 중의 1.0% 아가로스 겔에서 실행하였다. 분해된 PYD4 단편을 T4 DNA 리가제를 사용하여 XbaI 및 PacI 제한된 선형 pM1Ba107 플라스미드 (상기)에 결합하였다. 결합 생성물을 제조사의 지시에 따라 One Shot® TOP10 화학적 컴피턴트 이. 콜리 세포 (Invitrogen)에 형질전환하였다. 형질전환체를 2X YT + amp 플레이트에 도말하였고 37 °C에서 하룻밤 동안 배양하였다. 결과의 형질전환체 중 다수를 SnaBI 또는 EcoRI로 분해에 의해 스크리닝하였다. 원하는 밴드 크기를 수득하는 하나의 클론을 pMBin203로 지정하였다. 플라스미드 pMBin203은 다음 발현 카세트를 플랭크하는 NotI 부위를 함유한다: PDC 프로모터 및 PYD4 CDS의 종결자 업 및 다운스트림, URA3 프로모터, URA3 ORF, 및 URA3 종결자에 이어서 URA3 프로모터.

[0711] 약 4.1 kbp 단편을 (PDC 프로모터, PYD4 CDS, PDC 종결자, 및 URA3 선택 마커를 함유) 겔에서 잘라냈고 제조사의 지시에 따라 QIAQUICK® Gel Extraction Kit (Qiagen)를 사용하여 정제한 경우, 플라스미드 pMBin203를 NotI로 분해하였고 TAE 버퍼 중의 1.0% 아가로스 겔 상에서 분리하였다. 약 5.7 kbp 선형 플라스미드를 상기 설명된 바와 같이 정제한 경우, 플라스미드 pHJJ27 (adh9091 위치에 대한 5' 및 3' 상동 영역 함유; 도 21 참조)를 NotI으로 분해하였고, CIP를 처리하였고 TAE 버퍼 중의 1.0% 아가로스 겔 상에서 분리하였다. pMBin203의 단편을 상기 설명된 바와 같이 T4 DNA 리가제를 사용하여 NotI 제한된 pHJJ27에 결합하였다. 결합 생성물을 제조사의 지시에 따라 One Shot® TOP10 화학적 컴피턴트 이. 콜리 세포 (Invitrogen)에 형질전환하였다. 형질전환체를 2X YT + amp 플레이트에 도말하였고 37 °C에서 하룻밤 동안 배양하였다. 결과의 형질전환체 중 다수를

PstI로 분해에 의해 스크리닝하였다. 원하는 밴드 크기를 수득하는 하나의 클론을 pMBin204로 지정하였다. 플라스미드 pMBin204는 adh9091 위치에 대한 PYD4의 표적화를 허용한다.

- [0712] SEQ ID NO: 26의 아이. 오리엔탈리스 3-HPDH를 암호화하는 SEQ ID NO: 25의 뉴클레오타이드 서열을 NheI 및 PacI로 분해에 의해 플라스미드 pMBin190 (상기)으로부터 제거하였고 여기에 설명된 바와 같이 TBE 버퍼 중의 아가로스 겔 전기영동으로 정제하였다. 약 827 bp의 밴드를 겔에서 잘라냈고 제조사의 지시에 따라 NUCLEOSPIN® Extract II Kit (Macherey-Nagel)를 사용하여 정제하였다. 플라스미드 pMBin204 (상기)를 XbaI 및 PacI로 분해하였고 여기에 설명된 바와 같이 TBE 버퍼 중의 아가로스 겔 전기영동으로 정제하였다. 약 8.4 kbp의 밴드를 겔에서 잘라냈고 제조사의 지시에 따라 NUCLEOSPIN® Extract II Kit (Macherey-Nagel)를 사용하여 정제하였다.
- [0713] 상기 정제된 약 827 bp 아이. 오리엔탈리스 3-HPDH 유전자 생성물을 1 µL의 8.4 kbp 벡터, 10 µL의 827 bp 인서트, 10 mM ATP (New England Biolabs)가 들어있는 10X 결찰 버퍼 2 µL, 및 1 µL T4 리가제 (New England Biolabs)로 구성된, 20 µL의 총 반응량의 T4 리가제 (New England Biolabs)를 사용하여 8.4 kbp pMBin204 선형화 벡터에 결찰하였다. 반응물을 16 °C에서 18시간 동안 배양하였고 반응물의 10 µL 엘리퀀트를 제조사의 지시에 따라 One Shot TOP10 세포 (Invitrogen)에 형질전환하였다. 회수 기간 후, 형질전환 반응물의 2개의 100 µL 엘리퀀트를 mL 당 100 µg 앰피실린이 보충된 150 mM 2X YT 플레이트에 도말하였다. 플레이트를 37 °C에서 하룻밤 동안 배양하였다. 추정의 재조합 클론을 선택 플레이트로부터 선택하였고 플라스미드 DNA를 BIOROBOT® 9600 (Qiagen)를 사용하여 각각 제조하였다. 클론을 제한 분해로 분석하였고 올바른 제한 분해 패턴을 갖는 플라스미드를 pMcTs90으로 지정하였다.
- [0714] adh9091 위치에서 에스. 세레비시애 3-HPDH의 통합을 위한 플라스미드 구조물
- [0715] SEQ ID NO: 129의 에스. 세레비시애 3-HPDH를 암호화하는 826 bp 야생형 뉴클레오타이드 서열을 JGI69 게놈 DNA로부터 PCR 증폭하였고 유전자의 5' 말단에서 XbaI 부위 및 유전자의 3' 말단에서 PacI 부위로 수정하였다. 증폭 반응을 제조사의 지시에 따라 Platinum® Pfx DNA Polymerase (Invitrogen)를 사용하여 수행하였다. 마스터 PCR 반응물은 200 µL의 최종 부피에 1.125 µL의 에스. 세레비시애 게놈 DNA, 112.5 pM 각각의 프라이머 611191 및 611199, 1X Pfx 증폭 버퍼 (Invitrogen), 2 mM MgSO₄, 0.2 mM dNTP 믹스, 5 유닛 Platinum® Pfx DNA Polymerase (Invitrogen)를 함유하였다. 믹스를 8개의 튜브로 엘리퀀트하였고 구배 PCR을 수행하였다. 증폭 반응물을 95 °C에서 2분 동안 1 사이클; 각각 95 °C에서 30초 동안, 구배 40-55 °C에서 30초 동안, 및 72 °C에서 2분 동안 30 사이클; 및 72 °C에서 3분 동안 1 사이클로 프로그램된 EPPENDORF® MASTERCYCLER® (Eppendorf Scientific Inc.)로 수행하였다.
- [0716] 826 bp 야생형 YMR226c PCR 유전자 생성물을 여기에 설명된 바와 같이 TBE 버퍼 중의 1% 아가로스 겔 전기영동으로 정제하였다. 약 826 bp의 단편을 겔에서 잘라냈고 QIAQUICK® Gel Extraction Kit (Qiagen)를 사용하여 아가로스로부터 추출하였다. PCR 생성물을 37 °C에서 하룻밤 동안 XbaI 및 PacI로 분해하였고 QIAQUICK® PCR Purification Kit (Qiagen)를 사용하여 정제하였다.
- [0717] 플라스미드 pM1Ba100 (상기)를 XbaI 및 PacI로 분해한 후 이어서 CIP를 처리하였고 6.8 kbp 선형 단편을 생성하였다. 분해물을 제조사의 지시에 따라 QIAQUICK® PCR Purification Kit (Qiagen)를 사용하여 정제하였다.
- [0718] 826 bp YMR226c 정제된 및 분해된 pCR 단편을 pM1Ba100의 6.7 kbp 단편 1 µL, 1 µL 또는 7 µL의 826 bp YMR226c PCR 생성물, 10mM ATP (New England Biolabs)가 들어있는 10X 결찰 버퍼 1 µL, 및 1 µL T4 리가제 (New England Biolabs)로 구성된, 10 µL의 총 반응량의 T4 리가제 (New England Biolabs)를 사용하여 6.8 kbp pM1Ba100 선형화 벡터에 결찰하였다. 반응물을 16 °C에서 약 4시간 동안 배양하였고 전체 반응물을 제조사의 지시에 따라 Sure Chemically Competent Cell (Aglient)에 형질전환하였다. 형질전환체를 2X YT+amp 플레이트에 도말하였고 37 °C에서 하룻밤 동안 배양하였다. 결과의 형질전환체 중 다수를 XbaI 및 PacI를 사용하는 제한 분해에 의해 적절한 삽입에 대해 스크리닝하였다. 분해에 의해 올바른 클론을 DNA 시퀀싱으로 확인하였다. 올바른 분해된 밴드 크기 및 DNA 서열을 수득하는 클론을 pM1Ba101로 지정하였다.
- [0719] 플라스미드 pHJJ76-no ura (상기)를 NotI으로 분해한 후 이어서 CIP를 처리하였다. 선형 5.2 kbp 단편을 QIAQUICK® PCR Purification Kit (Qiagen)를 사용하여 정제하였다.
- [0720] YMR226c 발현 카세트를 NotI으로 분해에 의해 pM1Ba101로부터 잘라냈다. 약 3546 bp에서 밴드를 겔에서 잘라냈고 제조사의 지시에 따라 QIAQUICK® Gel Extraction Kit (Qiagen)를 사용하여 정제하였다.
- [0721] pM1Ba101의 3546 bp 정제된 단편을 pHJJ76-no ura의 5.2 kbp 단편 1 µL, pM1Ba101의 3546 bp 단편 1 µL 또는

5 μ L, 10mM ATP (New England Biolabs)가 들어있는 10X 결찰 버퍼 1 μ L, 및 1 μ L T4 리가제 (New England Biolabs)로 구성된, 10 μ L의 총 반응량의 T4 리가제 (New England Biolabs)를 사용하여 5.2 kbp pHJJ76-nura 선형화 벡터에 결찰하였다. 반응물을 16 $^{\circ}$ C에서 하룻밤 동안 배양하였고 반응물의 4 μ L 엘리퀼트를 제조사의 지시에 따라 ONE SHOT[®] TOP10 화학적 컴피턴트 이. 콜리 세포 (Invitrogen)에 형질전환하였다. 형질전환체를 2X YT+amp 플레이트에 도말하였고 37 $^{\circ}$ C에서 하룻밤 동안 배양하였다. 결과의 형질전환체 중 다수를 XbaI 및 KpnI를 사용하는 제한 분해에 의해 적절한 삽입에 대해 스크리닝하였다 올바른 분해된 밴드 크기를 수득하는 클론을 pM1Ba109로 지정하였다.

[0722] SEQ ID NO: 129의 에스. 세레비시애 3-HPDH를 암호화하는 야생형 뉴클레오티드 서열을 XbaI 및 PacI로 분해에 의해 플라스미드 pM1Ba109로부터 제거하였고 여기에 설명된 바와 같이 TBE 버퍼 중의 아가로스 겔 전기영동으로 정제하였다. 약 818 bp의 밴드를 겔에서 잘라냈고 제조사의 지시에 따라 NUCLEOSPIN[®] Extract II Kit (Macherey-Nagel)를 사용하여 정제하였다.

[0723] 정제된 약 818 bp 에스. 세레비시애 3-HPDH 유전자 생성물을 1 μ L의 8.4 kbp 벡터, 10 μ L의 818 bp 인서트, 10mM ATP (New England Biolabs)가 들어있는 10X 결찰 버퍼 2 μ L, 및 1 μ L T4 리가제 (New England Biolabs)로 구성된, 20 μ L의 총 반응량의 T4 리가제 (New England Biolabs)를 사용하여 상기 8.4 kbp pMBin204 선형화 벡터에 결찰하였다. 반응물을 16 $^{\circ}$ C에서 18시간 동안 배양하였고 반응물의 10 μ L 엘리퀼트를 제조사의 지시에 따라 One Shot TOP10 세포 (Invitrogen)에 형질전환하였다. 회수 기간 후, 형질전환 반응물의 2개의 100 μ L 엘리퀼트를 mL 당 100 μ g 앰피실린이 보충된 150 mM 2X YT 플레이트에 도말하였다. 플레이트를 37 $^{\circ}$ C에서 하룻밤 동안 배양하였다. 추정된 재조합 클론을 선택 플레이트로부터 선택하였고 플라스미드 DNA를 BIOROBOT[®] 9600 (Qiagen)을 사용하여 각각 제조하였다. 클론을 제한 분해로 분석하였고 올바른 제한 분해 패턴을 갖는 플라스미드를 pMcTs91로 지정하였다.

[0724] *adh9091* 위치에서 엠. 세둘라 3-HPDH의 통합을 위한 플라스미드 구조물

[0725] SEQ ID NO: 29의 엠. 세둘라 3-HPDH를 암호화하는 SEQ ID NO: 343의 아이. 오리엔탈리스 코돈-최적화된 뉴클레오티드 서열을 GeneArt[®]로 합성하였으며 플라스미드 11AAE2AP를 생성하였다. 합성 유전자를 XbaI 및 PacI로 플라스미드로부터 분해하였고 여기에 설명된 바와 같이 TBE 버퍼 중의 아가로스 겔 전기영동으로 정제하였다. 약 959 bp의 밴드를 겔에서 잘라냈고 제조사의 지시에 따라 NUCLEOSPIN[®] Extract II Kit (Macherey-Nagel)를 사용하여 정제하였다.

[0726] 정제된 약 959 bp 엠. 세둘라 3-HPDH 유전자 생성물을 1 μ L의 8.4 kbp 벡터, 16 μ L의 959 bp 인서트, 10mM ATP (New England Biolabs)가 들어있는 10X 결찰 버퍼 2 μ L, 및 1 μ L T4 리가제 (New England Biolabs)로 구성된, 20 μ L의 총 반응량의 T4 리가제 (New England Biolabs)를 사용하여 상기 8.4 kbp pMBin204 선형화 벡터에 결찰하였다. 반응물을 상온에서 1시간 동안 배양하였고 반응물의 10 μ L 엘리퀼트를 제조사의 지시에 따라 SoloPack Gold SuperCompetent 세포 (Agilent)에 형질전환하였다. 회수 기간 후, 형질전환 반응물의 2개의 100 μ L 엘리퀼트를 mL 당 100 μ g 앰피실린이 보충된 150 mM 2X YT 플레이트에 도말하였다. 플레이트를 37 $^{\circ}$ C에서 하룻밤 동안 배양하였다. 추정된 재조합 클론을 선택 플레이트로부터 선택하였고 플라스미드 DNABIOROBOT[®] 9600 (Qiagen)을 사용하여 각각 제조하였다. 클론을 제한 분해로 분석하였고 올바른 제한 분해 패턴을 갖는 플라스미드를 pMcTs76으로 지정하였다.

[0727] *adh9091* 위치에서 이. 콜리의 통합을 위한 플라스미드 구조물

[0728] SEQ ID NO: 27의 이. 콜리 3-HPDH를 암호화하는 SEQ ID NO: 143의 아이. 오리엔탈리스 코돈-최적화된 뉴클레오티드 서열을 GeneArt[®]로 합성하였으며 플라스미드 1045168를 생성하였다. 합성 유전자를 XbaI 및 PacI로 플라스미드로부터 분해하였고 여기에 설명된 바와 같이 TBE 버퍼 중의 아가로스 겔 전기영동으로 정제하였다. 약 761 bp의 밴드를 겔에서 잘라냈고 제조사의 지시에 따라 NUCLEOSPIN[®] Extract II Kit (Macherey-Nagel)를 사용하여 정제하였다.

[0729] 정제된 약 761 bp 이. 콜리 3-HPDH 유전자 생성물을 1 μ L의 8.4 kbp 벡터, 16 μ L의 761 bp 인서트, 10mM ATP (New England Biolabs)가 들어있는 10X 결찰 버퍼 2 μ L, 및 1 μ L T4 리가제 (New England Biolabs)로 구성된, 20 μ L의 총 반응량의 T4 리가제 (New England Biolabs)를 사용하여 상기 8.4 kbp pMBin204 선형화 벡터에 결찰하였다. 반응물을 상온에서 1시간 동안 배양하였고 반응물의 10 μ L 엘리퀼트를 제조사의 지시에 따라 SoloPack Gold SuperCompetent 세포 (Agilent)에 형질전환하였다. 회수 기간 후, 형질전환 반응물의 2개의 100

μ L 엘리퀀트를 mL 당 100 μ g 앰피실린이 보충된 150 mM 2X YT 플레이트에 도말하였다. 플레이트를 37 °C에서 하룻밤 동안 배양하였다. 주정의 재조합 클론을 선택 플레이트로부터 선택하였고 플라스미드 DNABIOROBOT® 9600 (Qiagen)을 사용하여 각각 제조하였다. 클론을 제한 분해로 분석하였고 올바른 제한 분해 패턴을 갖는 플라스미드를 pMcTs77로 지정하였다.

[0730] *McTs259의 adh9091에서 3-HPDH 단편의 통합*

[0731] 여기에 설명된 바와 같이 플라스미드 pMcTs76, pMcTs77, pMcTs91을 KpnI 및 ApaI로 분해하였고, 플라스미드 pMcTs90을 SacI 및 ApaI로 분해하였다. 결과의 분해 생성물을 TBE 버퍼 중의 1% 아가로스 겔 전기영동으로 정제하였고, pMcTs76의 5.5 kbp 단편, pMcTs77의 5.3 kbp 단편, pMcTs90의 5.4 kbp 단편, 및 pMcTs91의 5.4 kbp 단편을 겔에서 잘라냈고 제조사의 지시에 따라 NUCLEOSPIN® Extract II Kit (Macherey-Nagel)를 사용하여 정제하였다.

[0732] adh1202 위치에서 SEQ ID NO: 139의 비. 리체니포르미스 ADC를 암호화하는 뉴클레오타이드의 4개의 카피 및 SEQ ID NO: 26 (상기)의 3-HPDH를 암호화하는 네이티브 아이. 오리엔탈리스 유전자의 삭제를 발현하는 균주 McTs259를 분해된 pMcTs76, pMcTs77, pMcTs90, 또는 pMcTs91 DNA에 형질전환하였다. 올바른 위치 표적화 및 형질전환을 제조사의 지시에 따라 Phire® Plant Direct PCR kit (Finnzymes)를 사용하여 PCR로 확인하였다. adh9091 위치에서 통합을 확인하기 위해, 다음 프라이머 쌍을 사용하였다. 프라이머 0614627 및 0612909는 통합된 pMcTs76의 단편에 대한 약 3.47 kbp 밴드, 통합된 pMcTs77의 단편에 대한 약 3.27 kbp 밴드, 통합된 pMcTs90의 단편에 대한 약 3.34 kbp 밴드, 통합된 pMcTs91의 단편에 대한 약 3.33 kbp 밴드를 수득하였다; 프라이머 0612908 및 0614626은 약 1.97 kbp 밴드를 수득하였다. adh1202에서 통합을 체크하기 위해 다음 프라이머 쌍을 사용하였다. 프라이머 0611245 및 0612794는 약 3.0 kbp 밴드를 수득하였고 프라이머 0611815 및 0612795는 약 3.6 kbp 밴드를 수득하였다. 네이티브 아이. 오리엔탈리스 3-HPDH 유전자의 삭제를 체크하기 위해 다음 프라이머를 사용하였다. 프라이머 0613034 및 0613241은 약 1.4 kbp 밴드를 수득하였다. adh9091 위치에서 발현 카세트의 적절한 통합에 대해 예측된 밴드를 제공하고, adh1202에서 발현 카세트를 유지하고 아이. 오리엔탈리스 3-HPDH 위치에서 삭제를 유지하는 분리체를 저장하였고 표 29에 나타난 바와 같이 McTs261 (pMcTs76 단편), McTs263 (pMcTs77 단편), McTs267 (pMcTs90 단편), 및 McTs269 (pMcTs91 단편)로 지정하였다.

[0733] 표 29: 형질전환체 구조물

| 구조 플라스미드 | 유전자 | 유전자 근원 | 유전자 Product SEQ ID NO | 통합 구조 | 형질전환체 |
|----------|--------------------|------------|-----------------------|---------|---------|
| pMBin190 | 3-HPDH (YMR226c) | 아이. 오리엔탈리스 | 26 | pMcTs90 | McTs267 |
| 11AAE2AP | 3-HPDH (Msed_1993) | 엠. 세둘라 | 29 | pMcTs76 | McTs261 |
| pMIBa109 | 3-HPDH (YMR226c) | 에스. 세레비시애 | 129 | pMcTs91 | McTs269 |
| 1045168 | 3-HPDH (ydfG) | 이. 콜리 | 27 | pMcTs77 | McTs263 |

[0735] 형질전환체 균주를 상기 설명된 셰이크 플라스크 방법을 사용하여 3-HP 생산에 대해 테스트하였다. 이형 접합 형질전환체 McTs267, McTs269, 및 McTs263은 각각 g/L 건조 세포 중량 당 0.149 (+/- 0.024), 0.168 (+/- 0.052), 0.162 (+/- 0.018) g/L 3-HP를 생산하였다. 네이티브 균주 MeJi412는 g/L 건조 세포 중량 당 263 (+/- 0.026) g/L 3-HP를 생산하였고 3-HPDH 삭제 균주는 검출 가능한 3-HP를 생산하지 않았다. 이형 접합 형질전환체 McTs261은 이 실험으로 검출 가능한 3-HP를 생산하지 않았다. 이 결과들은 이형 접합 3-HPDH 형질전환체가 외인성 또는 내인성 3-HPDH 유전자 서열을 사용하여 3-HPDH 삭제 균주의 일부 3-HPDH 활성을 회복할 수 있다는 것을 제안한다.

[0736] 실시예 3B: 말레이트 경로 유전자를 발현하는 변형된 효모 균주

[0737] PEP, OAA, 및 말레이트 중간물을 이용하는 경로를 통해 3-HP를 생산하는 효모 세포는 경로에 수반되는 하나 이상의 효소를 발현함으로써 조작될 수 있다. 발현된 유전자는 하나 이상의 PPC, 말레이트 데히드로게나제, 및 말레이트 데카르복실라제 유전자를 포함할 수도 있다. 발현된 유전자는 숙주 세포에 비-네이티브인 근원 유전자로부터 유도될 수도 있다.

[0738] 실시예 3C: 말로네이트 세미알데히드 경로 유전자를 발현하는 변형된 효모 균주

- [0739] PEP, OAA, 및 말레이트 중간물을 이용하는 경로를 통해 3-HP를 생산하는 효모 세포는 경로에 수반되는 하나 이상의 효소를 발현함으로써 조작될 수 있다. 발현된 유전자는 PPC, 2-케토산 데카르복실라제, KGD, BCKA, 인돌피루베이트 데카르복실라제, 3-HPDH (말로네이트 세미알데히드 리덕타제 등), HIBADH, 및 4-히드록시부티레이트 데히드로게나제 유전자 중 하나 이상을 포함할 수도 있다. 발현된 유전자는 숙주 세포에 네이티브인 유전자로부터 유도될 수도 있거나, 그것들은 숙주 세포에 비-네이티브인 근원 유전자로부터 유도될 수도 있다.
- [0740] 실시예 3D: 말로닐-CoA 경로 유전자를 발현하는 변형된 효모 균주
- [0741] PEP, OAA, 말로닐-CoA, 및, 선택적으로, 말로네이트 세미알데히드 중간물을 이용하는 경로를 통해 3-HP를 생산하는 효모 세포는 경로에 수반되는 하나 이상의 효소를 발현함으로써 조작될 수 있다. 발현된 유전자는 PPC, OAA 포르마테리아제, 말로닐-CoA 리덕타제, CoA 아실화 말로네이트 세미알데히드 데히드로게나제, 3-HPDH (말로네이트 세미알데히드 리덕타제 등), HIBADH, 및 4-히드록시부티레이트 데히드로게나제 유전자 중 하나 이상을 포함할 수도 있다. 발현된 유전자는 숙주 세포에 네이티브인 유전자로부터 유도될 수도 있거나, 그것들은 숙주 세포에 비-네이티브인 근원 유전자로부터 유도될 수도 있다.
- [0742] 실시예 3E: 말로닐-CoA 경로 유전자를 발현하는 변형된 효모 균주
- [0743] 피루베이트, 아세틸-CoA, 말로닐-CoA, 및, 선택적으로, 말로네이트 세미알데히드 중간물을 이용하는 경로를 통해 3-HP를 생산하는 효모 세포는 경로에 수반되는 하나 이상의 효소를 발현함으로써 조작될 수 있다. 발현된 유전자는 PDH, 아세틸-CoA 카르복실라제, 말로닐-CoA 리덕타제, CoA 아실화 말로네이트 세미알데히드 데히드로게나제, 3-HPDH (말로네이트 세미알데히드 리덕타제 등), HIBADH, 및 4-히드록시부티레이트 데히드로게나제 유전자 중 하나 이상을 포함할 수도 있다. 발현된 유전자는 숙주 세포에 네이티브인 유전자로부터 유도될 수도 있거나, 그것들은 숙주 세포에 비-네이티브인 근원 유전자로부터 유도될 수도 있다.
- [0744] 실시예 3F: 알라닌 경로 유전자를 발현하는 변형된 효모 균주
- [0745] 피루베이트, 알라닌, β -알라닌, 및, 선택적으로, 말로네이트 세미알데히드, β -알라닐-CoA, 아크릴일-CoA, 및 3-HP-CoA 중간물을 이용하는 경로를 통해 3-HP를 생산하는 효모 세포는 경로에 수반되는 하나 이상의 효소를 발현함으로써 조작될 수 있다. 발현된 유전자는 알라닌 데히드로게나제, 피루베이트/알라닌 아미노트랜스퍼라제, 알라닌 2,3 아미노뮤타제, CoA 트랜스퍼라제, CoA 신테타제, β -알라닐-CoA 암모니아 리아제, 3-HP-CoA 데히드라타제, 3-HP-CoA 히드로라제, 3-히드록시이소부티릴-CoA 히드로라제, BAAT, 3-HPDH (말로네이트 세미알데히드 리덕타제 등), HIBADH, 및 4-히드록시부티레이트 데히드로게나제 유전자 중 하나 이상을 포함할 수도 있다. 발현된 유전자는 숙주 세포에 네이티브인 유전자로부터 유도될 수도 있거나, 그것들은 숙주 세포에 비-네이티브인 근원 유전자로부터 유도될 수도 있다.
- [0746] 실시예 3G: 락테이트 경로 유전자를 발현하는 변형된 효모 균주
- [0747] 피루베이트, 락테이트, 락틸-CoA, 아크릴일-CoA, 및 3-HP-CoA 중간물을 이용하는 경로를 통해 3-HP를 생산하는 효모 세포는 이 경로에 수반되는 하나 이상의 효소를 발현함으로써 조작될 수 있다. 발현된 유전자는 LDH, CoA 트랜스퍼라제, CoA 신테타제, 락틸-CoA 데히드라타제, 3-HP-CoA 데히드라타제, 3-HP-CoA 히드로라제, 및 3-히드록시이소부티릴-CoA 히드로라제 유전자 중 하나 이상을 포함할 수도 있다. 발현된 유전자는 숙주 세포에 네이티브인 유전자로부터 유도될 수도 있거나, 그것들은 숙주 세포에 비-네이티브인 근원 유전자로부터 유도될 수도 있다.
- [0748] 실시예 3H: 글리세롤 경로 유전자를 발현하는 변형된 효모 균주
- [0749] 글리세롤 및 3-HPA 중간물을 이용하는 경로를 통해 3-HP를 생산하는 효모 세포는 이 경로에 수반되는 하나 이상의 효소를 발현함으로써 조작될 수 있다. 발현된 유전자는 글리세롤 데히드라타제 및 알데히드 데히드로게나제 유전자 중 하나 이상을 포함할 수도 있다. 발현된 유전자는 숙주 세포에 네이티브인 유전자로부터 유도될 수도 있거나, 그것들은 숙주 세포에 비-네이티브인 근원 유전자로부터 유도될 수도 있다.
- [0750] 실시예 3I: β -알라닐 CoA 경로 유전자를 발현하는 변형된 효모 균주
- [0751] PEP 또는 피루베이트, β -알라닌, β -알라닐-CoA, 아크릴일-CoA, 3-HP-CoA, 및, 선택적으로 OAA, 아스파르테이트, 및 알라닌 중간물을 이용하는 경로를 통해 3-HP를 생산하는 효모 세포는 이 경로에 수반되는 하나 이상의 효소를 발현함으로써 조작될 수 있다. 발현된 유전자는 PPC, PYC, AAT, ADC, CoA 트랜스퍼라제, CoA 신테타제, β -알라닐-CoA 암모니아 리아제, 3-HP-CoA 데히드라타제, 3-HP-CoA 히드로라제, 3-히드록시이소부티릴-CoA 히드로라제, 알라닌 데히드로게나제, 피루베이트/알라닌 아미노트랜스퍼라제, 및 AAM 유전자 중 하나 이상을 포함할

수도 있다. 발현된 유전자는 숙주 세포에 네이티브인 유전자로부터 유도될 수도 있거나, 그것들은 숙주 세포에 비-네이티브인 근원 유전자로부터 유도될 수도 있다.

[0752] 일부 양태에서, 효모 세포 또는 이들의 사용 방법은 다음 번호가 매겨진 단락에 의해 설명될 수도 있다:

[0753] [B1] 활성 3-HP 발효 경로를 포함하는 유전적으로 변형된 효모 세포로서, 세포는 다음으로부터 선택된 하나 이상의 외인성 3-HP 경로 유전자를 포함한다:

[0754] 외인성 PPC 유전자;

[0755] 외인성 PYC 유전자;

[0756] 외인성 AAT 유전자;

[0757] 외인성 ADC 유전자;

[0758] 외인성 BAAT 또는 gabT 유전자; 및

[0759] 외인성 3-HPDH 유전자.

[0760] [B2] 외인성 AAT 유전자를 포함하는, 단락 B1의 유전적으로 변형된 효모 세포.

[0761] [B3] 외인성 PYC 유전자를 포함하는, 단락 B1 또는 B2의 유전적으로 변형된 효모 세포.

[0762] [B4] 외인성 ADC 유전자를 포함하는, 단락 B1 내지 단락 B3 중 어느 하나의 유전적으로 변형된 효모 세포.

[0763] [B5] 외인성 BAAT 또는 gabT 유전자를 포함하는, 단락 B1 내지 단락 B4 중 어느 하나의 유전적으로 변형된 효모 세포.

[0764] [B6] 외인성 3-HPDH 유전자를 포함하는, 단락 B1 내지 단락 B5 중 어느 하나의 유전적으로 변형된 효모 세포.

[0765] [B7] 다음을 포함하는, 단락 B1의 유전적으로 변형된 효모 세포:

[0766] 외인성 PYC 유전자;

[0767] 외인성 AAT 유전자;

[0768] 외인성 ADC 유전자;

[0769] 외인성 BAAT 또는 gabT 유전자; 및

[0770] 외인성 3-HPDH 유전자.

[0771] [B8] 외인성 PPC 유전자를 포함하는, 단락 B1 내지 단락 B7 중 어느 하나의 유전적으로 변형된 효모 세포.

[0772] [B9] 외인성 PPC 유전자가 SEQ ID NO: 2, 3, 4, 5, 6, 7, 및 8로부터 선택된 아미노산 서열과 적어도 60%, 예를 들어, 적어도 65%, 70%, 75%, 80%, 85%, 90%, 95%, 97%, 98%, 99%, 또는 100% 서열 동일성을 갖는 폴리펩티드를 암호화하는, 단락 B1 내지 단락 B7 중 어느 하나의 유전적으로 변형된 효모 세포.

[0773] [B10] SEQ ID NO: 2로부터 선택된 아미노산 서열과 적어도 60%, 예를 들어, 적어도 65%, 70%, 75%, 80%, 85%, 90%, 95%, 97%, 98%, 99%, 또는 100% 서열 동일성을 갖는 폴리펩티드를 암호화하는, 단락 B9의 유전적으로 변형된 효모 세포.

[0774] [B11] SEQ ID NO: 3으로부터 선택된 아미노산 서열과 적어도 60%, 예를 들어, 적어도 65%, 70%, 75%, 80%, 85%, 90%, 95%, 97%, 98%, 99%, 또는 100% 서열 동일성을 갖는 폴리펩티드를 암호화하는, 단락 B9의 유전적으로 변형된 효모 세포.

[0775] [B12] SEQ ID NO: 4로부터 선택된 아미노산 서열과 적어도 60%, 예를 들어, 적어도 65%, 70%, 75%, 80%, 85%, 90%, 95%, 97%, 98%, 99%, 또는 100% 서열 동일성을 갖는 폴리펩티드를 암호화하는, 단락 B9의 유전적으로 변형된 효모 세포.

[0776] [B13] SEQ ID NO: 5로부터 선택된 아미노산 서열과 적어도 60%, 예를 들어, 적어도 65%, 70%, 75%, 80%, 85%, 90%, 95%, 97%, 98%, 99%, 또는 100% 서열 동일성을 갖는 폴리펩티드를 암호화하는, 단락 B9의 유전적으로 변형된 효모 세포.

[0777] [B14] SEQ ID NO: 6으로부터 선택된 아미노산 서열과 적어도 60%, 예를 들어, 적어도 65%, 70%, 75%, 80%,

85%, 90%, 95%, 97%, 98%, 99%, 또는 100% 서열 동일성을 갖는 폴리펩티드를 암호화하는, 단락 B9의 유전적으로 변형된 효모 세포.

[0778] [B15] SEQ ID NO: 7로부터 선택된 아미노산 서열과 적어도 60%, 예를 들어, 적어도 65%, 70%, 75%, 80%, 85%, 90%, 95%, 97%, 98%, 99%, 또는 100% 서열 동일성을 갖는 폴리펩티드를 암호화하는, 단락 B9의 유전적으로 변형된 효모 세포.

[0779] [B16] SEQ ID NO: 8로부터 선택된 아미노산 서열과 적어도 60%, 예를 들어, 적어도 65%, 70%, 75%, 80%, 85%, 90%, 95%, 97%, 98%, 99%, 또는 100% 서열 동일성을 갖는 폴리펩티드를 암호화하는, 단락 B9의 유전적으로 변형된 효모 세포.

[0780] [B17] 외인성 PYC 유전자가 SEQ ID NO: 1의 뉴클레오티드 서열과 적어도 60%, 예를 들어, 적어도 65%, 70%, 75%, 80%, 85%, 90%, 95%, 97%, 98%, 99%, 또는 100% 서열 동일성을 갖는 뉴클레오티드 서열을 포함하는, 단락 B1 내지 단락 B16 중 어느 하나의 유전적으로 변형된 효모 세포.

[0781] [B18] AAT 유전자가 SEQ ID NO: 14, 15, 및 16의 아미노산 서열과 적어도 60%, 예를 들어, 적어도 65%, 70%, 75%, 80%, 85%, 90%, 95%, 97%, 98%, 99%, 또는 100% 서열 동일성을 갖는 폴리펩티드를 암호화하는, 단락 B1 내지 단락 B17 중 어느 하나의 유전적으로 변형된 효모 세포.

[0782] [B19] AAT 유전자가 SEQ ID NO: 14의 아미노산 서열과 적어도 60%, 예를 들어, 적어도 65%, 70%, 75%, 80%, 85%, 90%, 95%, 97%, 98%, 99%, 또는 100% 서열 동일성을 갖는 폴리펩티드를 암호화하는, 단락 B1 내지 단락 B17 중 어느 하나의 유전적으로 변형된 효모 세포.

[0783] [B20] AAT 유전자가 SEQ ID NO: 15의 아미노산 서열과 적어도 60%, 예를 들어, 적어도 65%, 70%, 75%, 80%, 85%, 90%, 95%, 97%, 98%, 99%, 또는 100% 서열 동일성을 갖는 폴리펩티드를 암호화하는, 단락 B1 내지 단락 B17 중 어느 하나의 유전적으로 변형된 효모 세포.

[0784] [B21] AAT 유전자가 SEQ ID NO: 16의 아미노산 서열과 적어도 60%, 예를 들어, 적어도 65%, 70%, 75%, 80%, 85%, 90%, 95%, 97%, 98%, 99%, 또는 100% 서열 동일성을 갖는 폴리펩티드를 암호화하는, 단락 B1 내지 단락 B17 중 어느 하나의 유전적으로 변형된 효모 세포.

[0785] [B22] AAT 유전자가 SEQ ID NO: 13의 뉴클레오티드 서열과 적어도 60%, 예를 들어, 적어도 65%, 70%, 75%, 80%, 85%, 90%, 95%, 97%, 98%, 99%, 또는 100% 서열 동일성을 갖는 뉴클레오티드 서열을 포함하는, 단락 B1 내지 단락 B21 중 어느 하나의 유전적으로 변형된 효모 세포

[0786] [B23] ADC 유전자가 SEQ ID NO: 17, 18, 133, 135, 137, 및 139로부터 선택된 아미노산 서열과 적어도 60%, 예를 들어, 적어도 65%, 70%, 75%, 80%, 85%, 90%, 95%, 97%, 98%, 99%, 또는 100% 서열 동일성을 갖는 폴리펩티드를 암호화하는, 단락 B1 내지 단락 B22 중 어느 하나의 유전적으로 변형된 효모 세포

[0787] [B24] ADC 유전자가 SEQ ID NO: 17의 아미노산 서열과 적어도 60%, 예를 들어, 적어도 65%, 70%, 75%, 80%, 85%, 90%, 95%, 97%, 98%, 99%, 또는 100% 서열 동일성을 갖는 폴리펩티드를 암호화하는, 단락 B23의 유전적으로 변형된 효모 세포.

[0788] [B25] ADC 유전자가 SEQ ID NO: 18의 아미노산 서열과 적어도 60%, 예를 들어, 적어도 65%, 70%, 75%, 80%, 85%, 90%, 95%, 97%, 98%, 99%, 또는 100% 서열 동일성을 갖는 폴리펩티드를 암호화하는, 단락 B23의 유전적으로 변형된 효모 세포.

[0789] [B26] ADC 유전자가 SEQ ID NO: 133의 아미노산 서열과 적어도 60%, 예를 들어, 적어도 65%, 70%, 75%, 80%, 85%, 90%, 95%, 97%, 98%, 99%, 또는 100% 서열 동일성을 갖는 폴리펩티드를 암호화하는, 단락 B23의 유전적으로 변형된 효모 세포.

[0790] [B27] ADC 유전자가 SEQ ID NO: 135의 아미노산 서열과 적어도 60%, 예를 들어, 적어도 65%, 70%, 75%, 80%, 85%, 90%, 95%, 97%, 98%, 99%, 또는 100% 서열 동일성을 갖는 폴리펩티드를 암호화하는, 단락 B23의 유전적으로 변형된 효모 세포.

[0791] [B28] ADC 유전자가 SEQ ID NO: 137의 아미노산 서열과 적어도 60%, 예를 들어, 적어도 65%, 70%, 75%, 80%, 85%, 90%, 95%, 97%, 98%, 99%, 또는 100% 서열 동일성을 갖는 폴리펩티드를 암호화하는, 단락 B23의 유전적으로 변형된 효모 세포.

- [0792] [B29] ADC 유전자가 SEQ ID NO: 139의 아미노산 서열과 적어도 60%, 예를 들어, 적어도 65%, 70%, 75%, 80%, 85%, 90%, 95%, 97%, 98%, 99%, 또는 100% 서열 동일성을 갖는 폴리펩티드를 암호화하는, 단락 B23의 유전적으로 변형된 효모 세포.
- [0793] [B30] ADC 유전자가 SEQ ID NO: 130, 131, 132, 134, 136, 및 138로부터 선택된 뉴클레오티드 서열과 적어도 60%, 예를 들어, 적어도 65%, 70%, 75%, 80%, 85%, 90%, 95%, 97%, 98%, 99%, 또는 100% 서열 동일성을 갖는 뉴클레오티드를 포함하는, 단락 B1 내지 단락 B29 중 어느 하나의 유전적으로 변형된 효모 세포.
- [0794] [B31] ADC 유전자가 SEQ ID NO: 130의 뉴클레오티드 서열과 적어도 60%, 예를 들어, 적어도 65%, 70%, 75%, 80%, 85%, 90%, 95%, 97%, 98%, 99%, 또는 100% 서열 동일성을 갖는 뉴클레오티드 서열을 포함하는, 단락 B30의 유전적으로 변형된 효모 세포.
- [0795] [B32] ADC 유전자가 SEQ ID NO: 131의 뉴클레오티드 서열과 적어도 60%, 예를 들어, 적어도 65%, 70%, 75%, 80%, 85%, 90%, 95%, 97%, 98%, 99%, 또는 100% 서열 동일성을 갖는 뉴클레오티드 서열을 포함하는, 단락 B30의 유전적으로 변형된 효모 세포.
- [0796] [B33] ADC 유전자가 SEQ ID NO: 132의 뉴클레오티드 서열과 적어도 60%, 예를 들어, 적어도 65%, 70%, 75%, 80%, 85%, 90%, 95%, 97%, 98%, 99%, 또는 100% 서열 동일성을 갖는 뉴클레오티드 서열을 포함하는, 단락 B30의 유전적으로 변형된 효모 세포.
- [0797] [B34] ADC 유전자가 SEQ ID NO: 134의 뉴클레오티드 서열과 적어도 60%, 예를 들어, 적어도 65%, 70%, 75%, 80%, 85%, 90%, 95%, 97%, 98%, 99%, 또는 100% 서열 동일성을 갖는 뉴클레오티드 서열을 포함하는, 단락 B30의 유전적으로 변형된 효모 세포.
- [0798] [B35] ADC 유전자가 SEQ ID NO: 136의 뉴클레오티드 서열과 적어도 60%, 예를 들어, 적어도 65%, 70%, 75%, 80%, 85%, 90%, 95%, 97%, 98%, 99%, 또는 100% 서열 동일성을 갖는 뉴클레오티드 서열을 포함하는, 단락 B30의 유전적으로 변형된 효모 세포.
- [0799] [B36] ADC 유전자가 SEQ ID NO: 138의 뉴클레오티드 서열과 적어도 60%, 예를 들어, 적어도 65%, 70%, 75%, 80%, 85%, 90%, 95%, 97%, 98%, 99%, 또는 100% 서열 동일성을 갖는 뉴클레오티드 서열을 포함하는, 단락 B30의 유전적으로 변형된 효모 세포.
- [0800] [B37] 외인성 BAAT 또는 gabT 유전자가 SEQ ID NO: 20, 21, 22, 23, 및 24로부터 선택된 아미노산 서열과 적어도 60%, 예를 들어, 적어도 65%, 70%, 75%, 80%, 85%, 90%, 95%, 97%, 98%, 99%, 또는 100% 서열 동일성을 갖는 폴리펩티드를 암호화하는, 단락 B1 내지 단락 B36 중 어느 하나의 유전적으로 변형된 효모 세포.
- [0801] [B38] 외인성 BAAT 또는 gabT 유전자가 SEQ ID NO: 20의 아미노산 서열과 적어도 60%, 예를 들어, 적어도 65%, 70%, 75%, 80%, 85%, 90%, 95%, 97%, 98%, 99%, 또는 100% 서열 동일성을 갖는 폴리펩티드를 암호화하는, 단락 B37의 유전적으로 변형된 효모 세포.
- [0802] [B39] 외인성 BAAT 또는 gabT 유전자가 SEQ ID NO: 21의 아미노산 서열과 적어도 60%, 예를 들어, 적어도 65%, 70%, 75%, 80%, 85%, 90%, 95%, 97%, 98%, 99%, 또는 100% 서열 동일성을 갖는 폴리펩티드를 암호화하는, 단락 B37의 유전적으로 변형된 효모 세포.
- [0803] [B40] 외인성 BAAT 또는 gabT 유전자가 SEQ ID NO: 22의 아미노산 서열과 적어도 60%, 예를 들어, 적어도 65%, 70%, 75%, 80%, 85%, 90%, 95%, 97%, 98%, 99%, 또는 100% 서열 동일성을 갖는 폴리펩티드를 암호화하는, 단락 B37의 유전적으로 변형된 효모 세포.
- [0804] [B41] 외인성 BAAT 또는 gabT 유전자가 SEQ ID NO: 23의 아미노산 서열과 적어도 60%, 예를 들어, 적어도 65%, 70%, 75%, 80%, 85%, 90%, 95%, 97%, 98%, 99%, 또는 100% 서열 동일성을 갖는 폴리펩티드를 암호화하는, 단락 B37의 유전적으로 변형된 효모 세포.
- [0805] [B42] 외인성 BAAT 또는 gabT 유전자가 SEQ ID NO: 24의 아미노산 서열과 적어도 60%, 예를 들어, 적어도 65%, 70%, 75%, 80%, 85%, 90%, 95%, 97%, 98%, 99%, 또는 100% 서열 동일성을 갖는 폴리펩티드를 암호화하는, 단락 B37의 유전적으로 변형된 효모 세포.
- [0806] [B43] BAAT 유전자 또는 gabT 유전자가 SEQ ID NO: 19, 140, 141, 및 142로부터 선택된 뉴클레오티드 서열과 적어도 60%, 예를 들어, 적어도 65%, 70%, 75%, 80%, 85%, 90%, 95%, 97%, 98%, 99%, 또는 100% 서열 동일성을

갖는 뉴클레오티드 서열을 포함하는, 단락 B1 내지 단락 B42 중 어느 하나의 유전적으로 변형된 효모 세포.

[0807] [B44] BAAT 유전자 또는 gabT 유전자가 SEQ ID NO: 19의 뉴클레오티드 서열과 적어도 60%, 예를 들어, 적어도 65%, 70%, 75%, 80%, 85%, 90%, 95%, 97%, 98%, 99%, 또는 100% 서열 동일성을 갖는 뉴클레오티드 서열을 포함하는, 단락 43의 유전적으로 변형된 효모 세포.

[0808] [B45] BAAT 유전자 또는 gabT 유전자가 SEQ ID NO: 140의 뉴클레오티드 서열과 적어도 60%, 예를 들어, 적어도 65%, 70%, 75%, 80%, 85%, 90%, 95%, 97%, 98%, 99%, 또는 100% 서열 동일성을 갖는 뉴클레오티드 서열을 포함하는, 단락 43의 유전적으로 변형된 효모 세포.

[0809] [B46] BAAT 유전자 또는 gabT 유전자가 SEQ ID NO: 141의 뉴클레오티드 서열과 적어도 60%, 예를 들어, 적어도 65%, 70%, 75%, 80%, 85%, 90%, 95%, 97%, 98%, 99%, 또는 100% 서열 동일성을 갖는 뉴클레오티드 서열을 포함하는, 단락 43의 유전적으로 변형된 효모 세포.

[0810] [B47] BAAT 유전자 또는 gabT 유전자가 SEQ ID NO: 142의 뉴클레오티드 서열과 적어도 60%, 예를 들어, 적어도 65%, 70%, 75%, 80%, 85%, 90%, 95%, 97%, 98%, 99%, 또는 100% 서열 동일성을 갖는 뉴클레오티드 서열을 포함하는, 단락 43의 유전적으로 변형된 효모 세포.

[0811] [B48] 상기 외인성 BAAT 유전자 또는 외인성 gabT는 BAAT 유전자이며, 이것은 또한 gabT 유전자인, 단락 B1 내지 단락 B47 중 어느 하나의 유전적으로 변형된 효모 세포.

[0812] [B49] 3-HPDH 유전자가 SEQ ID NO: 26, 27, 28, 29, 30, 31, 32, 33, 34, 및 129로 구성된 그룹으로부터 선택된 아미노산 서열과 적어도 60%, 예를 들어, 적어도 65%, 70%, 75%, 80%, 85%, 90%, 95%, 97%, 98%, 99%, 또는 100% 서열 동일성을 갖는 폴리펩티드를 암호화하는, 단락 B1 내지 단락 B48 중 어느 하나의 유전적으로 변형된 효모 세포.

[0813] [B50] 3-HPDH 유전자가 SEQ ID NO: 26의 아미노산 서열과 적어도 60%, 예를 들어, 적어도 65%, 70%, 75%, 80%, 85%, 90%, 95%, 97%, 98%, 99%, 또는 100% 서열 동일성을 갖는 폴리펩티드를 암호화하는, 단락 B49의 유전적으로 변형된 효모 세포.

[0814] [B51] 3-HPDH 유전자가 SEQ ID NO: 27의 아미노산 서열과 적어도 60%, 예를 들어, 적어도 65%, 70%, 75%, 80%, 85%, 90%, 95%, 97%, 98%, 99%, 또는 100% 서열 동일성을 갖는 폴리펩티드를 암호화하는, 단락 B49의 유전적으로 변형된 효모 세포.

[0815] [B52] 3-HPDH 유전자가 SEQ ID NO: 28의 아미노산 서열과 적어도 60%, 예를 들어, 적어도 65%, 70%, 75%, 80%, 85%, 90%, 95%, 97%, 98%, 99%, 또는 100% 서열 동일성을 갖는 폴리펩티드를 암호화하는, 단락 B49의 유전적으로 변형된 효모 세포.

[0816] [B53] 3-HPDH 유전자가 SEQ ID NO: 29의 아미노산 서열과 적어도 60%, 예를 들어, 적어도 65%, 70%, 75%, 80%, 85%, 90%, 95%, 97%, 98%, 99%, 또는 100% 서열 동일성을 갖는 폴리펩티드를 암호화하는, 단락 B49의 유전적으로 변형된 효모 세포.

[0817] [B54] 3-HPDH 유전자가 SEQ ID NO: 30의 아미노산 서열과 적어도 60%, 예를 들어, 적어도 65%, 70%, 75%, 80%, 85%, 90%, 95%, 97%, 98%, 99%, 또는 100% 서열 동일성을 갖는 폴리펩티드를 암호화하는, 단락 B49의 유전적으로 변형된 효모 세포.

[0818] [B55] 3-HPDH 유전자가 SEQ ID NO: 31의 아미노산 서열과 적어도 60%, 예를 들어, 적어도 65%, 70%, 75%, 80%, 85%, 90%, 95%, 97%, 98%, 99%, 또는 100% 서열 동일성을 갖는 폴리펩티드를 암호화하는, 단락 B49의 유전적으로 변형된 효모 세포.

[0819] [B56] 3-HPDH 유전자가 SEQ ID NO: 32의 아미노산 서열과 적어도 60%, 예를 들어, 적어도 65%, 70%, 75%, 80%, 85%, 90%, 95%, 97%, 98%, 99%, 또는 100% 서열 동일성을 갖는 폴리펩티드를 암호화하는, 단락 B49의 유전적으로 변형된 효모 세포.

[0820] [B57] 3-HPDH 유전자가 SEQ ID NO: 33의 아미노산 서열과 적어도 60%, 예를 들어, 적어도 65%, 70%, 75%, 80%, 85%, 90%, 95%, 97%, 98%, 99%, 또는 100% 서열 동일성을 갖는 폴리펩티드를 암호화하는, 단락 B49의 유전적으로 변형된 효모 세포.

[0821] [B58] 3-HPDH 유전자가 SEQ ID NO: 34의 아미노산 서열과 적어도 60%, 예를 들어, 적어도 65%, 70%, 75%, 80%,

85%, 90%, 95%, 97%, 98%, 99%, 또는 100% 서열 동일성을 갖는 폴리펩티드를 암호화하는, 단락 B49의 유전적으로 변형된 효모 세포.

[0822] [B59] 3-HPDH 유전자가 SEQ ID NO: 129의 아미노산 서열과 적어도 60%, 예를 들어, 적어도 65%, 70%, 75%, 80%, 85%, 90%, 95%, 97%, 98%, 99%, 또는 100% 서열 동일성을 갖는 폴리펩티드를 암호화하는, 단락 B49의 유전적으로 변형된 효모 세포.

[0823] [B60] 3-HPDH 유전자가 SEQ ID NO: 25, 143, 144, 및 343으로부터 선택된 뉴클레오티드 서열과 적어도 60%, 예를 들어, 적어도 65%, 70%, 75%, 80%, 85%, 90%, 95%, 97%, 98%, 99%, 또는 100% 서열 동일성을 갖는 뉴클레오티드 서열을 포함하는, 단락 B1 내지 단락 B59 중 어느 하나의 유전적으로 변형된 효모 세포.

[0824] [B61] 3-HPDH 유전자가 SEQ ID NO: 25의 뉴클레오티드 서열과 적어도 60%, 예를 들어, 적어도 65%, 70%, 75%, 80%, 85%, 90%, 95%, 97%, 98%, 99%, 또는 100% 서열 동일성을 갖는 뉴클레오티드 서열을 포함하는, 단락 B60의 유전적으로 변형된 효모 세포.

[0825] [B62] 3-HPDH 유전자가 SEQ ID NO: 143의 뉴클레오티드 서열과 적어도 60%, 예를 들어, 적어도 65%, 70%, 75%, 80%, 85%, 90%, 95%, 97%, 98%, 99%, 또는 100% 서열 동일성을 갖는 뉴클레오티드 서열을 포함하는, 단락 B60의 유전적으로 변형된 효모 세포.

[0826] [B63] 3-HPDH 유전자가 SEQ ID NO: 144의 뉴클레오티드 서열과 적어도 60%, 예를 들어, 적어도 65%, 70%, 75%, 80%, 85%, 90%, 95%, 97%, 98%, 99%, 또는 100% 서열 동일성을 갖는 뉴클레오티드 서열을 포함하는, 단락 B60의 유전적으로 변형된 효모 세포.

[0827] [B64] 3-HPDH 유전자가 SEQ ID NO: 343의 뉴클레오티드 서열과 적어도 60%, 예를 들어, 적어도 65%, 70%, 75%, 80%, 85%, 90%, 95%, 97%, 98%, 99%, 또는 100% 서열 동일성을 갖는 뉴클레오티드 서열을 포함하는, 단락 B60의 유전적으로 변형된 효모 세포.

[0828] [B65] 3-HPDH 유전자는 또한 HIBADH 유전자인, 단락 B1 내지 단락 B64 중 어느 하나의 유전적으로 변형된 효모 세포.

[0829] [B66] 3-HPDH 유전자는 또한 4-히드록시부티레이트 데히드로게나제 유전자인, 단락 B1 내지 단락 B65 중 어느 하나의 유전적으로 변형된 효모 세포.

[0830] [B67] PPC 유전자가 SEQ ID NO: 10, 11, 및 12로부터 선택된 아미노산 서열과 적어도 60%, 예를 들어, 적어도 65%, 70%, 75%, 80%, 85%, 90%, 95%, 97%, 98%, 99%, 또는 100% 서열 동일성을 갖는 폴리펩티드를 암호화하는, 단락 B1 내지 단락 B66 중 어느 하나의 유전적으로 변형된 효모 세포.

[0831] [B68] PPC 유전자가 SEQ ID NO: 10의 아미노산 서열과 적어도 60%, 예를 들어, 적어도 65%, 70%, 75%, 80%, 85%, 90%, 95%, 97%, 98%, 99%, 또는 100% 서열 동일성을 갖는 폴리펩티드를 암호화하는, 단락 B67의 유전적으로 변형된 효모 세포.

[0832] [B69] PPC 유전자가 SEQ ID NO: 11의 아미노산 서열과 적어도 60%, 예를 들어, 적어도 65%, 70%, 75%, 80%, 85%, 90%, 95%, 97%, 98%, 99%, 또는 100% 서열 동일성을 갖는 폴리펩티드를 암호화하는, 단락 B67의 유전적으로 변형된 효모 세포.

[0833] [B70] PPC 유전자가 SEQ ID NO: 12의 아미노산 서열과 적어도 60%, 예를 들어, 적어도 65%, 70%, 75%, 80%, 85%, 90%, 95%, 97%, 98%, 99%, 또는 100% 서열 동일성을 갖는 폴리펩티드를 암호화하는, 단락 B67의 유전적으로 변형된 효모 세포.

[0834] [B71] PPC 유전자가 SEQ ID NO: 9의 뉴클레오티드 서열과 적어도 60%, 예를 들어, 적어도 65%, 70%, 75%, 80%, 85%, 90%, 95%, 97%, 98%, 99%, 또는 100% 서열 동일성을 갖는 뉴클레오티드 서열을 포함하는, B1 내지 단락 B70 중 어느 하나의 유전적으로 변형된 효모 세포.

[0835] [B72] 상기 효모 세포는 크렙트리-음성인, B1 내지 단락 B71 중 어느 하나의 유전적으로 변형된 효모 세포.

[0836] [B73] 효모 세포는 이사트첸키아, 칸디다, 클루이베로미세스, 피치아, 쉬조사카로미세스, 토룰라스포라, 지고사카로미세스, 및 사카로미세스로부터 선택된 속에 속하는, B1 내지 단락 B72 중 어느 하나의 유전적으로 변형된 효모 세포.

[0837] [B74] 효모 세포는 아이. 오리엔탈리스/피. 페르멘탄스 클레이드 및 사카로미세스 클레이드로부터 선택된 클레

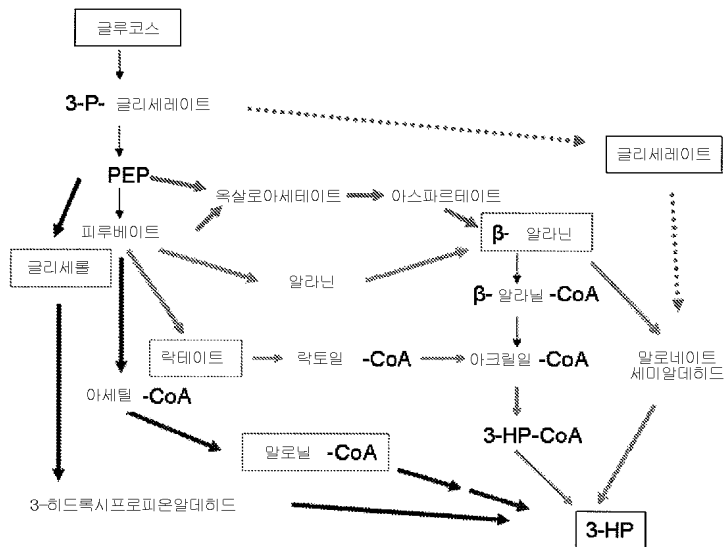
이드에 속하는, 단락 B73의 유전적으로 변형된 효모 세포.

- [0838] [B75] 효모 세포는 아이. 오리엔탈리스, 씨. 람비카, 및 에스. 볼테리로부터 선택되는, 단락 B73의 유전적으로 변형된 효모 세포.
- [0839] [B76] 상기 세포는 PDC, ADH, GAL6, CYB2A, CYB2B, GPD, GPP, ALD, 및 PCK 유전자로부터 선택된 네이티브 유전자의 하나 이상의 삭제 또는 붕괴를 추가로 포함하는, 단락 B1 내지 단락 B75 중 어느 하나의 유전적으로 변형된 효모 세포.
- [0840] [B77] 삭제 또는 붕괴 중 하나 이상이 외인성 3-HP 경로 유전자 중 하나 이상의 삽입으로부터 발생하는, 단락 B76의 유전적으로 변형된 효모 세포.
- [0841] [B78] 외인성 3-HP 경로 유전자 중 하나 이상이 하나 이상의 외인성 조절 요소에 작동 가능하게 결합되는, 단락 B1 내지 단락 B77 중 어느 하나의 유전적으로 변형된 효모 세포.
- [0842] [B79] 하나 이상의 조절 요소가 하나 이상의 3-HP 경로 유전자에 이질적인, 단락 B78의 유전적으로 변형된 효모 세포.
- [0843] [B80] 외인성 PYC 유전자가 PYC 유전자에 이질적인 외인성 프로모터에 작동 가능하게 결합되는, 단락 B1 내지 단락 B79 중 어느 하나의 유전적으로 변형된 효모 세포.
- [0844] [B81] 외인성 AAT 유전자가 AAT 유전자에 이질적인 외인성 프로모터에 작동 가능하게 결합되는, 단락 B1 내지 단락 B80 중 어느 하나의 유전적으로 변형된 효모 세포.
- [0845] [B82] 외인성 ADC 유전자가 ADC 유전자에 이질적인 외인성 프로모터에 작동 가능하게 결합되는, 단락 B1 내지 단락 B81 중 어느 하나의 유전적으로 변형된 효모 세포.
- [0846] [B83] 외인성 BAAT 또는 gabT 유전자가 BAAT 또는 gabT 유전자에 이질적인 외인성 프로모터에 작동 가능하게 결합되는, 단락 B1 내지 단락 B82 중 어느 하나의 유전적으로 변형된 효모 세포.
- [0847] [B84] 외인성 3-HPDH 유전자가 3-HPDH 유전자에 이질적인 외인성 프로모터에 작동 가능하게 결합되는, 단락 B1 내지 단락 B83 중 어느 하나의 유전적으로 변형된 효모 세포.
- [0848] [B85] 외인성 PPC 유전자가 PPC 유전자에 이질적인 외인성 프로모터에 작동 가능하게 결합되는, 단락 B1 내지 단락 B84 중 어느 하나의 유전적으로 변형된 효모 세포.
- [0849] [B86] 세포는 75 g/L 이상의 3-HP를 함유하는 배지에서 4 미만의 pH에서 성장할 수 있는, 단락 B1 내지 단락 B85 중 어느 하나의 유전적으로 변형된 효모 세포.
- [0850] [B87] 세포는 3-HP-저항성 효모 세포인, 단락 B1 내지 단락 B86 중 어느 하나의 유전적으로 변형된 효모 세포.
- [0851] [B88] 세포는 돌연변이 및/또는 선택을 겪었으며, 이러한 돌연변이 및/또는 선택된 세포는 같은 종의 야생형 세포보다 3-HP에 대하여 더 높은 정도의 저항성을 소유하는, 단락 B1 내지 단락 B87 중 어느 하나의 유전적으로 변형된 효모 세포.
- [0852] [B89] 세포는 하나 이상의 외인성 3-HP 경로 유전자를 가지고 유전적으로 변형되기 전에 돌연변이 및/또는 선택을 겪는, 단락 B88의 유전적으로 변형된 효모 세포.
- [0853] [B90] 세포는 젖산 또는 3-HP의 존재시 선택을 겪는, 단락 B88 또는 B89의 유전적으로 변형된 효모 세포.
- [0854] [B91] 선택은 케모스타트 선택인, 단락 B91의 유전적으로 변형된 효모 세포.
- [0855] [B92] (i) 적어도 하나의 탄소 근원을 포함하는 배지의 존재시 단락 B1 내지 단락 B91 중 어느 하나의 유전적으로 변형된 효모 세포를 배양하는 단계; 및
- [0856] (ii) 배양물로부터 3-HP를 분리하는 단계
- [0857] 를 포함하는 3-HP를 생산하는 방법.
- [0858] [B93] 상기 탄소 근원은 글루코스, 자일로스, 아라비노스, 수크로스, 프럭토스, 셀룰로스, 글루코스 올리고머, 및 글리세롤로부터 선택되는, 단락 B92의 방법.
- [0859] [B94] 배지는 5 미만, 예를 들어, 약 1.5 내지 약 4.5, 약 2.0 내지 약 4.0, 또는 약 2.0 내지 약 3.5의 범위

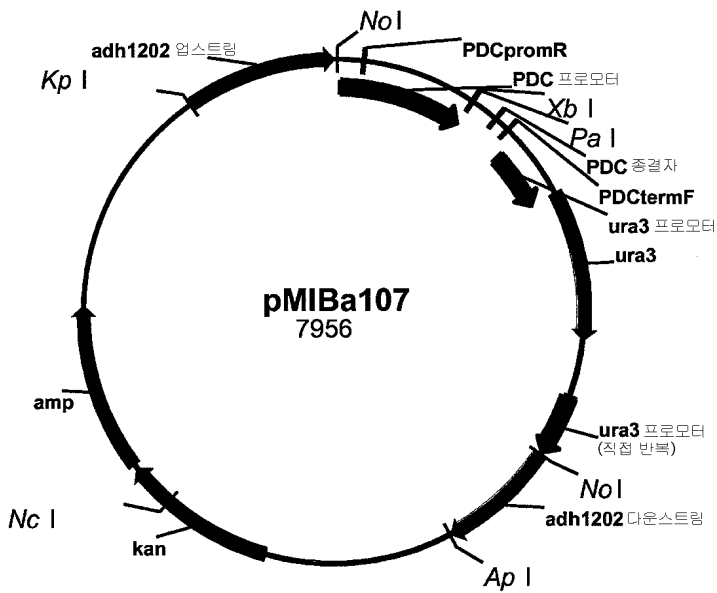
의 pH인, 단락 B92 또는 B93의 방법.

도면

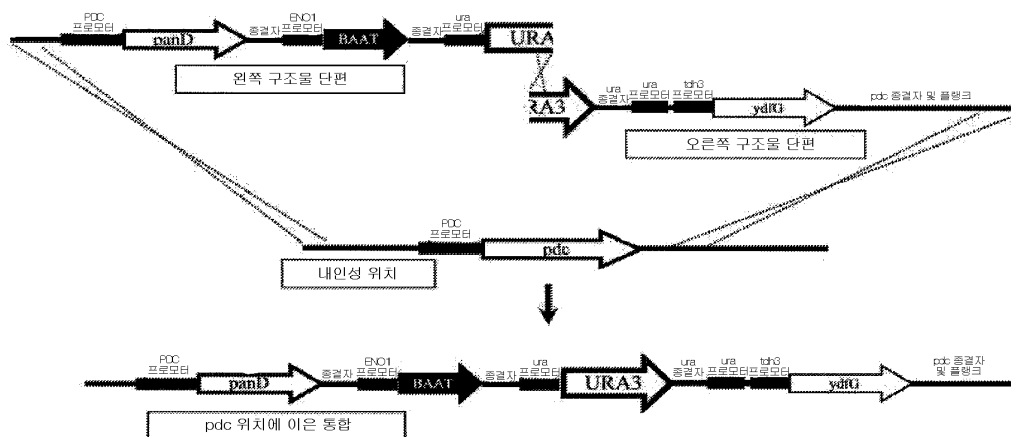
도면1



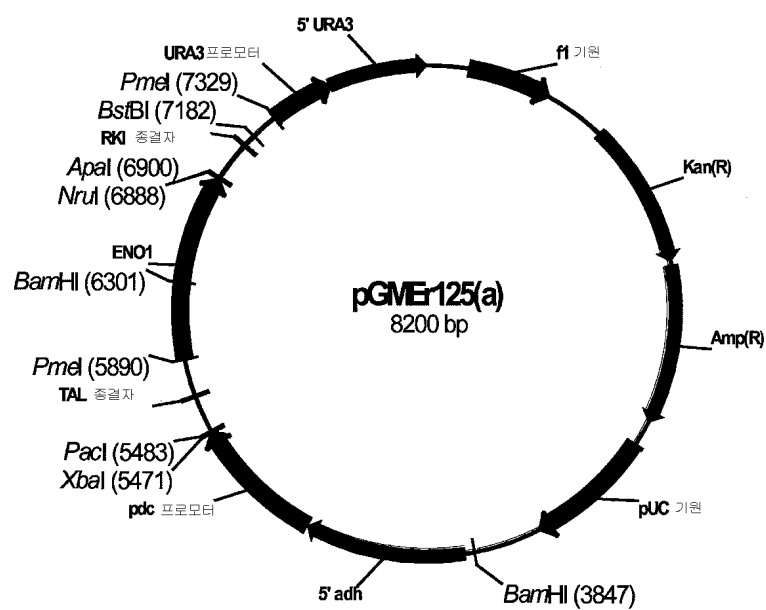
도면2



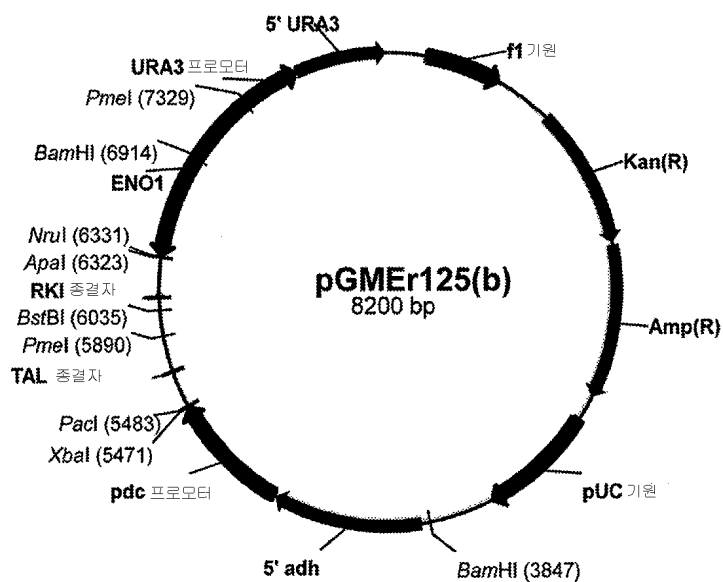
도면3



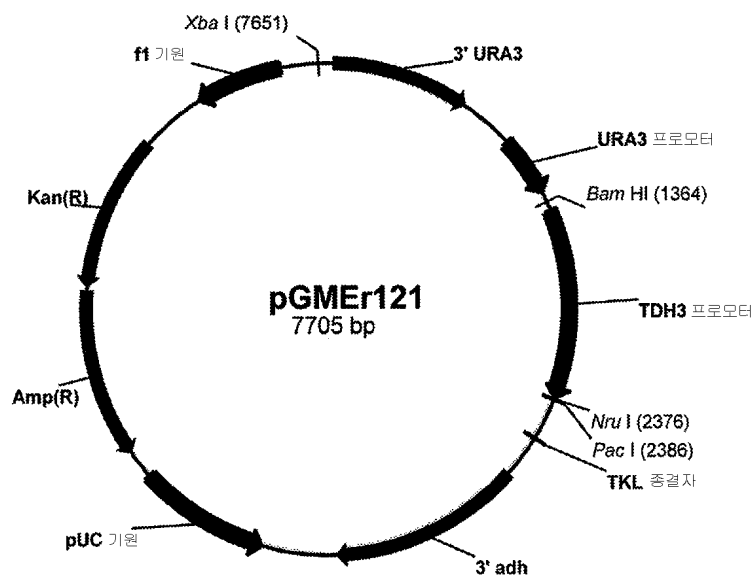
도면4



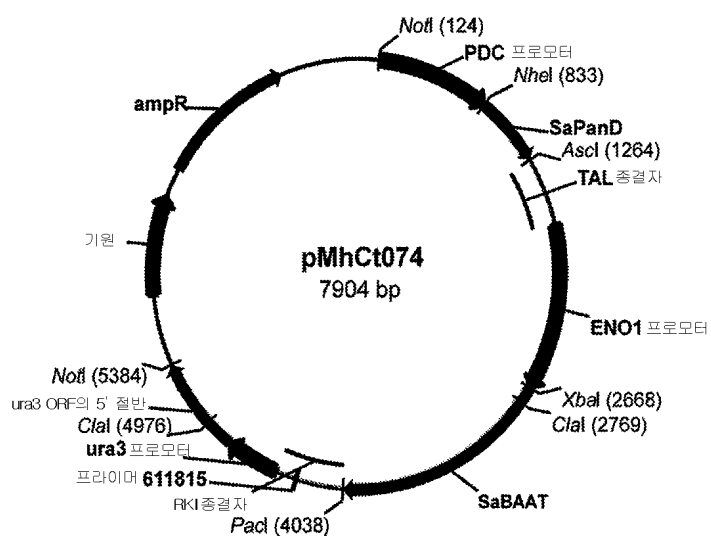
도면5



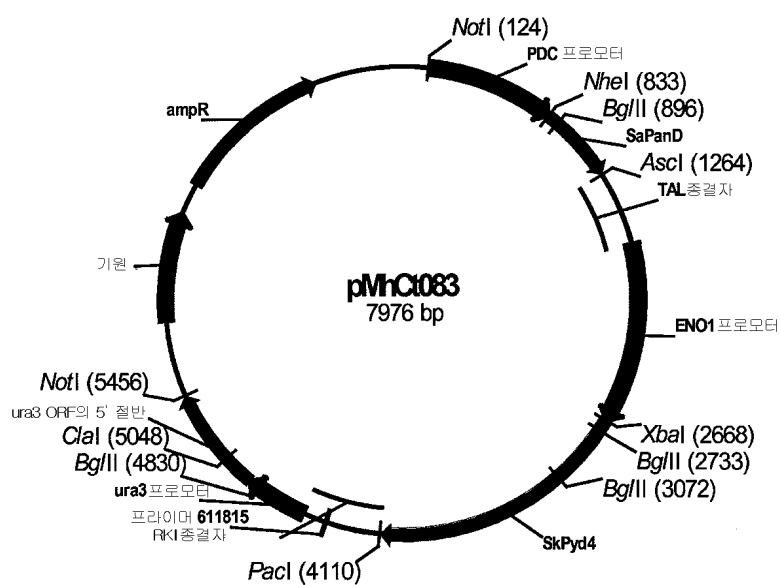
도면6



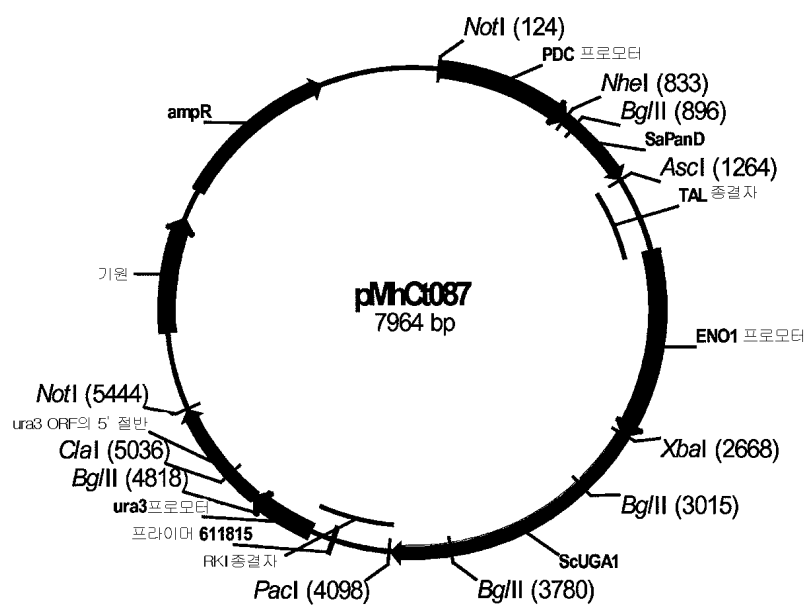
도면7



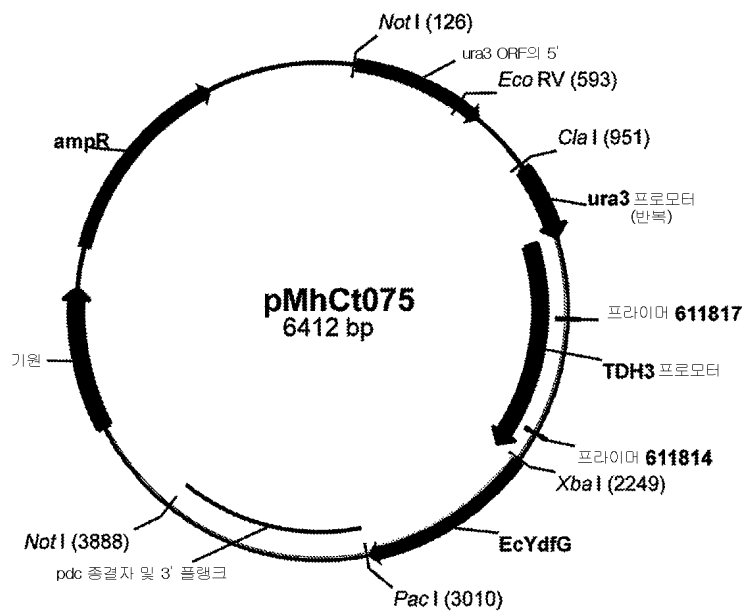
도면8



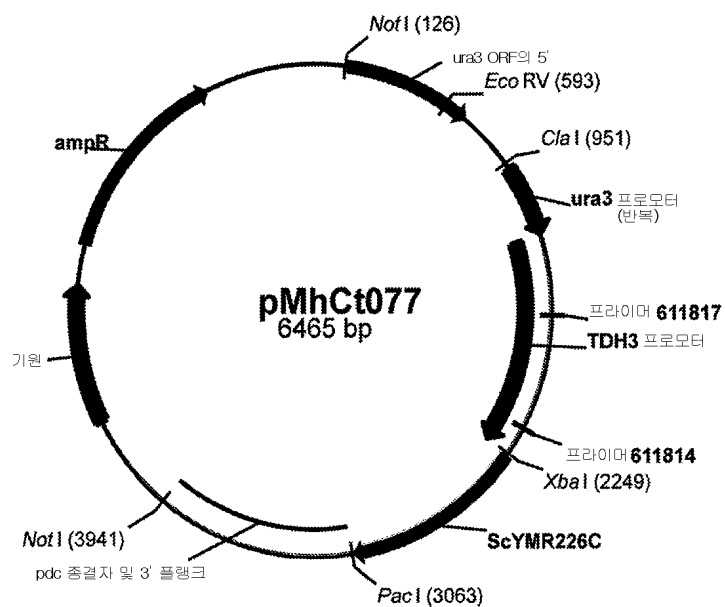
도면9



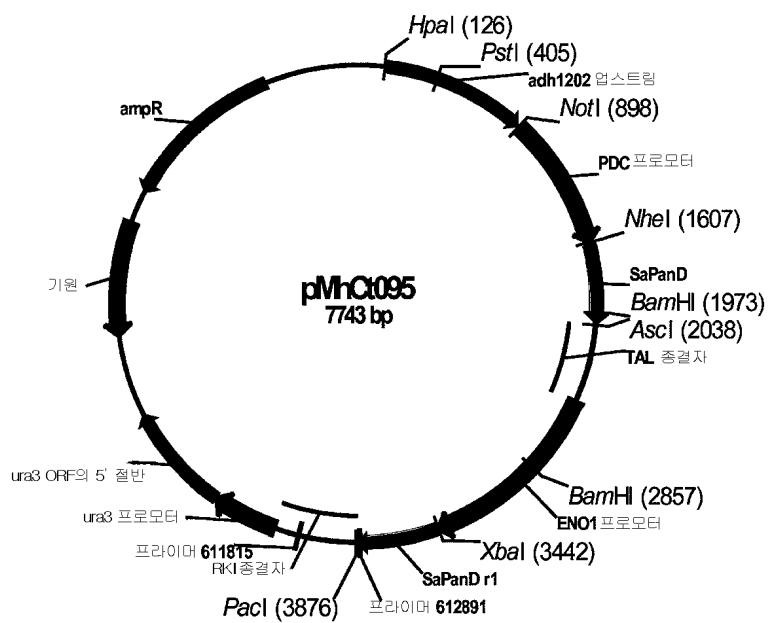
도면10



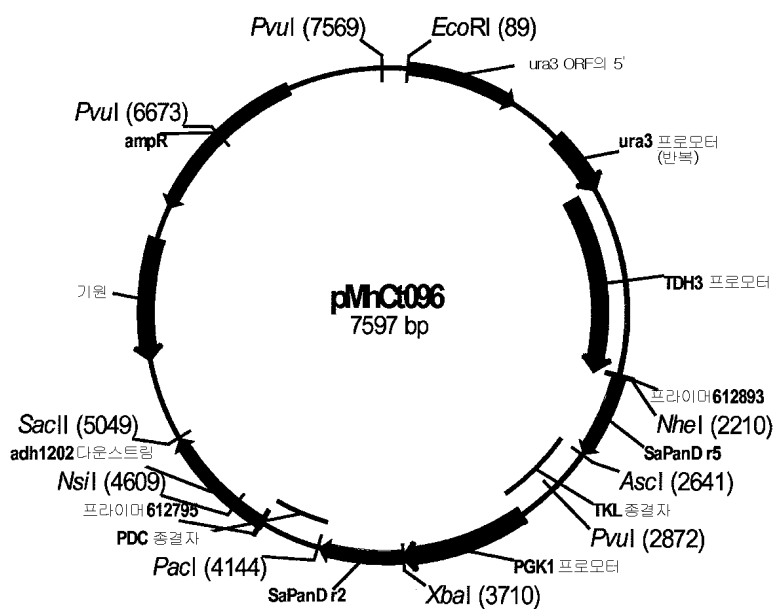
도면11



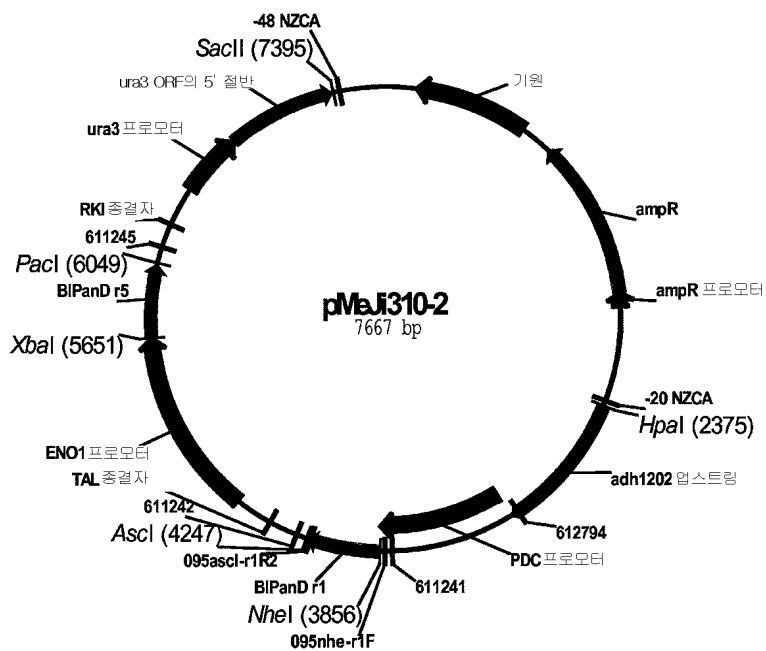
도면12



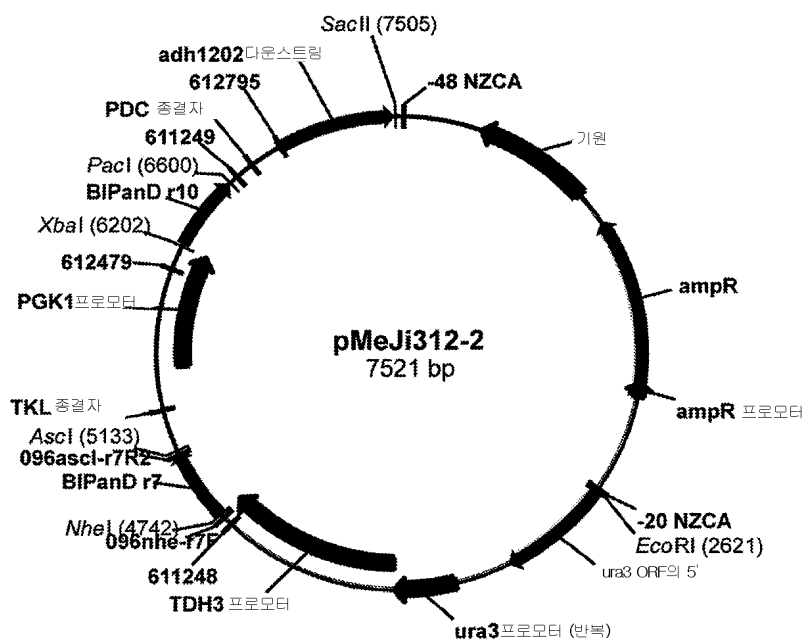
도면13



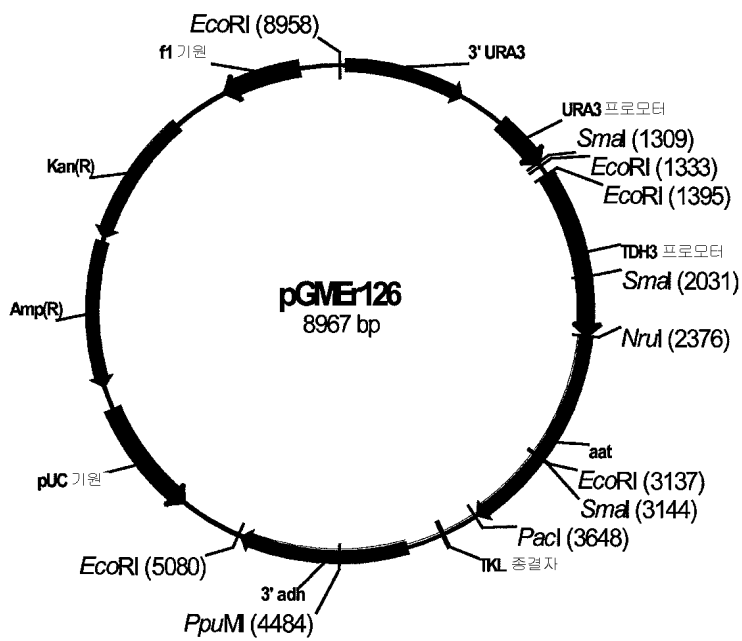
도면14



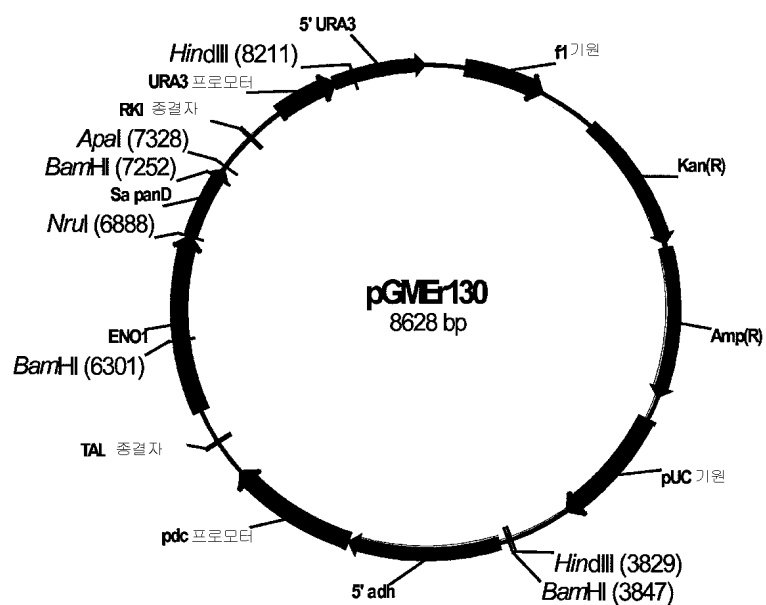
도면15



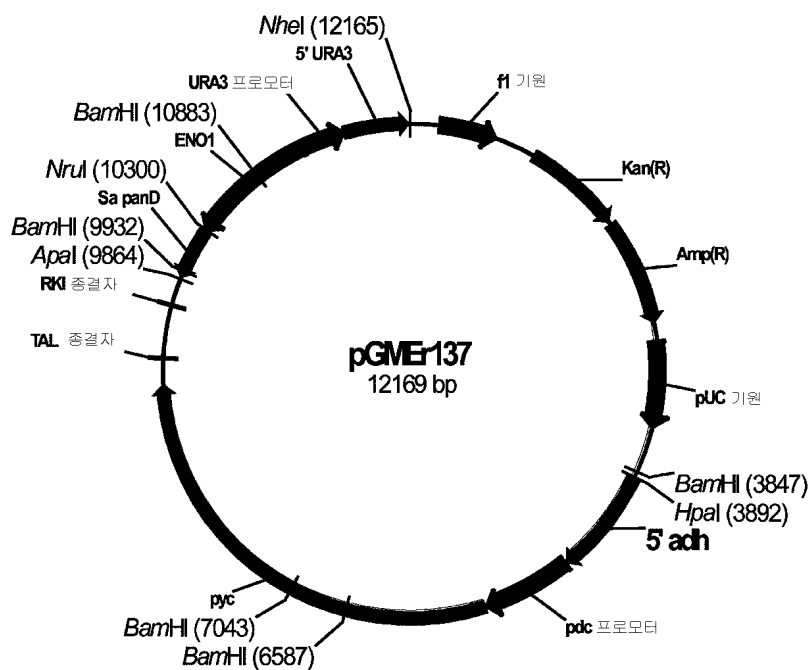
도면16



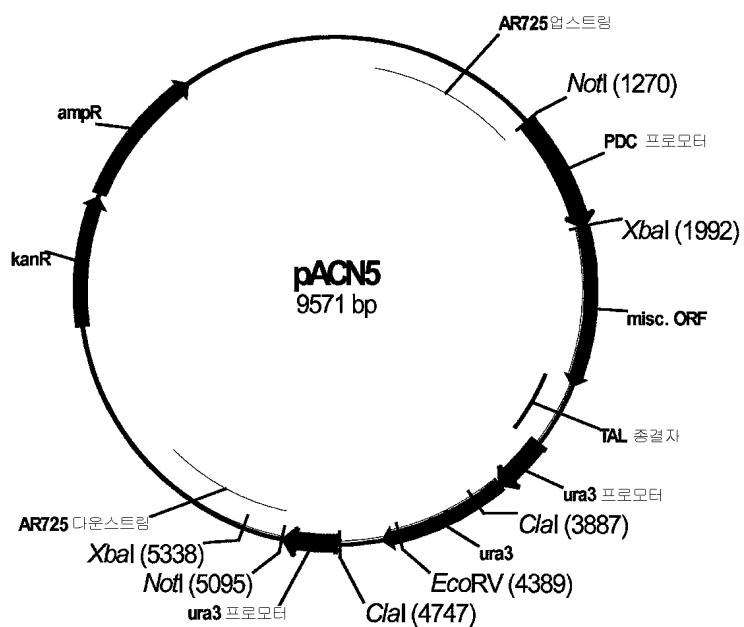
도면17



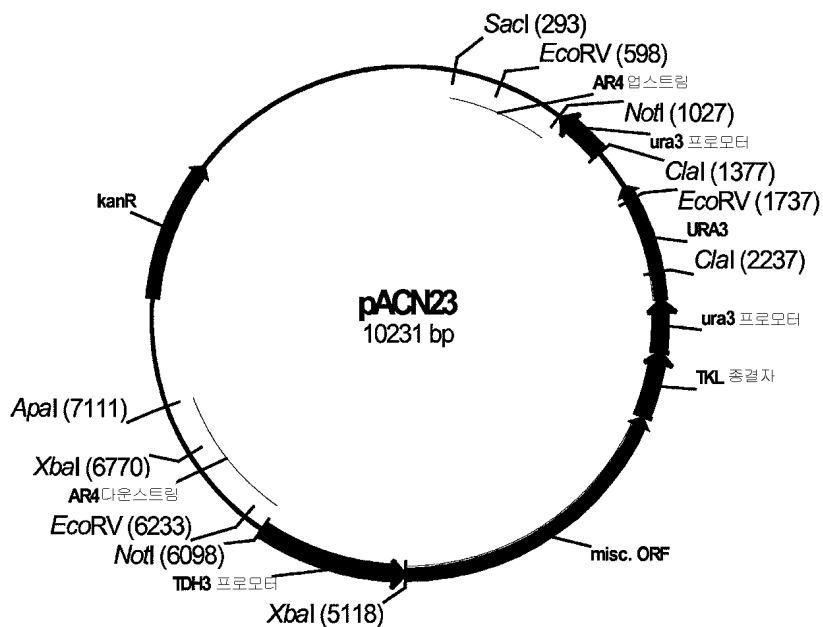
도면18



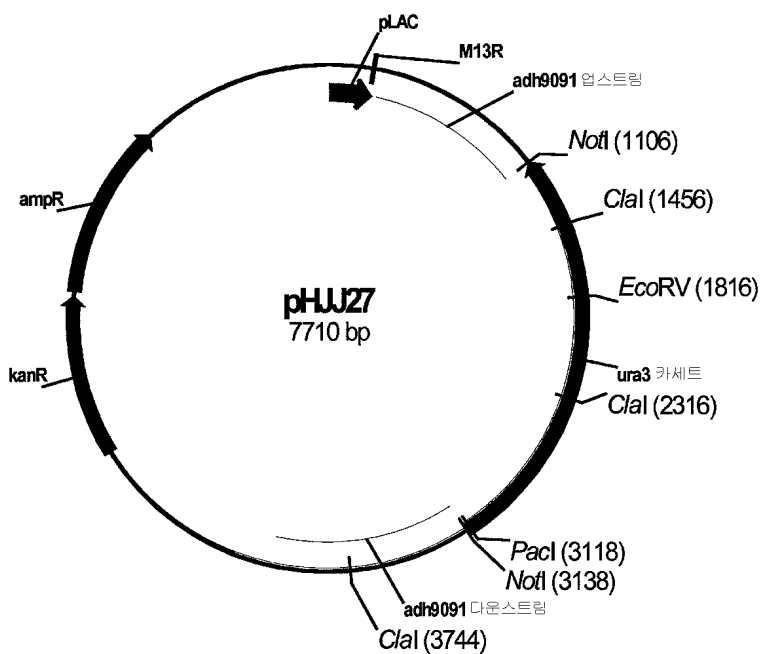
도면19



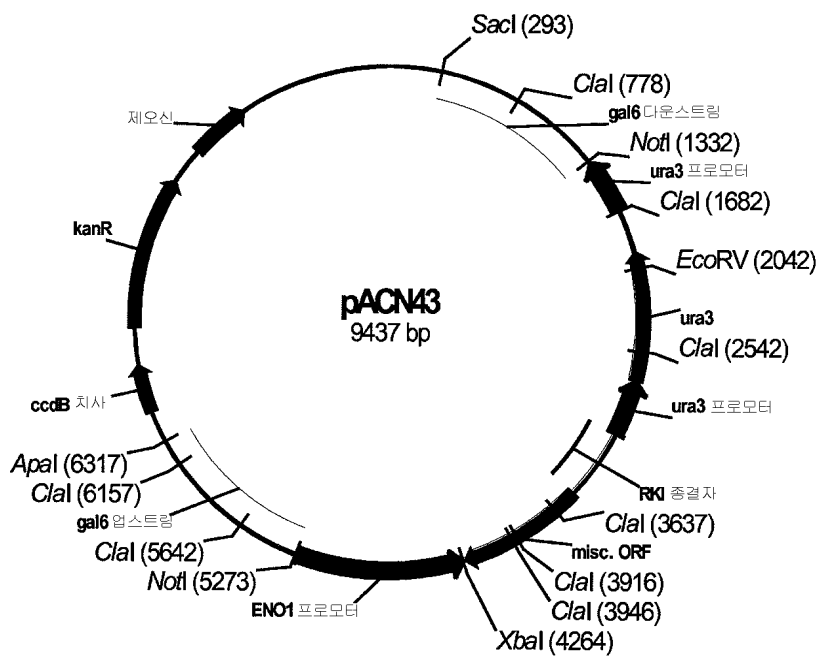
도면20



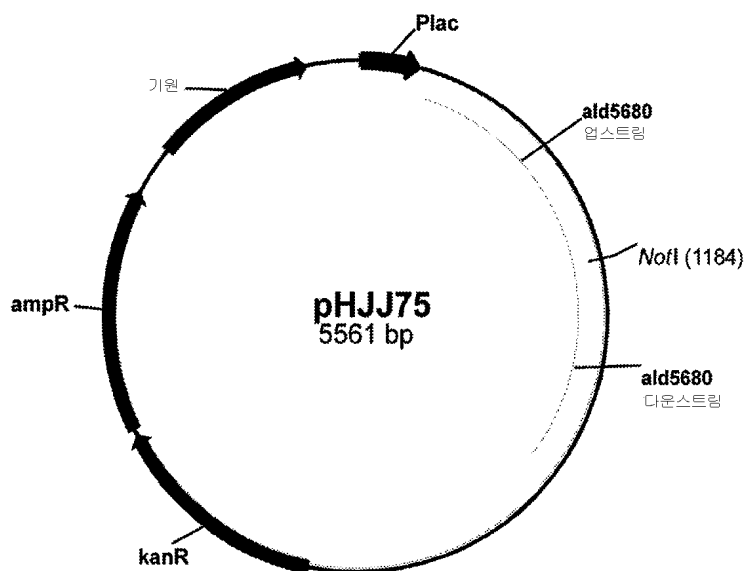
도면21



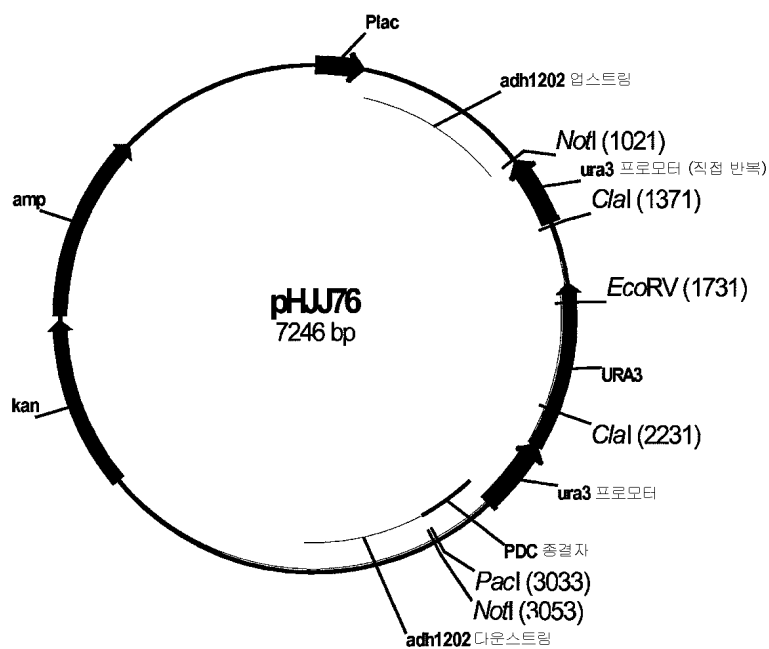
도면22



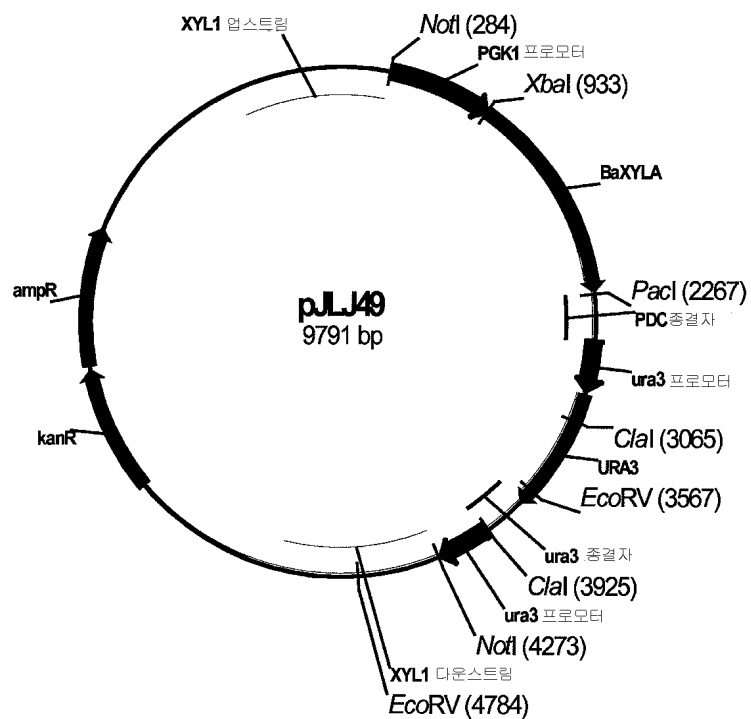
도면23



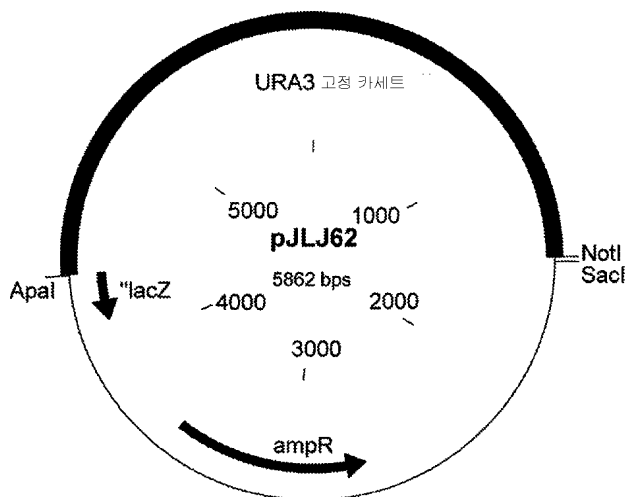
도면24



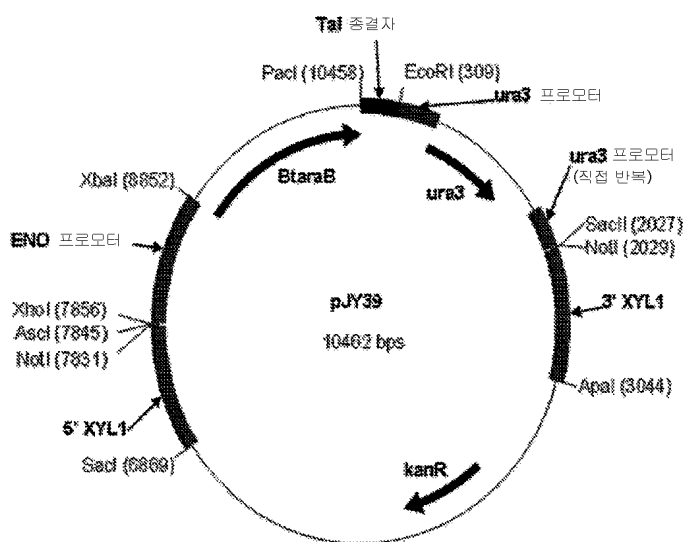
도면25



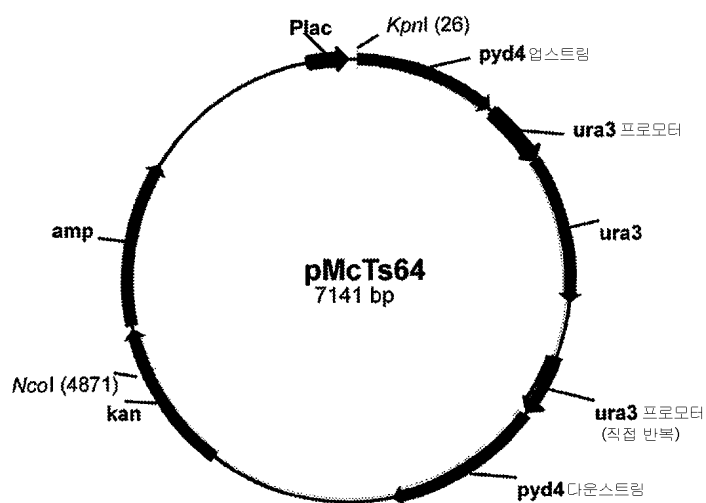
도면26



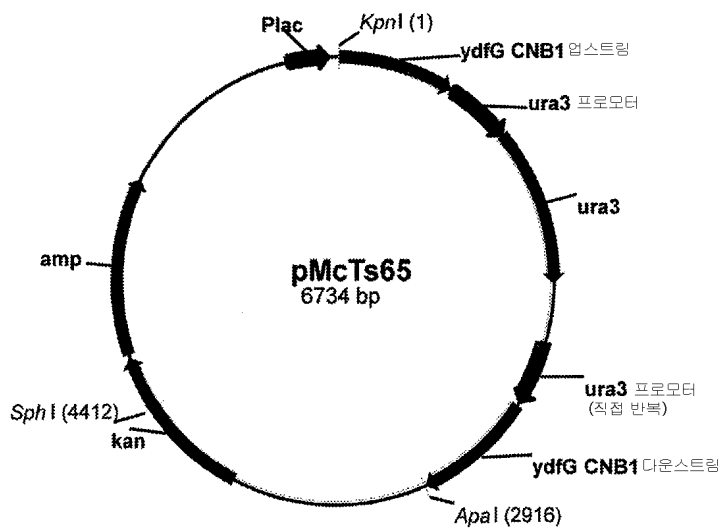
도면29



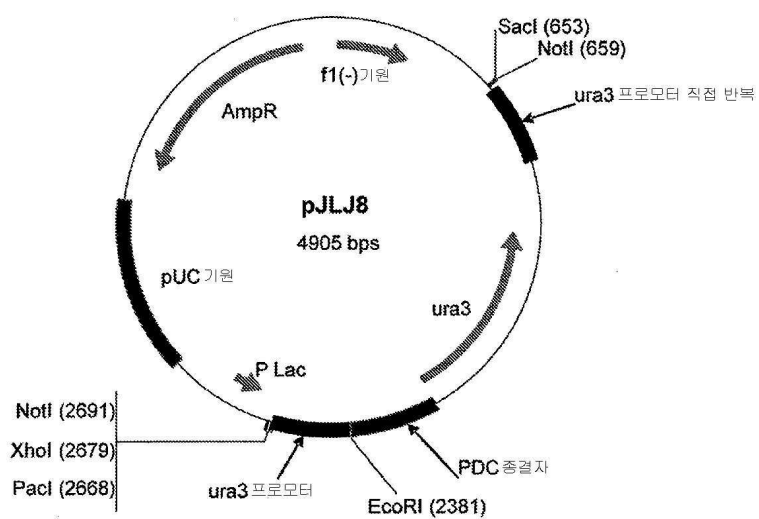
도면30



도면31



도면32



서열 목록

SEQUENCE LISTING

<110> Novozymes, Inc.
 Jessen, Holly
 Rush, Brian
 Huryta, Jeanette
 Mastel, Beth
 Berry, Alan
 Yaver, Debbie

<120> Compositions and Methods for 3-Hydroxypropionic Acid Production
 <130> 12084-WO-PCT
 <150> US 61/416,199
 <151> 2010-11-22
 <150> US 61/535,181
 <151> 2011-09-15
 <160> 351
 <170> PatentIn version 3.5
 <210> 1
 <211> 3543
 <212> DNA
 <213> *Issatchenkia orientalis*
 <400> 1

```

atgtcaactg tggaagatca ctctcccta cataaattga gaaaggaatc tgagattctt      60
tccaatgcaa acaaaatctt agtggctaata agagggtgaaa ttccaattag aattttcagg    120
tcagcccatg aattgtcaat gcatactgtg gcgatctatt cccatgaaga tcggttgtcc      180
atgcataagg tgaaggccga cgaggcttat gcaatcggta agactgggtca atattcgcca    240
gttcaagctt atctacaat tgacgaaatt atcaaaatag caaaggaaca tgatgtttcc      300
atgatccatc caggttatgg tttcttatct gaaaactccg aattcgcaaa gaaggttgaa     360
gaatccggta tgatttgggt tgggcctcct gctgaagtta ttgattctgt tggtgacaag     420

gtttctgcaa gaaatttggc aattaaatgt gacgttctctg ttgttctctgg taccgatggt    480
ccaattgaag acattgaaca ggctaaacag ttgttggaac aatatgggtta tcctgtcatt     540
ataaaggctg cathttgttg tgggtgtaga ggtatgagag ttgttagaga aggtgatgat      600
atagttgatg ctttccaaag agcgtcatct gaagcaaagt ctgccttttg taatggtact     660
tgttttattg aaagattttt ggataagcca aaacatattg aggttcaatt attggctgat      720
aattatggta acacaatcca tctctttgaa agagattggt ctgttcaaag aagacatcaa     780
aaggttgttg aaattgcacc tgccaaaact ttacctgttg aagttagaaa tgctatatta     840

aaggatgctg taacgttagc taaaaccgct aactatagaa atgctggtac tgcagaattt     900
ttagttgatt cccaaaacag acattatttt attgaaatta atccaagaat tcaagttgaa     960
catacaatta ctgaagaaat cacgggtgtt gatattgttg ccgctcaaat tcaaattgct    1020
gcagggtgat cattggaaca attgggtcta ttacaaaaca aaattacaac tagaggtttt    1080
    
```

| | |
|--|------|
| gcaattcaat gtagaattac aaccgaggat cctgctaaga attttgcccc agatacaggt | 1140 |
| aaaattgagg tttatagatc tgcagggtgt aacgggtgtca gattagatgg tggtaatggg | 1200 |
| tttgccgggtg ctgttatatc tcttcattat gactcgatgt tggttaaatg ttcaacatct | 1260 |
| ggttctaaact atgaaattgc cagaagaaag atgattagag ctttagttga atttagaatc | 1320 |
| agagggtgtca agaccaatat tcctttctta ttggcattgc taactcatcc agttttcatt | 1380 |
| tccgggtgatt gttggacaac ttttattgat gataccctt cgttattcga aatggtttct | 1440 |
| tcaaagaata gagcccaaaa attattggca tatattgggtg acttgtgtgt caatggttct | 1500 |
| tcaattaaag gtcaaatggg tttccctaaa ttgaacaagg aagcagaaat cccagatttg | 1560 |
| ttggatccaa atgatgaggt tattgatgtt tctaaacctt ctaccaatgg tctaagaccg | 1620 |
| tatctattaa agtatggacc agatgcgttt tccaaaaaag ttcgtgaatt cgatggttgt | 1680 |
| atgattatgg ataccacctg gagagatgca catcaatcat tattggctac aagagttaga | 1740 |
| actattgatt tactgagaat tgcaccaacg actagtcatg ccttacaaaa tgcatattgca | 1800 |
| ttagaatgtt ggggtggcgc aacatttgat gttgcgatga ggttcctcta tgaagatcct | 1860 |
| tgggagagat taagacaact tagaaaggca gttccaaata ttcctttcca aatgttattg | 1920 |
| agagggtgcta atgggtgtgc ttattcgtca ttacctgata atgcaattga tcattttgtt | 1980 |
| aagcaagcaa aggataatgg tgttgatatt ttcagagtct ttgatgcttt gaacgatttg | 2040 |
| gaacaattga aggttgggtg tgatgctgtc aagaaagccg gaggtgttgt tgaagctaca | 2100 |
| gtttgttact cagggtgatat gttaattcca ggtaaaaagt ataacttga ttattattta | 2160 |
| gagactgttg gaaagattgt ggaaatgggt acccatattt taggtattaa ggatatggct | 2220 |
| ggcacgttaa agccaaaggc tgctaagttg ttgattggct cgatcagatc aaaataccct | 2280 |
| gacttggtta tccatgtcca taccatgac tctgctggta ccggtatttc aacttatgtt | 2340 |
| gcatgcgcat tggcagggtc cgacattgtc gatttgtcaa tcaattcgat gtctggttta | 2400 |
| acctctcaac cttcaatgag tgcttttatt gctgctttag atgggtgatat cgaaactggt | 2460 |
| gttccagaac attttgcaag acaattagat gcatactggg cagaaatgag attgttatac | 2520 |
| tcatgtttcg aagccgactt gaagggaacca gaccagaag ttataaaca tgaaattcca | 2580 |
| ggtggacagt tgactaacct aatcttccaa gcccaacaag ttggtttggg tgaacaatgg | 2640 |
| gaagaaacta agaagaagta tgaagatgct aacatgttgt tgggtgatat tgtcaagggt | 2700 |
| accccaacct ccaaggttgt tgggtattta gcccaattta tggtttctaa taaattagaa | 2760 |
| aaagaagatg ttgaaaaact tgctaataaa ttagatttcc cagattcagt tcttgatttc | 2820 |
| tttgaaggat taatgggtac accatatggt ggattcccag agcctttgag aacaatgtc | 2880 |

atttccggca agagaagaaa attaaagggt agaccagggt tagaattaga acctttcaac 2940

ctcgaggaaa tcagagaaaa ttiggtttcc agatttgggtc caggtattac tgaatgtgat 3000

gttgcatctt ataacatgta tccaaagggt tacgagcaat atcgtaaggt ggttgaaaaa 3060

tatggtgatt tatctgtttt accaacaaaa gcatttttgg ctctccaac tattggtgaa 3120

gaagttcatg tggaaattga gcaaggtaag actttgatta ttaagttatt agccatttct 3180

gacttgtcta aatctcatgg tacaagagaa gtatactttg aattgaatgg tgaaatgaga 3240

aaggttacaa ttgaagataa aacagctgca attgagactg ttacaagagc aaaggctgac 3300

ggacacaatc caaatgaagt tgggtgcgcca atggctgggtg tcgttgttga agttagagtg 3360

aagcatggaa cagaagttaa gaagggtgat ccattagccg ttttgagtgc aatgaaaatg 3420

gaaatggtta tttctgctcc tgtagtgggt agggtcggtg aagttttgt caacgaaggc 3480

gattccgttg atatgggtga ttgcttgtg aaaattgcc aagatgaagc gccagcagct 3540

taa 3543

<210> 2

<211> 1180

<212> PRT

<213> Issatchenkia orientalis

<400> 2

Met Ser Thr Val Glu Asp His Ser Ser Leu His Lys Leu Arg Lys Glu

1 5 10 15

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

20 25 30

Glu Ile Pro Ile Arg Ile Phe Arg Ser Ala His Glu Leu Ser Met His

35 40 45

Thr Val Ala Ile Tyr Ser His Glu Asp Arg Leu Ser Met His Arg Leu

50 55 60

Lys Ala Asp Glu Ala Tyr Ala Ile Gly Lys Thr Gly Gln Tyr Ser Pro

65 70 75 80

Val Gln Ala Tyr Leu Gln Ile Asp Glu Ile Ile Lys Ile Ala Lys Glu

85 90 95

His Asp Val Ser Met Ile His Pro Gly Tyr Gly Phe Leu Ser Glu Asn

100 105 110

Ser Glu Phe Ala Lys Lys Val Glu Glu Ser Gly Met Ile Trp Val Gly
115 120 125

Pro Pro Ala Glu Val Ile Asp Ser Val Gly Asp Lys Val Ser Ala Arg
130 135 140

Asn Leu Ala Ile Lys Cys Asp Val Pro Val Val Pro Gly Thr Asp Gly
145 150 155 160

Pro Ile Glu Asp Ile Glu Gln Ala Lys Gln Phe Val Glu Gln Tyr Gly
165 170 175

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

Arg Val Val Arg Glu Gly Asp Asp Ile Val Asp Ala Phe Gln Arg Ala
195 200 205

Ser Ser Glu Ala Lys Ser Ala Phe Gly Asn Gly Thr Cys Phe Ile Glu
210 215 220

Arg Phe Leu Asp Lys Pro Lys His Ile Glu Val Gln Leu Leu Ala Asp
225 230 235 240

Asn Tyr Gly Asn Thr Ile His Leu Phe Glu Arg Asp Cys Ser Val Gln
245 250 255

Arg Arg His Gln Lys Val Val Glu Ile Ala Pro Ala Lys Thr Leu Pro
260 265 270

Val Glu Val Arg Asn Ala Ile Leu Lys Asp Ala Val Thr Leu Ala Lys
275 280 285

Thr Ala Asn Tyr Arg Asn Ala Gly Thr Ala Glu Phe Leu Val Asp Ser
290 295 300

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

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

Ile Gln Ile Ala Ala Gly Ala Ser Leu Glu Gln Leu Gly Leu Leu Gln
340 345 350

Asn Lys Ile Thr Thr Arg Gly Phe Ala Ile Gln Cys Arg Ile Thr Thr

355 360 365
 Glu Asp Pro Ala Lys Asn Phe Ala Pro Asp Thr Gly Lys Ile Glu Val
 370 375 380
 Tyr Arg Ser Ala Gly Gly Asn Gly Val Arg Leu Asp Gly Gly Asn Gly
 385 390 395 400

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

 Trp Thr Thr Phe Ile Asp Asp Thr Pro Ser Leu Phe Glu Met Val Ser
 465 470 475 480
 Ser Lys Asn Arg Ala Gln Lys Leu Leu Ala Tyr Ile Gly Asp Leu Cys
 485 490 495
 Val Asn Gly Ser Ser Ile Lys Gly Gln Ile Gly Phe Pro Lys Leu Asn
 500 505 510
 Lys Glu Ala Glu Ile Pro Asp Leu Leu Asp Pro Asn Asp Glu Val Ile
 515 520 525

 Asp Val Ser Lys Pro Ser Thr Asn Gly Leu Arg Pro Tyr Leu Leu Lys
 530 535 540
 Tyr Gly Pro Asp Ala Phe Ser Lys Lys Val Arg Glu Phe Asp Gly Cys
 545 550 555 560
 Met Ile Met Asp Thr Thr Trp Arg Asp Ala His Gln Ser Leu Leu Ala
 565 570 575
 Thr Arg Val Arg Thr Ile Asp Leu Leu Arg Ile Ala Pro Thr Thr Ser
 580 585 590

 His Ala Leu Gln Asn Ala Phe Ala Leu Glu Cys Trp Gly Gly Ala Thr
 595 600 605

Phe Asp Val Ala Met Arg Phe Leu Tyr Glu Asp Pro Trp Glu Arg Leu
 610 615 620
 Arg Gln Leu Arg Lys Ala Val Pro Asn Ile Pro Phe Gln Met Leu Leu
 625 630 635 640
 Arg Gly Ala Asn Gly Val Ala Tyr Ser Ser Leu Pro Asp Asn Ala Ile
 645 650 655

 Asp His Phe Val Lys Gln Ala Lys Asp Asn Gly Val Asp Ile Phe Arg
 660 665 670
 Val Phe Asp Ala Leu Asn Asp Leu Glu Gln Leu Lys Val Gly Val Asp
 675 680 685
 Ala Val Lys Lys Ala Gly Gly Val Val Glu Ala Thr Val Cys Tyr Ser
 690 695 700
 Gly Asp Met Leu Ile Pro Gly Lys Lys Tyr Asn Leu Asp Tyr Tyr Leu
 705 710 715 720

 Glu Thr Val Gly Lys Ile Val Glu Met Gly Thr His Ile Leu Gly Ile
 725 730 735
 Lys Asp Met Ala Gly Thr Leu Lys Pro Lys Ala Ala Lys Leu Leu Ile
 740 745 750
 Gly Ser Ile Arg Ser Lys Tyr Pro Asp Leu Val Ile His Val His Thr
 755 760 765
 His Asp Ser Ala Gly Thr Gly Ile Ser Thr Tyr Val Ala Cys Ala Leu
 770 775 780

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

 Gly Pro Asp Pro Glu Val Tyr Lys His Glu Ile Pro Gly Gly Gln Leu

| | | | |
|---|------|------|------|
| 850 | 855 | 860 | |
| Thr Asn Leu Ile Phe Gln Ala Gln Gln Val Gly Leu Gly Glu Gln Trp | | | |
| 865 | 870 | 875 | 880 |
| Glu Glu Thr Lys Lys Lys Tyr Glu Asp Ala Asn Met Leu Leu Gly Asp | | | |
| | 885 | 890 | 895 |
| Ile Val Lys Val Thr Pro Thr Ser Lys Val Val Gly Asp Leu Ala Gln | | | |
| | 900 | 905 | 910 |
| | | | |
| Phe Met Val Ser Asn Lys Leu Glu Lys Glu Asp Val Glu Lys Leu Ala | | | |
| | 915 | 920 | 925 |
| Asn Glu Leu Asp Phe Pro Asp Ser Val Leu Asp Phe Phe Glu Gly Leu | | | |
| | 930 | 935 | 940 |
| Met Gly Thr Pro Tyr Gly Gly Phe Pro Glu Pro Leu Arg Thr Asn Val | | | |
| 945 | 950 | 955 | 960 |
| Ile Ser Gly Lys Arg Arg Lys Leu Lys Gly Arg Pro Gly Leu Glu Leu | | | |
| | 965 | 970 | 975 |
| | | | |
| Glu Pro Phe Asn Leu Glu Glu Ile Arg Glu Asn Leu Val Ser Arg Phe | | | |
| | 980 | 985 | 990 |
| Gly Pro Gly Ile Thr Glu Cys Asp Val Ala Ser Tyr Asn Met Tyr Pro | | | |
| | 995 | 1000 | 1005 |
| Lys Val Tyr Glu Gln Tyr Arg Lys Val Val Glu Lys Tyr Gly Asp | | | |
| | 1010 | 1015 | 1020 |
| Leu Ser Val Leu Pro Thr Lys Ala Phe Leu Ala Pro Pro Thr Ile | | | |
| | 1025 | 1030 | 1035 |
| | | | |
| Gly Glu Glu Val His Val Glu Ile Glu Gln Gly Lys Thr Leu Ile | | | |
| | 1040 | 1045 | 1050 |
| Ile Lys Leu Leu Ala Ile Ser Asp Leu Ser Lys Ser His Gly Thr | | | |
| | 1055 | 1060 | 1065 |
| Arg Glu Val Tyr Phe Glu Leu Asn Gly Glu Met Arg Lys Val Thr | | | |
| | 1070 | 1075 | 1080 |
| Ile Glu Asp Lys Thr Ala Ala Ile Glu Thr Val Thr Arg Ala Lys | | | |
| | 1085 | 1090 | 1095 |

Ala Asp Gly His Asn Pro Asn Glu Val Gly Ala Pro Met Ala Gly
 1100 1105 1110
 Val Val Val Glu Val Arg Val Lys His Gly Thr Glu Val Lys Lys
 1115 1120 1125
 Gly Asp Pro Leu Ala Val Leu Ser Ala Met Lys Met Glu Met Val
 1130 1135 1140
 Ile Ser Ala Pro Val Ser Gly Arg Val Gly Glu Val Phe Val Asn
 1145 1150 1155

Glu Gly Asp Ser Val Asp Met Gly Asp Leu Leu Val Lys Ile Ala
 1160 1165 1170
 Lys Asp Glu Ala Pro Ala Ala
 1175 1180

<210> 3

<211> 1154

<212> PRT

<213> Rhodobacter sphaeroides

<400> 3

Met Ala Glu Phe Arg Lys Ile Leu Ile Ala Asn Arg Gly Glu Ile Ala
 1 5 10 15
 Ile Arg Val Met Arg Ala Ala Asn Glu Met Gly Lys Lys Thr Val Ala
 20 25 30

Val Tyr Ala Glu Glu Asp Lys Leu Ser Leu His Arg Phe Lys Ala Asp
 35 40 45
 Glu Ala Tyr Arg Ile Gly Glu Gly Leu Ser Pro Val Gly Ala Tyr Leu
 50 55 60
 Ser Ile Pro Glu Ile Ile Arg Val Ala Gln Met Ser Gly Ala Asp Ala
 65 70 75 80
 Ile His Pro Gly Tyr Gly Leu Leu Ser Glu Asn Pro Asp Phe Val Glu
 85 90 95

Ala Cys Asp Ala Ala Gly Ile Ala Phe Ile Gly Pro Lys Ala Glu Thr
 100 105 110
 Met Arg Ala Leu Gly Asp Lys Ala Ser Ala Arg Arg Val Ala Met Ala

115 120 125
Ala Gly Val Pro Val Ile Pro Ala Thr Glu Val Leu Gly Asp Asp Met
130 135 140
Glu Glu Ile Lys Arg Gln Ala Ala Glu Ile Gly Tyr Pro Leu Met Leu
145 150 155 160

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

Tyr His Leu Tyr Glu Arg Asp Cys Thr Val Gln Arg Arg Asn Gln Lys
225 230 235 240
Val Val Glu Arg Ala Pro Ala Pro Tyr Leu Thr Glu Glu Gln Arg Thr
245 250 255
Glu Ile Cys Glu Leu Gly Arg Arg Ile Cys Ala His Val Asn Tyr Glu
260 265 270
Cys Ala Gly Thr Val Glu Phe Leu Met Asp Met Asp Ser Glu Lys Phe
275 280 285

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

Pro Gln Asn Asn Phe Ile Pro Asp Tyr Gly Arg Leu Thr Ala Tyr Arg
355 360 365

Ser Ala Thr Gly Met Gly Ile Arg Leu Asp Gly Gly Thr Ala Tyr Ala
370 375 380

Gly Gly Val Ile Thr Arg Tyr Tyr Asp Ser Leu Leu Val Lys Val Thr
385 390 395 400

Ala Trp Ala Pro Thr Pro Glu Lys Ala Ile Ala Arg Met Asp Arg Ala
405 410 415

Leu Arg Glu Phe Arg Ile Arg Gly Val Ala Thr Asn Ile Ala Phe Val
420 425 430

Glu Asn Leu Leu Lys His Pro Ser Phe Leu Asp Tyr Ser Tyr Thr Thr
435 440 445

Lys Phe Ile Asp Thr Thr Pro Asp Leu Phe Asn Phe Lys Pro Arg Arg
450 455 460

Asp Arg Ala Thr Lys Ile Leu Thr Tyr Ile Ala Asp Ile Thr Val Asn
465 470 475 480

Gly His Pro Glu Thr Ala Gly Arg Val Arg Pro Ser Ala Glu Leu Lys
485 490 495

Asp Pro Lys Ala Pro Glu Pro Lys Gly Ala Pro Gln Pro Gly Thr Arg
500 505 510

Thr Leu Leu Glu Glu Lys Gly Pro Gln Ala Val Ala Asp Trp Met Ala
515 520 525

Ala Gln Thr Arg Val Leu Met Thr Asp Thr Thr Met Arg Asp Gly His
530 535 540

Gln Ser Leu Leu Ala Thr Arg Met Arg Ser Ile Asp Met Ile Lys Val
545 550 555 560

Thr Pro Ala Tyr Ala Ala Asn Leu Gly Gly Leu Phe Ser Val Glu Cys
565 570 575

Trp Gly Gly Ala Thr Phe Asp Val Ala Tyr Arg Phe Leu Gln Glu Cys
580 585 590

Pro Trp Gln Arg Leu Arg Asp Ile Arg Ala Arg Leu Pro Asn Val Met
595 600 605

Thr Gln Met Leu Leu Arg Ala Ser Asn Gly Val Gly Tyr Thr Asn Tyr

| | | | |
|---|-----|-----|-----|
| 610 | 615 | 620 | |
| Pro Asp Asn Val Val Gln Glu Phe Val Arg Gln Ala Ala Glu Thr Gly | | | |
| 625 | 630 | 635 | 640 |
| Val Asp Val Phe Arg Val Phe Asp Ser Leu Asn Trp Val Glu Asn Met | | | |
| | 645 | 650 | 655 |
| Arg Val Ala Met Asp Ala Val Ile Glu Ala Asn Lys Val Cys Glu Gly | | | |
| | 660 | 665 | 670 |
| Thr Ile Cys Tyr Thr Gly Asp Leu Leu Asp Pro Asp Arg Ser Lys Tyr | | | |
| | 675 | 680 | 685 |
| Asp Leu Asn Tyr Tyr Val Gly Met Gly Arg Ala Leu Arg Asp Ala Gly | | | |
| | 690 | 695 | 700 |
| Ala His Val Leu Gly Leu Lys Asp Met Ala Gly Leu Leu Lys Pro Ala | | | |
| 705 | 710 | 715 | 720 |
| Ala Ala Arg Val Leu Val Lys Ala Leu Lys Glu Glu Val Gly Leu Pro | | | |
| | 725 | 730 | 735 |
| Ile His Phe His Thr His Asp Thr Ser Gly Ile Ala Gly Ala Thr Val | | | |
| | 740 | 745 | 750 |
| Leu Ala Ala Cys Asp Ala Gly Val Asp Ala Val Asp Ala Ala Met Asp | | | |
| | 755 | 760 | 765 |
| Ala Phe Ser Gly Gly Thr Ser Gln Pro Cys Leu Gly Ser Ile Val Glu | | | |
| | 770 | 775 | 780 |
| Ala Leu Lys His Thr Asp Arg Asp Thr Gly Leu Asp Ile Ala Ala Ile | | | |
| 785 | 790 | 795 | 800 |
| Arg Glu Ile Ser Asp Tyr Trp Gly His Val Arg Gln Gln Tyr Ser Ala | | | |
| | 805 | 810 | 815 |
| Phe Glu Ser Gly Leu Pro Ser Pro Ala Ser Glu Val Tyr Leu His Glu | | | |
| | 820 | 825 | 830 |
| Met Pro Gly Gly Gln Phe Thr Asn Leu Lys Ala Gln Ala Arg Ser Met | | | |
| | 835 | 840 | 845 |
| Gly Leu Glu Glu Arg Trp Ser Glu Val Ala Gln Ala Tyr Ala Asp Ala | | | |
| | 850 | 855 | 860 |

Asn Arg Met Phe Gly Asp Ile Val Lys Val Thr Pro Ser Ser Lys Val
865 870 875 880
Val Gly Asp Met Ala Leu Met Met Val Ala Gln Gly Leu Thr Arg Glu
885 890 895
Glu Val Glu Asp Pro Glu Val Glu Val Ser Phe Pro Asp Ser Val Val
900 905 910
Asp Met Leu Lys Gly Asn Leu Gly Gln Pro His Gly Gly Trp Pro Glu
915 920 925

Pro Ile Leu Lys Lys Val Leu Lys Gly Glu Ala Pro Ser Thr Glu Arg
930 935 940
Pro Gly Ala His Leu Pro Pro Val Asp Ile Ala Ala Ala Arg Glu Lys
945 950 955 960
Leu Leu Ser Glu Ile Lys Gln Gly Asp Asp Asp Pro Leu Asp Thr Ala
965 970 975
Val Asp Ala Glu Asp Leu Asn Gly Tyr Leu Met Tyr Pro Lys Val Phe
980 985 990

Thr Asp Tyr Arg Ala Arg His Arg Ile Tyr Gly Pro Val Arg Thr Leu
995 1000 1005
Pro Thr Arg Thr Phe Phe Tyr Gly Met Glu Pro Gly Glu Glu Ile
1010 1015 1020
Ser Ala Glu Ile Asp Pro Gly Lys Thr Leu Glu Ile Arg Leu Ser
1025 1030 1035
Ala Val Gly Glu Thr Ser Asp Asp Gly Asp Ala Lys Val Phe Phe
1040 1045 1050

Glu Leu Asn Gly Gln Pro Arg Val Ile Arg Val Ala Asn Arg Ala
1055 1060 1065
Val Lys Ala Lys Thr Ala Thr Arg Pro Lys Ala Gln Asp Gly Asn
1070 1075 1080
Pro Ala His Val Gly Ala Pro Met Pro Gly Ser Val Ala Ser Val
1085 1090 1095
Ala Val Ser Ala Gly Gln Lys Val Lys Pro Gly Asp Leu Leu Val

1100 1105 1110
 Thr Ile Glu Ala Met Lys Met Glu Thr Gly Leu His Ala Asp Arg
 1115 1120 1125
 Ala Ala Thr Val Lys Ala Val His Val Gly Pro Gly Ala Gln Ile
 1130 1135 1140
 Glu Ala Lys Asp Leu Leu Val Glu Leu Glu Asp
 1145 1150
 <210> 4
 <211> 1154
 <212> PRT
 <213> Rhizobium etli
 <400> 4
 Met Pro Ile Ser Lys Ile Leu Val Ala Asn Arg Ser Glu Ile Ala Ile
 1 5 10 15

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

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

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

| | | | |
|---|-----|-----|-----|
| 145 | 150 | 155 | 160 |
| Tyr Pro Val Met Leu Lys Ala Ser Trp Gly Gly Gly Gly Arg Gly Met | | | |
| | 165 | 170 | 175 |
| Arg Val Ile Arg Ser Glu Ala Asp Leu Ala Lys Glu Val Thr Glu Ala | | | |
| | 180 | 185 | 190 |
| Lys Arg Glu Ala Met Ala Ala Phe Gly Lys Asp Glu Val Tyr Leu Glu | | | |
| | 195 | 200 | 205 |
| | | | |
| Lys Leu Val Glu Arg Ala Arg His Val Glu Ser Gln Ile Leu Gly Asp | | | |
| | 210 | 215 | 220 |
| Thr His Gly Asn Val Val His Leu Phe Glu Arg Asp Cys Ser Val Gln | | | |
| 225 | 230 | 235 | 240 |
| Arg Arg Asn Gln Lys Val Val Glu Arg Ala Pro Ala Pro Tyr Leu Ser | | | |
| | 245 | 250 | 255 |
| Glu Ala Gln Arg Gln Glu Leu Ala Ala Tyr Ser Leu Lys Ile Ala Gly | | | |
| | 260 | 265 | 270 |
| | | | |
| Ala Thr Asn Tyr Ile Gly Ala Gly Thr Val Glu Tyr Leu Met Asp Ala | | | |
| | 275 | 280 | 285 |
| Asp Thr Gly Lys Phe Tyr Phe Ile Glu Val Asn Pro Arg Ile Gln Val | | | |
| | 290 | 295 | 300 |
| Glu His Thr Val Thr Glu Val Val Thr Gly Ile Asp Ile Val Lys Ala | | | |
| 305 | 310 | 315 | 320 |
| Gln Ile His Ile Leu Asp Gly Ala Ala Ile Gly Thr Pro Gln Ser Gly | | | |
| | 325 | 330 | 335 |
| | | | |
| Val Pro Asn Gln Glu Asp Ile Arg Leu Asn Gly His Ala Leu Gln Cys | | | |
| | 340 | 345 | 350 |
| Arg Val Thr Thr Glu Asp Pro Glu His Asn Phe Ile Pro Asp Tyr Gly | | | |
| | 355 | 360 | 365 |
| Arg Ile Thr Ala Tyr Arg Ser Ala Ser Gly Phe Gly Ile Arg Leu Asp | | | |
| | 370 | 375 | 380 |
| Gly Gly Thr Ser Tyr Ser Gly Ala Ile Ile Thr Arg Tyr Tyr Asp Pro | | | |
| 385 | 390 | 395 | 400 |

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

 Gln Gln Val Lys Arg Gln Asp Arg Ala Thr Lys Leu Leu Thr Tyr Leu
 465 470 475 480
 Ala Asp Val Thr Val Asn Gly His Pro Glu Ala Lys Asp Arg Pro Lys
 485 490 495
 Pro Leu Glu Asn Ala Ala Arg Pro Val Val Pro Tyr Ala Asn Gly Asn
 500 505 510
 Gly Val Lys Asp Gly Thr Lys Gln Leu Leu Asp Thr Leu Gly Pro Lys
 515 520 525

 Lys Phe Gly Glu Trp Met Arg Asn Glu Lys Arg Val Leu Leu Thr Asp
 530 535 540
 Thr Thr Met Arg Asp Gly His Gln Ser Leu Leu Ala Thr Arg Met Arg
 545 550 555 560
 Thr Tyr Asp Ile Ala Arg Ile Ala Gly Thr Tyr Ser His Ala Leu Pro
 565 570 575
 Asn Leu Leu Ser Leu Glu Cys Trp Gly Gly Ala Thr Phe Asp Val Ser
 580 585 590

 Met Arg Phe Leu Thr Glu Asp Pro Trp Glu Arg Leu Ala Leu Ile Arg
 595 600 605
 Glu Gly Ala Pro Asn Leu Leu Leu Gln Met Leu Leu Arg Gly Ala Asn
 610 615 620
 Gly Val Gly Tyr Thr Asn Tyr Pro Asp Asn Val Val Lys Tyr Phe Val
 625 630 635 640
 Arg Gln Ala Ala Lys Gly Gly Ile Asp Leu Phe Arg Val Phe Asp Cys

| | | |
|---|-----|-----|
| 645 | 650 | 655 |
| Leu Asn Trp Val Glu Asn Met Arg Val Ser Met Asp Ala Ile Ala Glu | | |
| 660 | 665 | 670 |
| Glu Asn Lys Leu Cys Glu Ala Ala Ile Cys Tyr Thr Gly Asp Ile Leu | | |
| 675 | 680 | 685 |
| Asn Ser Ala Arg Pro Lys Tyr Asp Leu Lys Tyr Tyr Thr Asn Leu Ala | | |
| 690 | 695 | 700 |
| Val Glu Leu Glu Lys Ala Gly Ala His Ile Ile Ala Val Lys Asp Met | | |
| 705 | 710 | 715 |
| | | 720 |
| Ala Gly Leu Leu Lys Pro Ala Ala Ala Lys Val Leu Phe Lys Ala Leu | | |
| 725 | 730 | 735 |
| Arg Glu Ala Thr Gly Leu Pro Ile His Phe His Thr His Asp Thr Ser | | |
| 740 | 745 | 750 |
| Gly Ile Ala Ala Ala Thr Val Leu Ala Ala Val Glu Ala Gly Val Asp | | |
| 755 | 760 | 765 |
| Ala Val Asp Ala Ala Met Asp Ala Leu Ser Gly Asn Thr Ser Gln Pro | | |
| 770 | 775 | 780 |
| Cys Leu Gly Ser Ile Val Glu Ala Leu Ser Gly Ser Glu Arg Asp Pro | | |
| 785 | 790 | 795 |
| | | 800 |
| Gly Leu Asp Pro Ala Trp Ile Arg Arg Ile Ser Phe Tyr Trp Glu Ala | | |
| 805 | 810 | 815 |
| Val Arg Asn Gln Tyr Ala Ala Phe Glu Ser Asp Leu Lys Gly Pro Ala | | |
| 820 | 825 | 830 |
| Ser Glu Val Tyr Leu His Glu Met Pro Gly Gly Gln Phe Thr Asn Leu | | |
| 835 | 840 | 845 |
| Lys Glu Gln Ala Arg Ser Leu Gly Leu Glu Thr Arg Trp His Gln Val | | |
| 850 | 855 | 860 |
| Ala Gln Ala Tyr Ala Asp Ala Asn Gln Met Phe Gly Asp Ile Val Lys | | |
| 865 | 870 | 875 |
| | | 880 |
| Val Thr Pro Ser Ser Lys Val Val Gly Asp Met Ala Leu Met Met Val | | |
| 885 | 890 | 895 |

Ser Gln Asp Leu Thr Val Ala Asp Val Val Ser Pro Asp Arg Glu Val
900 905 910

Ser Phe Pro Glu Ser Val Val Ser Met Leu Lys Gly Asp Leu Gly Gln
915 920 925

Pro Pro Ser Gly Trp Pro Glu Ala Leu Gln Lys Lys Ala Leu Lys Gly
930 935 940

Glu Lys Pro Tyr Thr Val Arg Pro Gly Ser Leu Leu Lys Glu Ala Asp
945 950 955 960

Leu Asp Ala Glu Arg Lys Val Ile Glu Lys Lys Leu Glu Arg Glu Val
965 970 975

Ser Asp Phe Glu Phe Ala Ser Tyr Leu Met Tyr Pro Lys Val Phe Thr
980 985 990

Asp Phe Ala Leu Ala Ser Asp Thr Tyr Gly Pro Val Ser Val Leu Pro
995 1000 1005

Thr Pro Ala Tyr Phe Tyr Gly Leu Ala Asp Gly Glu Glu Leu Phe
1010 1015 1020

Ala Asp Ile Glu Lys Gly Lys Thr Leu Val Ile Val Asn Gln Ala
1025 1030 1035

Val Ser Ala Thr Asp Ser Gln Gly Met Val Thr Val Phe Phe Glu
1040 1045 1050

Leu Asn Gly Gln Pro Arg Arg Ile Lys Val Pro Asp Arg Ala His
1055 1060 1065

Gly Ala Thr Gly Ala Ala Val Arg Arg Lys Ala Glu Pro Gly Asn
1070 1075 1080

Ala Ala His Val Gly Ala Pro Met Pro Gly Val Ile Ser Arg Val
1085 1090 1095

Phe Val Ser Ser Gly Gln Ala Val Asn Ala Gly Asp Val Leu Val
1100 1105 1110

Ser Ile Glu Ala Met Lys Met Glu Thr Ala Ile His Ala Glu Lys
1115 1120 1125

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

1130 1135 1140
 Asp Ala Lys Asp Leu Leu Ala Val Tyr Gly Gly
 1145 1150
 <210> 5
 <211> 602

 <212> PRT
 <213> Pseudomonas fluorescens
 <400> 5
 Met Thr Lys Lys Ile Phe Val Thr Asp Thr Ile Leu Arg Asp Ala His
 1 5 10 15
 Gln Ser Leu Leu Ala Thr Arg Met Arg Thr Glu Asp Met Leu Pro Ile
 20 25 30
 Cys Asp Lys Leu Asp Lys Val Gly Tyr Trp Ser Leu Glu Cys Trp Gly
 35 40 45
 Gly Ala Thr Phe Asp Ala Cys Val Arg Phe Leu Lys Glu Asp Pro Trp

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

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

| | | | |
|---|---|-----|-----|
| 180 | 185 | 190 | |
| Leu Lys Ala Glu Gln Ser | Leu Pro Val Phe Ile His Ser His Asp Thr | | |
| 195 | 200 | 205 | |
| Ala Gly Leu Ala Ala Met Cys Gln Leu Lys Ala Ile Glu Asn Gly Ala | | | |
| 210 | 215 | 220 | |
| Asp His Ile Asp Thr Ala Ile Ser Ser Phe Ala Ser Gly Thr Ser His | | | |
| 225 | 230 | 235 | 240 |
| Pro Gly Thr Glu Ser Met Val Ala Ala Leu Lys Gly Thr Glu Phe Asp | | | |
| 245 | 250 | 255 | |
| Thr Gly Leu Asn Leu Glu Leu Leu Gln Glu Ile Gly Leu Tyr Phe Tyr | | | |
| 260 | 265 | 270 | |
| Ala Val Arg Lys Lys Tyr His Gln Phe Glu Ser Glu Phe Thr Ala Val | | | |
| 275 | 280 | 285 | |
| Asp Thr Arg Val Gln Val Asn Gln Val Pro Gly Gly Met Ile Ser Asn | | | |
| 290 | 295 | 300 | |
| Leu Ala Asn Gln Leu Lys Glu Gln Gly Ala Leu Asn Arg Met Gly Glu | | | |
| 305 | 310 | 315 | 320 |
| Val Leu Ala Glu Ile Pro Arg Val Arg Glu Asp Leu Gly Phe Pro Pro | | | |
| 325 | 330 | 335 | |
| Leu Val Thr Pro Thr Ser Gln Ile Val Gly Thr Gln Ala Phe Phe Asn | | | |
| 340 | 345 | 350 | |
| Val Leu Ala Gly Glu Arg Tyr Lys Thr Ile Thr Asn Glu Val Lys Leu | | | |
| 355 | 360 | 365 | |
| Tyr Leu Gln Gly Gly Tyr Gly Lys Ala Pro Gly Thr Val Asn Glu Lys | | | |
| 370 | 375 | 380 | |
| Leu Arg Arg Gln Ala Ile Gly Ser Glu Glu Val Ile Asp Val Arg Pro | | | |
| 385 | 390 | 395 | 400 |
| Ala Asp Leu Leu Lys Pro Glu Met Thr Lys Leu Arg Ala Asp Ile Gly | | | |
| 405 | 410 | 415 | |
| Ala Leu Ala Lys Ser Glu Glu Asp Val Leu Thr Phe Ala Met Phe Pro | | | |
| 420 | 425 | 430 | |

Asp Ile Gly Arg Lys Phe Leu Glu Glu Arg Ala Ala Gly Thr Leu Thr

435

440

445

Pro Glu Val Leu Leu Pro Ile Pro Glu Ala Gly Lys Val Ala Ser Ala

450

455

460

Gly Gly Glu Gly Val Pro Thr Glu Phe Val Ile Asp Val His Gly Glu

465

470

475

480

Thr Tyr Arg Val Asp Ile Thr Gly Val Gly Val Lys Ala Glu Gly Lys

485

490

495

Arg His Phe Tyr Leu Ser Ile Asp Gly Met Pro Glu Glu Val Val Phe

500

505

510

Glu Pro Leu Asn Glu Phe Val Gly Gly Gly Ser Ser Lys Arg Lys Gln

515

520

525

Ala Ser Ala Pro Gly His Val Ser Thr Thr Met Pro Gly Asn Ile Val

530

535

540

Asp Val Leu Val Lys Glu Gly Asp Thr Val Lys Ala Gly Gln Ala Val

545

550

555

560

Leu Ile Thr Glu Ala Met Lys Met Glu Thr Glu Val Gln Ala Ala Ile

565

570

575

Ala Gly Lys Val Thr Ala Ile His Val Ala Lys Gly Asp Arg Val Asn

580

585

590

Pro Gly Glu Ile Leu Ile Glu Ile Glu Gly

595

600

<210> 6

<211> 481

<212> PRT

<213> Pseudomonas fluorescens

<400> 6

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

1

5

10

15

Ile Ala Asn Arg Gly Glu Ile Ala Val Arg Ile Val Arg Ala Cys Ala

20

25

30

Glu Met Gly Ile Arg Ser Val Ala Ile Tyr Ser Asp Ala Asp Arg His
 35 40 45
 Ala Leu His Val Lys Arg Ala Asp Glu Ala His Ser Ile Gly Ala Glu
 50 55 60
 Pro Leu Ala Gly Tyr Leu Asn Pro Arg Lys Leu Val Asn Leu Ala Val
 65 70 75 80
 Glu Thr Gly Cys Asp Ala Leu His Pro Gly Tyr Gly Phe Leu Ser Glu
 85 90 95
 Asn Ala Glu Leu Ala Asp Ile Cys Ala Glu Arg Gly Ile Lys Phe Ile
 100 105 110
 Gly Pro Ser Ala Glu Val Ile Arg Arg Met Gly Asp Lys Thr Glu Ala
 115 120 125
 Arg Arg Ser Met Ile Lys Ala Gly Val Pro Val Thr Pro Gly Thr Glu
 130 135 140
 Gly Asn Val Ser Gly Ile Glu Glu Ala Leu Ser Glu Gly Asp Arg Ile
 145 150 155 160
 Gly Tyr Pro Val Met Leu Lys Ala Thr Ser Gly Gly Gly Gly Arg Gly
 165 170 175
 Ile Arg Arg Cys Asn Ser Arg Glu Glu Leu Glu Gln Asn Phe Pro Arg
 180 185 190
 Val Ile Ser Glu Ala Thr Lys Ala Phe Gly Ser Ala Glu Val Phe Leu
 195 200 205
 Glu Lys Cys Ile Val Asn Pro Lys His Ile Glu Ala Gln Ile Leu Gly
 210 215 220
 Asp Ser Phe Gly Asn Val Val His Leu Phe Glu Arg Asp Cys Ser Ile
 225 230 235 240
 Gln Arg Arg Asn Gln Lys Leu Ile Glu Ile Ala Pro Ser Pro Gln Leu
 245 250 255
 Thr Pro Glu Gln Arg Ala Tyr Ile Gly Asp Leu Ser Val Arg Ala Ala
 260 265 270
 Lys Ala Val Gly Tyr Glu Asn Ala Gly Thr Val Glu Phe Leu Leu Ala

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

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

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

 465 470 475 480
 Leu

<210> 7

<211> 1140

<212> PRT

<213> Corynebacterium glutamicum

<400> 7

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

| | | |
|---|-----|-----|
| 245 | 250 | 255 |
| Leu Asp Pro Glu Leu Arg Asp Arg Ile Cys Ala Asp Ala Val Lys Phe | | |
| 260 | 265 | 270 |
| Cys Arg Ser Ile Gly Tyr Gln Gly Ala Gly Thr Val Glu Phe Leu Val | | |
| 275 | 280 | 285 |
| Asp Glu Lys Gly Asn His Val Phe Ile Glu Met Asn Pro Arg Ile Gln | | |
| 290 | 295 | 300 |
| | | |
| Val Glu His Thr Val Thr Glu Glu Val Thr Glu Val Asp Leu Val Lys | | |
| 305 | 310 | 315 |
| Ala Gln Met Arg Leu Ala Ala Gly Ala Thr Leu Lys Glu Leu Gly Leu | | |
| 325 | 330 | 335 |
| Thr Gln Asp Lys Ile Lys Thr His Gly Ala Ala Leu Gln Cys Arg Ile | | |
| 340 | 345 | 350 |
| Thr Thr Glu Asp Pro Asn Asn Gly Phe Arg Pro Asp Thr Gly Thr Ile | | |
| 355 | 360 | 365 |
| | | |
| Thr Ala Tyr Arg Ser Pro Gly Gly Ala Gly Val Arg Leu Asp Gly Ala | | |
| 370 | 375 | 380 |
| Ala Gln Leu Gly Gly Glu Ile Thr Ala His Phe Asp Ser Met Leu Val | | |
| 385 | 390 | 395 |
| Lys Met Thr Cys Arg Gly Ser Asp Phe Glu Thr Ala Val Ala Arg Ala | | |
| 405 | 410 | 415 |
| Gln Arg Ala Leu Ala Glu Phe Thr Val Ser Gly Val Ala Thr Asn Ile | | |
| 420 | 425 | 430 |
| | | |
| Gly Phe Leu Arg Ala Leu Leu Arg Glu Glu Asp Phe Thr Ser Lys Arg | | |
| 435 | 440 | 445 |
| Ile Ala Thr Gly Phe Ile Ala Asp His Pro His Leu Leu Gln Ala Pro | | |
| 450 | 455 | 460 |
| Pro Ala Asp Asp Glu Gln Gly Arg Ile Leu Asp Tyr Leu Ala Asp Val | | |
| 465 | 470 | 475 |
| Thr Val Asn Lys Pro His Gly Val Arg Pro Lys Asp Val Ala Ala Pro | | |
| 485 | 490 | 495 |

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

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

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

Leu Tyr Thr Leu Asp Tyr Tyr Leu Lys Met Ala Glu Glu Ile Val Lys
690 695 700
Ser Gly Ala His Ile Leu Ala Ile Lys Asp Met Ala Gly Leu Leu Arg
705 710 715 720
Pro Ala Ala Val Thr Lys Leu Val Thr Ala Leu Arg Arg Glu Phe Asp
725 730 735
Leu Pro Val His Val His Thr His Asp Thr Ala Gly Gly Gln Leu Ala

740

745

750

Thr Tyr Phe Ala Ala Ala Gln Ala Gly Ala Asp Ala Val Asp Gly Ala

755

760

765

Ser Ala Pro Leu Ser Gly Thr Thr Ser Gln Pro Ser Leu Ser Ala Ile

770

775

780

Val Ala Ala Phe Ala His Thr Arg Arg Asp Thr Gly Leu Ser Leu Glu

785

790

795

800

Ala Val Ser Asp Leu Glu Pro Tyr Trp Glu Ala Val Arg Gly Leu Tyr

805

810

815

Leu Pro Phe Glu Ser Gly Thr Pro Gly Pro Thr Gly Arg Val Tyr Arg

820

825

830

His Glu Ile Pro Gly Gly Gln Leu Ser Asn Leu Arg Ala Gln Ala Thr

835

840

845

Ala Leu Gly Leu Ala Asp Arg Phe Glu Leu Ile Glu Asp Asn Tyr Ala

850

855

860

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

865

870

875

880

Lys Val Val Gly Asp Leu Ala Leu His Leu Val Gly Ala Gly Val Asp

885

890

895

Pro Ala Asp Phe Ala Ala Asp Pro Gln Lys Tyr Asp Ile Pro Asp Ser

900

905

910

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

915

920

925

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

930

935

940

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

945

950

955

960

Asp Asp Ser Lys Glu Arg Arg Asn Ser Leu Asn Arg Leu Leu Phe Pro

965

970

975

Lys Pro Thr Glu Glu Phe Leu Glu His Arg Arg Arg Phe Gly Asn Thr

980

985

990

Ser Ala Leu Asp Asp Arg Glu Phe Phe Tyr Gly Leu Val Glu Gly Arg
995 1000 1005

Glu Thr Leu Ile Arg Leu Pro Asp Val Arg Thr Pro Leu Leu Val
1010 1015 1020

Arg Leu Asp Ala Ile Ser Glu Pro Asp Asp Lys Gly Met Arg Asn
1025 1030 1035

Val Val Ala Asn Val Asn Gly Gln Ile Arg Pro Met Arg Val Arg
1040 1045 1050

Asp Arg Ser Val Glu Ser Val Thr Ala Thr Ala Glu Lys Ala Asp
1055 1060 1065

Ser Ser Asn Lys Gly His Val Ala Ala Pro Phe Ala Gly Val Val
1070 1075 1080

Thr Val Thr Val Ala Glu Gly Asp Glu Val Lys Ala Gly Asp Ala
1085 1090 1095

Val Ala Ile Ile Glu Ala Met Lys Met Glu Ala Thr Ile Thr Ala
1100 1105 1110

Ser Val Asp Gly Lys Ile Asp Arg Val Val Val Pro Ala Ala Thr
1115 1120 1125

Lys Val Glu Gly Gly Asp Leu Ile Val Val Val Ser
1130 1135 1140

<210> 8

<211> 1152

<212> PRT

<213> Sinorhizobium meliloti

<400> 8

Met Ser Ile Ser Lys Ile Leu Val Ala Asn Arg Ser Glu Ile Ala Ile
1 5 10 15

Arg Val Phe Arg Ala Ala Asn Glu Leu Gly Leu Lys Thr Val Ala Ile
20 25 30

Trp Ala Glu Glu Asp Lys Leu Ala Leu His Arg Phe Lys Ala Asp Glu
35 40 45

Ser Tyr Gln Val Gly Arg Gly Pro His Leu Pro Arg Asp Leu Gly Pro
50 55 60
Ile Met Ser Tyr Leu Ser Ile Asp Glu Val Ile Arg Val Ala Lys Leu
65 70 75 80
Ser Gly Ala Asp Ala Ile His Pro Gly Tyr Gly Leu Leu Ser Glu Ser
85 90 95
Pro Glu Phe Ala Glu Ala Cys Ala Ala Asn Gly Ile Thr Phe Ile Gly
100 105 110
Pro Lys Pro Glu Thr Met Arg Gln Leu Gly Asn Lys Val Ala Ala Arg
115 120 125
Asn Leu Ala Ile Ser Ile Gly Val Pro Val Val Pro Ala Thr Glu Pro
130 135 140
Leu Pro Asp Asp Pro Glu Glu Ile Lys Arg Leu Ala Glu Glu Ile Gly
145 150 155 160
Tyr Pro Val Met Leu Lys Ala Ser Trp Gly Gly Gly Gly Arg Gly Met
165 170 175
Arg Ala Ile Arg Asp Pro Lys Asp Leu Ile Arg Glu Val Thr Glu Ala
180 185 190
Lys Arg Glu Ala Lys Ala Ala Phe Gly Lys Asp Glu Val Tyr Leu Glu
195 200 205
Lys Leu Val Glu Arg Ala Arg His Val Glu Ser Gln Ile Leu Gly Asp
210 215 220
Thr His Gly Asn Val Val His Leu Phe Glu Arg Asp Cys Ser Ile Gln
225 230 235 240
Arg Arg Asn Gln Lys Val Val Glu Arg Ala Pro Ala Pro Tyr Leu Asn
245 250 255
Asp Ala Gln Arg Gln Glu Leu Ala Asp Tyr Ser Leu Lys Ile Ala Arg
260 265 270
Ala Thr Asn Tyr Ile Gly Ala Gly Thr Val Glu Tyr Leu Met Asp Ser
275 280 285
Asp Thr Gly Lys Phe Tyr Phe Ile Glu Val Asn Pro Arg Ile Gln Val

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

 355 360 365
 Arg Ile Thr Ala Tyr Arg Gly Ala Thr Gly Phe Gly Ile Arg Leu Asp
 370 375 380
 Gly Gly Thr Ala Tyr Ser Gly Ala Val Ile Thr Arg Tyr Tyr Asp Pro
 385 390 395 400
 Leu Leu Glu Lys Val Thr Ala Trp Ala Pro Asn Pro Gly Glu Ala Ile
 405 410 415
 Gln Arg Met Ile Arg Ala Leu Arg Glu Phe Arg Ile Arg Gly Val Ala

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

 485 490 495
 Pro Ser Asp Asp Ile Ala Ala Pro Val Val Pro Phe Thr Gly Gly Asp
 500 505 510
 Val Lys Pro Gly Thr Lys Gln Arg Leu Asp Gln Leu Gly Pro Lys Lys
 515 520 525
 Phe Ala Glu Trp Val Lys Ala Gln Pro Glu Val Leu Ile Thr Asp Thr
 530 535 540

Thr Met Arg Asp Gly His Gln Ser Leu Leu Ala Thr Arg Met Arg Thr

545 550 555 560

Tyr Asp Ile Ala Arg Ile Ala Gly Thr Tyr Ala Arg Ala Leu Pro Asn

565 570 575

Leu Phe Ser Leu Glu Cys Trp Gly Gly Ala Thr Phe Asp Val Ser Met

580 585 590

Arg Phe Leu Thr Glu Asp Pro Trp Glu Arg Leu Ala Met Val Arg Glu

595 600 605

Gly Ala Pro Asn Leu Leu Leu Gln Met Leu Leu Arg Gly Ala Asn Gly

610 615 620

Val Gly Tyr Lys Asn Tyr Pro Asp Asn Val Val Lys Tyr Phe Val Arg

625 630 635 640

Gln Ala Ala Lys Gly Gly Ile Asp Val Phe Arg Val Phe Asp Cys Leu

645 650 655

Asn Trp Val Glu Asn Met Arg Val Ala Met Asp Ala Val Ala Glu Glu

660 665 670

Asp Arg Ile Cys Glu Ala Ala Ile Cys Tyr Thr Gly Asp Ile Leu Asn

675 680 685

Ser Ala Arg Pro Lys Tyr Asp Leu Lys Tyr Tyr Thr Ala Leu Ala Ala

690 695 700

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

705 710 715 720

Gly Leu Leu Lys Pro Ala Ala Ala Arg Val Leu Phe Lys Ala Leu Lys

725 730 735

Glu Ala Thr Gly Leu Pro Ile His Phe His Thr His Asp Thr Ser Gly

740 745 750

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

755 760 765

Val Asp Ala Ala Met Asp Ala Leu Ser Gly Asn Thr Ser Gln Pro Cys

770 775 780

Leu Gly Ser Ile Val Glu Ala Leu Ser Gly Ser Glu Arg Asp Pro Gly

785 790 795 800
 Leu Asp Pro Glu Trp Ile Arg Arg Ile Ser Phe Tyr Trp Glu Ala Val

 805 810 815
 Arg His Gln Tyr Ala Ala Phe Glu Ser Asp Leu Lys Gly Pro Ala Ser
 820 825 830
 Glu Val Tyr Leu His Glu Met Pro Gly Gly Gln Phe Thr Asn Leu Lys
 835 840 845
 Glu Gln Ala Arg Ser Leu Gly Leu Glu Thr Arg Trp His Glu Val Ala
 850 855 860
 Gln Ala Tyr Ala Asp Ala Asn Arg Met Phe Gly Asp Ile Val Lys Val

 865 870 875 880
 Thr Pro Ser Ser Lys Val Val Gly Asp Met Ala Leu Met Met Val Ser
 885 890 895
 Gln Asp Leu Thr Val Ala Asp Val Glu Asn Pro Gly Lys Asp Ile Ala
 900 905 910
 Phe Pro Glu Ser Val Val Ser Met Leu Lys Gly Asp Leu Gly Gln Pro
 915 920 925
 Pro Gly Gly Trp Pro Glu Ala Leu Gln Lys Lys Ala Leu Lys Gly Glu

 930 935 940
 Glu Pro Tyr Asp Ala Arg Pro Gly Ser Leu Leu Glu Asp Ala Asp Leu
 945 950 955 960
 Asp Ala Glu Arg Lys Gly Ile Glu Glu Lys Leu Gly Arg Glu Val Thr
 965 970 975
 Asp Phe Glu Phe Ala Ser Tyr Leu Met Tyr Pro Lys Val Phe Thr Asp
 980 985 990
 Tyr Ala Val Ala Cys Glu Thr Tyr Gly Pro Val Ser Val Leu Pro Thr

 995 1000 1005
 Pro Ala Tyr Phe Tyr Gly Met Ala Pro Gly Glu Glu Leu Phe Ala
 1010 1015 1020
 Asp Ile Glu Lys Gly Lys Thr Leu Val Ile Leu Asn Gln Ala Gln
 1025 1030 1035

Gly Glu Ile Asp Glu Lys Gly Met Val Lys Met Phe Phe Glu Met
 1040 1045 1050
 Asn Gly Gln Pro Arg Ser Ile Lys Val Pro Asp Arg Asn Arg Gly
 1055 1060 1065
 Ala Ser Ala Ala Val Arg Arg Lys Ala Glu Ala Gly Asn Ala Ala
 1070 1075 1080
 His Leu Gly Ala Pro Met Pro Gly Val Ile Ser Thr Val Ala Val
 1085 1090 1095
 Ala Ser Gly Gln Ser Val Lys Ala Gly Asp Val Leu Leu Ser Ile
 1100 1105 1110
 Glu Ala Met Lys Met Glu Thr Ala Leu His Ala Glu Lys Asp Gly

1115 1120 1125
 Val Ile Ser Glu Val Leu Val Arg Ala Gly Asp Gln Ile Asp Ala
 1130 1135 1140
 Lys Asp Leu Leu Val Val Phe Gly Gly
 1145 1150

<210> 9

<211> 2652

<212> DNA

<213> Escherichia coli

<400> 9

| | |
|---|-----|
| atgaacgaac aatattccgc attgcgtagt aatgtcagta tgctcggcaa agtgctggga | 60 |
| gaaaccatca aggatgcgtt gggagaacac attcttgaac gcgtagaaac tatccgtaag | 120 |
| ttgtcgaaat cttcacgcgc tggcaatgat gctaaccgcc aggagttgct caccacctta | 180 |
| | |
| caaaatttgt cgaacgacga gctgctgccc gttgcgcgtg cgtttagtca gttcctgaac | 240 |
| ctggccaaca ccgccgagca ataccacagc atttcgccga aaggcgaagc tgccagcaac | 300 |
| ccggaagtga tcgcccgcac cctgcgtaaa ctgaaaaacc agccggaact gagcgaagac | 360 |
| accatcaaaa aagcagtgga atcgtgtcg ctggaactgg tcctcacggc tcaccaacc | 420 |
| gaaattaccc gtcgtacact gatccacaaa atggtggaag tgaacgcctg tttaaaacag | 480 |
| ctcgataaca aagatatcgc tgactacgaa cacaaccagc tgatgcgtcg cctgcgccag | 540 |
| ttgatgcccc agtcatggca taccgatgaa atccgtaagc tgcgtccaag cccggtagat | 600 |

| | |
|--|------|
| gaagccaaat ggggctttgc cgtagtggaa aacagcctgt ggcaaggcgt accaaattac | 660 |
| ctgcgcgaac tgaacgaaca actggaagag aacctcggct acaaactgcc cgtcgaattt | 720 |
| gttccggtcc gttttacttc gtggatgggc ggcgaccgac acggcaaccc gaacgtcact | 780 |
| gccgatatca cccgccacgt cctgctactc agccgctgga aagccaccga tttgttctg | 840 |
| aaagatatcc aggtgcttgg ttctgaactg tcgatggttg aagcgacccc tgaactgctg | 900 |
| gcgctggttg gcgaagaagg tgccgcagaa ccgtatcgct atctgatgaa aaacctgcgt | 960 |
| tctcgctga tggcgacaca ggcatggctg gaagcgcgcc tgaaaggcga agaactgcc | 1020 |
| aaaccagaag gcctgtctgac aaaaaacgaa gaactgtggg aaccgctcta cgcttctac | 1080 |
| cagtcacttc aggcgtgtgg catgggtatt atcgccaacg gcgatctgct cgacaccctg | 1140 |
| cgccgcgtga aatgtttcgg cgtaccgctg gtccgtattg atatccgtca ggagagcacg | 1200 |
| cgtcataccg aagcgctggg cgagctgacc cgctacctcg gtatcgcgca ctacgaaagc | 1260 |
| tggtcagagg ccgacaaaca ggcgttctg atccgcgaac tgaactccaa acgtccgctt | 1320 |
| ctgccgcgca actggcaacc aagcgccgaa acgcgcgaag tgctcgatac ctgccagggtg | 1380 |
| attgccgaag caccgcaagg ctccattgcc gcctacgtga tctcgatggc gaaaacgccg | 1440 |
| tccgacgtac tggctgtcca cctgctgctg aaagaagcgg gtatcgggtt tgcgatgccg | 1500 |
| gttgctccgc tgtttgaaac cctcgatgat ctgaacaacg ccaacgatgt catgaccag | 1560 |
| ctgtcaata ttgactgcta tctgtgctg attcaggcca aacagatggt gatgattggc | 1620 |
| tattccgact cagcaaaaga tgcgggagtg atggcagctt cctgggcgca atatcaggca | 1680 |
| caggatgcat taatcaaac ctgcgaaaaa gcgggtattg agctgacgtt gttccacggt | 1740 |
| cgcggcggtt ccattggtcg cggcgccgca cctgctcatg cggcgctgct gtcacaaccg | 1800 |
| ccaggaagcc tgaaaggcgg cctgcgcgta accgaacagg gcgagatgat ccgctttaa | 1860 |
| tatggtctgc cagaaatcac cgtcagcagc ctgtcgcttt ataccggggc gattctggaa | 1920 |
| gccaacctgc tgccaccgcc ggagccgaaa gagagctggc gtgcgattat ggatgaactg | 1980 |
| tcagtcactt cctgcgatgt ctaccgcggc tactacgtg aaaacaaaga tttgtgcct | 2040 |
| tacttccgct ccgtacgcc ggaacaagaa ctgggcaaac tgccgttggg ttcacgtccg | 2100 |
| gcgaaacgtc gccaaccggg cggcgctcag tcaactacgc ccattccgtg gatcttcgcc | 2160 |
| tggacgcaaa accgtctgat gtcctccgcc tggctgggtg caggtacggc gctgcaaaaa | 2220 |
| gtggtcgaag acggcaaaac gagcgagctg gaggtatgt gccgcgattg gccattcttc | 2280 |
| tcgacgcgtc tcggcatgct ggagatggtc ttcgcaaaag cagacctgtg gctggcggaa | 2340 |
| tactatgacc aacgcctggt agacaaagca ctgtggccgt taggtaaaga gttacgcaac | 2400 |
| ctgaagaag aagacatcaa agtgggtctg gcgattgcca acgattccca tctgatggcc | 2460 |

| | |
|---|------|
| gatctgccgt ggattgcaga gtctattcag ctacggaata ttacaccga cccgtgaac | 2520 |
| gtattgcagg ccgagttgct gcaccgtccc cgccaggcag aaaaagaagg ccaggaaccg | 2580 |
| gatcctcgcg tcgaacaagc gttaatggtc actattgccg ggattgcggc aggtatgcgt | 2640 |
| aataccggct aa | 2652 |

<210> 10

<211> 883

<212> PRT

<213> Escherichia coli

<400> 10

Met Asn Glu Gln Tyr Ser Ala Leu Arg Ser Asn Val Ser Met Leu Gly

1 5 10 15

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

20 25 30

Glu Arg Val Glu Thr Ile Arg Lys Leu Ser Lys Ser Ser Arg Ala Gly

35 40 45

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

50 55 60

Asn Asp Glu Leu Leu Pro Val Ala Arg Ala Phe Ser Gln Phe Leu Asn

65 70 75 80

Leu Ala Asn Thr Ala Glu Gln Tyr His Ser Ile Ser Pro Lys Gly Glu

85 90 95

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

100 105 110

Asn Gln Pro Glu Leu Ser Glu Asp Thr Ile Lys Lys Ala Val Glu Ser

115 120 125

Leu Ser Leu Glu Leu Val Leu Thr Ala His Pro Thr Glu Ile Thr Arg

130 135 140

Arg Thr Leu Ile His Lys Met Val Glu Val Asn Ala Cys Leu Lys Gln

145 150 155 160

Leu Asp Asn Lys Asp Ile Ala Asp Tyr Glu His Asn Gln Leu Met Arg

165 170 175

Arg Leu Arg Gln Leu Ile Ala Gln Ser Trp His Thr Asp Glu Ile Arg

180 185 190
Lys Leu Arg Pro Ser Pro Val Asp Glu Ala Lys Trp Gly Phe Ala Val

195 200 205
Val Glu Asn Ser Leu Trp Gln Gly Val Pro Asn Tyr Leu Arg Glu Leu

210 215 220
Asn Glu Gln Leu Glu Glu Asn Leu Gly Tyr Lys Leu Pro Val Glu Phe

225 230 235 240
Val Pro Val Arg Phe Thr Ser Trp Met Gly Gly Asp Arg Asp Gly Asn

245 250 255
Pro Asn Val Thr Ala Asp Ile Thr Arg His Val Leu Leu Leu Ser Arg

260 265 270
Trp Lys Ala Thr Asp Leu Phe Leu Lys Asp Ile Gln Val Leu Val Ser

275 280 285
Glu Leu Ser Met Val Glu Ala Thr Pro Glu Leu Leu Ala Leu Val Gly

290 295 300
Glu Glu Gly Ala Ala Glu Pro Tyr Arg Tyr Leu Met Lys Asn Leu Arg

305 310 315 320
Ser Arg Leu Met Ala Thr Gln Ala Trp Leu Glu Ala Arg Leu Lys Gly

325 330 335
Glu Glu Leu Pro Lys Pro Glu Gly Leu Leu Thr Gln Asn Glu Glu Leu

340 345 350
Trp Glu Pro Leu Tyr Ala Cys Tyr Gln Ser Leu Gln Ala Cys Gly Met

355 360 365
Gly Ile Ile Ala Asn Gly Asp Leu Leu Asp Thr Leu Arg Arg Val Lys

370 375 380
Cys Phe Gly Val Pro Leu Val Arg Ile Asp Ile Arg Gln Glu Ser Thr

385 390 395 400
Arg His Thr Glu Ala Leu Gly Glu Leu Thr Arg Tyr Leu Gly Ile Gly

405 410 415
Asp Tyr Glu Ser Trp Ser Glu Ala Asp Lys Gln Ala Phe Leu Ile Arg

420 425 430
 Glu Leu Asn Ser Lys Arg Pro Leu Leu Pro Arg Asn Trp Gln Pro Ser

 435 440 445
 Ala Glu Thr Arg Glu Val Leu Asp Thr Cys Gln Val Ile Ala Glu Ala
 450 455 460
 Pro Gln Gly Ser Ile Ala Ala Tyr Val Ile Ser Met Ala Lys Thr Pro
 465 470 475 480
 Ser Asp Val Leu Ala Val His Leu Leu Leu Lys Glu Ala Gly Ile Gly
 485 490 495
 Phe Ala Met Pro Val Ala Pro Leu Phe Glu Thr Leu Asp Asp Leu Asn

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

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

 625 630 635 640
 Ala Asn Leu Leu Pro Pro Pro Glu Pro Lys Glu Ser Trp Arg Arg Ile
 645 650 655
 Met Asp Glu Leu Ser Val Ile Ser Cys Asp Val Tyr Arg Gly Tyr Val
 660 665 670

Arg Glu Asn Lys Asp Phe Val Pro Tyr Phe Arg Ser Ala Thr Pro Glu
675 680 685

Gln Glu Leu Gly Lys Leu Pro Leu Gly Ser Arg Pro Ala Lys Arg Arg

690 695 700

Pro Thr Gly Gly Val Glu Ser Leu Arg Ala Ile Pro Trp Ile Phe Ala
705 710 715 720

Trp Thr Gln Asn Arg Leu Met Leu Pro Ala Trp Leu Gly Ala Gly Thr
725 730 735

Ala Leu Gln Lys Val Val Glu Asp Gly Lys Gln Ser Glu Leu Glu Ala
740 745 750

Met Cys Arg Asp Trp Pro Phe Phe Ser Thr Arg Leu Gly Met Leu Glu

755 760 765

Met Val Phe Ala Lys Ala Asp Leu Trp Leu Ala Glu Tyr Tyr Asp Gln
770 775 780

Arg Leu Val Asp Lys Ala Leu Trp Pro Leu Gly Lys Glu Leu Arg Asn
785 790 795 800

Leu Gln Glu Glu Asp Ile Lys Val Val Leu Ala Ile Ala Asn Asp Ser
805 810 815

His Leu Met Ala Asp Leu Pro Trp Ile Ala Glu Ser Ile Gln Leu Arg

820 825 830

Asn Ile Tyr Thr Asp Pro Leu Asn Val Leu Gln Ala Glu Leu Leu His
835 840 845

Arg Ser Arg Gln Ala Glu Lys Glu Gly Gln Glu Pro Asp Pro Arg Val
850 855 860

Glu Gln Ala Leu Met Val Thr Ile Ala Gly Ile Ala Ala Gly Met Arg
865 870 875 880

Asn Thr Gly

<210> 11

<211> 483

<212> PRT

<213> Methanobacterium thermoautotrophicum

<400> 11

Met Lys Val Pro Arg Cys Met Ser Thr Gln His Pro Asp Asn Val Asn
1 5 10 15

Pro Pro Phe Phe Ala Glu Glu Pro Glu Leu Gly Gly Glu Asp Glu Ile
20 25 30

Arg Glu Ala Tyr Tyr Val Phe Ser His Leu Gly Cys Asp Glu Gln Met
35 40 45

Trp Asp Cys Glu Gly Lys Glu Val Asp Asn Tyr Val Val Lys Lys Leu

50 55 60

Leu Thr Lys Tyr Gln Ala Phe Phe Arg Asp His Val Leu Gly Glu Asp
65 70 75 80

Leu Arg Leu Thr Leu Arg Val Pro Asn Pro Thr Val Glu Arg Ala Glu
85 90 95

Ala Lys Ile Leu Leu Glu Thr Leu Glu Ser Ile Pro Arg Ser Tyr Asp
100 105 110

Thr Ala Ser Leu Phe Tyr Gly Met Asp Ala Ala Pro Val Phe Glu Val

115 120 125

Ile Leu Pro Met Thr Ser Ser Ser Ser Cys Leu Asn Arg Ile His Ser
130 135 140

Tyr Tyr Leu Asp Phe Val Lys Gly Lys Glu Arg Leu Gln Leu Ala Asp
145 150 155 160

Gly Val Thr Val Lys Glu Trp Ile Gly Glu Phe Arg Pro Asp Glu Ile
165 170 175

Asn Val Ile Pro Leu Phe Glu Asp His Glu Gly Met Leu Asn Ala Ala

180 185 190

Lys Ile Thr Gly Glu Tyr Leu Asp Gly Lys Asp Ile Gln Glu Gln Arg
195 200 205

Val Phe Leu Ala Arg Ser Asp Pro Ala Met Asn Tyr Gly Met Ile Ser
210 215 220

Ala Thr Leu Leu Asn Arg Ile Ala Leu Ser Asp Phe Arg Asp Leu Glu
225 230 235 240

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

245 250 255

Pro Phe Arg Gly Asn Leu Arg Pro Asp Asn Val Glu Asp Val Thr Trp

260 265 270

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

275 280 285

Asp His Glu Pro Ser Asp Val Ile Arg Gly Ile Lys Lys Leu Arg Ser

290 295 300

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

305 310 315 320

Ile Ile Ser Ala Tyr Cys Arg Glu Tyr Arg Arg Gln Val Met Asp Leu

325 330 335

Val Asp Ile Ile Asn Arg Val Ala Arg Tyr Val Pro Gly Arg Arg Lys

340 345 350

Arg Lys Leu His Ile Gly Leu Phe Gly Tyr Ser Arg Ser Met Gly Asn

355 360 365

Val Ser Leu Pro Arg Ala Ile Thr Phe Thr Ala Ala Leu Tyr Ser Leu

370 375 380

Gly Val Pro Pro Glu Leu Leu Gly Phe Asn Ala Leu Ser Ser Gly Asp

385 390 395 400

Leu Glu Phe Ile Glu Glu Val Tyr Pro Gly Leu Gly Arg Asp Leu His

405 410 415

Asp Ala Ala Arg Tyr Ala Asn Pro Glu Ser Pro Phe Leu Ser Pro Glu

420 425 430

Val Lys Ser Ser Phe Glu Glu Tyr Leu Glu Pro Glu Tyr Asp Glu Gly

435 440 445

His Met Lys Thr Thr Glu Glu Ile Ile Arg Ala Leu Arg Ile Asn Arg

450 455 460

Thr Ala Asn Leu Gln Glu Leu Ile Leu Glu Ala Ala Ser Gln Arg Lys

465 470 475 480

Phe Leu Gly

<210> 12

<211> 537

<212> PRT

<213> Clostridium perfringens

<400> 12

Met Lys Ile Pro Cys Ser Met Met Thr Gln His Pro Asp Asn Val Glu

1 5 10 15

Thr Tyr Ile Ser Ile Gln Gln Glu Pro Ala Glu Ala Ile Lys Gly Leu

20 25 30

Thr Pro Gln Asp Lys Gly Gly Leu Gly Ile Glu Glu Val Met Ile Asp

35 40 45

Phe Glu Gly Lys Leu Thr Pro Tyr His Gln Thr Ser Gln Ile Ala Leu

50 55 60

Gly Leu Ile Ser Asn Gly Ile Ile Pro Gly Lys Asp Val Arg Val Thr

65 70 75 80

Pro Arg Ile Pro Asn Ala Asn Lys Glu Ser Val Phe Arg Gln Leu Met

85 90 95

Ser Ile Met Ser Ile Ile Glu Thr Asn Val Gln Ser Lys Glu Leu Thr

100 105 110

Gly Thr Pro Ala Ile Ser Glu Val Val Val Pro Met Ile Glu Thr Gly

115 120 125

Lys Glu Ile Ser Glu Phe Gln Asp Arg Val Asn Ser Val Val Asp Met

130 135 140

Gly Asn Lys Asn Tyr Lys Thr Lys Leu Asp Leu Asn Ser Val Arg Ile

145 150 155 160

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

165 170 175

Leu Asp Glu His Tyr Glu Ile Glu Lys Ser Lys Gly His Ile Leu Lys

180 185 190

Asp Leu Arg Ile Met Ile Ala Arg Ser Asp Thr Ala Met Ser Tyr Gly

195

200

205

Leu Ile Ser Gly Val Leu Ser Val Leu Met Ala Val Asp Gly Ala Tyr

210

215

220

Lys Trp Gly Glu Lys His Gly Val Thr Ile Ser Pro Ile Leu Gly Cys

225

230

235

240

Gly Ser Leu Pro Phe Arg Gly His Phe Ser Glu Glu Asn Ile Asp Glu

245

250

255

Ile Leu Ala Thr Tyr Ser Gly Ile Lys Thr Phe Thr Phe Gln Ser Ala

260

265

270

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

275

280

285

Leu Lys Glu Lys Ile Ala Gln Ser Lys Pro Arg Asn Phe Ser Glu Glu

290

295

300

Asp Lys Asp Leu Met Lys Glu Phe Ile Gly Ile Cys Ser Lys His Tyr

305

310

315

320

Leu Gln Thr Phe Leu Lys Val Ile Asp Thr Val Ser Phe Val Ser Asp

325

330

335

Phe Ile Pro Lys Asn Arg Asp Arg Leu Thr Lys Ala Lys Thr Gly Leu

340

345

350

Glu Tyr Asn Arg Glu Val Ala Asn Leu Asp Asn Val Ala Asp Leu Val

355

360

365

Lys Asp Glu Val Leu Lys Gln Glu Ile Leu Ser Ile Asp Asn Ser Lys

370

375

380

Glu Tyr Ala Val Pro Arg Ala Ile Ser Phe Thr Gly Ala Met Tyr Thr

385

390

395

400

Leu Gly Met Pro Pro Glu Leu Met Gly Met Gly Arg Ala Leu Asn Glu

405

410

415

Ile Lys Thr Lys Tyr Gly Gln Glu Gly Ile Asp Lys Leu Leu Glu Ile

420

425

430

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

435

440

445

Gly Gly Val Ser Lys Lys Ile Ile Asp Glu Glu Ala Arg Gln Glu Tyr
450 455 460

Lys Glu Asp Met Lys Tyr Val Asn Glu Ile Leu Asn Leu Gly Leu Asp
465 470 475 480

Tyr Asp Phe Leu Asn Glu Asn Glu Phe Tyr His Thr Leu Leu Lys Thr
485 490 495

Thr Lys Pro Ile Ile Met His Leu Met Gly Leu Glu Glu Asn Val Met
500 505 510

Arg Asn Ser Thr Glu Glu Leu Lys Ile Leu Asn Glu Trp Ile Val Arg
515 520 525

Met Gly Lys Val Arg Gly Ser Ile Gly
530 535

<210> 13

<211> 1260

<212> DNA

<213> Issatchenkia orientalis

<400> 13

atgtccagag gcttctttac tgagaacatt acgcaattgc caccagaccc tttgttttgt 60
cttaaggcca ggttcagcaa tgactcacgt gaaaacaagg tcgatttagg tattggggca 120
tatagggacg acaacggtaa gccatggatc ttaccatctg tcaggttggc cgaaaacttg 180
attcagaact cccagacta caaccatgag tacctaccaa tcggtggact tgctgatttc 240
acttctgctg cggcaagagt tgtatttga ggcgattcta aagccatttc gcaaaaccgt 300

cttgtctcca tccagagttt gtcaggtaca ggtgccttac atgttctgg tctatttacc 360
aagcgccaat acaagtctct tgatggcact tccgaagacc ctctaata tctatcgga 420
cctacatggg ccaaccacgt tcaaatcttt gaagttattg gtctcaagcc tgtattctat 480
ccatattggc atgccgcaag caagaccttg gatctgaagg gctacttaaa ggcaataaac 540
gatgtccag aagggtcgtt tttgtattg catgcaacgg ctcataaccc tactggtttg 600
gatccaacac aagaacaatg gatggagatt ttggccgcta taagtccaa aaagcatctg 660
ccattatttg attgtgcata tcagggtttc acctccgggt ctctagatag agatgcttgg 720

gctgttcgag aagctgtcaa caatgacaag tacgaattcc cggaattat tgtctgtcaa 780
tcgtttgcga aaaatgttgg catgtatggt gaacggattg gtgcagttca tattgttcta 840

cctgaatcag acgcttcctt aaacagcgcc atcttctccc aattgcaaaa gacaatcaga 900
 tcggagattt ccaatccacc aggatacggg gcaaagattg tgtctaaagt tttgaacact 960
 ccggaacttt acaaacagtg ggagcaagat ttgatcacca tgtcttcgag aatcactgca 1020
 atgagaaaagg agctagtaaa tgagctcgag cgtcttggaa cccctggcac ttggagacac 1080
 atcaccgagc aacagggtat gtttctctt actggtttga acccggagca ggttgccaag 1140

ctagagaagg agcatggtgt ttatcttggt cgtatgggac gtgcaagtat tgcaggcctc 1200
 aacatgggaa acgtcaagta tggtagcaag gccattgact ctgtcgtgag agacctttag 1260

<210> 14

<211> 419

<212> PRT

<213> Issatchenkia orientalis

<400> 14

Met Ser Arg Gly Phe Phe Thr Glu Asn Ile Thr Gln Leu Pro Pro Asp

1 5 10 15

Pro Leu Phe Gly Leu Lys Ala Arg Phe Ser Asn Asp Ser Arg Glu Asn

20 25 30

Lys Val Asp Leu Gly Ile Gly Ala Tyr Arg Asp Asp Asn Gly Lys Pro

35 40 45

Trp Ile Leu Pro Ser Val Arg Leu Ala Glu Asn Leu Ile Gln Asn Ser

50 55 60

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

65 70 75 80

Thr Ser Ala Ala Ala Arg Val Val Phe Gly Gly Asp Ser Lys Ala Ile

85 90 95

Ser Gln Asn Arg Leu Val Ser Ile Gln Ser Leu Ser Gly Thr Gly Ala

100 105 110

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

115 120 125

Gly Thr Ser Glu Asp Pro Leu Ile Tyr Leu Ser Glu Pro Thr Trp Ala

130 135 140

Asn His Val Gln Ile Phe Glu Val Ile Gly Leu Lys Pro Val Phe Tyr

145 150 155 160

Pro Tyr Trp His Ala Ala Ser Lys Thr Leu Asp Leu Lys Gly Tyr Leu

165 170 175

Lys Ala Ile Asn Asp Ala Pro Glu Gly Ser Val Phe Val Leu His Ala

180 185 190

Thr Ala His Asn Pro Thr Gly Leu Asp Pro Thr Gln Glu Gln Trp Met

195 200 205

Glu Ile Leu Ala Ala Ile Ser Ala Lys Lys His Leu Pro Leu Phe Asp

210 215 220

Cys Ala Tyr Gln Gly Phe Thr Ser Gly Ser Leu Asp Arg Asp Ala Trp

225 230 235 240

Ala Val Arg Glu Ala Val Asn Asn Asp Lys Tyr Glu Phe Pro Gly Ile

245 250 255

Ile Val Cys Gln Ser Phe Ala Lys Asn Val Gly Met Tyr Gly Glu Arg

260 265 270

Ile Gly Ala Val His Ile Val Leu Pro Glu Ser Asp Ala Ser Leu Asn

275 280 285

Ser Ala Ile Phe Ser Gln Leu Gln Lys Thr Ile Arg Ser Glu Ile Ser

290 295 300

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

305 310 315 320

Pro Glu Leu Tyr Lys Gln Trp Glu Gln Asp Leu Ile Thr Met Ser Ser

325 330 335

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

340 345 350

Gly Thr Pro Gly Thr Trp Arg His Ile Thr Glu Gln Gln Gly Met Phe

355 360 365

Ser Phe Thr Gly Leu Asn Pro Glu Gln Val Ala Lys Leu Glu Lys Glu

370 375 380

His Gly Val Tyr Leu Val Arg Ser Gly Arg Ala Ser Ile Ala Gly Leu

385 390 395 400

Asn Met Gly Asn Val Lys Tyr Val Ala Lys Ala Ile Asp Ser Val Val

405 410 415

Arg Asp Leu

<210> 15

<211> 418

<212> PRT

<213> *Saccharomyces cerevisiae*

<400> 15

Met Ser Ala Thr Leu Phe Asn Asn Ile Glu Leu Leu Pro Pro Asp Ala

1 5 10 15

Leu Phe Gly Ile Lys Gln Arg Tyr Gly Gln Asp Gln Arg Ala Thr Lys

20 25 30

Val Asp Leu Gly Ile Gly Ala Tyr Arg Asp Asp Asn Gly Lys Pro Trp

35 40 45

Val Leu Pro Ser Val Lys Ala Ala Glu Lys Leu Ile His Asn Asp Ser

50 55 60

Ser Tyr Asn His Glu Tyr Leu Gly Ile Thr Gly Leu Pro Ser Leu Thr

65 70 75 80

Ser Asn Ala Ala Lys Ile Ile Phe Gly Thr Gln Ser Asp Ala Phe Gln

85 90 95

Glu Asp Arg Val Ile Ser Val Gln Ser Leu Ser Gly Thr Gly Ala Leu

100 105 110

His Ile Ser Ala Lys Phe Phe Ser Lys Phe Phe Pro Asp Lys Leu Val

115 120 125

Tyr Leu Ser Lys Pro Thr Trp Ala Asn His Met Ala Ile Phe Glu Asn

130 135 140

Gln Gly Leu Lys Thr Ala Thr Tyr Pro Tyr Trp Ala Asn Glu Thr Lys

145 150 155 160

Ser Leu Asp Leu Asn Gly Phe Leu Asn Ala Ile Gln Lys Ala Pro Glu

165 170 175

Gly Ser Ile Phe Val Leu His Ser Cys Ala His Asn Pro Thr Gly Leu

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

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

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

 370 375 380
 Ser Gly Arg Ala Ser Ile Ala Gly Leu Asn Gln Gly Asn Val Glu Tyr
 385 390 395 400
 Val Ala Lys Ala Ile Asp Glu Val Val Arg Phe Tyr Thr Ile Glu Ala
 405 410 415
 Lys Leu

<210> 16

<211> 396

<212> PRT

<213> Escherichia coli

<400> 16

Met Phe Glu Asn Ile Thr Ala Ala Pro Ala Asp Pro Ile Leu Gly Leu

1 5 10 15

Ala Asp Leu Phe Arg Ala Asp Glu Arg Pro Gly Lys Ile Asn Leu Gly

20 25 30

Ile Gly Val Tyr Lys Asp Glu Thr Gly Lys Thr Pro Val Leu Thr Ser

35 40 45

Val Lys Lys Ala Glu Gln Tyr Leu Leu Glu Asn Glu Thr Thr Lys Asn

50 55 60

Tyr Leu Gly Ile Asp Gly Ile Pro Glu Phe Gly Arg Cys Thr Gln Glu

65 70 75 80

Leu Leu Phe Gly Lys Gly Ser Ala Leu Ile Asn Asp Lys Arg Ala Arg

85 90 95

Thr Ala Gln Thr Pro Gly Gly Thr Gly Ala Leu Arg Val Ala Ala Asp

100 105 110

Phe Leu Ala Lys Asn Thr Ser Val Lys Arg Val Trp Val Ser Asn Pro

115 120 125

Ser Trp Pro Asn His Lys Ser Val Phe Asn Ser Ala Gly Leu Glu Val

130 135 140

Arg Glu Tyr Ala Tyr Tyr Asp Ala Glu Asn His Thr Leu Asp Phe Asp

145 150 155 160

Ala Leu Ile Asn Ser Leu Asn Glu Ala Gln Ala Gly Asp Val Val Leu

165 170 175

Phe His Gly Cys Cys His Asn Pro Thr Gly Ile Asp Pro Thr Leu Glu

180 185 190

Gln Trp Gln Thr Leu Ala Gln Leu Ser Val Glu Lys Gly Trp Leu Pro

195 200 205

Leu Phe Asp Phe Ala Tyr Gln Gly Phe Ala Arg Gly Leu Glu Glu Asp

210 215 220
Ala Glu Gly Leu Arg Ala Phe Ala Ala Met His Lys Glu Leu Ile Val
225 230 235 240
Ala Ser Ser Tyr Ser Lys Asn Phe Gly Leu Tyr Asn Glu Arg Val Gly
245 250 255
Ala Cys Thr Leu Val Ala Ala Asp Ser Glu Thr Val Asp Arg Ala Phe
260 265 270

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

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

<210> 17

<

211> 139

<212> PRT

<213> Streptomyces avermitilis

<400> 17

Met Leu Arg Thr Met Phe Lys Ser Lys Ile His Arg Ala Thr Val Thr
1 5 10 15
Gln Ala Asp Leu His Tyr Val Gly Ser Val Thr Ile Asp Ala Asp Leu
20 25 30

Leu Asp Ala Ala Asp Leu Leu Pro Gly Glu Leu Val His Ile Val Asp
35 40 45
Ile Thr Asn Gly Ala Arg Leu Glu Thr Tyr Val Ile Glu Gly Glu Arg
50 55 60
Gly Ser Gly Val Val Gly Ile Asn Gly Ala Ala Ala His Leu Val His
65 70 75 80
Pro Gly Asp Leu Val Ile Ile Ile Ser Tyr Ala Gln Val Ser Asp Ala
85 90 95
Glu Ala Arg Ala Leu Arg Pro Arg Val Val His Val Asp Arg Asp Asn
100 105 110
Arg Val Val Ala Leu Gly Ala Asp Pro Ala Glu Pro Val Pro Gly Ser

115 120 125
Asp Gln Ala Arg Ser Pro Gln Ala Val Thr Ala
130 135

<210> 18

<211> 127

<212> PRT

<213> Clostridium acetobutylicum

<400> 18

Met His Leu Asn Met Leu Lys Ser Lys Ile His Arg Ala Thr Val Val
1 5 10 15
Gln Ala Asp Leu Asn Tyr Val Gly Ser Ile Thr Ile Asp Arg Asn Leu
20 25 30
Met Asp Lys Ala Asn Ile Leu Glu Tyr Glu Lys Val Glu Ile Ala Asn

35 40 45
Ile Asn Asn Gly Ala Arg Phe Glu Thr Tyr Val Ile Ala Gly Glu Ala
50 55 60
Gly Ser Gly Ile Ile Cys Leu Asn Gly Ala Ala Ala Arg Cys Ala Gln
65 70 75 80
Ala Gly Asp Lys Val Ile Ile Met Cys Tyr Cys Ser Leu Thr Pro Glu
85 90 95

Glu Ala Ser Glu His Arg Pro Lys Val Val Phe Val Asn Asp Asp Asn

100 105 110
Ser Ile Ser Asn Val Thr Glu Tyr Glu Lys His Gly Thr Ile Gly

115 120 125

<210> 19

<211> 1404

<212> DNA

<213> Issatchenkia orientalis

<400> 19

atgtctatta gtgaaaaata ttttctcaa gaacctcaat ctctgaatt gaaaactgca 60

attccaggtc ctcaatcaaa ggcaaagctc gaggaattat ctgctgtcta tgatacaaag 120

gctgcatatt ttgttaccga ctactacaaa tctcttggtta actatattgt ggatgcagat 180

ggcaacaagc tactagattc ttattgccaa atctcttcta tcgcattggg ttacaataat 240

ccagcattat taaaagtagc acattctgat gaaatgacag ttgctttatg taacagacct 300

gctttggcat gttttccatc cactgattac tatgaaatac taaagaaggg attgttgtcc 360

gttgctcaa agggattaga taaggtttgt actgcacaca cgggatctga tgccaatgaa 420

atggcattta aggctgcatt tttgtttcaa gcaagtaaga agagagggtga caaacattt 480

accagcgaag agctggaatc tgtcatggag aacaagttgc caggcacctc tgacatggtt 540

atcctgtcat ttgaaaaagg gttccatggt agattgtttg gatctttatc taccactaga 600

tctaaagcta ttcacaaact ggatattcct gcgtttgaat ggccaaaggc tccattccct 660

cagttaaagt atcctctgga tcaattccaa gctgaaaaca aagcagaaga agaaagatgt 720

ttgaaggctt tagaggaaat tattgtcaac tctcttgcca aaattgcagc tgcaatcatt 780

gaaccggctc aatctgaagg tggatgataat catgcttcac cagaattctt ccaaggattt 840

agagaaatca ccaaaaagca cgggtgtcatt cttattgttg atgaagtca aacaggaggt 900

ggtgcttctg gtaagatgtg gttacatgaa cactatggca ttgtcccaga catcatgact 960

ttttctaaaa aaatgcaaaa tgcaggtttc tttttcagt aagcaggtct tgctggggac 1020

caaccattca gacaattcaa tacctgggtc ggtgatccat caaaagctct aattgcaaga 1080

accataattg aagaaattaa agataagaac ctattgacta gtgttaccga aacagggtgac 1140

tacctatatt caaagctcga agcaatttca gcaaagtatg acaaaatgat caacttgaga 1200

ggtaaggga gaggtttctt tattgcattt gatgccccaa caccggagtt aagaacaaaa 1260

tttattgctg aatgtaagaa attaggttta aacattggtg gatgcggtga acaagggttt 1320

agattgagac ctgcattagt ttttgaaaag aagcatgctg atatcttagc ctccattatt 1380

gatcaagctt tttccaaaat ttaa 1404

<210> 20

<211> 467

<212> PRT

<213> Issatchenkia orientalis

<400> 20

Met Ser Ile Ser Glu Lys Tyr Phe Pro Gln Glu Pro Gln Ser Pro Glu

1 5 10 15

Leu Lys Thr Ala Ile Pro Gly Pro Gln Ser Lys Ala Lys Leu Glu Glu

20 25 30

Leu Ser Ala Val Tyr Asp Thr Lys Ala Ala Tyr Phe Val Thr Asp Tyr

35 40 45

Tyr Lys Ser Leu Gly Asn Tyr Ile Val Asp Ala Asp Gly Asn Lys Leu

50 55 60

Leu Asp Ser Tyr Cys Gln Ile Ser Ser Ile Ala Leu Gly Tyr Asn Asn

65 70 75 80

Pro Ala Leu Leu Lys Val Ala His Ser Asp Glu Met Thr Val Ala Leu

85 90 95

Cys Asn Arg Pro Ala Leu Ala Cys Phe Pro Ser Thr Asp Tyr Tyr Glu

100 105 110

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

115 120 125

Val Cys Thr Ala His Thr Gly Ser Asp Ala Asn Glu Met Ala Phe Lys

130 135 140

Ala Ala Phe Leu Phe Gln Ala Ser Lys Lys Arg Gly Asp Lys Pro Phe

145 150 155 160

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

165 170 175

Ser Asp Met Val Ile Leu Ser Phe Glu Lys Gly Phe His Gly Arg Leu

180 185 190

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

 Ala Ala Ile Ile Glu Pro Val Gln Ser Glu Gly Gly Asp Asn His Ala
 260 265 270
 Ser Pro Glu Phe Phe Gln Gly Ile Arg Glu Ile Thr Lys Lys His Gly
 275 280 285
 Val Ile Leu Ile Val Asp Glu Val Gln Thr Gly Gly Gly Ala Ser Gly
 290 295 300
 Lys Met Trp Leu His Glu His Tyr Gly Ile Val Pro Asp Ile Met Thr
 305 310 315 320

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

 Lys Leu Glu Ala Ile Ser Ala Lys Tyr Asp Lys Met Ile Asn Leu Arg
 385 390 395 400
 Gly Lys Gly Arg Gly Phe Phe Ile Ala Phe Asp Ala Pro Thr Pro Glu
 405 410 415
 Leu Arg Asn Lys Phe Ile Ala Glu Cys Lys Lys Leu Gly Leu Asn Ile
 420 425 430
 Gly Gly Cys Gly Glu Gln Gly Val Arg Leu Arg Pro Ala Leu Val Phe

435

440

445

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

450

455

460

Ser Lys Ile

465

<210> 21

<211> 475

<212> PRT

<213> Saccharomyces kluyveri

<400> 21

Met Pro Ser Tyr Ser Val Ala Glu Leu Tyr Tyr Pro Asp Glu Pro Thr

1

5

10

15

Glu Pro Lys Ile Ser Thr Ser Ser Tyr Pro Gly Pro Lys Ala Lys Gln

20

25

30

Glu Leu Glu Lys Leu Ser Asn Val Phe Asp Thr Arg Ala Ala Tyr Leu

35

40

45

Leu Ala Asp Tyr Tyr Lys Ser Arg Gly Asn Tyr Ile Val Asp Gln Asp

50

55

60

Gly Asn Val Leu Leu Asp Val Tyr Ala Gln Ile Ser Ser Ile Ala Leu

65

70

75

80

Gly Tyr Asn Asn Pro Glu Ile Leu Lys Val Ala Lys Ser Asp Ala Met

85

90

95

Ser Val Ala Leu Ala Asn Arg Pro Ala Leu Ala Cys Phe Pro Ser Asn

100

105

110

Asp Tyr Gly Gln Leu Leu Glu Asp Gly Leu Leu Lys Ala Ala Pro Gln

115

120

125

Gly Gln Asp Lys Ile Trp Thr Ala Leu Ser Gly Ser Asp Ala Asn Glu

130

135

140

Thr Ala Phe Lys Ala Cys Phe Met Tyr Gln Ala Ala Lys Lys Arg Asn

145

150

155

160

Gly Arg Ser Phe Ser Thr Glu Glu Leu Glu Ser Val Met Asp Asn Gln

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

 225 230 235 240
 Glu Ser Ser Cys Ile Glu Lys Phe Ser Gln Ile Val Gln Glu Trp Gln
 245 250 255
 Gly Lys Ile Ala Ala Val Ile Ile Glu Pro Ile Gln Ser Glu Gly Gly
 260 265 270
 Asp Asn His Ala Ser Ser Asp Phe Phe Gln Lys Leu Arg Glu Ile Thr
 275 280 285
 Ile Glu Asn Gly Ile Leu Met Ile Val Asp Glu Val Gln Thr Gly Val

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

 355 360 365
 Tyr Glu Glu Ile Val Lys His Asp Leu Val Thr Arg Thr Ala Glu Val
 370 375 380
 Gly Asn Tyr Leu Phe Asn Arg Leu Glu Lys Leu Phe Glu Gly Lys Asn
 385 390 395 400
 Tyr Ile Gln Asn Leu Arg Gly Lys Gly Gln Gly Thr Tyr Ile Ala Phe
 405 410 415

Asp Phe Gly Thr Ser Ser Glu Arg Asp Ser Phe Leu Ser Arg Leu Arg

420 425 430

Cys Asn Gly Ala Asn Val Ala Gly Cys Gly Asp Ser Ala Val Arg Leu

435 440 445

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

450 455 460

Ile Phe Asp Lys Thr Leu Arg Gln Leu Tyr Gly

465 470 475

<210> 22

<211> 451

<212> PRT

<213> Streptomyces avermitilis

<400> 22

Met Thr Pro Gln Pro Asn Pro Gln Val Gly Ala Ala Val Lys Ala Ala

1 5 10 15

Asp Arg Ala His Val Phe His Ser Trp Ser Ala Gln Glu Leu Ile Asp

20 25 30

Pro Leu Ala Val Ala Gly Ala Glu Gly Ser Tyr Phe Trp Asp Tyr Asp

35 40 45

Gly Arg Arg Tyr Leu Asp Phe Thr Ser Gly Leu Val Phe Thr Asn Ile

50 55 60

Gly Tyr Gln His Pro Lys Val Val Ala Ala Ile Gln Glu Gln Ala Ala

65 70 75 80

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

85 90 95

Ala Ala Arg Leu Ile Ala Glu Arg Thr Pro Gly Asp Leu Asp Lys Ile

100 105 110

Phe Phe Thr Asn Gly Gly Ala Asp Ala Ile Glu His Ala Val Arg Met

115 120 125

Ala Arg Ile His Thr Gly Arg Pro Lys Val Leu Ser Ala Tyr Arg Ser

130 135 140

Tyr His Gly Gly Thr Gln Gln Ala Val Asn Ile Thr Gly Asp Pro Arg
 145 150 155 160
 Arg Trp Ala Ser Asp Ser Ala Ser Ala Gly Val Val His Phe Trp Ala
 165 170 175
 Pro Tyr Leu Tyr Arg Ser Arg Phe Tyr Ala Glu Thr Glu Gln Gln Glu
 180 185 190
 Cys Glu Arg Ala Leu Glu His Leu Glu Thr Thr Ile Ala Phe Glu Gly

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

 260 265 270
 Pro Asp Leu Met Thr Phe Ala Lys Gly Val Asn Ser Gly Tyr Val Pro
 275 280 285
 Leu Gly Gly Val Ala Ile Ser Gly Lys Ile Ala Glu Thr Phe Gly Lys
 290 295 300
 Arg Ala Tyr Pro Gly Gly Leu Thr Tyr Ser Gly His Pro Leu Ala Cys
 305 310 315 320
 Ala Ala Ala Val Ala Thr Ile Asn Val Met Ala Glu Glu Gly Val Val

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

385 390 395 400
 Phe Gly Ala Ala Ala Lys Ala Asn Gly Leu Trp Pro Phe Ile Asn Met
 405 410 415
 Asn Arg Thr His Val Val Pro Pro Cys Asn Val Thr Glu Ala Glu Ala
 420 425 430
 Lys Glu Gly Leu Ala Ala Leu Asp Ala Ala Leu Ser Val Ala Asp Glu
 435 440 445
 Tyr Thr Val
 450
 <210> 23
 <211> 419

 <212> PRT
 <213> Streptomyces avermitilis
 <400> 23
 Met Ser Ala Leu Ser Pro His Leu Arg Gln Ala Thr Pro Val Val Ala
 1 5 10 15
 Val Arg Gly Glu Gly Val His Leu Tyr Gly Glu Asp Gly Arg Arg Tyr
 20 25 30
 Leu Asp Phe Thr Ala Gly Ile Gly Val Thr Ser Thr Gly His Cys His
 35 40 45
 Pro Arg Val Val Ala Ala Ala Gln Glu Gln Ala Gly Thr Leu Val His

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

 115 120 125
 Arg Thr Val Ala Ala Ala Ala Met Thr Thr Ser Gly Thr Arg Phe Arg
 130 135 140

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

385 390 395 400
 Val Asp Glu Gly Leu Arg Ala Trp Ser Ala Ala Val Glu Val Gly Val
 405 410 415
 Pro Ala Arg

 <210> 24
 <211> 471
 <212> PRT
 <213> *Saccharomyces cerevisiae*
 <400> 24
 Met Ser Ile Cys Glu Gln Tyr Tyr Pro Glu Glu Pro Thr Lys Pro Thr
 1 5 10 15

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

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

 Ala Ala Phe Ile Tyr Tyr Arg Ala Lys Gln Arg Gly Tyr Asp Ala Asp
 145 150 155 160
 Phe Ser Glu Lys Glu Asn Leu Ser Val Met Asp Asn Asp Ala Pro Gly
 165 170 175

Ala Pro His Leu Ala Val Leu Ser Phe Lys Arg Ala Phe His Gly Arg
180 185 190

Leu Phe Ala Ser Gly Ser Thr Thr Cys Ser Lys Pro Ile His Lys Leu
195 200 205

Asp Phe Pro Ala Phe His Trp Pro His Ala Glu Tyr Pro Ser Tyr Gln
210 215 220

Tyr Pro Leu Asp Glu Asn Ser Asp Ala Asn Arg Lys Glu Asp Asp His
225 230 235 240

Cys Leu Ala Ile Val Glu Glu Leu Ile Lys Thr Trp Ser Ile Pro Val
245 250 255

Ala Ala Leu Ile Ile Glu Pro Ile Gln Ser Glu Gly Gly Asp Asn His
260 265 270

Ala Ser Lys Tyr Phe Leu Gln Lys Leu Arg Asp Ile Thr Leu Lys Tyr
275 280 285

Asn Val Val Tyr Ile Ile Asp Glu Val Gln Thr Gly Val Gly Ala Thr
290 295 300

Gly Lys Leu Trp Cys His Glu Tyr Ala Asp Ile Gln Pro Pro Val Asp
305 310 315 320

Leu Val Thr Phe Ser Lys Lys Phe Gln Ser Ala Gly Tyr Phe Phe His
325 330 335

Asp Pro Lys Phe Ile Pro Asn Lys Pro Tyr Arg Gln Phe Asn Thr Trp
340 345 350

Cys Gly Glu Pro Ala Arg Met Ile Ile Ala Gly Ala Ile Gly Gln Glu
355 360 365

Ile Ser Asp Lys Lys Leu Thr Glu Gln Cys Ser Arg Val Gly Asp Tyr
370 375 380

Leu Phe Lys Lys Leu Glu Gly Leu Gln Lys Lys Tyr Pro Glu Asn Phe
385 390 395 400

Gln Asn Leu Arg Gly Lys Gly Arg Gly Thr Phe Ile Ala Trp Asp Leu
405 410 415

Pro Thr Gly Glu Lys Arg Asp Leu Leu Leu Lys Lys Leu Lys Leu Asn

| | | | |
|--|-----|-----|-----|
| 420 | 425 | 430 | |
| Gly Cys Asn Val Gly Gly Cys Ala Val His Ala Val Arg Leu Arg Pro | | | |
| 435 | 440 | 445 | |
| Ser Leu Thr Phe Glu Glu Lys His Ala Asp Ile Phe Ile Glu Ala Leu | | | |
| 450 | 455 | 460 | |
| | | | |
| Ala Lys Ser Val Asn Glu Leu | | | |
| 465 | 470 | | |
| <210> 25 | | | |
| <211> 813 | | | |
| <212> DNA | | | |
| <213> Issatchenkia orientalis | | | |
| <400> 25 | | | |
| atgttttgta atatttccca aagacttgca ggcaagaaca tcctaattac aggtgcgtcc | | | 60 |
| actggtatcg gataccatac agcaaagtat ttgcagaag ctgcaaatgg agacttgaag | | | 120 |
| ttggttttgg ctgcaagaag aaaggagaag ctggaggcac taaaggcaga cttgcttgcc | | | 180 |
| aagiatccat ccatcaaagt ccatattgag agtttggatg tctccaaaac ggaaaccatt | | | 240 |
| gcacctttct taaaagggtt acctgaggaa ttttcaattg tcgacgtgtt ggtcaacaat | | | 300 |
| | | | |
| gcaggtaagg cgcttggttt ggatccaatt ggctctgtcg atccaaagga cgtggatgaa | | | 360 |
| atgttccaga ccaatgtttt gggtatgatt caattgacc agttggttgt acagcaaag | | | 420 |
| aaggagagaa actccgggga cattgtccaa ctaggttcag tggctggtag aaaccatac | | | 480 |
| ccaggtgggtg gtatctactg tgcctccaag gccgcattga gatcttttac acatgtattg | | | 540 |
| agagaggaat tgattaatac caagattaga gtgattgaaa tcgagcctgg aaatgttgca | | | 600 |
| actgaggaat ttcttttgac cagattcaaa ggtgataagt ccaaggccga aaaggtctat | | | 660 |
| gagggaaccg agccattgta tggtaaccgat attgcagaat tgattctatt tgcagtttct | | | 720 |
| | | | |
| agacctcaaa aactgttat tgcagaaaca ctgtttttg ctagtaacca agcttctgct | | | 780 |
| taccatattt tcagaggatc attagataaa tag | | | 813 |
| <210> 26 | | | |
| <211> 270 | | | |
| <212> PRT | | | |
| <213> Issatchenkia orientalis | | | |
| <400> 26 | | | |
| Met Phe Gly Asn Ile Ser Gln Arg Leu Ala Gly Lys Asn Ile Leu Ile | | | |

1 5 10 15
 Thr Gly Ala Ser Thr Gly Ile Gly Tyr His Thr Ala Lys Tyr Phe Ala
 20 25 30
 Glu Ala Ala Asn Gly Asp Leu Lys Leu Val Leu Ala Ala Arg Arg Lys

 35 40 45
 Glu Lys Leu Glu Ala Leu Lys Ala Asp Leu Leu Ala Lys Tyr Pro Ser
 50 55 60
 Ile Lys Val His Ile Glu Ser Leu Asp Val Ser Lys Thr Glu Thr Ile
 65 70 75 80
 Ala Pro Phe Leu Lys Gly Leu Pro Glu Glu Phe Ser Ile Val Asp Val
 85 90 95
 Leu Val Asn Asn Ala Gly Lys Ala Leu Gly Leu Asp Pro Ile Gly Ser

 100 105 110
 Val Asp Pro Lys Asp Val Asp Glu Met Phe Gln Thr Asn Val Leu Gly
 115 120 125
 Met Ile Gln Leu Thr Gln Leu Val Val Gln Gln Met Lys Glu Arg Asn
 130 135 140
 Ser Gly Asp Ile Val Gln Leu Gly Ser Val Ala Gly Arg Asn Pro Tyr
 145 150 155 160
 Pro Gly Gly Gly Ile Tyr Cys Ala Ser Lys Ala Ala Leu Arg Ser Phe

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

 225 230 235 240
 Arg Pro Gln Asn Thr Val Ile Ala Glu Thr Leu Val Phe Ala Ser Asn
 245 250 255

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

260 265 270

<210> 27

<211> 248

<212> PRT

<213> Escherichia coli

<400> 27

Met Ile Val Leu Val Thr Gly Ala Thr Ala Gly Phe Gly Glu Cys Ile

1 5 10 15

Thr Arg Arg Phe Ile Gln Gln Gly His Lys Val Ile Ala Thr Gly Arg

20 25 30

Arg Gln Glu Arg Leu Gln Glu Leu Lys Asp Glu Leu Gly Asp Asn Leu

35 40 45

Tyr Ile Ala Gln Leu Asp Val Arg Asn Arg Ala Ala Ile Glu Glu Met

50 55 60

Leu Ala Ser Leu Pro Ala Glu Trp Cys Asn Ile Asp Ile Leu Val Asn

65 70 75 80

Asn Ala Gly Leu Ala Leu Gly Met Glu Pro Ala His Lys Ala Ser Val

85 90 95

Glu Asp Trp Glu Thr Met Ile Asp Thr Asn Asn Lys Gly Leu Val Tyr

100 105 110

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

115 120 125

Ile Ile Asn Ile Gly Ser Thr Ala Gly Ser Trp Pro Tyr Ala Gly Gly

130 135 140

Asn Val Tyr Gly Ala Thr Lys Ala Phe Val Arg Gln Phe Ser Leu Asn

145 150 155 160

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

165 170 175

Pro Gly Leu Val Gly Gly Thr Glu Phe Ser Asn Val Arg Phe Lys Gly

180 185 190

Asp Asp Gly Lys Ala Glu Lys Thr Tyr Gln Asn Thr Val Ala Leu Thr

195

200

205

Pro Glu Asp Val Ser Glu Ala Val Trp Trp Val Ser Thr Leu Pro Ala

210

215

220

His Val Asn Ile Asn Thr Leu Glu Met Met Pro Val Thr Gln Ser Tyr

225

230

235

240

Ala Gly Leu Asn Val His Arg Gln

245

<210> 28

<211> 298

<212> PRT

<213> *Alcaligenes faecalis*

<400> 28

Met Ser Asn Thr Ile Ala Phe Ile Gly Leu Gly His Met Gly Lys Pro

1

5

10

15

Met Ala Leu Asn Leu Leu Lys Ala Gly His Ser Leu Asn Val Phe Asp

20

25

30

Leu Asn Ala Gln Ala Met Gln Glu Leu Gln Ala Ala Gly Ala Gln Val

35

40

45

Gly Glu Ser Ala Val Gln Ile Ala Gln Asp Ala Gln Met Val Phe Thr

50

55

60

Met Leu Pro Ala Gly Arg His Val Arg Gln Val Tyr Glu Gly Glu Asn

65

70

75

80

Gly Leu Leu Gln Thr Val Ala Pro Gly Thr Val Leu Val Asp Cys Ser

85

90

95

Thr Ile Asp Ala Gln Thr Ser Gln Asp Leu Ala Ala Lys Ala Ser Lys

100

105

110

Leu Gly Leu Phe Met Leu Asp Ala Pro Val Ser Gly Gly Thr Gly Gly

115

120

125

Ala Ile Ala Gly Thr Leu Thr Phe Met Val Gly Gly Glu Asp Gln Ala

130

135

140

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

145 150 155 160
 His Ala Gly Lys Ala Gly Ala Gly Gln Val Ala Lys Ile Cys Asn Asn
 165 170 175
 Met Leu Leu Gly Ile Leu Met Ala Gly Thr Ala Glu Ala Leu Ala Leu
 180 185 190
 Gly Val Ala His Gly Leu Asp Pro Ala Val Leu Ser Thr Ile Met Ala
 195 200 205

Arg Ser Ser Gly Arg Asn Trp Ala Thr Glu Leu Tyr Asn Pro Trp Pro
 210 215 220
 Gly Val Met Pro Asp Val Pro Ala Ser Arg Asp Tyr Gln Gly Gly Phe
 225 230 235 240
 Ala Thr Gly Leu Met Leu Lys Asp Leu Gly Leu Ala Ala Asp Ala Ala
 245 250 255
 Val Ser Gln Asn Ser Ala Thr Pro Leu Gly Glu Leu Ala Arg Asn Leu
 260 265 270

Phe Ala Leu His Ala Ala Gln Gly Gln Asn Ala Gly Leu Asp Phe Ser
 275 280 285
 Ser Ile Leu Asn Leu Tyr Arg Gln Lys His
 290 295

<210> 29

<211> 314

<212> PRT

<213> Metallosphaera sedula

<400> 29

Met Thr Glu Lys Val Ser Val Val Gly Ala Gly Val Ile Gly Val Gly
 1 5 10 15
 Trp Ala Thr Leu Phe Ala Ser Lys Gly Tyr Ser Val Ser Leu Tyr Thr
 20 25 30

Glu Lys Lys Glu Thr Leu Asp Lys Gly Ile Glu Lys Leu Arg Asn Tyr
 35 40 45
 Val Gln Val Met Lys Asn Asn Ser Gln Ile Thr Glu Asp Val Asn Thr
 50 55 60

Val Ile Ser Arg Val Ser Pro Thr Thr Asn Leu Asp Glu Ala Val Arg
65 70 75 80
Gly Ala Asn Phe Val Ile Glu Ala Val Ile Glu Asp Tyr Asp Ala Lys
85 90 95

Lys Lys Ile Phe Gly Tyr Leu Asp Ser Val Leu Asp Lys Glu Val Ile
100 105 110
Leu Ala Ser Ser Thr Ser Gly Leu Leu Ile Thr Glu Val Gln Lys Ala
115 120 125
Met Ser Lys His Pro Glu Arg Ala Val Ile Ala His Pro Trp Asn Pro
130 135 140
Pro His Leu Leu Pro Leu Val Glu Ile Val Pro Gly Glu Lys Thr Ser
145 150 155 160

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

Leu Arg Trp Ala Phe Met Gly Pro Phe Leu Thr Tyr His Leu Gly Gly
225 230 235 240
Gly Glu Gly Gly Leu Glu Tyr Phe Phe Asn Arg Gly Phe Gly Tyr Gly
245 250 255
Ala Asn Glu Trp Met His Thr Leu Ala Lys Tyr Asp Lys Phe Pro Tyr
260 265 270
Thr Gly Val Thr Lys Ala Ile Gln Gln Met Lys Glu Tyr Ser Phe Ile
275 280 285

Lys Gly Lys Thr Phe Gln Glu Ile Ser Lys Trp Arg Asp Glu Lys Leu
290 295 300
Leu Lys Val Tyr Lys Leu Val Trp Glu Lys

| | | | | | | | | | | | | | | | | |
|-------|--------------------|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|--|
| 305 | 310 | | | | | | | | | | | | | | | |
| <210> | 30 | | | | | | | | | | | | | | | |
| <211> | 295 | | | | | | | | | | | | | | | |
| <212> | PRT | | | | | | | | | | | | | | | |
| <213> | Pseudomonas putida | | | | | | | | | | | | | | | |
| <400> | 30 | | | | | | | | | | | | | | | |
| Met | Arg | Ile | Ala | Phe | Ile | Gly | Leu | Gly | Asn | Met | Gly | Ala | Pro | Met | Ala | |
| 1 | | | 5 | | | | | 10 | | | | | 15 | | | |
| Arg | Asn | Leu | Ile | Lys | Ala | Gly | His | Gln | Leu | Asn | Leu | Phe | Asp | Leu | Asn | |
| | | | 20 | | | | | 25 | | | | | 30 | | | |
| Lys | Ala | Val | Leu | Ala | Glu | Leu | Ala | Glu | Leu | Gly | Gly | Gln | Ile | Ser | Pro | |
| | | 35 | | | | | 40 | | | | | 45 | | | | |
| Ser | Pro | Lys | Asp | Ala | Ala | Ala | Asn | Ser | Glu | Leu | Val | Ile | Thr | Met | Leu | |
| | 50 | | | | | 55 | | | | | 60 | | | | | |
| Pro | Ala | Ala | Ala | His | Val | Arg | Ser | Val | Tyr | Leu | Asn | Glu | Asp | Gly | Val | |
| 65 | | | | 70 | | | | | | 75 | | | | | 80 | |
| Leu | Ala | Gly | Ile | Arg | Pro | Gly | Thr | Pro | Thr | Val | Asp | Cys | Ser | Thr | Ile | |
| | | | | 85 | | | | | | 90 | | | | 95 | | |
| Asp | Pro | Gln | Thr | Ala | Arg | Asp | Val | Ser | Lys | Ala | Ala | Ala | Ala | Lys | Gly | |
| | | 100 | | | | | | 105 | | | | | 110 | | | |
| Val | Asp | Met | Gly | Asp | Ala | Pro | Val | Ser | Gly | Gly | Thr | Gly | Gly | Ala | Ala | |
| | | 115 | | | | | 120 | | | | | 125 | | | | |
| Ala | Gly | Thr | Leu | Thr | Phe | Met | Val | Gly | Ala | Ser | Thr | Glu | Leu | Phe | Ala | |
| | 130 | | | | | 135 | | | | | 140 | | | | | |
| Ser | Leu | Lys | Pro | Val | Leu | Glu | Gln | Met | Gly | Arg | Asn | Ile | Val | His | Cys | |
| 145 | | | | 150 | | | | | | 155 | | | | | 160 | |
| Gly | Glu | Val | Gly | Thr | Gly | Gln | Ile | Ala | Lys | Ile | Cys | Asn | Asn | Leu | Leu | |
| | | | 165 | | | | | | 170 | | | | 175 | | | |
| Leu | Gly | Ile | Ser | Met | Ile | Gly | Val | Ser | Glu | Ala | Met | Ala | Leu | Gly | Asn | |
| | | 180 | | | | | | 185 | | | | | 190 | | | |
| Ala | Leu | Gly | Ile | Asp | Thr | Lys | Val | Leu | Ala | Gly | Ile | Ile | Asn | Ser | Ser | |
| | 195 | | | | | 200 | | | | | | 205 | | | | |

Thr Gly Arg Cys Trp Ser Ser Asp Thr Tyr Asn Pro Trp Pro Gly Ile
210 215 220

Ile Glu Thr Ala Pro Ala Ser Arg Gly Tyr Thr Gly Gly Phe Gly Ala
225 230 235 240

Glu Leu Met Leu Lys Asp Leu Gly Leu Ala Thr Glu Ala Ala Arg Gln
245 250 255

Ala His Gln Pro Val Ile Leu Gly Ala Val Ala Gln Gln Leu Tyr Gln
260 265 270

Ala Met Ser Leu Arg Gly Glu Gly Gly Lys Asp Phe Ser Ala Ile Val
275 280 285

Glu Gly Tyr Arg Lys Lys Asp
290 295

<210> 31
<211> 295
<212> PRT
<213> Pseudomonas putida
<400> 31

Met Arg Ile Ala Phe Ile Gly Leu Gly Asn Met Gly Ala Pro Met Ala
1 5 10 15

Arg Asn Leu Ile Lys Ala Gly His Gln Leu Asn Leu Phe Asp Leu Asn
20 25 30

Lys Thr Val Leu Ala Glu Leu Ala Glu Leu Gly Gly Gln Ile Ser Pro
35 40 45

Ser Pro Lys Asp Ala Ala Ala Ser Ser Glu Leu Val Ile Thr Met Leu
50 55 60

Pro Ala Ala Ala His Val Arg Ser Val Tyr Leu Asn Asp Asp Gly Val
65 70 75 80

Leu Ala Gly Ile Arg Pro Gly Thr Pro Thr Val Asp Cys Ser Thr Ile
85 90 95

Asp Pro Gln Thr Ala Arg Asp Val Ser Lys Ala Ala Ala Ala Lys Gly
100 105 110

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

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

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

<210> 32

<211> 298

<212> PRT

<213> Pseudomonas aeruginosa

<400> 32

Met Thr Asp Ile Ala Phe Leu Gly Leu Gly Asn Met Gly Gly Pro Met
1 5 10 15
Ala Ala Asn Leu Leu Lys Ala Gly His Arg Val Asn Val Phe Asp Leu

| | | |
|---|-----|-----|
| 20 | 25 | 30 |
| Gln Pro Lys Ala Val Leu Gly Leu Val Glu Gln Gly Ala Gln Gly Ala | | |
| 35 | 40 | 45 |
| Asp Ser Ala Leu Gln Cys Cys Glu Gly Ala Glu Val Val Ile Ser Met | | |
| 50 | 55 | 60 |
| | | |
| Leu Pro Ala Gly Gln His Val Glu Ser Leu Tyr Leu Gly Asp Asp Gly | | |
| 65 | 70 | 75 |
| Leu Leu Ala Arg Val Ala Gly Lys Pro Leu Leu Ile Asp Cys Ser Thr | | |
| 85 | 90 | 95 |
| Ile Ala Pro Glu Thr Ala Arg Lys Val Ala Glu Ala Ala Ala Ala Lys | | |
| 100 | 105 | 110 |
| Gly Leu Thr Leu Leu Asp Ala Pro Val Ser Gly Gly Val Gly Gly Ala | | |
| 115 | 120 | 125 |
| | | |
| Arg Ala Gly Thr Leu Ser Phe Ile Val Gly Gly Pro Ala Glu Gly Phe | | |
| 130 | 135 | 140 |
| Ala Arg Ala Arg Pro Val Leu Glu Asn Met Gly Arg Asn Ile Phe His | | |
| 145 | 150 | 155 |
| Ala Gly Asp His Gly Ala Gly Gln Val Ala Lys Ile Cys Asn Asn Met | | |
| 165 | 170 | 175 |
| Leu Leu Gly Ile Leu Met Ala Gly Thr Ala Glu Ala Leu Ala Leu Gly | | |
| 180 | 185 | 190 |
| | | |
| Val Lys Asn Gly Leu Asp Pro Ala Val Leu Ser Glu Val Met Lys Gln | | |
| 195 | 200 | 205 |
| Ser Ser Gly Gly Asn Trp Ala Leu Asn Leu Tyr Asn Pro Trp Pro Gly | | |
| 210 | 215 | 220 |
| Val Met Pro Gln Ala Pro Ala Ser Asn Gly Tyr Ala Gly Gly Phe Gln | | |
| 225 | 230 | 235 |
| Val Arg Leu Met Asn Lys Asp Leu Gly Leu Ala Leu Ala Asn Ala Gln | | |
| 245 | 250 | 255 |
| | | |
| Ala Val Gln Ala Ser Thr Pro Leu Gly Ala Leu Ala Arg Asn Leu Phe | | |
| 260 | 265 | 270 |

Ser Leu His Ala Gln Ala Asp Ala Glu His Glu Gly Leu Asp Phe Ser

275

280

285

Ser Ile Gln Lys Leu Tyr Arg Gly Lys Asp

290

295

<210> 33

<211> 382

<212> PRT

<213> Ralstonia eutropha

<400> 33

Met Ala Phe Ile Tyr Tyr Leu Thr His Ile His Leu Asp Phe Gly Ala

1

5

10

15

Val Ser Leu Leu Lys Ser Glu Cys Glu Arg Ile Gly Ile Arg Arg Pro

20

25

30

Leu Leu Val Thr Asp Lys Gly Val Val Ala Ala Gly Val Ala Gln Arg

35

40

45

Ala Ile Asp Ala Met Gln Gly Leu Gln Val Ala Val Phe Asp Glu Thr

50

55

60

Pro Ser Asn Pro Thr Glu Ala Met Val Arg Lys Ala Ala Ala Gln Tyr

65

70

75

80

Arg Glu Ala Gly Cys Asp Gly Leu Val Ala Val Gly Gly Gly Ser Ser

85

90

95

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

100

105

110

Leu Thr Thr Tyr Ala Thr Ile Glu Gly Gly Ser Ala Arg Ile Thr Asp

115

120

125

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

130

135

140

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

145

150

155

160

Gly Phe His Ser Trp His Leu Leu Pro Lys Ser Ala Val Cys Asp Pro

165

170

175

Glu Leu Thr Leu Gly Leu Pro Ala Gly Leu Thr Ala Ala Thr Gly Met

180 185 190
 Asp Ala Ile Ala His Cys Ile Glu Thr Phe Leu Ala Pro Ala Phe Asn
 195 200 205

Pro Pro Ala Asp Gly Ile Ala Leu Asp Gly Leu Glu Arg Gly Trp Gly
 210 215 220
 His Ile Glu Arg Ala Thr Arg Asp Gly Gln Asp Arg Asp Ala Arg Leu
 225 230 235 240
 Asn Met Met Ser Ala Ser Met Gln Gly Ala Met Ala Phe Gln Lys Gly
 245 250 255
 Leu Gly Cys Val His Ser Leu Ser His Pro Leu Gly Gly Leu Lys Ile
 260 265 270

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

Thr Gly Leu Arg Gln Met Gly Val Thr Glu Asp Met Phe Asp Lys Val
 340 345 350
 Ile Ala Gly Ala Leu Val Asp His Cys His Lys Thr Asn Pro Lys Glu
 355 360 365
 Ala Ser Ala Ala Asp Tyr Arg Arg Met Leu Glu Gln Ser Met
 370 375 380

<210> 34

<211> 371

<212> PRT

<213> Clostridium kluyveri

<400> 34

Met Lys Leu Leu Lys Leu Ala Pro Asp Val Tyr Lys Phe Asp Thr Ala

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

 65 70 75 80
 Arg Ile Ile Ala Val Gly Gly Gly Ser Val Ile Asp Ile Ala Lys Ile
 85 90 95
 Leu Ser Leu Lys Tyr Thr Asp Asp Ser Leu Asp Leu Phe Glu Gly Lys
 100 105 110
 Val Pro Leu Val Lys Asn Lys Glu Leu Ile Ile Val Pro Thr Thr Cys
 115 120 125
 Gly Thr Gly Ser Glu Val Thr Asn Val Ser Val Ala Glu Leu Lys Arg

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

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

Tyr Pro Ile Gly Gly Asn Tyr His Val Pro His Gly Glu Ala Asn Tyr

260 265 270
 Leu Phe Phe Thr Glu Ile Phe Lys Thr Tyr Tyr Glu Lys Asn Pro Asn
 275 280 285
 Gly Lys Ile Lys Asp Val Asn Lys Leu Leu Ala Gly Ile Leu Lys Cys
 290 295 300
 Asp Glu Ser Glu Ala Tyr Asp Ser Leu Ser Gln Leu Leu Asp Lys Leu
 305 310 315 320
 Leu Ser Arg Lys Pro Leu Arg Glu Tyr Gly Met Lys Glu Glu Glu Ile

325 330 335
 Glu Thr Phe Ala Asp Ser Val Ile Glu Gly Gln Gln Arg Leu Leu Val
 340 345 350
 Asn Asn Tyr Glu Pro Phe Ser Arg Glu Asp Ile Val Asn Thr Tyr Lys
 355 360 365
 Lys Leu Tyr

370

<210> 35

<211> 538

<212> PRT

<213> Mannheimia succiniciproducens

<400> 35

Met Thr Asp Leu Asn Gln Leu Thr Gln Glu Leu Gly Ala Leu Gly Ile

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

Val Trp Trp Thr Ser Glu Lys Val Lys Asn Asp Asn Lys Pro Met Ser
85 90 95

Gln Asp Thr Trp Asn Ser Leu Lys Gly Leu Val Ala Asp Gln Leu Ser
100 105 110

Gly Lys Arg Leu Phe Val Val Asp Ala Phe Cys Gly Ala Asn Lys Asp
115 120 125

Thr Arg Leu Ala Val Arg Val Val Thr Glu Val Ala Trp Gln Ala His
130 135 140

Phe Val Thr Asn Met Phe Ile Arg Pro Ser Ala Glu Glu Leu Lys Gly
145 150 155 160

Phe Lys Pro Asp Phe Val Val Met Asn Gly Ala Lys Cys Thr Asn Pro
165 170 175

Asn Trp Lys Glu Gln Gly Leu Asn Ser Glu Asn Phe Val Ala Phe Asn
180 185 190

Ile Thr Glu Gly Val Gln Leu Ile Gly Gly Thr Trp Tyr Gly Gly Glu
195 200 205

Met Lys Lys Gly Met Phe Ser Met Met Asn Tyr Phe Leu Pro Leu Arg
210 215 220

Gly Ile Ala Ser Met His Cys Ser Ala Asn Val Gly Lys Asp Gly Asp
225 230 235 240

Thr Ala Ile Phe Phe Gly Leu Ser Gly Thr Gly Lys Thr Thr Leu Ser
245 250 255

Thr Asp Pro Lys Arg Gln Leu Ile Gly Asp Asp Glu His Gly Trp Asp
260 265 270

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

Asn Leu Ser Ala Glu Asn Glu Pro Asp Ile Tyr Gly Ala Ile Lys Arg
290 295 300

Asp Ala Leu Leu Glu Asn Val Val Val Leu Asp Asn Gly Asp Val Asp
305 310 315 320

Tyr Ala Asp Gly Ser Lys Thr Glu Asn Thr Arg Val Ser Tyr Pro Ile

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

385 390 395 400
Pro Thr Phe Ser Ala Cys Phe Gly Ala Ala Phe Leu Ser Leu His Pro
405 410 415
Thr Gln Tyr Ala Glu Val Leu Val Lys Arg Met Gln Glu Ser Gly Ala
420 425 430
Glu Ala Tyr Leu Val Asn Thr Gly Trp Asn Gly Thr Gly Lys Arg Ile
435 440 445
Ser Ile Lys Asp Thr Arg Gly Ile Ile Asp Ala Ile Leu Asp Gly Ser

450 455 460
Ile Asp Lys Ala Glu Met Gly Ser Leu Pro Ile Phe Asp Phe Ser Ile
465 470 475 480
Pro Lys Ala Leu Pro Gly Val Asn Pro Ala Ile Leu Asp Pro Arg Asp
485 490 495
Thr Tyr Ala Asp Lys Ala Gln Trp Glu Glu Lys Ala Gln Asp Leu Ala
500 505 510
Gly Arg Phe Val Lys Asn Phe Glu Lys Tyr Thr Gly Thr Ala Glu Gly

515 520 525
Gln Ala Leu Val Ala Ala Gly Pro Lys Ala

530 535

<210> 36

<211> 532

<212> PRT

<213> Anaerobiospirillum succiniciproducens

<400> 36

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

| | | |
|---|-----|-----|
| 245 | 250 | 255 |
| Arg Leu Leu Ile Gly Asp Asp Glu His Gly Trp Asp Asp Asp Gly Val | | |
| 260 | 265 | 270 |
| Phe Asn Phe Glu Gly Gly Cys Tyr Ala Lys Val Ile Asn Leu Ser Lys | | |
| 275 | 280 | 285 |
| | | |
| Glu Asn Glu Pro Asp Ile Trp Gly Ala Ile Lys Arg Asn Ala Leu Leu | | |
| 290 | 295 | 300 |
| Glu Asn Val Thr Val Asp Ala Asn Gly Lys Val Asp Phe Ala Asp Lys | | |
| 305 | 310 | 315 |
| Ser Val Thr Glu Asn Thr Arg Val Ser Tyr Pro Ile Phe His Ile Lys | | |
| 325 | 330 | 335 |
| Asn Ile Val Lys Pro Val Ser Lys Ala Pro Ala Ala Lys Arg Val Ile | | |
| 340 | 345 | 350 |
| | | |
| Phe Leu Ser Ala Asp Ala Phe Gly Val Leu Pro Pro Val Ser Ile Leu | | |
| 355 | 360 | 365 |
| Ser Lys Glu Gln Thr Lys Tyr Tyr Phe Leu Ser Gly Phe Thr Ala Lys | | |
| 370 | 375 | 380 |
| Leu Ala Gly Thr Glu Arg Gly Ile Thr Glu Pro Thr Pro Thr Phe Ser | | |
| 385 | 390 | 395 |
| Ser Cys Phe Gly Ala Ala Phe Leu Thr Leu Pro Pro Thr Lys Tyr Ala | | |
| 405 | 410 | 415 |
| | | |
| Glu Val Leu Val Lys Arg Met Glu Ala Ser Gly Ala Lys Ala Tyr Leu | | |
| 420 | 425 | 430 |
| Val Asn Thr Gly Trp Asn Gly Thr Gly Lys Arg Ile Ser Ile Lys Asp | | |
| 435 | 440 | 445 |
| Thr Arg Gly Ile Ile Asp Ala Ile Leu Asp Gly Ser Ile Asp Thr Ala | | |
| 450 | 455 | 460 |
| Asn Thr Ala Thr Ile Pro Tyr Phe Asn Phe Thr Val Pro Thr Glu Leu | | |
| 465 | 470 | 475 |
| | | |
| Lys Gly Val Asp Thr Lys Ile Leu Asp Pro Arg Asn Thr Tyr Ala Asp | | |
| 485 | 490 | 495 |

Ala Ser Glu Trp Glu Val Lys Ala Lys Asp Leu Ala Glu Arg Phe Gln
 500 505 510
 Lys Asn Phe Lys Lys Phe Glu Ser Leu Gly Gly Asp Leu Val Lys Ala
 515 520 525
 Gly Pro Gln Leu

530

<210> 37

<211> 538

<212> PRT

<213> Actinobacillus succinogenes

<400

> 37

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

Ser Pro Lys Asp Lys Tyr Ile Val Cys Asp Glu Thr Thr Lys Asp Thr
 65 70 75 80
 Val Trp Trp Asn Ser Glu Ala Ala Lys Asn Asp Asn Lys Pro Met Thr
 85 90 95
 Gln Glu Thr Trp Lys Ser Leu Arg Glu Leu Val Ala Lys Gln Leu Ser
 100 105 110
 Gly Lys Arg Leu Phe Val Val Glu Gly Tyr Cys Gly Ala Ser Glu Lys
 115 120 125

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

| | | | |
|---|-----|-----|-----|
| 165 | 170 | 175 | |
| Asn Trp Lys Glu Gln Gly Leu Asn Ser Glu Asn Phe Val Ala Phe Asn | | | |
| 180 | 185 | 190 | |
| Ile Thr Glu Gly Ile Gln Leu Ile Gly Gly Thr Trp Tyr Gly Gly Glu | | | |
| 195 | 200 | 205 | |
| Met Lys Lys Gly Met Phe Ser Met Met Asn Tyr Phe Leu Pro Leu Lys | | | |
| 210 | 215 | 220 | |
| Gly Val Ala Ser Met His Cys Ser Ala Asn Val Gly Lys Asp Gly Asp | | | |
| 225 | 230 | 235 | 240 |
| Val Ala Ile Phe Phe Gly Leu Ser Gly Thr Gly Lys Thr Thr Leu Ser | | | |
| 245 | 250 | 255 | |
| Thr Asp Pro Lys Arg Gln Leu Ile Gly Asp Asp Glu His Gly Trp Asp | | | |
| 260 | 265 | 270 | |
| Glu Ser Gly Val Phe Asn Phe Glu Gly Gly Cys Tyr Ala Lys Thr Ile | | | |
| 275 | 280 | 285 | |
| Asn Leu Ser Gln Glu Asn Glu Pro Asp Ile Tyr Gly Ala Ile Arg Arg | | | |
| 290 | 295 | 300 | |
| Asp Ala Leu Leu Glu Asn Val Val Val Arg Ala Asp Gly Ser Val Asp | | | |
| 305 | 310 | 315 | 320 |
| Phe Asp Asp Gly Ser Lys Thr Glu Asn Thr Arg Val Ser Tyr Pro Ile | | | |
| 325 | 330 | 335 | |
| Tyr His Ile Asp Asn Ile Val Arg Pro Val Ser Lys Ala Gly His Ala | | | |
| 340 | 345 | 350 | |
| Thr Lys Val Ile Phe Leu Thr Ala Asp Ala Phe Gly Val Leu Pro Pro | | | |
| 355 | 360 | 365 | |
| Val Ser Lys Leu Thr Pro Glu Gln Thr Glu Tyr Tyr Phe Leu Ser Gly | | | |
| 370 | 375 | 380 | |
| Phe Thr Ala Lys Leu Ala Gly Thr Glu Arg Gly Val Thr Glu Pro Thr | | | |
| 385 | 390 | 395 | 400 |
| Pro Thr Phe Ser Ala Cys Phe Gly Ala Ala Phe Leu Ser Leu His Pro | | | |
| 405 | 410 | 415 | |

Ile Gln Tyr Ala Asp Val Leu Val Glu Arg Met Lys Ala Ser Gly Ala

420 425 430

Glu Ala Tyr Leu Val Asn Thr Gly Trp Asn Gly Thr Gly Lys Arg Ile

435 440 445

Ser Ile Lys Asp Thr Arg Gly Ile Ile Asp Ala Ile Leu Asp Gly Ser

450 455 460

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

465 470 475 480

Pro Lys Ala Leu Pro Gly Val Asp Pro Ala Ile Leu Asp Pro Arg Asp

485 490 495

Thr Tyr Ala Asp Lys Ala Gln Trp Gln Val Lys Ala Glu Asp Leu Ala

500 505 510

Asn Arg Phe Val Lys Asn Phe Val Lys Tyr Thr Ala Asn Pro Glu Ala

515 520 525

Ala Lys Leu Val Gly Ala Gly Pro Lys Ala

530 535

<210> 38

<211> 618

<212> PRT

<213> Ralstonia eutropha

<400> 38

Met Asn His Pro Ser Met Gln Gly Thr Thr Ala Leu Asn Val Pro Ala

1 5 10 15

Trp Val Arg Asn Gln Lys Leu Val Ala Trp Val Ala Glu Ile Ala Ala

20 25 30

Leu Thr Lys Pro Glu Arg Ile His Trp Cys Asp Gly Ser Gln Glu Glu

35 40 45

Tyr Asp Arg Leu Cys Glu Gln Met Val Ala Ala Gly Thr Leu Lys Arg

50 55 60

Leu Asn Pro Ala Lys Arg Lys Asn Ser Tyr Leu Ala Leu Ser Asp Pro

65 70 75 80

Ser Asp Val Ala Arg Val Glu Asp Arg Thr Phe Ile Cys Ser Gln Lys

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

Asn Ala Met Ala Thr Leu Lys Glu Asn Val Ile Phe Thr Asn Val Ala
 340 345 350

 Leu Thr Asp Asp Gly Asp Val Trp Trp Glu Gly Met Thr Lys Glu Ala
 355 360 365
 Pro Ala His Leu Thr Asp Trp Gln Gly Lys Asp Trp Thr Pro Glu Ile
 370 375 380
 Ala Lys Ala Thr Gly Ala Lys Ala Ala His Pro Asn Ala Arg Phe Thr
 385 390 395 400
 Ala Pro Ala Ser Gln Cys Pro Ser Ile Asp Glu Asn Trp Asp Asn Pro
 405 410 415

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

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

 Glu Gly Ala Glu His Val Phe Gly Thr Ser Pro Arg Tyr Glu Asp Leu
 545 550 555 560
 Asn Trp Ser Gly Val Glu Phe Ser Val Ala Gln Phe Thr Gln Val Thr
 565 570 575
 Ser Ile Asp Ala Asp Ala Trp Lys Gln Glu Leu Ala Leu His Asp Glu

580 585 590
 Leu Phe Thr Gln Leu Lys His Asn Leu Pro Gln Ala Leu Ala Glu Ala
 595 600 605

Arg Ala Ala Leu Gly Lys Arg Leu Glu Gly
 610 615

<210> 39

<211> 540

<212> PRT

<213> Escherichia coli

<400> 39

Met Arg Val Asn Asn Gly Leu Thr Pro Gln Glu Leu Glu Ala Tyr Gly
 1 5 10 15
 Ile Ser Asp Val His Asp Ile Val Tyr Asn Pro Ser Tyr Asp Leu Leu
 20 25 30
 Tyr Gln Glu Glu Leu Asp Pro Ser Leu Thr Gly Tyr Glu Arg Gly Val
 35 40 45

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

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

| | | |
|---|-----|-----|
| 165 | 170 | 175 |
| Asn Pro Gln Trp Lys Glu Gln Gly Leu Asn Ser Glu Asn Phe Val Ala | | |
| 180 | 185 | 190 |
| Phe Asn Leu Thr Glu Arg Met Gln Leu Ile Gly Gly Thr Trp Tyr Gly | | |
| 195 | 200 | 205 |
| Gly Glu Met Lys Lys Gly Met Phe Ser Met Met Asn Tyr Leu Leu Pro | | |
| 210 | 215 | 220 |
| Leu Lys Gly Ile Ala Ser Met His Cys Ser Ala Asn Val Gly Glu Lys | | |
| 225 | 230 | 235 |
| | | 240 |
| | | |
| Gly Asp Val Ala Val Phe Phe Gly Leu Ser Gly Thr Gly Lys Thr Thr | | |
| 245 | 250 | 255 |
| Leu Ser Thr Asp Pro Lys Arg Arg Leu Ile Gly Asp Asp Glu His Gly | | |
| 260 | 265 | 270 |
| Trp Asp Asp Asp Gly Val Phe Asn Phe Glu Gly Gly Cys Tyr Ala Lys | | |
| 275 | 280 | 285 |
| Thr Ile Lys Leu Ser Lys Glu Ala Glu Pro Glu Ile Tyr Asn Ala Ile | | |
| 290 | 295 | 300 |
| | | |
| Arg Arg Asp Ala Leu Leu Glu Asn Val Thr Val Arg Glu Asp Gly Thr | | |
| 305 | 310 | 315 |
| Ile Asp Phe Asp Asp Gly Ser Lys Thr Glu Asn Thr Arg Val Ser Tyr | | |
| 325 | 330 | 335 |
| Pro Ile Tyr His Ile Asp Asn Ile Val Lys Pro Val Ser Lys Ala Gly | | |
| 340 | 345 | 350 |
| His Ala Thr Lys Val Ile Phe Leu Thr Ala Asp Ala Phe Gly Val Leu | | |
| 355 | 360 | 365 |
| | | |
| Pro Pro Val Ser Arg Leu Thr Ala Asp Gln Thr Gln Tyr His Phe Leu | | |
| 370 | 375 | 380 |
| Ser Gly Phe Thr Ala Lys Leu Ala Gly Thr Glu Arg Gly Ile Thr Glu | | |
| 385 | 390 | 395 |
| Pro Thr Pro Thr Phe Ser Ala Cys Phe Gly Ala Ala Phe Leu Ser Leu | | |
| 405 | 410 | 415 |

His Pro Thr Gln Tyr Ala Glu Val Leu Val Lys Arg Met Gln Ala Ala
420 425 430

Gly Ala Gln Ala Tyr Leu Val Asn Thr Gly Trp Asn Gly Thr Gly Lys
435 440 445

Arg Ile Ser Ile Lys Asp Thr Arg Ala Ile Ile Asp Ala Ile Leu Asn
450 455 460

Gly Ser Leu Asp Asn Ala Glu Thr Phe Thr Leu Pro Met Phe Asn Leu
465 470 475 480

Ala Ile Pro Thr Glu Leu Pro Gly Val Asp Thr Lys Ile Leu Asp Pro
485 490 495

Arg Asn Thr Tyr Ala Ser Pro Glu Gln Trp Gln Glu Lys Ala Glu Thr
500 505 510

Leu Ala Lys Leu Phe Ile Asp Asn Phe Asp Lys Tyr Thr Asp Thr Pro
515 520 525

Ala Gly Ala Ala Leu Val Ala Ala Gly Pro Lys Leu
530 535 540

<210> 40

<211> 236

<212> PRT

<213> Lactococcus lactis

<400> 40

Met Ser Glu Ile Thr Gln Leu Phe Gln Tyr Asn Thr Leu Gly Ala Leu
1 5 10 15

Met Ala Gly Leu Tyr Glu Gly Thr Met Thr Ile Gly Glu Leu Leu Lys
20 25 30

His Gly Asp Leu Gly Ile Gly Thr Leu Asp Ser Ile Asp Gly Glu Leu
35 40 45

Ile Val Leu Asp Gly Lys Ala Tyr Gln Ala Lys Gly Asp Lys Thr Ile
50 55 60

Val Glu Leu Thr Asp Asp Ile Lys Val Pro Tyr Ala Ala Val Val Pro
65 70 75 80

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

Gln Pro Glu Tyr Thr Glu Glu Asn Ile Lys Gly Thr Ile Val Gly Ile
145 150 155 160
Trp Thr Pro Glu Met Phe His Gly Val Ser Val Ala Gly Tyr His Leu
165 170 175
His Phe Ile Ser Glu Asp Phe Thr Phe Gly Gly His Val Leu Asp Phe
180 185 190
Ile Ile Asp Asn Gly Thr Val Glu Ile Gly Ala Ile Asp Gln Leu Asn
195 200 205

Gln Ser Phe Pro Val Gln Asp Arg Lys Phe Leu Phe Ala Asp Leu Asp
210 215 220
Ile Glu Ala Leu Lys Lys Asp Ile Asp Val Ala Glu
225 230 235

<210> 41

<211> 239

<212> PRT

<213> Streptococcus thermophilus

<400> 41

Met Ser Glu Ala Ile Lys Leu Phe Gln Tyr Asn Thr Leu Gly Ala Leu
1 5 10 15
Met Ala Gly Leu Tyr Gly Gly Thr Leu Thr Val Gly Glu Leu Leu Glu

20 25 30
His Gly Asp Leu Gly Leu Gly Thr Leu Asp Ser Ile Asp Gly Glu Leu
35 40 45
Ile Val Leu Asp Gly Lys Ala Tyr Gln Ala Lys Gly Ser Glu Gly Lys

50 55 60
 Val Glu Val Val Glu Val Ser Pro Asp Glu Lys Val Pro Tyr Ala Ala
 65 70 75 80
 Val Val Pro His Gln Ala Glu Val Ile Phe Arg Gln Arg Tyr Glu Met

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

 145 150 155 160
 Val Gly Ile Trp Thr Pro Glu Met Phe His Gly Val Ser Val Ala Gly
 165 170 175
 Tyr His Leu His Phe Ile Ser Asp Asp Leu Thr Phe Gly Gly His Val
 180 185 190
 Met Asp Phe Val Ile Glu Asn Gly Ile Ile Glu Val Gly Pro Val Asp
 195 200 205
 Gln Leu Asp Gln Arg Phe Pro Val Gln Asp Arg Gln Tyr Leu Phe Ala

 210 215 220
 Lys Phe Asn Val Asp Glu Met Arg Lys Asp Ile Thr Lys Ala Glu
 225 230 235
 <210> 42
 <211> 285
 <212> PRT
 <213> Brevibacillus brevis
 <400> 42
 Met Lys Lys Asn Ile Ile Thr Ser Ile Thr Ser Leu Ala Leu Val Ala
 1 5 10 15
 Gly Leu Ser Leu Thr Ala Phe Ala Ala Thr Thr Ala Thr Val Pro Ala

20

25

30

Pro Pro Ala Lys Gln Glu Ser Lys Pro Ala Val Ala Ala Asn Pro Ala

35

40

45

Pro Lys Asn Val Leu Phe Gln Tyr Ser Thr Ile Asn Ala Leu Met Leu

50

55

60

Gly Gln Phe Glu Gly Asp Leu Thr Leu Lys Asp Leu Lys Leu Arg Gly

65

70

75

80

Asp Met Gly Leu Gly Thr Ile Asn Asp Leu Asp Gly Glu Met Ile Gln

85

90

95

Met Gly Thr Lys Phe Tyr Gln Ile Asp Ser Thr Gly Lys Leu Ser Glu

100

105

110

Leu Pro Glu Ser Val Lys Thr Pro Phe Ala Val Thr Thr His Phe Glu

115

120

125

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

130

135

140

Thr Lys Met Leu Glu Glu Lys Phe Glu Asn Lys Asn Val Phe Tyr Ala

145

150

155

160

Val Lys Leu Thr Gly Thr Phe Lys Met Val Lys Ala Arg Thr Val Pro

165

170

175

Lys Gln Thr Arg Pro Tyr Pro Gln Leu Thr Glu Val Thr Lys Lys Gln

180

185

190

Ser Glu Phe Glu Phe Lys Asn Val Lys Gly Thr Leu Ile Gly Phe Tyr

195

200

205

Thr Pro Asn Tyr Ala Ala Ala Leu Asn Val Pro Gly Phe His Leu His

210

215

220

Phe Ile Thr Glu Asp Lys Thr Ser Gly Gly His Val Leu Asn Leu Gln

225

230

235

240

Phe Asp Asn Ala Asn Leu Glu Ile Ser Pro Ile His Glu Phe Asp Val

245

250

255

Gln Leu Pro His Thr Asp Asp Phe Ala His Ser Asp Leu Thr Gln Val

260

265

270

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

275 280 285

<210> 43

<211> 259

<212> PRT

<213> Enterobacter aerogenes

<400> 43

Met Met Met His Ser Ser Ala Cys Asp Cys Glu Ala Ser Leu Cys Glu

1 5 10 15

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

20 25 30

Thr Ser Leu Met Ser Ala Leu Leu Ser Gly Val Tyr Val Gly Glu Thr

35 40 45

Thr Ile Ala Asp Leu Leu Ala His Gly Asp Phe Gly Leu Gly Thr Phe

50 55 60

Asn Glu Leu Asp Gly Glu Met Ile Ala Phe Ser Ser Gln Val Tyr Gln

65 70 75 80

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

85 90 95

Pro Phe Ala Val Met Thr Trp Phe Gln Pro Gln Tyr Arg Lys Thr Phe

100 105 110

Asn Gly Pro Val Ser Arg Gln Gln Ile His Asp Val Ile Asp Gln Gln

115 120 125

Ile Pro Ser Asp Asn Leu Phe Cys Val Arg Ile Asp Gly Asn Phe Arg

130 135 140

His Ala His Thr Arg Thr Val Pro Arg Gln Thr Pro Pro Tyr Arg Ala

145 150 155 160

Met Thr Asp Val Leu Asp Asp Gln Pro Val Phe Arg Phe Asn Gln Arg

165 170 175

Glu Gly Val Leu Val Gly Phe Arg Thr Pro Gln His Met Gln Gly Ile

180 185 190

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

245 250 255

Val Glu Asn

<210> 44

<211> 1231

<212> PRT

<213> Mycobacterium tuberculosis

<400> 44

Met Ala Asn Ile Ser Ser Pro Phe Gly Gln Asn Glu Trp Leu Val Glu
 1 5 10 15
 Glu Met Tyr Arg Lys Phe Arg Asp Asp Pro Ser Ser Val Asp Pro Ser
 20 25 30
 Trp His Glu Phe Leu Val Asp Tyr Ser Pro Glu Pro Thr Ser Gln Pro

35 40 45
 Ala Ala Glu Pro Thr Arg Val Thr Ser Pro Leu Val Ala Glu Arg Ala
 50 55 60
 Ala Ala Ala Ala Pro Gln Ala Pro Pro Lys Pro Ala Asp Thr Ala Ala
 65 70 75 80
 Ala Gly Asn Gly Val Val Ala Ala Leu Ala Ala Lys Thr Ala Val Pro
 85 90 95
 Pro Pro Ala Glu Gly Asp Glu Val Ala Val Leu Arg Gly Ala Ala Ala

100 105 110
 Ala Val Val Lys Asn Met Ser Ala Ser Leu Glu Val Pro Thr Ala Thr
 115 120 125
 Ser Val Arg Ala Val Pro Ala Lys Leu Leu Ile Asp Asn Arg Ile Val

130 135 140
 Ile Asn Asn Gln Leu Lys Arg Thr Arg Gly Gly Lys Ile Ser Phe Thr
 145 150 155 160
 His Leu Leu Gly Tyr Ala Leu Val Gln Ala Val Lys Lys Phe Pro Asn

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

 225 230 235 240
 Arg Asp Gly Lys Leu Thr Thr Glu Asp Phe Ala Gly Val Thr Ile Ser
 245 250 255
 Leu Thr Asn Pro Gly Thr Ile Gly Thr Val His Ser Val Pro Arg Leu
 260 265 270
 Met Pro Gly Gln Gly Ala Ile Ile Gly Val Gly Ala Met Glu Tyr Pro
 275 280 285
 Ala Glu Phe Gln Gly Ala Ser Glu Glu Arg Ile Ala Glu Leu Gly Ile

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

 355 360 365
 Asn Ala Arg Val Met Asn Leu Ile Ala Ala Tyr Arg Asn Arg Gly His
 370 375 380

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

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

 675 680 685
 Ala Pro Glu Tyr Ser Arg Ser Ser Glu Tyr Cys Thr Asp Val Ala Lys
 690 695 700
 Met Ile Gly Ala Pro Ile Phe His Val Asn Gly Asp Asp Pro Glu Ala
 705 710 715 720
 Cys Val Trp Val Ala Arg Leu Ala Val Asp Phe Arg Gln Arg Phe Lys
 725 730 735
 Lys Asp Val Val Ile Asp Met Leu Cys Tyr Arg Arg Arg Gly His Asn

 740 745 750
 Glu Gly Asp Asp Pro Ser Met Thr Asn Pro Tyr Val Tyr Asp Val Val
 755 760 765
 Asp Thr Lys Arg Gly Ala Arg Lys Ser Tyr Thr Glu Ala Leu Ile Gly
 770 775 780
 Arg Gly Asp Ile Ser Met Lys Glu Ala Glu Asp Ala Leu Arg Asp Tyr
 785 790 795 800
 Gln Gly Gln Leu Glu Arg Val Phe Asn Glu Val Arg Glu Leu Glu Lys

 805 810 815
 His Gly Val Gln Pro Ser Glu Ser Val Glu Ser Asp Gln Met Ile Pro
 820 825 830
 Ala Gly Leu Ala Thr Ala Val Asp Lys Ser Leu Leu Ala Arg Ile Gly
 835 840 845
 Asp Ala Phe Leu Ala Leu Pro Asn Gly Phe Thr Ala His Pro Arg Val
 850 855 860
 Gln Pro Val Leu Glu Lys Arg Arg Glu Met Ala Tyr Glu Gly Lys Ile

 865 870 875 880

Asp Trp Ala Phe Gly Glu Leu Leu Ala Leu Gly Ser Leu Val Ala Glu
 885 890 895
 Gly Lys Leu Val Arg Leu Ser Gly Gln Asp Ser Arg Arg Gly Thr Phe
 900 905 910
 Ser Gln Arg His Ser Val Leu Ile Asp Arg His Thr Gly Glu Glu Phe
 915 920 925
 Thr Pro Leu Gln Leu Leu Ala Thr Asn Ser Asp Gly Ser Pro Thr Gly
 930 935 940
 Gly Lys Phe Leu Val Tyr Asp Ser Pro Leu Ser Glu Tyr Ala Ala Val
 945 950 955 960
 Gly Phe Glu Tyr Gly Tyr Thr Val Gly Asn Pro Asp Ala Val Val Leu
 965 970 975
 Trp Glu Ala Gln Phe Gly Asp Phe Val Asn Gly Ala Gln Ser Ile Ile
 980 985 990
 Asp Glu Phe Ile Ser Ser Gly Glu Ala Lys Trp Gly Gln Leu Ser Asn
 995 1000 1005
 Val Val Leu Leu Leu Pro His Gly His Glu Gly Gln Gly Pro Asp
 1010 1015 1020
 His Thr Ser Ala Arg Ile Glu Arg Phe Leu Gln Leu Trp Ala Glu
 1025 1030 1035
 Gly Ser Met Thr Ile Ala Met Pro Ser Thr Pro Ser Asn Tyr Phe
 1040 1045 1050
 His Leu Leu Arg Arg His Ala Leu Asp Gly Ile Gln Arg Pro Leu
 1055 1060 1065
 Ile Val Phe Thr Pro Lys Ser Met Leu Arg His Lys Ala Ala Val
 1070 1075 1080
 Ser Glu Ile Lys Asp Phe Thr Glu Ile Lys Phe Arg Ser Val Leu
 1085 1090 1095
 Glu Glu Pro Thr Tyr Glu Asp Gly Ile Gly Asp Arg Asn Lys Val
 1100 1105 1110
 Ser Arg Ile Leu Leu Thr Ser Gly Lys Leu Tyr Tyr Glu Leu Ala

1115 1120 1125
Ala Arg Lys Ala Lys Asp Asn Arg Asn Asp Leu Ala Ile Val Arg
1130 1135 1140
Leu Glu Gln Leu Ala Pro Leu Pro Arg Arg Arg Leu Arg Glu Thr
1145 1150 1155
Leu Asp Arg Tyr Glu Asn Val Lys Glu Phe Phe Trp Val Gln Glu
1160 1165 1170
Glu Pro Ala Asn Gln Gly Ala Trp Pro Arg Phe Gly Leu Glu Leu

1175 1180 1185
Pro Glu Leu Leu Pro Asp Lys Leu Ala Gly Ile Lys Arg Ile Ser
1190 1195 1200
Arg Arg Ala Met Ser Ala Pro Ser Ser Gly Ser Ser Lys Val His
1205 1210 1215
Ala Val Glu Gln Gln Glu Ile Leu Asp Glu Ala Phe Gly
1220 1225 1230

<210> 45

<211> 985

<212> PRT

<213> Bradyrhizobium japonicum

<400> 45

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

Trp Glu Arg Ala Asn Trp Pro Leu Thr Pro Gln Asp Asp Leu Thr Ser
65 70 75 80
Ala Leu Asp Gly Asn Trp Ala Glu Val Glu Lys Ala Val Gly Gly Lys
85 90 95

Ile Ala Ala Lys Ala Gln Ala Lys Gly Ala Asp Ile Ser Ser Ala Asp
100 105 110

Leu Leu Gln Ala Thr Arg Asp Ser Val Arg Ala Leu Met Leu Ile Arg
115 120 125

Ser Tyr Arg Met Arg Gly His Phe His Ala Lys Leu Asp Pro Leu Gly
130 135 140

Ile Glu Ala Pro Arg Asn Arg Glu Glu Leu Asp Pro Arg Thr Tyr Gly
145 150 155 160

Phe Ser Glu Ala Asp Phe Asp Arg Lys Ile Phe Leu Asp His Val Leu
165 170 175

Gly Leu Glu Tyr Gly Thr Leu Arg Glu Ile Thr Ala Ile Cys Glu Arg
180 185 190

Thr Tyr Cys Gln Thr Leu Gly Val Glu Phe Met His Ile Ser Asn Ala
195 200 205

Ala Gln Lys Ala Trp Ile Gln Glu Arg Ile Glu Gly Pro Asp Lys Glu
210 215 220

Ile Ser Phe Thr Arg Glu Gly Arg Arg Ala Ile Leu Thr Lys Leu Val
225 230 235 240

Glu Ala Glu Gly Phe Glu Lys Phe Cys Asp Thr Lys Phe Thr Gly Thr
245 250 255

Lys Arg Phe Gly Leu Asp Gly Ala Glu Ser Leu Ile Pro Ala Leu Glu
260 265 270

Gln Ile Ile Lys Arg Gly Gly Asn Leu Gly Val Lys Glu Ile Val Leu
275 280 285

Gly Met Pro His Arg Gly Arg Leu Asn Val Leu Thr Gln Val Met Gly
290 295 300

Lys Ala His Arg Ala Leu Phe His Glu Phe Lys Gly Gly Ser Ala Asn
305 310 315 320

Pro Asp Ala Val Glu Gly Ser Gly Asp Val Lys Tyr His Leu Gly Ala
325 330 335

Ser Ser Asp Arg Glu Phe Asp Gly Asn Arg Ile His Leu Ser Leu Thr

340 345 350
 Ala Asn Pro Ser His Leu Glu Ile Val Asp Pro Val Val Leu Gly Lys
 355 360 365
 Val Arg Ala Lys Gln Asp Gln His Gly Asp Pro Pro Asp Met Arg Ile
 370 375 380

 Ser Val Met Pro Leu Leu Met His Gly Asp Ala Ala Phe Ala Gly Gln
 385 390 395 400
 Gly Val Val Ala Glu Cys Phe Gly Leu Ser Asp Leu Lys Gly Tyr Arg
 405 410 415
 Thr Gly Gly Ser Val His Phe Ile Val Asn Asn Gln Ile Gly Phe Thr
 420 425 430
 Thr Tyr Pro Arg Tyr Ser Arg Ser Ser Pro Tyr Pro Ser Asp Val Ala
 435 440 445

 Lys Met Ile Asp Ala Pro Ile Phe His Val Asn Gly Asp Asp Pro Glu
 450 455 460
 Ala Val Val Phe Ala Ala Lys Val Ala Thr Glu Phe Arg Gln Lys Phe
 465 470 475 480
 His Lys Pro Val Val Ile Asp Met Phe Cys Tyr Arg Arg His Gly His
 485 490 495
 Asn Glu Gly Asp Glu Pro Ala Phe Thr Gln Pro Val Met Tyr Lys Lys
 500 505 510

 Ile Ala Ala His Pro Ser Thr Leu Glu Leu Tyr Ala Arg Arg Leu Ile
 515 520 525
 Ser Glu Gly Val Met Thr Glu Gly Glu Val Asp Lys Ala Lys Ala Asp
 530 535 540
 Trp Arg Ala Arg Leu Asp Ala Glu Phe Glu Ala Gly Thr Ser Tyr Lys
 545 550 555 560
 Pro Asn Lys Ala Asp Trp Leu Asp Gly Lys Trp Ala Gly Phe Lys Ile
 565 570 575

 Ala Asp Gln Glu Glu Asp Ala Arg Arg Gly Val Thr Gly Val Asp Ile
 580 585 590

Thr Ala Leu Lys Asp Ile Gly Arg Lys Ile Thr Lys Val Pro Asp Gly
 595 600 605
 Phe Arg Val His Arg Thr Ile Gln Arg Phe Leu Glu Asn Arg Ser Lys
 610 615 620
 Ala Ile Asp Ser Gly Ala Gly Ile Asp Trp Ala Thr Gly Glu Ala Leu
 625 630 635 640

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

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

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

 Glu Glu Leu Ala Lys Gly Thr Thr Phe His Arg Ile Leu Tyr Asp Asp

835 840 845
 Ala Gln Met Leu Pro Thr Asp Ala Ile Lys Leu Val Pro Asp Glu Lys
 850 855 860
 Ile Arg Arg Ile Val Leu Cys Ser Gly Lys Val Tyr Tyr Asp Leu Tyr
 865 870 875 880
 Glu Glu Arg Glu Lys Arg Gly Ile Asp Asp Ile Tyr Leu Met Arg Val
 885 890 895

 Glu Gln Leu Tyr Pro Val Pro Leu Lys Ala Leu Val Ala Glu Leu Ser
 900 905 910
 Arg Phe Lys Lys Ala Glu Val Val Trp Cys Gln Glu Glu Pro Arg Asn
 915 920 925
 Met Gly Ala Trp His Phe Ile Glu Pro Tyr Leu Glu Trp Val Leu Asn
 930 935 940
 Gln Val Asn Gly Val Ser Arg Arg Pro Arg Tyr Val Gly Arg Ala Ala
 945 950 955 960

 Ser Ala Ala Thr Ala Thr Gly Leu Met Ser Lys His Gln Ala Gln Leu
 965 970 975
 Lys Ala Phe Leu Asp Glu Ala Leu Ser
 980 985

 <210> 46
 <211> 995
 <212> PRT
 <213> Rhizobium loti
 <400> 46
 Met Ala Arg Gln Asp Gln Thr Asn Asp Gln Phe Ser Leu Thr Ser Phe
 1 5 10 15
 Leu Tyr Gly Gly Asn Ala Asp Tyr Ile Asp Ala Leu Tyr Ala Ala Tyr
 20 25 30

 Glu Asp Asp Pro Ala Ser Val Asn Pro Glu Trp Gln Glu Phe Phe Ala
 35 40 45
 Gly Leu Lys Asp Asp Ala Gly Asp Val Arg Arg Asn Ala Lys Gly Ala
 50 55 60

Ser Trp Ala Lys Pro Ser Trp Pro Leu Gln Ala Asn Gly Glu Leu Val
65 70 75 80

Ser Ala Leu Asp Gly Asn Trp Gly Ile Val Glu Lys His Leu Glu Lys
85 90 95

Lys Val Lys Asp Lys Ala Val Thr Asn Gly Val Val Leu Ser Asp Ala
100 105 110

Asp Val His Gln Ala Thr Arg Asp Ser Val Arg Ala Ile Met Met Ile
115 120 125

Arg Ala Tyr Arg Met Arg Gly His Leu His Ala Asn Leu Asp Pro Leu
130 135 140

Gly Ile Ala Lys Pro Leu Glu Asp Tyr Asn Glu Leu Ser Pro Glu Asn
145 150 155 160

Tyr Gly Phe Thr Ala Ala Asp Tyr Asp Arg Pro Ile Phe Leu Asp Asn
165 170 175

Val Leu Gly Leu Glu Phe Gly Thr Ile Arg Gln Met Leu Glu Ile Leu
180 185 190

Thr Arg Thr Tyr Cys Ser Thr Leu Gly Val Glu Phe Met His Ile Ser
195 200 205

Asp Pro Glu Glu Lys Ala Trp Ile Gln Ala Arg Ile Glu Gly Ala Asp
210 215 220

Lys Glu Ile Ser Phe Thr Asn Thr Gly Lys Lys Ala Ile Leu Gln Lys
225 230 235 240

Leu Val Glu Ala Glu Gly Phe Glu Gln Phe Ile Asp Val Lys Tyr Lys
245 250 255

Gly Thr Lys Arg Phe Gly Leu Asp Gly Gly Glu Ala Leu Ile Pro Ala
260 265 270

Leu Glu Gln Ile Val Lys Arg Gly Gly Gln Leu Gly Met Lys Glu Ile
275 280 285

Val Leu Gly Met Ala His Arg Gly Arg Leu Asn Val Leu Ser Gln Val
290 295 300

Met Ala Lys Pro His Arg Ala Ile Phe His Glu Phe Lys Gly Gly Ser

305 310 315 320
 Ala Ala Pro Asp Glu Val Glu Gly Ser Gly Asp Val Lys Tyr His Leu
 325 330 335
 Gly Ala Ser Ser Asp Arg Glu Phe Asp Gly Asn Lys Val His Leu Ser
 340 345 350

 Leu Thr Ala Asn Pro Ser His Leu Glu Ile Val Asp Pro Val Val Met
 355 360 365
 Gly Lys Ala Arg Ala Lys Gln Asp Tyr Leu Phe Gly Arg Gly Arg Glu
 370 375 380
 Glu Ile Val Pro Leu Glu Glu Arg Ala Lys Val Leu Pro Leu Leu Leu
 385 390 395 400
 His Gly Asp Ala Ala Phe Ala Gly Gln Gly Val Ile Ala Glu Ile Leu
 405 410 415

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

 Val Ala Ile Glu Phe Arg Met Lys Phe His Lys Pro Val Val Val Asp
 485 490 495
 Met Phe Cys Tyr Arg Arg Phe Gly His Asn Glu Gly Asp Glu Pro Ala
 500 505 510
 Phe Thr Gln Pro Ile Met Tyr Arg Asn Ile Arg Thr His Lys Thr Thr
 515 520 525
 Val Gln Ile Tyr Ala Asp Arg Leu Ile Ala Glu Gly His Ile Thr Gln
 530 535 540

 Ala Glu Leu Asp Gln Met Lys Ala Asp Trp Arg Ala His Leu Glu Ser
 545 550 555 560

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

Lys Lys Leu Thr Glu Val Pro Lys Gly Phe Glu Ala His Lys Thr Ile
610 615 620
Ile Arg Phe Leu Glu Asn Arg Arg Glu Ala Ile Glu Ser Gly Glu Gly
625 630 635 640
Ile Asp Trp Ser Thr Ala Glu Ala Leu Ala Phe Gly Ala Ile Leu Leu
645 650 655
Asp Gly Asn Pro Ile Arg Leu Ser Gly Gln Asp Ser Glu Arg Gly Thr
660 665 670

Phe Ser Gln Arg His Ser Val Leu Tyr Asp Gln Arg Asp Glu Thr Arg
675 680 685
Tyr Ile Pro Leu Asn Asn Leu Ser Ala Ala Gln Ala Gly Tyr Glu Val
690 695 700
Ile Asn Ser Met Leu Ser Glu Glu Ala Val Leu Gly Phe Glu Tyr Gly
705 710 715 720
Tyr Ser Leu Ala Glu Pro Lys Ala Leu Thr Leu Trp Glu Ala Gln Phe
725 730 735

Gly Asp Phe Ala Asn Gly Ala Gln Val Val Phe Asp Gln Phe Ile Ser
740 745 750
Ser Gly Glu Arg Lys Trp Leu Arg Met Ser Gly Leu Val Cys Leu Leu
755 760 765
Pro His Gly Tyr Glu Gly Gln Gly Pro Glu His Ser Ser Ala Arg Leu
770 775 780
Glu Arg Phe Leu Gln Leu Cys Ala Glu Asp Asn Met Gln Val Ala Asn
785 790 795 800

Cys Thr Thr Pro Ala Asn Tyr Phe His Ile Leu Arg Arg Gln Leu Lys

805 810 815
 Arg Asp Phe Arg Lys Pro Leu Ile Leu Met Thr Pro Lys Ser Leu Leu
 820 825 830
 Arg His Lys Arg Ala Val Ser Thr Leu Pro Glu Ile Ser Gly Glu Ser
 835 840 845
 Ser Phe His Arg Leu Leu Trp Asp Asp Ala Gln Leu Leu Pro Asn Gln
 850 855 860

Pro Ile Lys Leu Thr Lys Asp Ser Lys Ile Arg Arg Val Val Leu Cys
 865 870 875 880
 Ser Gly Lys Val Tyr Tyr Asp Leu Tyr Glu Glu Arg Glu Lys Arg Gly
 885 890 895
 Ile Asn Asp Ile Tyr Leu Leu Arg Val Glu Gln Leu Tyr Pro Phe Pro
 900 905 910
 Ala Lys Ala Leu Ile Thr Glu Leu Ser Arg Phe Arg Asn Ala Glu Met
 915 920 925

Val Trp Cys Gln Glu Glu Pro Lys Asn Met Gly Ala Trp Ser Phe Ile
 930 935 940
 Asp Pro Tyr Leu Glu Trp Val Leu Ala His Ile Asp Ala Lys His Gln
 945 950 955 960
 Arg Val Arg Tyr Thr Gly Arg Pro Ala Ala Ala Ser Pro Ala Thr Gly
 965 970 975
 Leu Met Ser Lys His Leu Ala Gln Leu Ala Ala Leu Leu Glu Asp Ala
 980 985 990

Leu Gly Glu
 995

<210> 47

<211> 547

<212> PRT

<213> Lactococcus lactis

<400> 47

Met Tyr Thr Val Gly Asp Tyr Leu Leu Asp Arg Leu His Glu Leu Gly
 1 5 10 15

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

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

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

 370 375 380
 Ser Thr Ile Phe Leu Lys Ser Asn Ser Arg Phe Ile Gly Gln Pro Leu
 385 390 395 400
 Trp Gly Ser Ile Gly Tyr Thr Phe Pro Ala Ala Leu Gly Ser Gln Ile
 405 410 415
 Ala Asp Lys Glu Ser Arg His Leu Leu Phe Ile Gly Asp Gly Ser Leu
 420 425 430
 Gln Leu Thr Val Gln Glu Leu Gly Leu Ser Ile Arg Glu Lys Leu Asn

 435 440 445
 Pro Ile Cys Phe Ile Ile Asn Asn Asp Gly Tyr Thr Val Glu Arg Glu
 450 455 460
 Ile His Gly Pro Thr Gln Ser Tyr Asn Asp Ile Pro Met Trp Asn Tyr
 465 470 475 480
 Ser Lys Leu Pro Glu Thr Phe Gly Ala Thr Glu Asp Arg Val Val Ser
 485 490 495
 Lys Ile Val Arg Thr Glu Asn Glu Phe Val Ser Val Met Lys Glu Ala

 500 505 510

Gln Ala Asp Val Asn Arg Met Tyr Trp Ile Glu Leu Val Leu Glu Lys

515

520

525

Glu Asp Ala Pro Lys Leu Leu Lys Lys Met Gly Lys Leu Phe Ala Glu

530

535

540

Gln Asn Lys

545

<210> 48

<211> 1728

<212> DNA

<213> Issatchenkia orientalis

<400> 48

atgactgaca aaatctccct aggtacttat ctgtttgaaa agttaaagga agcaggctct 60

tattccatct ttggtgttcc tgggtgattc aatttggcat tgttggacca cgtcaaggaa 120

gttgaaggca ttagatgggt cggtaacgct aacgagttga atgccggcta cgaagctgat 180

ggttatgcaa gaatcaatgg atttgcattc ctaatcacca cctttggtgt cggatgaattg 240

tctgccgtca atgccattgc aggttcttat gctgaacacg tccattgat ccatattgtt 300

ggtatgcctt ccttgtctgc tatgaagaac aacttgttgt tacaccatac cttgggtgac 360

acaagattcg acaacttcac cgaatgtca aagaaaatca gtgcaaaggt tgaaattgtt 420

tacgatttgg aatcagctcc aaaattaatt aataacttga ttgaaaccgc ttatcacaca 480

aagagaccag tctacttggg acttctctcc aactttgctg atgaattggt tccagcggca 540

ttagttaagg aaaacaagtt acatttagaa gaacctctaa acaaccccggt tgctgaagaa 600

gaattcattc ataacgttgt tgaatggtc aagaaggcag aaaaaccaat cattctcgtt 660

gacgcttggt ctgcaagaca taacatttct aaggaagtga gagagttggc taaattgact 720

aaattccctg tcttcaccac cccaatgggt aaatctactg ttgatgaaga tgatgaagaa 780

ttctttggct tatacttggg ttctctatct gctccagatg ttaaggacat tgttggccca 840

accgattgta tcttatcctt aggtgggttta ccttctgatt tcaacaccgg ttccttctca 900

tatggttaca ccactaagaa tgtcgttgaa ttccattcca actactgtaa attcaaatct 960

gcaacttatg aaaacttgat gatgaagggc gcagtccaaa gattgatcag cgaattgaag 1020

aatattaagt attccaatgt ctcaacttta tctccaccaa aatctaaatt tgcttacgaa 1080

tctgcaaagg ttgtccaga aggtatcatc actcaagatt acctgtggaa gagattatct 1140

tacttcttaa agccaagaga tatcattgtc actgaaactg gtacttcctc ctttgggtgc 1200

ttggctaccc acttaccaag agattcaaag tctatctccc aagtcttatg gggttccatt 1260

ggtttctcct taccagctgc agttggtgct gcatttgctg ctgaagatgc acacaaacaa 1320

actggcgaac aagaaagaag aactgttttg tttattggtg atggttcttt acaattgact 1380

gtccaatcaa tctcagatgc tgcaagatgg aacatcaagc catacatctt catcttaaac 1440

aacagagggtt acactatcga aaagttgatc cacggctgctc atgaggacta caaccaaatt 1500

caaccatggg atcaccaatt gttattgaag ctctttgctg acaagacca atatgaaaac 1560

catgttggtta aatccgctaa ggacttggac gctttgatga aggatgaagc attcaacaag 1620

gaagataaga ttagagtcac tgaattatc ttggatgaat tcgatgctcc agaaatcttg 1680

gttgctcaag ctaaattatc tgatgaaatc aactctaaag ccgcttaa 1728

<210> 49

<211> 575

<212> PRT

<213> Issatchenkia orientalis

<400> 49

Met Thr Asp Lys Ile Ser Leu Gly Thr Tyr Leu Phe Glu Lys Leu Lys

1 5 10 15

Glu Ala Gly Ser Tyr Ser Ile Phe Gly Val Pro Gly Asp Phe Asn Leu

20 25 30

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

35 40 45

Asn Ala Asn Glu Leu Asn Ala Gly Tyr Glu Ala Asp Gly Tyr Ala Arg

50 55 60

Ile Asn Gly Phe Ala Ser Leu Ile Thr Thr Phe Gly Val Gly Glu Leu

65 70 75 80

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

85 90 95

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

100 105 110

Leu Leu His His Thr Leu Gly Asp Thr Arg Phe Asp Asn Phe Thr Glu

115 120 125

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

130 135 140

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

385 390 395 400
 Leu Ala Thr His Leu Pro Arg Asp Ser Lys Ser Ile Ser Gln Val Leu
 405 410 415
 Trp Gly Ser Ile Gly Phe Ser Leu Pro Ala Ala Val Gly Ala Ala Phe
 420 425 430
 Ala Ala Glu Asp Ala His Lys Gln Thr Gly Glu Gln Glu Arg Arg Thr

 435 440 445
 Val Leu Phe Ile Gly Asp Gly Ser Leu Gln Leu Thr Val Gln Ser Ile
 450 455 460
 Ser Asp Ala Ala Arg Trp Asn Ile Lys Pro Tyr Ile Phe Ile Leu Asn
 465 470 475 480
 Asn Arg Gly Tyr Thr Ile Glu Lys Leu Ile His Gly Arg His Glu Asp
 485 490 495
 Tyr Asn Gln Ile Gln Pro Trp Asp His Gln Leu Leu Leu Lys Leu Phe

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

<210> 50

<211> 563

<212> PRT

<213> *Saccharomyces cerevisiae*

<400> 50

Met Ser Glu Ile Thr Leu Gly Lys Tyr Leu Phe Glu Arg Leu Lys Gln
 1 5 10 15
 Val Asn Val Asn Thr Val Phe Gly Leu Pro Gly Asp Phe Asn Leu Ser
 20 25 30

Leu Leu Asp Lys Ile Tyr Glu Val Glu Gly Met Arg Trp Ala Gly Asn
 35 40 45

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

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

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

 Pro Ala Phe Val Thr Pro Met Gly Lys Gly Ser Ile Asp Glu Gln His
 245 250 255
 Pro Arg Tyr Gly Gly Val Tyr Val Gly Thr Leu Ser Lys Pro Glu Val
 260 265 270
 Lys Glu Ala Val Glu Ser Ala Asp Leu Ile Leu Ser Val Gly Ala Leu

| | | |
|---|-----|-----|
| 275 | 280 | 285 |
| Leu Ser Asp Phe Asn Thr Gly Ser Phe Ser Tyr Ser Tyr Lys Thr Lys | | |
| 290 | 295 | 300 |
| | | |
| Asn Ile Val Glu Phe His Ser Asp His Met Lys Ile Arg Asn Ala Thr | | |
| 305 | 310 | 315 |
| Phe Pro Gly Val Gln Met Lys Phe Val Leu Gln Lys Leu Leu Thr Thr | | |
| 325 | 330 | 335 |
| Ile Ala Asp Ala Ala Lys Gly Tyr Lys Pro Val Ala Val Pro Ala Arg | | |
| 340 | 345 | 350 |
| Thr Pro Ala Asn Ala Ala Val Pro Ala Ser Thr Pro Leu Lys Gln Glu | | |
| 355 | 360 | 365 |
| | | |
| Trp Met Trp Asn Gln Leu Gly Asn Phe Leu Gln Glu Gly Asp Val Val | | |
| 370 | 375 | 380 |
| Ile Ala Glu Thr Gly Thr Ser Ala Phe Gly Ile Asn Gln Thr Thr Phe | | |
| 385 | 390 | 395 |
| Pro Asn Asn Thr Tyr Gly Ile Ser Gln Val Leu Trp Gly Ser Ile Gly | | |
| 405 | 410 | 415 |
| Phe Thr Thr Gly Ala Thr Leu Gly Ala Ala Phe Ala Ala Glu Glu Ile | | |
| 420 | 425 | 430 |
| | | |
| Asp Pro Lys Lys Arg Val Ile Leu Phe Ile Gly Asp Gly Ser Leu Gln | | |
| 435 | 440 | 445 |
| Leu Thr Val Gln Glu Ile Ser Thr Met Ile Arg Trp Gly Leu Lys Pro | | |
| 450 | 455 | 460 |
| Tyr Leu Phe Val Leu Asn Asn Asp Gly Tyr Thr Ile Glu Lys Leu Ile | | |
| 465 | 470 | 475 |
| His Gly Pro Lys Ala Gln Tyr Asn Glu Ile Gln Gly Trp Asp His Leu | | |
| 485 | 490 | 495 |
| | | |
| Ser Leu Leu Pro Thr Phe Gly Ala Lys Asp Tyr Glu Thr His Arg Val | | |
| 500 | 505 | 510 |
| Ala Thr Thr Gly Glu Trp Asp Lys Leu Thr Gln Asp Lys Ser Phe Asn | | |
| 515 | 520 | 525 |

Asp Asn Ser Lys Ile Arg Met Ile Glu Ile Met Leu Pro Val Phe Asp
 530 535 540
 Ala Pro Gln Asn Leu Val Glu Gln Ala Lys Leu Thr Ala Ala Thr Asn
 545 550 555 560

Ala Lys Gln

<210> 51

<211> 563

<212> PRT

<213> Kluyveromyces lactis

<400> 51

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

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

115 120 125
 Ser Ser Asn Ile Ser Glu Thr Thr Ala Met Ile Thr Asp Ile Asn Thr
 130 135 140
 Ala Pro Ala Glu Ile Asp Arg Cys Ile Arg Thr Thr Tyr Val Ser Gln
 145 150 155 160
 Arg Pro Val Tyr Leu Gly Leu Pro Ala Asn Leu Val Asp Leu Thr Val

165 170 175
 Pro Ala Ser Leu Leu Asp Thr Pro Ile Asp Leu Ser Leu Lys Pro Asn

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

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

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

 370 375 380
 Ile Thr Glu Thr Gly Thr Ser Ala Phe Gly Ile Asn Gln Thr His Phe
 385 390 395 400
 Pro Asn Asn Thr Tyr Gly Ile Ser Gln Val Leu Trp Gly Ser Ile Gly
 405 410 415

Phe Thr Thr Gly Ala Thr Leu Gly Ala Ala Phe Ala Ala Glu Glu Ile
420 425 430
Asp Pro Lys Lys Arg Val Ile Leu Phe Ile Gly Asp Gly Ser Leu Gln
435 440 445
Leu Thr Val Gln Glu Ile Ser Thr Met Ile Arg Trp Gly Leu Lys Pro
450 455 460
Tyr Leu Phe Val Leu Asn Asn Asp Gly Tyr Thr Ile Glu Arg Leu Ile
465 470 475 480
His Gly Glu Thr Ala Gln Tyr Asn Cys Ile Gln Asn Trp Gln His Leu
485 490 495
Glu Leu Leu Pro Thr Phe Gly Ala Lys Asp Tyr Glu Ala Val Arg Val
500 505 510
Ser Thr Thr Gly Glu Trp Asn Lys Leu Thr Thr Asp Glu Lys Phe Gln
515 520 525
Asp Asn Thr Arg Ile Arg Leu Ile Glu Val Met Leu Pro Thr Met Asp
530 535 540
Ala Pro Ser Asn Leu Val Lys Gln Ala Gln Leu Thr Ala Ala Thr Asn
545 550 555 560
Ala Lys Asn

<210> 52

<211> 568

<212> PRT

<213> *Zymomonas mobilis*

<400> 52

Met Ser Tyr Thr Val Gly Thr Tyr Leu Ala Glu Arg Leu Val Gln Ile
1 5 10 15
Gly Leu Lys His His Phe Ala Val Ala Gly Asp Tyr Asn Leu Val Leu
20 25 30
Leu Asp Asn Leu Leu Leu Asn Lys Asn Met Glu Gln Val Tyr Cys Cys
35 40 45

Asn Glu Leu Asn Cys Gly Phe Ser Ala Glu Gly Tyr Ala Arg Ala Lys

50 55 60

Gly Ala Ala Ala Ala Val Val Thr Tyr Ser Val Gly Ala Leu Ser Ala

65 70 75 80

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

85 90 95

Ile Ser Gly Ala Pro Asn Asn Asn Asp His Ala Ala Gly His Val Leu

100 105 110

His His Ala Leu Gly Lys Thr Asp Tyr His Tyr Gln Leu Glu Met Ala

115 120 125

Lys Asn Ile Thr Ala Ala Ala Glu Ala Ile Tyr Thr Pro Glu Glu Ala

130 135 140

Pro Ala Lys Ile Asp His Val Ile Lys Thr Ala Leu Arg Glu Lys Lys

145 150 155 160

Pro Val Tyr Leu Glu Ile Ala Cys Asn Ile Ala Ser Met Pro Cys Ala

165 170 175

Ala Pro Gly Pro Ala Ser Ala Leu Phe Asn Asp Glu Ala Ser Asp Glu

180 185 190

Ala Ser Leu Asn Ala Ala Val Glu Glu Thr Leu Lys Phe Ile Ala Asn

195 200 205

Arg Asp Lys Val Ala Val Leu Val Gly Ser Lys Leu Arg Ala Ala Gly

210 215 220

Ala Glu Glu Ala Ala Val Lys Phe Ala Asp Ala Leu Gly Gly Ala Val

225 230 235 240

Ala Thr Met Ala Ala Ala Lys Ser Phe Phe Pro Glu Glu Asn Pro His

245 250 255

Tyr Ile Gly Thr Ser Trp Gly Glu Val Ser Tyr Pro Gly Val Glu Lys

260 265 270

Thr Met Lys Glu Ala Asp Ala Val Ile Ala Leu Ala Pro Val Phe Asn

275 280 285

Asp Tyr Ser Thr Thr Gly Trp Thr Asp Ile Pro Asp Pro Lys Lys Leu

290 295 300
 Val Leu Ala Glu Pro Arg Ser Val Val Val Asn Gly Ile Arg Phe Pro

 305 310 315 320
 Ser Val His Leu Lys Asp Tyr Leu Thr Arg Leu Ala Gln Lys Val Ser

 325 330 335
 Lys Lys Thr Gly Ala Leu Asp Phe Phe Lys Ser Leu Asn Ala Gly Glu

 340 345 350
 Leu Lys Lys Ala Ala Pro Ala Asp Pro Ser Ala Pro Leu Val Asn Ala

 355 360 365
 Glu Ile Ala Arg Gln Val Glu Ala Leu Leu Thr Pro Asn Thr Thr Val

 370 375 380
 Ile Ala Glu Thr Gly Asp Ser Trp Phe Asn Ala Gln Arg Met Lys Leu
 385 390 395 400
 Pro Asn Gly Ala Arg Val Glu Tyr Glu Met Gln Trp Gly His Ile Gly

 405 410 415
 Trp Ser Val Pro Ala Ala Phe Gly Tyr Ala Val Gly Ala Pro Glu Arg

 420 425 430
 Arg Asn Ile Leu Met Val Gly Asp Gly Ser Phe Gln Leu Thr Ala Gln

 435 440 445
 Glu Val Ala Gln Met Val Arg Leu Lys Leu Pro Val Ile Ile Phe Leu

 450 455 460
 Ile Asn Asn Tyr Gly Tyr Thr Ile Glu Val Met Ile His Asp Gly Pro
 465 470 475 480
 Tyr Asn Asn Ile Lys Asn Trp Asp Tyr Ala Gly Leu Met Glu Val Phe

 485 490 495
 Asn Gly Asn Gly Gly Tyr Asp Ser Gly Ala Gly Lys Gly Leu Lys Ala

 500 505 510
 Lys Thr Gly Gly Glu Leu Ala Glu Ala Ile Lys Val Ala Leu Ala Asn

 515 520 525
 Thr Asp Gly Pro Thr Leu Ile Glu Cys Phe Ile Gly Arg Glu Asp Cys

 530 535 540

Thr Glu Glu Leu Val Lys Trp Gly Lys Arg Val Ala Ala Ala Asn Ser
 545 550 555 560
 Arg Lys Pro Val Asn Lys Leu Leu
 565

<210> 53

<211> 557

<212> PRT

<213> Acetobacter pasteurianus

<400> 53

Met Thr Tyr Thr Val Gly Met Tyr Leu Ala Glu Arg Leu Val Gln Ile
 1 5 10 15
 Gly Leu Lys His His Phe Ala Val Gly Gly Asp Tyr Asn Leu Val Leu
 20 25 30
 Leu Asp Gln Leu Leu Leu Asn Lys Asp Met Lys Gln Ile Tyr Cys Cys
 35 40 45
 Asn Glu Leu Asn Cys Gly Phe Ser Ala Glu Gly Tyr Ala Arg Ser Asn

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

115 120 125
 Arg Gln Val Thr Cys Ala Ala Glu Ser Ile Thr Asp Ala His Ser Ala
 130 135 140
 Pro Ala Lys Ile Asp His Val Ile Arg Thr Ala Leu Arg Glu Arg Lys
 145 150 155 160
 Pro Ala Tyr Leu Asp Ile Ala Cys Asn Ile Ala Ser Glu Pro Cys Val
 165 170 175

Arg Pro Gly Pro Val Ser Ser Leu Leu Ser Glu Pro Glu Ile Asp His

180 185 190

Thr Ser Leu Lys Ala Ala Val Asp Ala Thr Val Ala Leu Leu Lys Asn

195 200 205

Arg Pro Ala Pro Val Met Leu Leu Gly Ser Lys Leu Arg Ala Ala Asn

210 215 220

Ala Leu Ala Ala Thr Glu Thr Leu Ala Asp Lys Leu Gln Cys Ala Val

225 230 235 240

Thr Ile Met Ala Ala Ala Lys Gly Phe Phe Pro Glu Asp His Ala Gly

245 250 255

Phe Arg Gly Leu Tyr Trp Gly Glu Val Ser Asn Pro Gly Val Gln Glu

260 265 270

Leu Val Glu Thr Ser Asp Ala Leu Leu Cys Ile Ala Pro Val Phe Asn

275 280 285

Asp Tyr Ser Thr Val Gly Trp Ser Gly Met Pro Lys Gly Pro Asn Val

290 295 300

Ile Leu Ala Glu Pro Asp Arg Val Thr Val Asp Gly Arg Ala Tyr Asp

305 310 315 320

Gly Phe Thr Leu Arg Ala Phe Leu Gln Ala Leu Ala Glu Lys Ala Pro

325 330 335

Ala Arg Pro Ala Ser Ala Gln Lys Ser Ser Val Pro Thr Cys Ser Leu

340 345 350

Thr Ala Thr Ser Asp Glu Ala Gly Leu Thr Asn Asp Glu Ile Val Arg

355 360 365

His Ile Asn Ala Leu Leu Thr Ser Asn Thr Thr Leu Val Ala Glu Thr

370 375 380

Gly Asp Ser Trp Phe Asn Ala Met Arg Met Thr Leu Ala Gly Ala Arg

385 390 395 400

Val Glu Leu Glu Met Gln Trp Gly His Ile Gly Trp Ser Val Pro Ser

405 410 415

Ala Phe Gly Asn Ala Met Gly Ser Gln Asp Arg Gln His Val Val Met

420 425 430
 Val Gly Asp Gly Ser Phe Gln Leu Thr Ala Gln Glu Val Ala Gln Met

 435 440 445
 Val Arg Tyr Glu Leu Pro Val Ile Ile Phe Leu Ile Asn Asn Arg Gly
 450 455 460
 Tyr Val Ile Glu Ile Ala Ile His Asp Gly Pro Tyr Asn Tyr Ile Lys
 465 470 475 480
 Asn Trp Asp Tyr Ala Gly Leu Met Glu Val Phe Asn Ala Gly Glu Gly
 485 490 495
 His Gly Leu Gly Leu Lys Ala Thr Thr Pro Lys Glu Leu Thr Glu Ala

 500 505 510
 Ile Ala Arg Ala Lys Ala Asn Thr Arg Gly Pro Thr Leu Ile Glu Cys
 515 520 525
 Gln Ile Asp Arg Thr Asp Cys Thr Asp Met Leu Val Gln Trp Gly Arg
 530 535 540
 Lys Val Ala Ser Thr Asn Ala Arg Lys Thr Thr Leu Ala
 545 550 555
 <210> 54
 <211> 528
 <212> PRT
 <213> Pseudomonas putida
 <400>
 54
 Met Ala Ser Val His Gly Thr Thr Tyr Glu Leu Leu Arg Arg Gln Gly
 1 5 10 15
 Ile Asp Thr Val Phe Gly Asn Pro Gly Ser Asn Glu Leu Pro Phe Leu
 20 25 30
 Lys Asp Phe Pro Glu Asp Phe Arg Tyr Ile Leu Ala Leu Gln Glu Ala
 35 40 45
 Cys Val Val Gly Ile Ala Asp Gly Tyr Ala Gln Ala Ser Arg Lys Pro
 50 55 60

 Ala Phe Ile Asn Leu His Ser Ala Ala Gly Thr Gly Asn Ala Met Gly

| | | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|--|
| 65 | | | | | 70 | | | | 75 | | | | 80 | | | |
| Ala | Leu | Ser | Asn | Ala | Trp | Asn | Ser | His | Ser | Pro | Leu | Ile | Val | Thr | Ala | |
| | | | | 85 | | | | 90 | | | | 95 | | | | |
| Gly | Gln | Gln | Thr | Arg | Ala | Met | Ile | Gly | Val | Glu | Ala | Leu | Leu | Thr | Asn | |
| | | | | 100 | | | | 105 | | | | 110 | | | | |
| Val | Asp | Ala | Ala | Asn | Leu | Pro | Arg | Pro | Leu | Val | Lys | Trp | Ser | Tyr | Glu | |
| | | | | 115 | | | | 120 | | | | 125 | | | | |
| | | | | | | | | | | | | | | | | |
| Pro | Ala | Ser | Ala | Ala | Glu | Val | Pro | His | Ala | Met | Ser | Arg | Ala | Ile | His | |
| 130 | | | | | | | | 135 | | | | 140 | | | | |
| Met | Ala | Ser | Met | Ala | Pro | Gln | Gly | Pro | Val | Tyr | Leu | Ser | Val | Pro | Tyr | |
| 145 | | | | 150 | | | | 155 | | | | 160 | | | | |
| Asp | Asp | Trp | Asp | Lys | Asp | Ala | Asp | Pro | Gln | Ser | His | His | Leu | Phe | Asp | |
| | | | | 165 | | | | 170 | | | | 175 | | | | |
| Arg | His | Val | Ser | Ser | Ser | Val | Arg | Leu | Asn | Asp | Gln | Asp | Leu | Asp | Ile | |
| 180 | | | | | | | | 185 | | | | 190 | | | | |
| | | | | | | | | | | | | | | | | |
| Leu | Val | Lys | Ala | Leu | Asn | Ser | Ala | Ser | Asn | Pro | Ala | Ile | Val | Leu | Gly | |
| 195 | | | | | | | | 200 | | | | 205 | | | | |
| Pro | Asp | Val | Asp | Ala | Ala | Asn | Ala | Asn | Ala | Asp | Cys | Val | Met | Leu | Ala | |
| 210 | | | | 215 | | | | 220 | | | | | | | | |
| Glu | Arg | Leu | Lys | Ala | Pro | Val | Trp | Val | Ala | Pro | Ser | Ala | Pro | Arg | Cys | |
| 225 | | | | 230 | | | | 235 | | | | 240 | | | | |
| Pro | Phe | Pro | Thr | Arg | His | Pro | Cys | Phe | Arg | Gly | Leu | Met | Pro | Ala | Gly | |
| | | | | 245 | | | | 250 | | | | 255 | | | | |
| | | | | | | | | | | | | | | | | |
| Ile | Ala | Ala | Ile | Ser | Gln | Leu | Leu | Glu | Gly | His | Asp | Val | Val | Leu | Val | |
| 260 | | | | | | | | 265 | | | | 270 | | | | |
| Ile | Gly | Ala | Pro | Val | Phe | Arg | Tyr | His | Gln | Tyr | Asp | Pro | Gly | Gln | Tyr | |
| 275 | | | | | | | | 280 | | | | 285 | | | | |
| Leu | Lys | Pro | Gly | Thr | Arg | Leu | Ile | Ser | Val | Thr | Cys | Asp | Pro | Leu | Glu | |
| 290 | | | | 295 | | | | 300 | | | | | | | | |
| Ala | Ala | Arg | Ala | Pro | Met | Gly | Asp | Ala | Ile | Val | Ala | Asp | Ile | Gly | Ala | |
| 305 | | | | 310 | | | | 315 | | | | 320 | | | | |

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

Gln Arg Leu Asn Met Arg Asn Pro Gly Ser Tyr Tyr Phe Cys Ala Ala
385 390 395 400
Gly Gly Leu Gly Phe Ala Leu Pro Ala Ala Ile Gly Val Gln Leu Ala
405 410 415
Glu Pro Glu Arg Gln Val Ile Ala Val Ile Gly Asp Gly Ser Ala Asn
420 425 430
Tyr Ser Ile Ser Ala Leu Trp Thr Ala Ala Gln Tyr Asn Ile Pro Thr
435 440 445

Ile Phe Val Ile Met Asn Asn Gly Thr Tyr Gly Ala Leu Arg Trp Phe
450 455 460
Ala Gly Val Leu Glu Ala Glu Asn Val Pro Gly Leu Asp Val Pro Gly
465 470 475 480
Ile Asp Phe Arg Ala Leu Ala Lys Gly Tyr Gly Val Gln Ala Leu Lys
485 490 495
Ala Asp Asn Leu Glu Gln Leu Lys Gly Ser Leu Gln Glu Ala Leu Ser
500 505 510

Ala Lys Gly Pro Val Leu Ile Glu Val Ser Thr Val Ser Pro Val Lys
515 520 525

<210> 55

<211> 528

<212> PRT

<213> *Pseudomonas aeruginosa*

<400> 55

Met Lys Thr Val His Ser Ala Ser Tyr Glu Ile Leu Arg Arg His Gly

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

Ile Ala Gly Ile Ser Arg Leu Leu Asp Gly His Asp Leu Ile Leu Val
260 265 270

Val Gly Ala Pro Val Phe Arg Tyr His Gln Phe Ala Pro Gly Asp Tyr
275 280 285

Leu Pro Ala Gly Ala Glu Leu Val Gln Val Thr Cys Asp Pro Gly Glu
290 295 300

Ala Ala Arg Ala Pro Met Gly Asp Ala Leu Val Gly Asp Ile Ala Leu
305 310 315 320

Thr Leu Glu Ala Leu Leu Glu Gln Val Arg Pro Ser Ala Arg Pro Leu
325 330 335

Pro Glu Ala Leu Pro Arg Pro Pro Ala Leu Ala Glu Glu Gly Gly Pro
340 345 350

Leu Arg Pro Glu Thr Val Phe Asp Val Ile Asp Ala Leu Ala Pro Arg
355 360 365

Asp Ala Ile Phe Val Lys Glu Ser Thr Ser Thr Val Thr Ala Phe Trp
370 375 380

Gln Arg Val Glu Met Arg Glu Pro Gly Ser Tyr Phe Phe Pro Ala Ala
385 390 395 400

Gly Gly Leu Gly Phe Gly Leu Pro Ala Ala Val Gly Ala Gln Leu Ala
405 410 415

Gln Pro Arg Arg Gln Val Ile Gly Ile Ile Gly Asp Gly Ser Ala Asn
420 425 430

Tyr Gly Ile Thr Ala Leu Trp Ser Ala Ala Gln Tyr Arg Val Pro Ala
435 440 445

Val Phe Ile Ile Leu Lys Asn Gly Thr Tyr Gly Ala Leu Arg Trp Phe
450 455 460

Ala Gly Val Leu Glu Val Pro Asp Ala Pro Gly Leu Asp Val Pro Gly
465 470 475 480

Leu Asp Phe Cys Ala Ile Ala Arg Gly Tyr Gly Val Glu Ala Leu His
485 490 495

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

| | | |
|---|-----|-----|
| 500 | 505 | 510 |
| Ala Asp Arg Pro Val Leu Ile Glu Val Pro Thr Gln Thr Ile Glu Pro | | |
| 515 | 520 | 525 |
| <210> 56 | | |
| <211> 526 | | |
| <212> PRT | | |
| <213> Pseudomonas stutzeri | | |
| <400> 56 | | |
| Met Ala Ser Val His Ser Ile Thr Tyr Glu Leu Leu Arg Arg Gln Gly | | |
| 1 | 5 | 10 |
| | | 15 |
| Ile Asp Thr Val Phe Gly Asn Pro Gly Ser Asn Glu Leu Pro Phe Leu | | |
| 20 | 25 | 30 |
| Lys Asp Phe Pro Glu Asp Phe Arg Tyr Ile Leu Ala Leu Gln Glu Ala | | |
| 35 | 40 | 45 |
| Cys Val Val Gly Ile Ala Asp Gly Tyr Ala Gln Ala Ser Arg Lys Pro | | |
| 50 | 55 | 60 |
| Ala Phe Ile Asn Leu His Ser Ala Ala Gly Thr Gly Asn Ala Met Gly | | |
| 65 | 70 | 75 |
| | | 80 |
| Ala Met Ser Asn Ala Trp Asn Cys His Ser Pro Leu Ile Val Thr Ala | | |
| 85 | 90 | 95 |
| Gly Gln Gln Asn Arg Ala Met Ile Gly Val Glu Ala Leu Leu Thr Asn | | |
| 100 | 105 | 110 |
| Val Asp Ala Ala Ser Leu Pro Arg Pro Leu Val Lys Trp Ser Tyr Glu | | |
| 115 | 120 | 125 |
| Pro Ala Ser Ala Ala Glu Val Pro His Ala Met Ser Arg Ala Ile His | | |
| 130 | 135 | 140 |
| Met Ala Ser Met Ala Pro Arg Gly Pro Val Tyr Leu Ser Val Pro Tyr | | |
| 145 | 150 | 155 |
| Asp Asp Trp Asp Lys Glu Ala Asp Pro Gln Ser His His Leu Tyr Asp | | |
| 165 | 170 | 175 |
| Arg Ser Val Asn Ser Ala Val Arg Leu Asn Asp Gln Asp Leu Glu Val | | |
| 180 | 185 | 190 |

Leu Val Glu Ala Leu Asn Ser Ala Ser Asn Pro Ala Ile Val Leu Gly
 195 200 205

 Pro Asp Val Asp Ser Ala Asn Ala Asn Ala Asp Cys Val Thr Leu Ala
 210 215 220
 Glu Arg Leu Lys Ala Pro Val Trp Val Ala Pro Ser Ala Pro Arg Cys
 225 230 235 240
 Pro Phe Pro Thr Arg His Pro Cys Phe Arg Gly Leu Met Pro Ala Gly
 245 250 255
 Ile Ala Ala Ile Ser Gln Leu Leu Glu Gly His Asp Val Val Leu Val
 260 265 270

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

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

 Gly Gly Leu Gly Phe Ala Leu Pro Ala Ala Ile Gly Val Gln Leu Ala
 405 410 415
 Glu Pro Asp Arg Gln Val Ile Ala Val Ile Gly Asp Gly Ser Ala Asn
 420 425 430
 Tyr Ser Ile Ser Ala Leu Trp Thr Ala Ala His Tyr Asn Ile Pro Ala

435 440 445
 Ile Phe Leu Ile Met Asn Asn Gly Thr Tyr Gly Ala Leu Arg Trp Phe
 450 455 460

Ala Gly Val Leu Glu Ala Glu Asn Val Pro Gly Leu Asp Val Pro Gly
 465 470 475 480
 Ile Asp Phe Cys Ala Ile Ala Lys Gly Tyr Gly Ile Pro Ala Leu Lys
 485 490 495
 Ala Asp Asn Leu Glu Gln Leu Lys Gly Ser Ile His Glu Ala Leu Ser
 500 505 510
 Ala Lys Gly Pro Val Leu Ile Glu Val Ser Thr Val Ser Leu
 515 520 525

<210> 57

<211> 528

<212> PRT

<213> Pseudomonas fluorescens

<400> 57

Met Lys Thr Val His Ser Ala Ser Tyr Asp Ile Leu Arg Gln Gln Gly
 1 5 10 15
 Leu Thr Thr Val Phe Gly Asn Pro Gly Ser Asn Glu Leu Pro Phe Leu
 20 25 30
 Lys Gly Phe Pro Glu Asp Phe Arg Tyr Ile Leu Gly Leu His Glu Gly
 35 40 45
 Ala Val Val Gly Met Ala Asp Gly Phe Ala Leu Ala Ser Gly Gln Pro

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

| | | | |
|---|-------------------------------------|-----|-----|
| 115 | 120 | 125 | |
| Pro Ala Cys Ala Glu Asp Val | Pro Arg Ala Leu Ser Gln Ala Ile His | | |
| 130 | 135 | 140 | |
| Met Ala Asn Gln Ala Pro Lys Gly Pro Val Tyr Leu Ser Ile Pro Tyr | | | |
| 145 | 150 | 155 | 160 |
| Asp Asp Trp Ala Arg Pro Ala Pro Ala Gly Val Glu His Leu Ala Arg | | | |
| 165 | 170 | 175 | |
| Arg Gln Val Ala Thr Ala Gly Leu Pro Ser Ala Ala Gln Leu Arg Ser | | | |
| 180 | 185 | 190 | |
| Leu Val Gln Arg Leu Ala Ala Ala Arg Asn Pro Val Leu Val Leu Gly | | | |
| 195 | 200 | 205 | |
| Pro Asp Val Asp Gly Ser Arg Ser Asn His Leu Ala Val Gln Leu Ala | | | |
| 210 | 215 | 220 | |
| Glu Lys Leu Arg Met Pro Ala Trp Val Ala Pro Ser Ala Ser Arg Cys | | | |
| 225 | 230 | 235 | 240 |
| Pro Phe Pro Thr Arg His Pro Ser Phe Arg Gly Val Leu Pro Ala Ala | | | |
| 245 | 250 | 255 | |
| Ile Ala Gly Ile Ser Arg Cys Leu Ala Asp His Asp Leu Ile Leu Val | | | |
| 260 | 265 | 270 | |
| Val Gly Ala Pro Val Phe Arg Tyr His Gln Phe Ala Pro Gly Asp Tyr | | | |
| 275 | 280 | 285 | |
| Leu Pro Ala Gly Thr Glu Leu Leu His Ile Thr Cys Asp Pro Gly Glu | | | |
| 290 | 295 | 300 | |
| Ala Ala Arg Ala Pro Met Gly Asp Ala Leu Val Gly Asp Ile Val Glu | | | |
| 305 | 310 | 315 | 320 |
| Thr Leu Gln Ala Leu Val Trp Ala Leu Pro Asp Cys Asp Arg Pro Gln | | | |
| 325 | 330 | 335 | |
| Pro Gln Ala Leu Pro Pro Ala Ala Pro Val Glu Glu Leu Gly Gly Leu | | | |
| 340 | 345 | 350 | |
| Leu Arg Pro Glu Thr Val Phe Asp Val Ile Asp Glu Leu Ala Pro Lys | | | |
| 355 | 360 | 365 | |

Asp Ala Ile Tyr Val Lys Glu Ser Thr Ser Thr Val Gly Ala Phe Trp

370

375

380

Gln Arg Val Glu Met Arg Glu Pro Gly Ser Tyr Tyr Phe Pro Ala Ala

385

390

395

400

Gly Gly Leu Gly Phe Gly Leu Pro Ala Ala Val Gly Val Gln Leu Ala

405

410

415

Arg Pro Glu Arg Arg Val Ile Gly Val Ile Gly Asp Gly Ser Ala Asn

420

425

430

Tyr Gly Ile Thr Ala Leu Trp Thr Ala Ala Gln Tyr Gln Ile Pro Val

435

440

445

Val Phe Ile Ile Leu Lys Asn Gly Thr Tyr Gly Ala Leu Arg Trp Phe

450

455

460

Ala Gly Val Leu Gln Val Ser Asp Ala Pro Gly Leu Asp Val Pro Gly

465

470

475

480

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

485

490

495

Ala Asn Thr Arg Glu Ala Phe Ala Gln Ala Leu Ser Glu Ala Leu Ala

500

505

510

Gly Asp Arg Pro Val Leu Ile Glu Val Pro Thr Leu Thr Ile Glu Pro

515

520

525

<210> 58

<211> 1220

<212> PRT

<213> Chloroflexus aurantiacus

<400> 58

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

1

5

10

15

Ile Ala Leu Ile Thr Gly Gly Ala Gly Asn Ile Gly Ser Glu Leu Thr

20

25

30

Arg Arg Phe Leu Ala Glu Gly Ala Thr Val Ile Ile Ser Gly Arg Asn

35

40

45

Arg Ala Lys Leu Thr Ala Leu Ala Glu Met Gln Ala Glu Ala Gly Val
 50 55 60
 Pro Ala Lys Arg Ile Asp Leu Glu Val Met Asp Gly Ser Asp Pro Val
 65 70 75 80
 Ala Val Arg Ala Gly Ile Glu Ala Ile Val Ala Arg His Gly Gln Ile
 85 90 95

 Asp Ile Leu Val Asn Asn Ala Gly Ser Ala Gly Ala Gln Arg Arg Leu
 100 105 110
 Ala Glu Ile Pro Leu Thr Glu Ala Glu Leu Gly Pro Gly Ala Glu Glu
 115 120 125
 Thr Leu His Ala Ser Ile Ala Asn Leu Leu Gly Met Gly Trp His Leu
 130 135 140
 Met Arg Ile Ala Ala Pro His Met Pro Val Gly Ser Ala Val Ile Asn
 145 150 155 160

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

 Lys Gly Arg Pro Glu Gly Asp Thr Ala His His Phe Leu Asn Thr Met
 225 230 235 240
 Arg Leu Cys Arg Ala Asn Asp Gln Gly Ala Leu Glu Arg Arg Phe Pro
 245 250 255
 Ser Val Gly Asp Val Ala Asp Ala Ala Val Phe Leu Ala Ser Ala Glu
 260 265 270
 Ser Ala Ala Leu Ser Gly Glu Thr Ile Glu Val Thr His Gly Met Glu
 275 280 285

 Leu Pro Ala Cys Ser Glu Thr Ser Leu Leu Ala Arg Thr Asp Leu Arg

290 295 300
 Thr Ile Asp Ala Ser Gly Arg Thr Thr Leu Ile Cys Ala Gly Asp Gln
 305 310 315 320
 Ile Glu Glu Val Met Ala Leu Thr Gly Met Leu Arg Thr Cys Gly Ser
 325 330 335
 Glu Val Ile Ile Gly Phe Arg Ser Ala Ala Ala Leu Ala Gln Phe Glu
 340 345 350

 Gln Ala Val Asn Glu Ser Arg Arg Leu Ala Gly Ala Asp Phe Thr Pro
 355 360 365
 Pro Ile Ala Leu Pro Leu Asp Pro Arg Asp Pro Ala Thr Ile Asp Ala
 370 375 380
 Val Phe Asp Trp Ala Gly Glu Asn Thr Gly Gly Ile His Trp Ile Leu
 385 390 395 400
 Pro Ala Thr Ser His Glu Pro Ala Pro Cys Val Ile Glu Val Asp Asp
 405 410 415

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

 Gln Leu Ile Arg Val Trp Arg His Glu Ala Glu Leu Asp Tyr Gln Arg
 485 490 495
 Ala Ser Ala Ala Gly Asp His Val Leu Pro Pro Val Trp Ala Asn Gln
 500 505 510
 Ile Val Arg Phe Ala Asn Arg Ser Leu Glu Gly Leu Glu Phe Ala Cys
 515 520 525
 Ala Trp Thr Ala Gln Leu Leu His Ser Gln Arg His Ile Asn Glu Ile
 530 535 540

Thr Leu Asn Ile Pro Ala Asn Ile Ser Ala Thr Thr Gly Ala Arg Ser
 545 550 555 560
 Ala Ser Val Gly Trp Ala Glu Ser Leu Ile Gly Leu His Leu Gly Lys
 565 570 575
 Val Ala Leu Ile Thr Gly Gly Ser Ala Gly Ile Gly Gly Gln Ile Gly
 580 585 590
 Arg Leu Leu Ala Leu Ser Gly Ala Arg Val Met Leu Ala Ala Arg Asp
 595 600 605

Arg His Lys Leu Glu Gln Met Gln Ala Met Ile Gln Ser Glu Leu Ala
 610 615 620
 Glu Val Gly Tyr Thr Asp Val Glu Asp Arg Val His Ile Ala Pro Gly
 625 630 635 640
 Cys Asp Val Ser Ser Glu Ala Gln Leu Ala Asp Leu Val Glu Arg Thr
 645 650 655
 Leu Ser Ala Phe Gly Thr Val Asp Tyr Leu Ile Asn Asn Ala Gly Ile
 660 665 670

Ala Gly Val Glu Glu Met Val Ile Asp Met Pro Val Glu Gly Trp Arg
 675 680 685
 His Thr Leu Phe Ala Asn Leu Ile Ser Asn Tyr Ser Leu Met Arg Lys
 690 695 700
 Leu Ala Pro Leu Met Lys Lys Gln Gly Ser Gly Tyr Ile Leu Asn Val
 705 710 715 720
 Ser Ser Tyr Phe Gly Gly Glu Lys Asp Ala Ala Ile Pro Tyr Pro Asn
 725 730 735

Arg Ala Asp Tyr Ala Val Ser Lys Ala Gly Gln Arg Ala Met Ala Glu
 740 745 750
 Val Phe Ala Arg Phe Leu Gly Pro Glu Ile Gln Ile Asn Ala Ile Ala
 755 760 765
 Pro Gly Pro Val Glu Gly Asp Arg Leu Arg Gly Thr Gly Glu Arg Pro
 770 775 780
 Gly Leu Phe Ala Arg Arg Ala Arg Leu Ile Leu Glu Asn Lys Arg Leu

785 790 795 800

Asn Glu Leu His Ala Ala Leu Ile Ala Ala Ala Arg Thr Asp Glu Arg

805 810 815

Ser Met His Glu Leu Val Glu Leu Leu Leu Pro Asn Asp Val Ala Ala

820 825 830

Leu Glu Gln Asn Pro Ala Ala Pro Thr Ala Leu Arg Glu Leu Ala Arg

835 840 845

Arg Phe Arg Ser Glu Gly Asp Pro Ala Ala Ser Ser Ser Ser Ala Leu

850 855 860

Leu Asn Arg Ser Ile Ala Ala Lys Leu Leu Ala Arg Leu His Asn Gly

865 870 875 880

Gly Tyr Val Leu Pro Ala Asp Ile Phe Ala Asn Leu Pro Asn Pro Pro

885 890 895

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

900 905 910

Arg Asp Gly Ile Met Gly Met Leu Tyr Leu Gln Arg Met Pro Thr Glu

915 920 925

Phe Asp Val Ala Met Ala Thr Val Tyr Tyr Leu Ala Asp Arg Asn Val

930 935 940

Ser Gly Glu Thr Phe His Pro Ser Gly Gly Leu Arg Tyr Glu Arg Thr

945 950 955 960

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

965 970 975

Glu Leu Val Gly Ser Thr Val Tyr Leu Ile Gly Glu His Leu Thr Glu

980 985 990

His Leu Asn Leu Leu Ala Met Tyr Leu Glu Arg Tyr Gly Ala Arg Gln

995 1000 1005

Val Trp Ile Val Glu Thr Glu Thr Gly Ala Glu Thr Met Arg Arg

1010 1015 1020

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

1025 1030 1035

| | | | |
|---------|--------------------------|---------------------|-------------|
| Ala Gly | Asp Gln Ile Glu Ala | Ala Ile Asp Gln Ala | Ile Thr Arg |
| 1040 | 1045 | 1050 | |
| | | | |
| Tyr Gly | Arg Pro Gly Pro Val | Val Cys Thr Pro Phe | Arg Pro Leu |
| 1055 | 1060 | 1065 | |
| Pro Thr | Val Pro Leu Val Gly | Arg Lys Asp Ser Asp | Trp Ser Thr |
| 1070 | 1075 | 1080 | |
| Val Leu | Ser Glu Ala Glu Phe | Ala Glu Leu Cys Glu | His Gln Leu |
| 1085 | 1090 | 1095 | |
| Thr His | His Phe Arg Val Ala | Arg Lys Ile Ala Leu | Ser Asp Gly |
| 1100 | 1105 | 1110 | |
| | | | |
| Ala Ser | Leu Ala Leu Val Thr | Pro Glu Thr Thr Ala | Thr Ser Thr |
| 1115 | 1120 | 1125 | |
| Thr Glu | Gln Phe Ala Leu Ala | Asn Phe Ile Lys Thr | Thr Leu His |
| 1130 | 1135 | 1140 | |
| Ala Phe | Thr Ala Thr Ile Gly | Val Glu Ser Glu Arg | Thr Ala Gln |
| 1145 | 1150 | 1155 | |
| Arg Ile | Leu Ile Asn Gln Val | Asp Leu Thr Arg Arg | Ala Arg Ala |
| 1160 | 1165 | 1170 | |
| | | | |
| Glu Glu | Pro Arg Asp Pro His | Glu Arg Gln Gln Glu | Leu Glu Arg |
| 1175 | 1180 | 1185 | |
| Phe Ile | Glu Ala Val Leu Leu | Val Thr Ala Pro Leu | Pro Pro Glu |
| 1190 | 1195 | 1200 | |
| Ala Asp | Thr Arg Tyr Ala Gly | Arg Ile His Arg Gly | Arg Ala Ile |
| 1205 | 1210 | 1215 | |
| Thr Val | | | |
| 1220 | | | |
| | | | |
| <210> | 59 | | |
| | | | |
| <211> | 1229 | | |
| | | | |
| <212> | PRT | | |
| | | | |
| <213> | Roseiflexus castenholzii | | |
| | | | |
| <400> | 59 | | |

Met Ser Thr Val Arg Arg Leu Glu Gly Lys Val Ala Leu Ile Thr Gly

1 5 10 15
Gly Ala Gly Asn Ile Gly Glu Val Ile Thr Arg Arg Phe Leu Ala Glu

20 25 30
Gly Ala Thr Val Val Ile Thr Gly Arg Asn Ala Glu Lys Leu Ala Val

35 40 45
Tyr Arg Arg Arg Leu Ile Asp Glu Glu Arg Val Ala Pro Glu Arg Val

50 55 60
Val Ala Leu Arg Met Asp Gly Ser Asp Ile Ala Gln Val Arg Ala Gly

65 70 75 80
Val Ala Gln Ile Val His Gly Gly Thr Asp Val Pro Ile Pro Leu His

85 90 95
Arg Ile Asp Ile Leu Val Asn Asn Ala Gly Ser Ala Gly Pro Arg Arg

100 105 110
Arg Leu Val Asp Ile Pro Leu Glu Pro Ser Glu Val Gln Pro Pro Asp

115 120 125
Ser Glu Thr Leu Ala Gln Ala Val Gly Asn Leu Val Gly Ile Thr Trp

130 135 140
Asn Leu Thr Arg Ala Ala Ala Pro His Met Pro Ser Gly Ser Ser Val

145 150 155 160
Ile Asn Ile Ser Thr Ile Phe Ser Arg Thr Asp Tyr Tyr Gly Arg Ile

165 170 175
Ala Tyr Val Ala Pro Lys Ala Ala Leu Asn Ala Leu Ser Asp Gly Leu

180 185 190
Ala Arg Glu Leu Gly Val Arg Gly Ile Arg Val Asn Thr Ile Tyr Pro

195 200 205
Gly Pro Ile Glu Ser Glu Arg Ile Tyr Thr Met Phe Gln Ala Met Asp

210 215 220
Ala Leu Lys Gly Gln Pro Glu Gly Asp Thr Ala Ser Gly Phe Leu Arg

225 230 235 240
Met Met Arg Leu Ser Arg Ile Asp Gln Asn Gly Glu Val Val Lys Arg

| | | |
|---|-----|-----|
| 245 | 250 | 255 |
| Phe Pro Ser Pro Val Asp Val Ala Asn Thr Ala Val Phe Leu Ala Ser | | |
| 260 | 265 | 270 |
| Asp Glu Ser Ala Ala Phe Thr Gly His Ala Phe Glu Val Thr His Gly | | |
| 275 | 280 | 285 |
| Met Glu Val Pro Thr Glu Ser Arg Thr Thr Phe Val Ser Arg Pro Gly | | |
| 290 | 295 | 300 |
| Leu Arg Ser Val Asp Ala Thr Gly Lys Val Ile Leu Ile Cys Ala Gly | | |
| 305 | 310 | 315 |
| Asp Gln Val Asp Asp Ala Val Ala Leu Ala Asp Thr Leu Arg Ser Cys | | |
| 325 | 330 | 335 |
| Arg Ala Thr Val Val Ile Gly Phe Arg Asp Pro Arg Ala Leu Glu Lys | | |
| 340 | 345 | 350 |
| Ala Ser Val Leu Leu Arg Glu Pro Arg His Ala Leu Ala Ala Asp Met | | |
| 355 | 360 | 365 |
| Tyr Gly Arg Pro Thr Met Thr Ala Glu Ala Arg Leu Val Arg Leu Asp | | |
| 370 | 375 | 380 |
| Pro Leu Asp Pro Arg Ala Ala Ala Gln Thr Leu Glu Gln Ile His Ala | | |
| 385 | 390 | 395 |
| Glu Leu Gly Ala Ile His His Ala Val Val Leu Pro Gly Gln Ser Arg | | |
| 405 | 410 | 415 |
| His Ala Pro Ser Ala Ser Leu Ile Glu Val Asp Asp Gln Val Val Glu | | |
| 420 | 425 | 430 |
| Arg Phe Leu His Gln Glu Leu Val Gly Thr Ile Ala Leu Ala Arg Glu | | |
| 435 | 440 | 445 |
| Leu Ala Arg Phe Trp Glu Glu Tyr Pro Ser Gly Ser Ser Met His Arg | | |
| 450 | 455 | 460 |
| Val Leu Phe Val Ser Asn Pro Asp Asp Gln Gln Gly Asn Gln Tyr Ser | | |
| 465 | 470 | 475 |
| His Ile Leu Arg Ala Ala Val Glu Gln Leu Val Arg Val Trp Arg His | | |
| 485 | 490 | 495 |

Glu Ser Glu Tyr Asp Ser Val Asn Pro Ala His Gln Gln Glu Gly Gln
 500 505 510
 Ser Ser Ala Ala Val Trp Ala Asn Gln Leu Ile Arg Tyr Val Asn Asn
 515 520 525
 Glu Met Ala Asn Leu Asp Phe Thr Cys Ala Trp Val Ala Lys Leu Leu
 530 535 540
 Gly Ser Asp Arg Arg Ile Ala Glu Ile Asn Leu Tyr Leu Pro Glu Glu
 545 550 555 560
 Ile Val Gly Thr Ile Gly Val His Asn Pro Gly Phe Gly Trp Ala Glu
 565 570 575
 Ser Leu Phe Gly Leu His Met Gly Lys Val Ala Leu Ile Thr Gly Gly
 580 585 590
 Ser Ala Gly Ile Gly Gly Gln Ile Gly Arg Leu Leu Ala Leu Ser Gly
 595 600 605
 Ala His Val Met Leu Ala Ala Arg Asn Ala Asp Gln Leu Glu Gln Met
 610 615 620
 Arg Ala Ser Ile Val Arg Glu Val Arg Asp Ala Ser Tyr Pro Asp Ala
 625 630 635 640
 Glu Ser Arg Val Ala Ile Phe Pro Gly Ser Asp Val Ser Asp Ile Asp
 645 650 655
 Gly Leu Glu Arg Leu Val Asn His Thr Val Arg Val Phe Gly Lys Val
 660 665 670
 Asp Tyr Leu Ile Asn Asn Ala Gly Ile Ala Gly Ala Glu Glu Met Val
 675 680 685
 Ile Asp Met Pro Val Asp Ala Trp Arg His Thr Leu Arg Ala Asn Leu
 690 695 700
 Ile Ser Asn Tyr Ala Leu Leu Arg Arg Leu Ala Pro Gln Met Lys Ala
 705 710 715 720
 Ala Gly Gly Ala Tyr Val Leu Asn Val Ser Ser Tyr Phe Gly Gly Glu
 725 730 735
 Lys Tyr Val Ala Ile Pro Tyr Pro Asn Arg Ser Asp Tyr Ala Val Ser

740 745 750
 Lys Ala Gly Gln Arg Ala Met Val Glu Ser Leu Ala Arg Phe Leu Gly
 755 760 765
 Pro Glu Ile Gln Ile Asn Ala Ile Ala Pro Gly Pro Val Glu Gly Glu

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

 835 840 845
 Ala Pro Leu Arg Arg Leu Ala Thr Met Leu Arg Glu Thr Ser Asp Ala
 850 855 860
 Gly Gly Ser Ala Gln Ser Tyr Leu Met Asn Ala Thr Ile Ala Arg Lys
 865 870 875 880
 Leu Leu Asn Arg Leu Glu Asn Gly Gly Tyr Ile Thr Leu His Asp Arg
 885 890 895
 Arg Ala Leu Thr Val Glu Pro Pro Glu Pro Phe Phe Thr Glu Ala Gln

 900 905 910
 Ile Glu Arg Glu Ala Ile Lys Val Arg Asp Gly Ile Leu Gly Met Leu
 915 920 925
 His Leu Gln Arg Met Pro Thr Glu Phe Asp Val Ala Leu Ala Thr Val
 930 935 940
 Phe Tyr Leu Ala Asp Arg Asn Val Thr Gly Glu Thr Phe His Pro Ser
 945 950 955 960
 Gly Gly Leu Arg Phe Glu Arg Thr Val Thr Glu Gly Glu Leu Phe Gly

 965 970 975
 Lys Pro Gly Gln Gln Arg Leu Glu Arg Leu Lys Gly Ser Val Val Tyr
 980 985 990

Leu Ile Gly Glu His Leu Arg Gln His Leu Val Leu Leu Ala Arg Thr
 995 1000 1005
 Phe Leu Asp Glu Ile His Val Ala Arg Val Val Leu Leu Thr Glu
 1010 1015 1020
 Thr Thr Gln Ala Ala Thr Asp Leu Ala Ala Glu Leu Ser Asp Tyr
 1025 1030 1035
 Glu Ala Ala Gly Arg Phe Val Val Ile Pro Thr Cys Gly Asp Ile
 1040 1045 1050
 Glu Gly Gly Ile Asp Arg Ala Met Ala Glu Tyr Gly Arg Pro Gly
 1055 1060 1065
 Pro Val Ile Ser Thr Pro Phe Arg Pro Leu Pro Asp Arg Ala Leu
 1070 1075 1080
 Ser Ala Arg Asn Gly Asp Trp Ser Ser Val Leu Thr Thr Ala Glu
 1085 1090 1095
 Phe Glu Glu Leu Val Glu Gln Gln Ile Thr His His Phe Arg Val
 1100 1105 1110
 Ala Arg Lys Ala Gly Leu Ile Glu Gly Ala Asn Val Thr Leu Val
 1115 1120 1125
 Thr Pro Pro Thr Ser Ala Arg Ser Thr Ser Glu Glu Phe Ala Leu
 1130 1135 1140
 Ala Asn Phe Val Lys Thr Thr Leu His Ala Leu Thr Ala Thr Ala
 1145 1150 1155
 Gly Ala Glu Ser Glu Arg Thr Val Pro His Val Pro Val Asn Gln
 1160 1165 1170
 Val Asp Leu Thr Arg Arg Ala Arg Ser Glu Glu Pro Arg Thr Pro
 1175 1180 1185
 Ser Glu Glu Glu Glu Glu Leu Gln Arg Phe Val Asn Ala Val Leu
 1190 1195 1200
 Leu Thr Ser Ala Pro Leu Pro Thr Pro Leu Glu Ser Arg Tyr Arg
 1205 1210 1215
 Ala Arg Ile Tyr Arg Gly Asn Ala Ile Thr Val

1220 1225

<210> 60

<211> 1217

<212> PRT

<213> Erythrobacter sp.

<400> 60

Met Ser Lys Glu Gly Asn Ala Ala Lys Gly Arg Leu Glu Gly Lys Val

1 5 10 15

Ala Leu Ile Thr Gly Ala Ala Gly Asn Leu Gly Asn Glu Ile Ser Arg

20 25 30

Ala Phe Ala Arg Glu Gly Ala Phe Val Val Met Thr Gly Arg Thr Glu

35 40 45

Glu Arg Ile Ser Ala Ala Arg Glu Gln Leu Ile Ala Asp Thr Gly Val

50 55 60

Ala Pro Glu Arg Ile Asp Thr Ala Val Leu Asp Gly Gly Asn Pro Asp

65 70 75 80

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

85 90 95

Asp Ile Leu Ile Asn Asn Ala Gly Ser Ala Gly Pro Lys Gln Pro Leu

100 105 110

His Asn Val Pro Leu Ser Pro Gln Glu Met Glu Ala Cys Gly Asp Thr

115 120 125

Glu Thr Val Arg Asp Ala Met Leu Asn Ile Leu Gly Val Thr Trp Asn

130 135 140

Met Ala Arg Ile Val Ala Pro Met Met Pro Val Gly Gly Ala Met Val

145 150 155 160

Asn Ile Ser Thr Ile Phe Ser His Thr Arg Tyr Tyr Gly Arg Thr Ala

165 170 175

Tyr Val Val Pro Lys Ala Ala Leu Asn Ala Leu Ser Asn Gln Leu Ala

180 185 190

Ser Glu Leu Gly Pro Arg Gly Ile Arg Val Asn Thr Val Phe Pro Gly

195 200 205

Pro Ile Glu Ser Asp Arg Ile Arg Thr Val Phe Ala Ala Met Asp Glu
 210 215 220
 Val Gln Ser Gln Pro Lys Asp Thr Thr Ala Asn Tyr Phe Thr Gly Arg

 225 230 235 240
 Met Ala Leu Thr Arg Ser Val Asn Gly Lys Val Asp Gly Lys Pro Leu
 245 250 255
 Pro Asn Pro Lys Asp Ile Ala Gly Thr Cys Leu Phe Leu Ala Ser Glu
 260 265 270
 Glu Ala Ala Gly Ile Ala Gly Glu Glu Val Asp Val Thr His Gly Leu
 275 280 285
 Ser Ala Asn Arg Thr Ser Ala Ser Thr Tyr Met Thr Arg Pro Ser Met

 290 295 300
 Arg Ser Leu Asp Gly Ala Gly Leu Asn Ile Phe Ile Val Ser Gly Glu
 305 310 315 320
 Asn Trp Asp Asp Ala Leu Val Ala Ala His Thr Leu Ile Gly Ser Gly
 325 330 335
 Ala Lys Val Arg Leu Gly Leu Ala Arg Asn Ala Asp Val Ala Gln Ala
 340 345 350
 Asn Ala Arg Leu Lys Ala Gln Gly Ile Gly Glu Glu Leu Thr Val Thr

 355 360 365
 Arg Phe Asn Arg Ala Glu Pro Asp Ala Met Glu Asp Ala Leu Ala Ala
 370 375 380
 Phe Ser Gly Asp Val Asp Gly Ala Ile Thr Gly Ala Ile Ile Leu Pro
 385 390 395 400
 Val Lys Pro Ser Gly His Phe Thr Gly Ser Leu Leu Ala Ala Asp Asp
 405 410 415
 Asp Thr Val Thr Lys Phe Met Asp Thr Glu Leu Val Gly Ala Ile Ala

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

450 455 460
 Asn Ser Phe Ala Ser Val Leu Ser Ala Gly Ile Thr Gln Leu Ile Arg
 465 470 475 480
 Ile Trp Arg Asp Glu Glu Arg Val Gln Ala Gly Asn Gly Ser Thr Glu

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

 545 550 555 560
 Thr Gly Leu His Leu Gly Lys Val Ala Phe Ile Thr Gly Gly Ser Ala
 565 570 575
 Gly Ile Gly Gly Gln Val Ala Arg Leu Leu Ala Leu Ala Gly Ala Lys
 580 585 590
 Val Met Met Val Ala Arg Arg Glu Ser Glu Leu Val Ala Ala Arg Asp
 595 600 605
 Arg Ile Val Gly Glu Leu Gln Asp Ile Gly Phe Ala Gly Val Glu Arg

 610 615 620
 Arg Val Lys Tyr Met Ala Asp Ile Asp Val Ser Asp Phe Ala Ser Leu
 625 630 635 640
 Asp Lys Ala Val Asp Ala Thr Leu Glu Glu Phe Gly Arg Ile Asp Tyr
 645 650 655
 Leu Ile Asn Asn Ala Gly Val Ala Gly Ala Glu Asp Met Val Ile Asp
 660 665 670
 Met Glu Pro Glu Ala Trp Arg Phe Thr Leu Asp Ala Asn Leu Ile Ser

 675 680 685
 Asn Tyr His Leu Met Gln Arg Val Val Pro Leu Met Lys Glu Gln Gly
 690 695 700

Ser Gly Tyr Val Leu Asn Val Ser Ser Tyr Phe Gly Gly Glu Lys Phe
 705 710 715 720
 Leu Ala Val Ala Tyr Pro Asn Arg Ala Asp Tyr Gly Leu Ser Lys Ala
 725 730 735
 Gly Gln Arg Ala Met Val Glu Ala Phe Ser Pro Phe Leu Gly Pro Glu
 740 745 750
 Val Gln Cys Asn Ala Ile Ala Pro Gly Pro Val Asp Gly Asp Arg Leu
 755 760 765
 Ser Gly Thr Gly Gly Lys Pro Gly Leu Phe Gln Arg Arg Ala Lys Leu
 770 775 780
 Ile Leu Glu Asn Lys Arg Leu Asn Ala Val Tyr Ser Ala Val Ile His
 785 790 795 800
 Ala Ile Arg Glu Gly Gly Asp Ala Ala Lys Ile Leu Thr Arg Leu Ser
 805 810 815
 Arg Asn Ser Thr Ser Thr Leu Ser His Asp Ala Glu Ala Pro Glu Glu
 820 825 830
 Leu Arg Lys Leu Ala Leu Asp Phe Ala Ser Gln Gly Asp Gly Leu Cys
 835 840 845
 Thr Trp Asp Gln Tyr Leu Leu Thr Asp Ala Met Ala Gln Arg Leu Leu
 850 855 860
 Val Arg Leu Gln Leu Gly Gly Phe Leu Leu Gly Ser Asn Glu Trp Ala
 865 870 875 880
 Ser Leu Ser Ser Ser Glu Gln Thr Trp Leu Lys Leu Ser Pro Pro Asp
 885 890 895
 Asp Lys Pro Phe Leu Pro Ala Ala Gln Val Asp Lys Val Ala Asn Gly
 900 905 910
 Val Gly Lys Gly Val Ile Ser Gln Leu His Leu Gly Ala Met Pro Thr
 915 920 925
 Glu Ala Glu Val Ala Gln Ala Thr Val Phe Phe Leu Ala Asp Arg Ala
 930 935 940
 Val Ser Gly Glu Thr Phe Met Pro Ser Gly Gly Leu Arg Val Glu Arg

945 950 955 960
 Ser Asn Thr Glu Arg Glu Met Phe Gly Ser Pro Lys Gln Glu Arg Ile
 965 970 975
 Asp Lys Met Lys Gly Lys Thr Val Trp Ile Ile Gly Glu His Leu Ser
 980 985 990
 Asp Tyr Val Ala Ala Thr Ile Glu Glu Leu Val Ser Gly Cys Gly Val

 995 1000 1005
 Ala Lys Val Val Leu Ile Ala Lys Asp Lys Ser Gly Glu Lys Ala
 1010 1015 1020
 Val Arg Asp Gln Leu Pro Asn Asp Leu Ser Lys Asp Ala Leu Glu
 1025 1030 1035
 Val Leu Ile Ala Gly Asp Gly Leu Glu Glu Ala Met Asp Glu Ala
 1040 1045 1050
 Leu Gly His Trp Gly Lys Pro Thr Thr Val Leu Ser Met Pro Gly

 1055 1060 1065
 Glu Pro Leu Pro Asp His Leu Phe Glu Gly Gly Asn Pro Leu Ser
 1070 1075 1080
 Thr Lys Asp Phe Ala His Met Val Glu Ala Asn Ile Thr Arg His
 1085 1090 1095
 Tyr Arg Val Thr Arg Lys Ala Ser Leu Tyr Asp Gly Cys Gln Val
 1100 1105 1110
 Val Leu Val Ser Pro Asp Val Pro Tyr Gly Ser Asp Gly Pro Gly

 1115 1120 1125
 Val Ala Leu Ala Asn Phe Val Lys Thr Ser Leu His Ala Phe Thr
 1130 1135 1140
 Ala Thr Val Ala Val Glu Asn Glu Arg Leu Val His Asp Val Pro
 1145 1150 1155
 Val Asn Gln Ile Asn Leu Thr Arg Arg Val Ser Ser Glu Glu Pro
 1160 1165 1170
 Arg Asp Ala Asp Glu His Ala Glu Glu Leu Arg Arg Phe Thr Arg

 1175 1180 1185

Ala Val Leu Leu Val Gly Ala Pro Leu Pro Asp Ala Gln Asp Ser

1190 1195 1200

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

1205 1210 1215

<210> 61

<211> 357

<212> PRT

<213> Metallosphaera sedula

<400> 61

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

1 5 10 15

Gly Ile Glu Tyr Val Arg Met Leu Ala Asp His Pro Tyr Ile Lys Pro

20 25 30

Thr Tyr Leu Ala Gly Lys Gly Ser Val Gly Lys Pro Tyr Gly Glu Ile

35 40 45

Val Arg Trp Gln Thr Val Gly Asn Val Pro Lys Glu Val Ala Asn Gln

50 55 60

Glu Val Lys Pro Thr Asp Pro Lys Leu Met Asp Asp Val Asp Ile Ile

65 70 75 80

Phe Ser Pro Leu Pro Gln Gly Ala Ala Gly Pro Val Glu Glu Gln Phe

85 90 95

Ala Lys Leu Gly Phe Asn Val Ile Ser Asn Ser Pro Asp His Arg Phe

100 105 110

Asp Met Asp Val Pro Met Ile Ile Pro Glu Val Asn Pro His Thr Val

115 120 125

Thr Leu Ile Asp Glu Gln Arg Lys Arg Arg Asp Trp Lys Gly Phe Ile

130 135 140

Val Thr Thr Pro Leu Cys Thr Ala Gln Gly Ala Ala Ile Pro Leu Thr

145 150 155 160

Pro Ile Tyr Gln Asn Phe Lys Met Ser Gly Val Met Ile Thr Thr Met

165 170 175

Gln Ser Leu Ser Gly Ala Gly Tyr Pro Gly Ile Ala Ser Leu Asp Ile

180 185 190
Val Asp Asn Ala Leu Pro Leu Gly Asp Gly Tyr Asp Ala Lys Thr Val
195 200 205

Lys Glu Ile Thr Arg Ile Leu Ser Glu Val Lys Arg Asn Val Gln Glu
210 215 220
Pro Gly Val Asn Glu Ile Thr Leu Asp Ala Thr Thr His Arg Ile Ala
225 230 235 240
Thr Ile His Gly His Tyr Glu Val Ala Tyr Val Thr Phe Lys Glu Asp
245 250 255
Thr Asp Val Arg Lys Val Met Glu Ser Met Glu Ser Phe Lys Gly Glu
260 265 270

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

Ala Ala Gly Gly Gly Val Leu Thr Ala Glu Leu Leu Val Glu Lys Gly
340 345 350
Tyr Ile Asp Lys Arg
355

<210> 62

<211> 356

<212> PRT

<213> Sulfolobus tokodaii

<400> 62

Met Arg Arg Thr Leu Lys Ala Ala Ile Leu Gly Ala Thr Gly Leu Val
1 5 10 15
Gly Ile Glu Tyr Val Arg Met Leu Ser Asn His Pro Tyr Ile Lys Pro

20

25

30

Ala Tyr Leu Ala Gly Lys Gly Ser Val Gly Lys Pro Tyr Gly Glu Val

35

40

45

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

50

55

60

Glu Ile Lys Pro Thr Asp Pro Lys Leu Met Asp Asp Val Asp Ile Ile

65

70

75

80

Phe Ser Pro Leu Pro Gln Gly Ala Ala Gly Pro Val Glu Glu Gln Phe

85

90

95

Ala Lys Glu Gly Phe Pro Val Ile Ser Asn Ser Pro Asp His Arg Phe

100

105

110

Asp Pro Asp Val Pro Leu Leu Val Pro Glu Leu Asn Pro His Thr Ile

115

120

125

Ser Leu Ile Asp Glu Gln Arg Lys Arg Arg Glu Trp Lys Gly Phe Ile

130

135

140

Val Thr Thr Pro Leu Cys Thr Ala Gln Gly Ala Ala Ile Pro Leu Gly

145

150

155

160

Ala Ile Phe Lys Asp Tyr Lys Met Asp Gly Ala Phe Ile Thr Thr Ile

165

170

175

Gln Ser Leu Ser Gly Ala Gly Tyr Pro Gly Ile Pro Ser Leu Asp Val

180

185

190

Val Asp Asn Ile Leu Pro Leu Gly Asp Gly Tyr Asp Ala Lys Thr Ile

195

200

205

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

210

215

220

Pro Lys Leu Glu Asp Val Ser Leu Ala Ala Thr Thr His Arg Ile Ala

225

230

235

240

Thr Ile His Gly His Tyr Glu Val Leu Tyr Val Ser Phe Lys Glu Glu

245

250

255

Thr Ala Ala Glu Lys Val Lys Glu Thr Leu Glu Asn Phe Arg Gly Glu

260

265

270

Pro Gln Asp Leu Lys Leu Pro Thr Ala Pro Ser Lys Pro Ile Ile Val
275 280 285

Met Asn Glu Asp Thr Arg Pro Gln Val Tyr Phe Asp Arg Trp Ala Gly
290 295 300

Asp Ile Pro Gly Met Ser Val Val Val Gly Arg Leu Lys Gln Val Asn
305 310 315 320

Lys Arg Met Ile Arg Leu Val Ser Leu Ile His Asn Thr Val Arg Gly
325 330 335

Ala Ala Gly Gly Gly Ile Leu Ala Ala Glu Leu Leu Val Glu Lys Gly
340 345 350

Tyr Ile Glu Lys
355

<210> 63
<211> 482
<212> PRT
<213> *Saccharomyces cerevisiae*
<400> 63

Met Ser Ala Phe Val Arg Val Val Pro Arg Ile Ser Arg Ser Ser Val
1 5 10 15

Leu Thr Arg Ser Leu Arg Leu Gln Leu Arg Cys Tyr Ala Ser Tyr Pro
20 25 30

Glu His Thr Ile Ile Gly Met Pro Ala Leu Ser Pro Thr Met Thr Gln
35 40 45

Gly Asn Leu Ala Ala Trp Thr Lys Lys Glu Gly Asp Gln Leu Ser Pro
50 55 60

Gly Glu Val Ile Ala Glu Ile Glu Thr Asp Lys Ala Gln Met Asp Phe
65 70 75 80

Glu Phe Gln Glu Asp Gly Tyr Leu Ala Lys Ile Leu Val Pro Glu Gly
85 90 95

Thr Lys Asp Ile Pro Val Asn Lys Pro Ile Ala Val Tyr Val Glu Asp
100 105 110

Lys Ala Asp Val Pro Ala Phe Lys Asp Phe Lys Leu Glu Asp Ser Gly
 115 120 125
 Ser Asp Ser Lys Thr Ser Thr Lys Ala Gln Pro Ala Glu Pro Gln Ala
 130 135 140
 Glu Lys Lys Gln Glu Ala Pro Ala Glu Glu Thr Lys Thr Ser Ala Pro
 145 150 155 160
 Glu Ala Lys Lys Ser Asp Val Ala Ala Pro Gln Gly Arg Ile Phe Ala
 165 170 175

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

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

 Leu Ser Ile Asn Asp Leu Leu Val Lys Ala Ile Thr Val Ala Ala Lys
 305 310 315 320
 Arg Val Pro Asp Ala Asn Ala Tyr Trp Leu Pro Asn Glu Asn Val Ile
 325 330 335
 Arg Lys Phe Lys Asn Val Asp Val Ser Val Ala Val Ala Thr Pro Thr
 340 345 350
 Gly Leu Leu Thr Pro Ile Val Lys Asn Cys Glu Ala Lys Gly Leu Ser

355

360

365

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

370

375

380

Lys Leu Ala Pro Glu Glu Phe Gln Gly Gly Thr Ile Cys Ile Ser Asn

385

390

395

400

Met Gly Met Asn Asn Ala Val Asn Met Phe Thr Ser Ile Ile Asn Pro

405

410

415

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

420

425

430

Glu Asp Ala Ala Ala Glu Asn Gly Phe Ser Phe Asp Asn Gln Val Thr

435

440

445

Ile Thr Gly Thr Phe Asp His Arg Thr Ile Asp Gly Ala Lys Gly Ala

450

455

460

Glu Phe Met Lys Glu Leu Lys Thr Val Ile Glu Asn Pro Leu Glu Met

465

470

475

480

Leu Leu

<210> 64

<211> 420

<212> PRT

<213> *Saccharomyces cerevisiae*

<400> 64

Met Leu Ala Ala Ser Phe Lys Arg Gln Pro Ser Gln Leu Val Arg Gly

1

5

10

15

Leu Gly Ala Val Leu Arg Thr Pro Thr Arg Ile Gly His Val Arg Thr

20

25

30

Met Ala Thr Leu Lys Thr Thr Asp Lys Lys Ala Pro Glu Asp Ile Glu

35

40

45

Gly Ser Asp Thr Val Gln Ile Glu Leu Pro Glu Ser Ser Phe Glu Ser

50

55

60

Tyr Met Leu Glu Pro Pro Asp Leu Ser Tyr Glu Thr Ser Lys Ala Thr

65 70 75 80
 Leu Leu Gln Met Tyr Lys Asp Met Val Ile Ile Arg Arg Met Glu Met
 85 90 95
 Ala Cys Asp Ala Leu Tyr Lys Ala Lys Lys Ile Arg Gly Phe Cys His
 100 105 110
 Leu Ser Val Gly Gln Glu Ala Ile Ala Val Gly Ile Glu Asn Ala Ile
 115 120 125
 Thr Lys Leu Asp Ser Ile Ile Thr Ser Tyr Arg Cys His Gly Phe Thr

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

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

 260 265 270
 Lys Val Asn Gly Met Asp Ile Leu Ala Val Tyr Gln Ala Ser Lys Phe
 275 280 285
 Ala Lys Asp Trp Cys Leu Ser Gly Lys Gly Pro Leu Val Leu Glu Tyr
 290 295 300
 Glu Thr Tyr Arg Tyr Gly Gly His Ser Met Ser Asp Pro Gly Thr Thr
 305 310 315 320

Tyr Arg Thr Arg Asp Glu Ile Gln His Met Arg Ser Lys Asn Asp Pro

325 330 335

Ile Ala Gly Leu Lys Met His Leu Ile Asp Leu Gly Ile Ala Thr Glu

340 345 350

Ala Glu Val Lys Ala Tyr Asp Lys Ser Ala Arg Lys Tyr Val Asp Glu

355 360 365

Gln Val Glu Leu Ala Asp Ala Ala Pro Pro Pro Glu Ala Lys Leu Ser

370 375 380

Ile Leu Phe Glu Asp Val Tyr Val Lys Gly Thr Glu Thr Pro Thr Leu

385 390 395 400

Arg Gly Arg Ile Pro Glu Asp Thr Trp Asp Phe Lys Lys Gln Gly Phe

405 410 415

Ala Ser Arg Asp

420

<210> 65

<211> 366

<212> PRT

<213> *Saccharomyces cerevisiae*

<400> 65

Met Phe Ser Arg Leu Pro Thr Ser Leu Ala Arg Asn Val Ala Arg Arg

1 5 10 15

Ala Pro Thr Ser Phe Val Arg Pro Ser Ala Ala Ala Ala Ala Leu Arg

20 25 30

Phe Ser Ser Thr Lys Thr Met Thr Val Arg Glu Ala Leu Asn Ser Ala

35 40 45

Met Ala Glu Glu Leu Asp Arg Asp Asp Asp Val Phe Leu Ile Gly Glu

50 55 60

Glu Val Ala Gln Tyr Asn Gly Ala Tyr Lys Val Ser Lys Gly Leu Leu

65 70 75 80

Asp Arg Phe Gly Glu Arg Arg Val Val Asp Thr Pro Ile Thr Glu Tyr

85 90 95

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

 145 150 155 160
 Gly Ala Gln His Ser Gln Asp Phe Ser Pro Trp Tyr Gly Ser Ile Pro
 165 170 175
 Gly Leu Lys Val Leu Val Pro Tyr Ser Ala Glu Asp Ala Arg Gly Leu
 180 185 190
 Leu Lys Ala Ala Ile Arg Asp Pro Asn Pro Val Val Phe Leu Glu Asn
 195 200 205
 Glu Leu Leu Tyr Gly Glu Ser Phe Glu Ile Ser Glu Glu Ala Leu Ser

 210 215 220
 Pro Glu Phe Thr Leu Pro Tyr Lys Ala Lys Ile Glu Arg Glu Gly Thr
 225 230 235 240
 Asp Ile Ser Ile Val Thr Tyr Thr Arg Asn Val Gln Phe Ser Leu Glu
 245 250 255
 Ala Ala Glu Ile Leu Gln Lys Lys Tyr Gly Val Ser Ala Glu Val Ile
 260 265 270
 Asn Leu Arg Ser Ile Arg Pro Leu Asp Thr Glu Ala Ile Ile Lys Thr

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

340 345 350
 Pro Thr Ile Val Lys Ala Val Lys Glu Val Leu Ser Ile Glu
 355 360 365
 <210> 66
 <211> 499
 <212> PRT
 <213> *Saccharomyces cerevisiae*
 <400> 66
 Met Leu Arg Ile Arg Ser Leu Leu Asn Asn Lys Arg Ala Phe Ser Ser
 1 5 10 15
 Thr Val Arg Thr Leu Thr Ile Asn Lys Ser His Asp Val Val Ile Ile
 20 25 30
 Gly Gly Gly Pro Ala Gly Tyr Val Ala Ala Ile Lys Ala Ala Gln Leu
 35 40 45
 Gly Phe Asn Thr Ala Cys Val Glu Lys Arg Gly Lys Leu Gly Gly Thr
 50 55 60
 Cys Leu Asn Val Gly Cys Ile Pro Ser Lys Ala Leu Leu Asn Asn Ser
 65 70 75 80
 His Leu Phe His Gln Met His Thr Glu Ala Gln Lys Arg Gly Ile Asp
 85 90 95
 Val Asn Gly Asp Ile Lys Ile Asn Val Ala Asn Phe Gln Lys Ala Lys
 100 105 110
 Asp Asp Ala Val Lys Gln Leu Thr Gly Gly Ile Glu Leu Leu Phe Lys
 115 120 125
 Lys Asn Lys Val Thr Tyr Tyr Lys Gly Asn Gly Ser Phe Glu Asp Glu
 130 135 140
 Thr Lys Ile Arg Val Thr Pro Val Asp Gly Leu Glu Gly Thr Val Lys
 145 150 155 160
 Glu Asp His Ile Leu Asp Val Lys Asn Ile Ile Val Ala Thr Gly Ser
 165 170 175
 Glu Val Thr Pro Phe Pro Gly Ile Glu Ile Asp Glu Glu Lys Ile Val
 180 185 190

Ser Ser Thr Gly Ala Leu Ser Leu Lys Glu Ile Pro Lys Arg Leu Thr
195 200 205

Ile Ile Gly Gly Gly Ile Ile Gly Leu Glu Met Gly Ser Val Tyr Ser
210 215 220

Arg Leu Gly Ser Lys Val Thr Val Val Glu Phe Gln Pro Gln Ile Gly
225 230 235 240

Ala Ser Met Asp Gly Glu Val Ala Lys Ala Thr Gln Lys Phe Leu Lys
245 250 255

Lys Gln Gly Leu Asp Phe Lys Leu Ser Thr Lys Val Ile Ser Ala Lys
260 265 270

Arg Asn Asp Asp Lys Asn Val Val Glu Ile Val Val Glu Asp Thr Lys
275 280 285

Thr Asn Lys Gln Glu Asn Leu Glu Ala Glu Val Leu Leu Val Ala Val
290 295 300

Gly Arg Arg Pro Tyr Ile Ala Gly Leu Gly Ala Glu Lys Ile Gly Leu
305 310 315 320

Glu Val Asp Lys Arg Gly Arg Leu Val Ile Asp Asp Gln Phe Asn Ser
325 330 335

Lys Phe Pro His Ile Lys Val Val Gly Asp Val Thr Phe Gly Pro Met
340 345 350

Leu Ala His Lys Ala Glu Glu Glu Gly Ile Ala Ala Val Glu Met Leu
355 360 365

Lys Thr Gly His Gly His Val Asn Tyr Asn Asn Ile Pro Ser Val Met
370 375 380

Tyr Ser His Pro Glu Val Ala Trp Val Gly Lys Thr Glu Glu Gln Leu
385 390 395 400

Lys Glu Ala Gly Ile Asp Tyr Lys Ile Gly Lys Phe Pro Phe Ala Ala
405 410 415

Asn Ser Arg Ala Lys Thr Asn Gln Asp Thr Glu Gly Phe Val Lys Ile
420 425 430

Leu Ile Asp Ser Lys Thr Glu Arg Ile Leu Gly Ala His Ile Ile Gly

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

Leu Ser Glu Ala Phe Lys Glu Ala Asn Met Ala Ala Tyr Asp Lys Ala
 485 490 495
 Ile His Cys

<210> 67

<211> 887

<212> PRT

<213> Escherichia coli

<400> 67

Met Ser Glu Arg Phe Pro Asn Asp Val Asp Pro Ile Glu Thr Arg Asp
 1 5 10 15
 Trp Leu Gln Ala Ile Glu Ser Val Ile Arg Glu Glu Gly Val Glu Arg
 20 25 30
 Ala Gln Tyr Leu Ile Asp Gln Leu Leu Ala Glu Ala Arg Lys Gly Gly

35 40 45
 Val Asn Val Ala Ala Gly Thr Gly Ile Ser Asn Tyr Ile Asn Thr Ile
 50 55 60
 Pro Val Glu Glu Gln Pro Glu Tyr Pro Gly Asn Leu Glu Leu Glu Arg
 65 70 75 80
 Arg Ile Arg Ser Ala Ile Arg Trp Asn Ala Ile Met Thr Val Leu Arg
 85 90 95
 Ala Ser Lys Lys Asp Leu Glu Leu Gly Gly His Met Ala Ser Phe Gln

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

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

385 390 395 400
 Ala Glu Gly Lys Asn Ile Ala His Gln Val Lys Lys Met Asn Met Asp
 405 410 415
 Gly Val Arg His Ile Arg Asp Arg Phe Asn Val Pro Val Ser Asp Ala

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

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

 545 550 555 560
 Glu Lys Gly Gln Ile Leu Gln Glu Gly Ile Asn Glu Leu Gly Ala Gly
 565 570 575
 Cys Ser Trp Leu Ala Ala Ala Thr Ser Tyr Ser Thr Asn Asn Leu Pro
 580 585 590
 Met Ile Pro Phe Tyr Ile Tyr Tyr Ser Met Phe Gly Phe Gln Arg Ile
 595 600 605
 Gly Asp Leu Cys Trp Ala Ala Gly Asp Gln Gln Ala Arg Gly Phe Leu

 610 615 620
 Ile Gly Gly Thr Ser Gly Arg Thr Thr Leu Asn Gly Glu Gly Leu Gln
 625 630 635 640

His Glu Asp Gly His Ser His Ile Gln Ser Leu Thr Ile Pro Asn Cys
645 650 655

Ile Ser Tyr Asp Pro Ala Tyr Ala Tyr Glu Val Ala Val Ile Met His
660 665 670

Asp Gly Leu Glu Arg Met Tyr Gly Glu Lys Gln Glu Asn Val Tyr Tyr
675 680 685

Tyr Ile Thr Thr Leu Asn Glu Asn Tyr His Met Pro Ala Met Pro Glu
690 695 700

Gly Ala Glu Glu Gly Ile Arg Lys Gly Ile Tyr Lys Leu Glu Thr Ile
705 710 715 720

Glu Gly Ser Lys Gly Lys Val Gln Leu Leu Gly Ser Gly Ser Ile Leu
725 730 735

Arg His Val Arg Glu Ala Ala Glu Ile Leu Ala Lys Asp Tyr Gly Val
740 745 750

Gly Ser Asp Val Tyr Ser Val Thr Ser Phe Thr Glu Leu Ala Arg Asp
755 760 765

Gly Gln Asp Cys Glu Arg Trp Asn Met Leu His Pro Leu Glu Thr Pro
770 775 780

Arg Val Pro Tyr Ile Ala Gln Val Met Asn Asp Ala Pro Ala Val Ala
785 790 795 800

Ser Thr Asp Tyr Met Lys Leu Phe Ala Glu Gln Val Arg Thr Tyr Val
805 810 815

Pro Ala Asp Asp Tyr Arg Val Leu Gly Thr Asp Gly Phe Gly Arg Ser
820 825 830

Asp Ser Arg Glu Asn Leu Arg His His Phe Glu Val Asp Ala Ser Tyr
835 840 845

Val Val Val Ala Ala Leu Gly Glu Leu Ala Lys Arg Gly Glu Ile Asp
850 855 860

Lys Lys Val Val Ala Asp Ala Ile Ala Lys Phe Asn Ile Asp Ala Asp
865 870 875 880

Lys Val Asn Pro Arg Leu Ala

885

<210> 68

<211> 630

<212> PRT

<213> Escherichia coli

<400> 68

Met Ala Ile Glu Ile Lys Val Pro Asp Ile Gly Ala Asp Glu Val Glu

1 5 10 15

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

20 25 30

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

35 40 45

Pro Gln Ala Gly Ile Val Lys Glu Ile Lys Val Ser Val Gly Asp Lys

50 55 60

Thr Gln Thr Gly Ala Leu Ile Met Ile Phe Asp Ser Ala Asp Gly Ala

65 70 75 80

Ala Asp Ala Ala Pro Ala Gln Ala Glu Glu Lys Lys Glu Ala Ala Pro

85 90 95

Ala Ala Ala Pro Ala Ala Ala Ala Ala Lys Asp Val Asn Val Pro Asp

100 105 110

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

115 120 125

Asp Lys Val Glu Ala Glu Gln Ser Leu Ile Thr Val Glu Gly Asp Lys

130 135 140

Ala Ser Met Glu Val Pro Ala Pro Phe Ala Gly Thr Val Lys Glu Ile

145 150 155 160

Lys Val Asn Val Gly Asp Lys Val Ser Thr Gly Ser Leu Ile Met Val

165 170 175

Phe Glu Val Ala Gly Glu Ala Gly Ala Ala Ala Pro Ala Ala Lys Gln

180 185 190

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

195 200 205

Val Asn Val Pro Asp Ile Gly Gly Asp Glu Val Glu Val Thr Glu Val
210 215 220

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

225 230 235 240

Val Glu Gly Asp Lys Ala Ser Met Glu Val Pro Ala Pro Phe Ala Gly
245 250 255

Val Val Lys Glu Leu Lys Val Asn Val Gly Asp Lys Val Lys Thr Gly
260 265 270

Ser Leu Ile Met Ile Phe Glu Val Glu Gly Ala Ala Pro Ala Ala Ala
275 280 285

Pro Ala Lys Gln Glu Ala Ala Ala Pro Ala Pro Ala Ala Lys Ala Glu

290 295 300

Ala Pro Ala Ala Ala Pro Ala Ala Lys Ala Glu Gly Lys Ser Glu Phe
305 310 315 320

Ala Glu Asn Asp Ala Tyr Val His Ala Thr Pro Leu Ile Arg Arg Leu
325 330 335

Ala Arg Glu Phe Gly Val Asn Leu Ala Lys Val Lys Gly Thr Gly Arg
340 345 350

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

355 360 365

Ile Lys Arg Ala Glu Ala Ala Pro Ala Ala Thr Gly Gly Gly Ile Pro
370 375 380

Gly Met Leu Pro Trp Pro Lys Val Asp Phe Ser Lys Phe Gly Glu Ile
385 390 395 400

Glu Glu Val Glu Leu Gly Arg Ile Gln Lys Ile Ser Gly Ala Asn Leu
405 410 415

Ser Arg Asn Trp Val Met Ile Pro His Val Thr His Phe Asp Lys Thr

420 425 430

Asp Ile Thr Glu Leu Glu Ala Phe Arg Lys Gln Gln Asn Glu Glu Ala
435 440 445

Ala Lys Arg Lys Leu Asp Val Lys Ile Thr Pro Val Val Phe Ile Met

450 455 460
Lys Ala Val Ala Ala Ala Leu Glu Gln Met Pro Arg Phe Asn Ser Ser
465 470 475 480
Leu Ser Glu Asp Gly Gln Arg Leu Thr Leu Lys Lys Tyr Ile Asn Ile

485 490 495
Gly Val Ala Val Asp Thr Pro Asn Gly Leu Val Val Pro Val Phe Lys
500 505 510
Asp Val Asn Lys Lys Gly Ile Ile Glu Leu Ser Arg Glu Leu Met Thr
515 520 525
Ile Ser Lys Lys Ala Arg Asp Gly Lys Leu Thr Ala Gly Glu Met Gln
530 535 540
Gly Gly Cys Phe Thr Ile Ser Ser Ile Gly Gly Leu Gly Thr Thr His

545 550 555 560
Phe Ala Pro Ile Val Asn Ala Pro Glu Val Ala Ile Leu Gly Val Ser
565 570 575
Lys Ser Ala Met Glu Pro Val Trp Asn Gly Lys Glu Phe Val Pro Arg
580 585 590
Leu Met Leu Pro Ile Ser Leu Ser Phe Asp His Arg Val Ile Asp Gly
595 600 605
Ala Asp Gly Ala Arg Phe Ile Thr Ile Ile Asn Asn Thr Leu Ser Asp

610 615 620
Ile Arg Arg Leu Val Met
625 630
<210> 69
<211> 474
<212> PRT
<213> Escherichia coli
<400> 69
Met Ser Thr Glu Ile Lys Thr Gln Val Val Val Leu Gly Ala Gly Pro
1 5 10 15
Ala Gly Tyr Ser Ala Ala Phe Arg Cys Ala Asp Leu Gly Leu Glu Thr
20 25 30

Val Ile Val Glu Arg Tyr Asn Thr Leu Gly Gly Val Cys Leu Asn Val

35

40

45

Gly Cys Ile Pro Ser Lys Ala Leu Leu His Val Ala Lys Val Ile Glu

50

55

60

Glu Ala Lys Ala Leu Ala Glu His Gly Ile Val Phe Gly Glu Pro Lys

65

70

75

80

Thr Asp Ile Asp Lys Ile Arg Thr Trp Lys Glu Lys Val Ile Asn Gln

85

90

95

Leu Thr Gly Gly Leu Ala Gly Met Ala Lys Gly Arg Lys Val Lys Val

100

105

110

Val Asn Gly Leu Gly Lys Phe Thr Gly Ala Asn Thr Leu Glu Val Glu

115

120

125

Gly Glu Asn Gly Lys Thr Val Ile Asn Phe Asp Asn Ala Ile Ile Ala

130

135

140

Ala Gly Ser Arg Pro Ile Gln Leu Pro Phe Ile Pro His Glu Asp Pro

145

150

155

160

Arg Ile Trp Asp Ser Thr Asp Ala Leu Glu Leu Lys Glu Val Pro Glu

165

170

175

Arg Leu Leu Val Met Gly Gly Gly Ile Ile Gly Leu Glu Met Gly Thr

180

185

190

Val Tyr His Ala Leu Gly Ser Gln Ile Asp Val Val Glu Met Phe Asp

195

200

205

Gln Val Ile Pro Ala Ala Asp Lys Asp Ile Val Lys Val Phe Thr Lys

210

215

220

Arg Ile Ser Lys Lys Phe Asn Leu Met Leu Glu Thr Lys Val Thr Ala

225

230

235

240

Val Glu Ala Lys Glu Asp Gly Ile Tyr Val Thr Met Glu Gly Lys Lys

245

250

255

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

260

265

270

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

275 280 285
 Val Asp Asp Arg Gly Phe Ile Arg Val Asp Lys Gln Leu Arg Thr Asn

 290 295 300
 Val Pro His Ile Phe Ala Ile Gly Asp Ile Val Gly Gln Pro Met Leu
 305 310 315 320
 Ala His Lys Gly Val His Glu Gly His Val Ala Ala Glu Val Ile Ala
 325 330 335
 Gly Lys Lys His Tyr Phe Asp Pro Lys Val Ile Pro Ser Ile Ala Tyr
 340 345 350
 Thr Glu Pro Glu Val Ala Trp Val Gly Leu Thr Glu Lys Glu Ala Lys

 355 360 365
 Glu Lys Gly Ile Ser Tyr Glu Thr Ala Thr Phe Pro Trp Ala Ala Ser
 370 375 380
 Gly Arg Ala Ile Ala Ser Asp Cys Ala Asp Gly Met Thr Lys Leu Ile
 385 390 395 400
 Phe Asp Lys Glu Ser His Arg Val Ile Gly Gly Ala Ile Val Gly Thr
 405 410 415
 Asn Gly Gly Glu Leu Leu Gly Glu Ile Gly Leu Ala Ile Glu Met Gly

 420 425 430
 Cys Asp Ala Glu Asp Ile Ala Leu Thr Ile His Ala His Pro Thr Leu
 435 440 445
 His Glu Ser Val Gly Leu Ala Ala Glu Val Phe Glu Gly Ser Ile Thr
 450 455 460
 Asp Leu Pro Asn Pro Lys Ala Lys Lys Lys
 465 470
 <210> 70
 <211> 371
 <212> PRT
 <213> Bacillus subtilis
 <400> 70
 Met Ala Ala Lys Thr Lys Lys Ala Ile Val Asp Ser Lys Lys Gln Phe

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

 65 70 75 80
 Thr Ala Gly Gln Glu Ala Ser Gln Ile Ala Thr His Phe Ala Leu Glu
 85 90 95
 Lys Glu Asp Phe Val Leu Pro Gly Tyr Arg Asp Val Pro Gln Leu Ile
 100 105 110
 Trp His Gly Leu Pro Leu Tyr Gln Ala Phe Leu Phe Ser Arg Gly His
 115 120 125
 Phe Arg Gly Asn Gln Met Pro Asp Asp Val Asn Ala Leu Ser Pro Gln

 130 135 140
 Ile Ile Ile Gly Ala Gln Tyr Ile Gln Thr Ala Gly Val Ala Leu Gly
 145 150 155 160
 Leu Lys Lys Arg Gly Lys Lys Ala Val Ala Ile Thr Tyr Thr Gly Asp
 165 170 175
 Gly Gly Ala Ser Gln Gly Asp Phe Tyr Glu Gly Ile Asn Phe Ala Gly
 180 185 190
 Ala Tyr Lys Ala Pro Ala Ile Phe Val Val Gln Asn Asn Arg Tyr Ala

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

Asn Gly Glu Gly Pro Thr Leu Ile Glu Thr Leu Thr Phe Arg Tyr Gly

260 265 270
Pro His Thr Met Ala Gly Asp Asp Pro Thr Lys Tyr Arg Thr Lys Glu
275 280 285
Ile Glu Asn Glu Trp Glu Gln Lys Asp Pro Leu Val Arg Phe Arg Ala
290 295 300
Phe Leu Glu Asn Lys Gly Leu Trp Ser Glu Glu Glu Glu Ala Lys Val
305 310 315 320
Ile Glu Asp Ala Lys Glu Glu Ile Lys Gln Ala Ile Lys Lys Ala Asp

325 330 335
Ala Glu Pro Lys Gln Lys Val Thr Asp Leu Met Lys Ile Met Tyr Glu
340 345 350
Lys Met Pro His Asn Leu Glu Glu Gln Phe Glu Ile Tyr Thr Gln Lys
355 360 365
Glu Ser Lys

370

<210> 71

<211> 325

<212> PRT

<213> Bacillus subtilis

<400> 71

Met Ala Gln Met Thr Met Ile Gln Ala Ile Thr Asp Ala Leu Arg Thr
1 5 10 15

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

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

Val Val Ile Pro Ser Thr Pro Tyr Asp Ala Lys Gly Leu Leu Ile Ser
145 150 155 160
Ala Ile Arg Asp Asn Asp Pro Val Val Phe Leu Glu His Met Lys Leu
165 170 175
Tyr Arg Ser Phe Arg Gln Glu Val Pro Glu Glu Glu Tyr Thr Ile Glu
180 185 190
Leu Gly Lys Ala Asp Val Lys Arg Glu Gly Thr Asp Leu Ser Ile Ile
195 200 205

Thr Tyr Gly Ala Met Val His Glu Ser Leu Lys Ala Ala Asp Glu Leu
210 215 220
Glu Lys Asp Gly Ile Ser Ala Glu Val Val Asp Leu Arg Thr Val Ser
225 230 235 240
Pro Leu Asp Ile Asp Thr Ile Ile Ala Ser Val Glu Lys Thr Gly Arg
245 250 255
Ala Ile Val Val Gln Glu Ala Gln Lys Gln Ala Gly Ile Ala Ala Asn
260 265 270

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

<210> 72

<211> 442

<212> PRT

<213> *Bacillus subtilis*

<400> 72

Met Ala Phe Glu Phe Lys Leu Pro Asp Ile Gly Glu Gly Ile His Glu

1 5 10 15

Gly Glu Ile Val Lys Trp Phe Val Lys Pro Asn Asp Glu Val Asp Glu

20 25 30

Asp Asp Val Leu Ala Glu Val Gln Asn Asp Lys Ala Val Val Glu Ile

35 40 45

Pro Ser Pro Val Lys Gly Lys Val Leu Glu Leu Lys Val Glu Glu Gly

50 55 60

Thr Val Ala Thr Val Gly Gln Thr Ile Ile Thr Phe Asp Ala Pro Gly

65 70 75 80

Tyr Glu Asp Leu Gln Phe Lys Gly Ser Asp Glu Ser Asp Asp Ala Lys

85 90 95

Thr Glu Ala Gln Val Gln Ser Thr Ala Glu Ala Gly Gln Asp Val Ala

100 105 110

Lys Glu Glu Gln Ala Gln Glu Pro Ala Lys Ala Thr Gly Ala Gly Gln

115 120 125

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

130 135 140

Ser Val Arg Lys Tyr Ala Arg Glu Lys Gly Val Asp Ile Arg Lys Val

145 150 155 160

Thr Gly Ser Gly Asn Asn Gly Arg Val Val Lys Glu Asp Ile Asp Ser

165 170 175

Phe Val Asn Gly Gly Ala Gln Glu Ala Ala Pro Gln Glu Thr Ala Ala

180 185 190

Pro Gln Glu Thr Ala Ala Lys Pro Ala Ala Ala Pro Ala Pro Glu Gly

195 200 205

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

 Lys Gln Val Ala Ala Asp Gln Gly Ile Lys Leu Thr Tyr Leu Pro Tyr
 260 265 270
 Val Val Lys Ala Leu Thr Ser Ala Leu Lys Lys Phe Pro Val Leu Asn
 275 280 285
 Thr Ser Ile Asp Asp Lys Thr Asp Glu Val Ile Gln Lys His Tyr Phe
 290 295 300
 Asn Ile Gly Ile Ala Ala Asp Thr Glu Lys Gly Leu Leu Val Pro Val
 305 310 315 320

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

 Ile Gly Arg Ile Ala Glu Lys Ala Ile Val Arg Asp Gly Glu Ile Val
 385 390 395 400
 Ala Ala Pro Val Leu Ala Leu Ser Leu Ser Phe Asp His Arg Met Ile
 405 410 415
 Asp Gly Ala Thr Ala Gln Asn Ala Leu Asn His Ile Lys Arg Leu Leu
 420 425 430
 Asn Asp Pro Gln Leu Ile Leu Met Glu Ala
 435 440

<210> 73

<211> 470

<212>

PRT

<213> Bacillus subtilis

<400> 73

Met Val Val Gly Asp Phe Pro Ile Glu Thr Asp Thr Leu Val Ile Gly

1 5 10 15

Ala Gly Pro Gly Gly Tyr Val Ala Ala Ile Arg Ala Ala Gln Leu Gly

20 25 30

Gln Lys Val Thr Val Val Glu Lys Ala Thr Leu Gly Gly Val Cys Leu

35 40 45

Asn Val Gly Cys Ile Pro Ser Lys Ala Leu Ile Asn Ala Gly His Arg

50 55 60

Tyr Glu Asn Ala Lys His Ser Asp Asp Met Gly Ile Thr Ala Glu Asn

65 70 75 80

Val Thr Val Asp Phe Thr Lys Val Gln Glu Trp Lys Ala Ser Val Val

85 90 95

Asn Lys Leu Thr Gly Gly Val Ala Gly Leu Leu Lys Gly Asn Lys Val

100 105 110

Asp Val Val Lys Gly Glu Ala Tyr Phe Val Asp Ser Asn Ser Val Arg

115 120 125

Val Met Asp Glu Asn Ser Ala Gln Thr Tyr Thr Phe Lys Asn Ala Ile

130 135 140

Ile Ala Thr Gly Ser Arg Pro Ile Glu Leu Pro Asn Phe Lys Tyr Ser

145 150 155 160

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

165 170 175

Lys Lys Leu Val Val Ile Gly Gly Gly Tyr Ile Gly Thr Glu Leu Gly

180 185 190

Thr Ala Tyr Ala Asn Phe Gly Thr Glu Leu Val Ile Leu Glu Gly Gly

195 200 205

Asp Glu Ile Leu Pro Gly Phe Glu Lys Gln Met Ser Ser Leu Val Thr

210 215 220

Arg Arg Leu Lys Lys Lys Gly Asn Val Glu Ile His Thr Asn Ala Met
 225 230 235 240
 Ala Lys Gly Val Glu Glu Arg Pro Asp Gly Val Thr Val Thr Phe Glu
 245 250 255

 Val Lys Gly Glu Glu Lys Thr Val Asp Ala Asp Tyr Val Leu Ile Thr
 260 265 270
 Val Gly Arg Arg Pro Asn Thr Asp Glu Leu Gly Leu Glu Gln Val Gly
 275 280 285
 Ile Glu Met Thr Asp Arg Gly Ile Val Lys Thr Asp Lys Gln Cys Arg
 290 295 300
 Thr Asn Val Pro Asn Ile Tyr Ala Ile Gly Asp Ile Ile Glu Gly Pro
 305 310 315 320

 Pro Leu Ala His Lys Ala Ser Tyr Glu Gly Lys Ile Ala Ala Glu Ala
 325 330 335
 Ile Ala Gly Glu Pro Ala Glu Ile Asp Tyr Leu Gly Ile Pro Ala Val
 340 345 350
 Val Phe Ser Glu Pro Glu Leu Ala Ser Val Gly Tyr Thr Glu Ala Gln
 355 360 365
 Ala Lys Glu Glu Gly Leu Asp Ile Val Ala Ala Lys Phe Pro Phe Ala
 370 375 380

 Ala Asn Gly Arg Ala Leu Ser Leu Asn Glu Thr Asp Gly Phe Met Lys
 385 390 395 400
 Leu Ile Thr Arg Lys Glu Asp Gly Leu Val Ile Gly Ala Gln Ile Ala
 405 410 415
 Gly Ala Ser Ala Ser Asp Met Ile Ser Glu Leu Ser Leu Ala Ile Glu
 420 425 430
 Gly Gly Met Thr Ala Glu Asp Ile Ala Met Thr Ile His Ala His Pro
 435 440 445

 Thr Leu Gly Glu Ile Thr Met Glu Ala Ala Glu Val Ala Ile Gly Ser
 450 455 460
 Pro Ile His Ile Val Lys

| | | | | | | | | | | | | | | | | |
|-------|--------------------------|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|--|
| 465 | 470 | | | | | | | | | | | | | | | |
| <210> | 74 | | | | | | | | | | | | | | | |
| <211> | 2123 | | | | | | | | | | | | | | | |
| <212> | PRT | | | | | | | | | | | | | | | |
| <213> | Saccharomyces cerevisiae | | | | | | | | | | | | | | | |
| <400> | 74 | | | | | | | | | | | | | | | |
| Met | Arg | Ser | Ile | Arg | Lys | Trp | Ala | Tyr | Glu | Thr | Phe | Asn | Asp | Glu | Lys | |
| 1 | | | | 5 | | | | | 10 | | | | | 15 | | |
| Ile | Ile | Gln | Phe | Val | Val | Met | Ala | Thr | Pro | Asp | Asp | Leu | His | Ala | Asn | |
| | | | | 20 | | | | 25 | | | | | 30 | | | |
| Ser | Glu | Tyr | Ile | Arg | Met | Ala | Asp | Gln | Tyr | Val | Gln | Val | Pro | Gly | Gly | |
| | | | | 35 | | | 40 | | | | | 45 | | | | |
| Thr | Asn | Asn | Asn | Asn | Tyr | Ala | Asn | Ile | Asp | Leu | Ile | Leu | Asp | Val | Ala | |
| | 50 | | | | | 55 | | | | | 60 | | | | | |
| Glu | Gln | Thr | Asp | Val | Asp | Ala | Val | Trp | Ala | Gly | Trp | Gly | His | Ala | Ser | |
| 65 | | | | | 70 | | | | | 75 | | | | | 80 | |
| Glu | Asn | Pro | Cys | Leu | Pro | Glu | Leu | Leu | Ala | Ser | Ser | Gln | Arg | Lys | Ile | |
| | | | | 85 | | | | | 90 | | | | | 95 | | |
| Leu | Phe | Ile | Gly | Pro | Pro | Gly | Arg | Ala | Met | Arg | Ser | Leu | Gly | Asp | Lys | |
| | | | | 100 | | | | 105 | | | | | 110 | | | |
| Ile | Ser | Ser | Thr | Ile | Val | Ala | Gln | Ser | Ala | Lys | Ile | Pro | Cys | Ile | Pro | |
| | | | | 115 | | | | 120 | | | | | 125 | | | |
| Trp | Ser | Gly | Ser | His | Ile | Asp | Thr | Ile | His | Ile | Asp | Asn | Lys | Thr | Asn | |
| | 130 | | | | | 135 | | | | | 140 | | | | | |
| Phe | Val | Ser | Val | Pro | Asp | Asp | Val | Tyr | Val | Arg | Gly | Cys | Cys | Ser | Ser | |
| 145 | | | | | 150 | | | | | 155 | | | | | 160 | |
| Pro | Glu | Asp | Ala | Leu | Glu | Lys | Ala | Lys | Leu | Ile | Gly | Phe | Pro | Val | Met | |
| | | | | 165 | | | | | 170 | | | | | 175 | | |
| Ile | Lys | Ala | Ser | Glu | Gly | Gly | Gly | Gly | Lys | Gly | Ile | Arg | Arg | Val | Asp | |
| | | | | 180 | | | | 185 | | | | | 190 | | | |
| Asn | Glu | Asp | Asp | Phe | Ile | Ala | Leu | Tyr | Arg | Gln | Ala | Val | Asn | Glu | Thr | |
| | 195 | | | | | 200 | | | | | | 205 | | | | |

Pro Gly Ser Pro Met Phe Val Met Lys Val Val Thr Asp Ala Arg His
210 215 220

Leu Glu Val Gln Leu Leu Ala Asp Gln Tyr Gly Thr Asn Ile Thr Leu
225 230 235 240

Phe Gly Arg Asp Cys Ser Ile Gln Arg Arg His Gln Lys Ile Ile Glu
245 250 255

Glu Ala Pro Val Thr Ile Thr Lys Pro Glu Thr Phe Gln Arg Met Glu
260 265 270

Arg Ala Ala Ile Arg Leu Gly Glu Leu Val Gly Tyr Val Ser Ala Gly
275 280 285

Thr Val Glu Tyr Leu Tyr Ser Pro Lys Asp Asp Lys Phe Tyr Phe Leu
290 295 300

Glu Leu Asn Pro Arg Leu Gln Val Glu His Pro Thr Thr Glu Met Ile
305 310 315 320

Ser Gly Val Asn Leu Pro Ala Thr Gln Leu Gln Ile Ala Met Gly Ile
325 330 335

Pro Met His Met Ile Ser Asp Ile Arg Lys Leu Tyr Gly Leu Asp Pro
340 345 350

Thr Gly Thr Ser Tyr Ile Asp Phe Lys Asn Leu Lys Arg Pro Ser Pro
355 360 365

Lys Gly His Cys Ile Ser Cys Arg Ile Thr Ser Glu Asp Pro Asn Glu
370 375 380

Gly Phe Lys Pro Ser Thr Gly Lys Ile His Glu Leu Asn Phe Arg Ser
385 390 395 400

Ser Ser Asn Val Trp Gly Tyr Phe Ser Val Gly Asn Asn Gly Ala Ile
405 410 415

His Ser Phe Ser Asp Ser Gln Phe Gly His Ile Phe Ala Val Gly Asn
420 425 430

Asp Arg Gln Asp Ala Lys Gln Asn Met Val Leu Ala Leu Lys Asp Phe
435 440 445

Ser Ile Arg Gly Glu Phe Lys Thr Pro Ile Glu Tyr Leu Ile Glu Leu

450 455 460
 Leu Glu Thr Arg Asp Phe Glu Ser Asn Asn Ile Ser Thr Gly Trp Leu
 465 470 475 480

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

Asp Asn Asn Arg Tyr Leu Phe Asn Val Ala Gln Ser Ser Glu Glu Gln
 545 550 555 560
 Phe Ile Leu Ser Ile Asn Lys Ser Gln Cys Glu Val Asn Val Gln Lys
 565 570 575
 Leu Ser Ser Asp Cys Leu Leu Ile Ser Val Asp Gly Lys Cys His Thr
 580 585 590
 Val Tyr Trp Lys Asp Asp Ile Arg Gly Thr Arg Leu Ser Ile Asp Ser
 595 600 605

Asn Thr Ile Phe Leu Glu Ala Glu Leu Asn Pro Thr Gln Val Ile Ser
 610 615 620
 Pro Thr Pro Gly Lys Leu Val Lys Tyr Leu Val Arg Ser Gly Asp His
 625 630 635 640
 Val Phe Ala Gly Gln Gln Tyr Ala Glu Ile Glu Ile Met Lys Met Gln
 645 650 655
 Met Pro Leu Val Ala Lys Ser Asp Gly Val Ile Glu Leu Leu Arg Gln
 660 665 670

Pro Gly Ser Ile Ile Glu Ala Gly Asp Val Ile Ala Lys Leu Thr Leu
 675 680 685
 Asp Ser Pro Ser Lys Ala Asn Glu Ser Ser Leu Tyr Arg Gly Glu Leu
 690 695 700

Pro Val Leu Gly Pro Pro Leu Ile Glu Gly Ser Arg Pro Asn His Lys
705 710 715 720

Leu Arg Val Leu Ile Asn Arg Leu Glu Asn Ile Leu Asn Gly Tyr His
725 730 735

Glu Asn Ser Gly Ile Glu Thr Thr Leu Lys Glu Leu Ile Lys Ile Leu
740 745 750

Arg Asp Gly Arg Leu Pro Tyr Ser Glu Trp Asp Ser Gln Ile Ser Thr
755 760 765

Val Arg Asn Arg Leu Pro Arg Gln Leu Asn Glu Gly Leu Gly Asn Leu
770 775 780

Val Lys Lys Ser Val Ser Phe Pro Ala Lys Glu Leu His Lys Leu Met
785 790 795 800

Lys Arg Tyr Leu Glu Glu Asn Thr Asn Asp His Val Val Tyr Val Ala
805 810 815

Leu Gln Pro Leu Leu Lys Ile Ser Glu Arg Tyr Ser Glu Gly Leu Ala
820 825 830

Asn His Glu Cys Glu Ile Phe Leu Lys Leu Ile Lys Lys Tyr Tyr Ala
835 840 845

Val Glu Lys Ile Phe Glu Asn His Asp Ile His Glu Glu Arg Asn Leu
850 855 860

Leu Asn Leu Arg Arg Lys Asp Leu Thr Asn Leu Lys Lys Ile Leu Cys
865 870 875 880

Ile Ser Leu Ser His Ala Asn Val Val Ala Lys Asn Lys Leu Val Thr
885 890 895

Ala Ile Leu His Glu Tyr Glu Pro Leu Cys Gln Asp Ser Ser Lys Met
900 905 910

Ser Leu Lys Phe Arg Ala Val Ile His Asp Leu Ala Ser Leu Glu Ser
915 920 925

Lys Trp Ala Lys Glu Val Ala Val Lys Ala Arg Ser Val Leu Leu Arg
930 935 940

Gly Ile Phe Pro Pro Ile Lys Lys Arg Lys Glu His Ile Lys Thr Leu

| | | | |
|---|------|------|------|
| 945 | 950 | 955 | 960 |
| Leu Gln Leu His Ile Lys Asp Thr Gly Ala Glu Asn Ile His Ser Arg | | | |
| | 965 | 970 | 975 |
| Asn Ile Tyr Ser Cys Met Arg Asp Phe Gly Asn Leu Ile His Ser Asn | | | |
| | 980 | 985 | 990 |
| | | | |
| Leu Ile Gln Leu Gln Asp Leu Phe Phe Phe Phe Gly His Gln Asp Thr | | | |
| | 995 | 1000 | 1005 |
| Ala Leu Ser Ser Ile Ala Ser Glu Ile Tyr Ala Arg Tyr Ala Tyr | | | |
| | 1010 | 1015 | 1020 |
| Gly Asn Tyr Gln Leu Lys Ser Ile Lys Ile His Lys Gly Ala Pro | | | |
| | 1025 | 1030 | 1035 |
| Asp Leu Leu Met Ser Trp Gln Phe Ser Ser Leu Arg Asn Tyr Leu | | | |
| | 1040 | 1045 | 1050 |
| | | | |
| Val Asn Ser Asp Gly Glu Ser Asp Glu Phe Thr Lys Leu Ser Lys | | | |
| | 1055 | 1060 | 1065 |
| Pro Pro Ser Thr Ser Gly Lys Ser Ser Ala Asn Ser Phe Gly Leu | | | |
| | 1070 | 1075 | 1080 |
| Leu Val Asn Met Arg Ala Leu Glu Ser Leu Glu Lys Thr Leu Asp | | | |
| | 1085 | 1090 | 1095 |
| Glu Val Tyr Glu Gln Ile His Ile Pro Glu Glu Arg Leu Ser Ser | | | |
| | 1100 | 1105 | 1110 |
| | | | |
| Gly Glu Asn Ser Leu Ile Val Asn Ile Leu Ser Pro Ile Arg Tyr | | | |
| | 1115 | 1120 | 1125 |
| Arg Ser Glu Asn Asp Leu Ile Lys Thr Leu Lys Ile Lys Leu His | | | |
| | 1130 | 1135 | 1140 |
| Glu Asn Glu Arg Gly Leu Ser Lys Leu Lys Val Asn Arg Ile Thr | | | |
| | 1145 | 1150 | 1155 |
| Phe Ala Phe Ile Ala Ala Asn Ala Pro Ala Val Lys Phe Tyr Ser | | | |
| | 1160 | 1165 | 1170 |
| | | | |
| Phe Asp Gly Thr Thr Tyr Asp Glu Ile Ser Gln Ile Arg Asn Met | | | |
| | 1175 | 1180 | 1185 |

| | | |
|-----------------------------|---------------------|-------------|
| Asp Pro Ser Tyr Glu Ala Pro | Leu Glu Leu Gly Lys | Met Ser Asn |
| 1190 | 1195 | 1200 |
| Tyr Lys Ile Arg Ser Leu Pro | Thr Tyr Asp Ser Ser | Ile Arg Ile |
| 1205 | 1210 | 1215 |
| Phe Glu Gly Ile Ser Lys Phe | Thr Pro Leu Asp Lys | Arg Phe Phe |
| 1220 | 1225 | 1230 |
| Val Arg Lys Ile Ile Asn Ser | Phe Met Tyr Asn Asp | Gln Lys Thr |
| 1235 | 1240 | 1245 |
| Thr Glu Glu Asn Leu Lys Ala | Glu Ile Asn Ala Gln | Val Val Tyr |
| 1250 | 1255 | 1260 |
| Met Leu Glu His Leu Gly Ala | Val Asp Ile Ser Asn | Ser Asp Leu |
| 1265 | 1270 | 1275 |
| Asn His Ile Phe Leu Ser Phe | Asn Thr Val Leu Asn | Ile Pro Val |
| 1280 | 1285 | 1290 |
| His Arg Leu Glu Glu Ile Val | Ser Thr Ile Leu Lys | Thr His Glu |
| 1295 | 1300 | 1305 |
| Thr Arg Leu Phe Gln Glu Arg | Ile Thr Asp Val Glu | Ile Cys Ile |
| 1310 | 1315 | 1320 |
| Ser Val Glu Cys Leu Glu Thr | Lys Lys Pro Ala Pro | Leu Arg Leu |
| 1325 | 1330 | 1335 |
| Leu Ile Ser Asn Lys Ser Gly | Tyr Val Val Lys Ile | Glu Thr Tyr |
| 1340 | 1345 | 1350 |
| Tyr Glu Lys Ile Gly Lys Asn | Gly Asn Leu Ile Leu | Glu Pro Cys |
| 1355 | 1360 | 1365 |
| Ser Glu Gln Ser His Tyr Ser | Gln Lys Ser Leu Ser | Leu Pro Tyr |
| 1370 | 1375 | 1380 |
| Ser Val Lys Asp Trp Leu Gln | Pro Lys Arg Tyr Lys | Ala Gln Phe |
| 1385 | 1390 | 1395 |
| Met Gly Thr Thr Tyr Val Tyr | Asp Phe Pro Gly Leu | Phe His Gln |
| 1400 | 1405 | 1410 |
| Ala Ala Ile Gln Gln Trp Lys | Arg Tyr Phe Pro Lys | His Lys Leu |

| | | |
|---|------|------|
| 1415 | 1420 | 1425 |
| Asn Asp Ser Phe Phe Ser Trp Val Glu Leu Ile Glu Gln Asn Gly | | |
| 1430 | 1435 | 1440 |
| Asn Leu Ile Lys Val Asn Arg Glu Pro Gly Leu Asn Asn Ile Gly | | |
| 1445 | 1450 | 1455 |
| Met Val Ala Phe Glu Ile Met Val Gln Thr Pro Glu Tyr Pro Glu | | |
| 1460 | 1465 | 1470 |
| | | |
| Gly Arg Asn Met Ile Val Ile Ser Asn Asp Ile Thr Tyr Asn Ile | | |
| 1475 | 1480 | 1485 |
| Gly Ser Phe Gly Pro Arg Glu Asp Leu Phe Phe Asp Arg Val Thr | | |
| 1490 | 1495 | 1500 |
| Asn Tyr Ala Arg Glu Arg Gly Ile Pro Arg Ile Tyr Leu Ala Ala | | |
| 1505 | 1510 | 1515 |
| Asn Ser Gly Ala Lys Leu Gly Ile Ala Glu Glu Leu Ile Pro Leu | | |
| 1520 | 1525 | 1530 |
| | | |
| Phe Arg Val Ala Trp Asn Asp Pro Ser Asp Pro Thr Lys Gly Phe | | |
| 1535 | 1540 | 1545 |
| Gln Tyr Leu Tyr Leu Ala Pro Lys Asp Met Gln Leu Leu Lys Asp | | |
| 1550 | 1555 | 1560 |
| Ser Gly Lys Gly Asn Ser Val Val Val Glu His Lys Met Val Tyr | | |
| 1565 | 1570 | 1575 |
| Gly Glu Glu Arg Tyr Ile Ile Lys Ala Ile Val Gly Phe Glu Glu | | |
| 1580 | 1585 | 1590 |
| | | |
| Gly Leu Gly Val Glu Cys Leu Gln Gly Ser Gly Leu Ile Ala Gly | | |
| 1595 | 1600 | 1605 |
| Ala Thr Ser Lys Ala Tyr Arg Asp Ile Phe Thr Ile Thr Ala Val | | |
| 1610 | 1615 | 1620 |
| Thr Cys Arg Ser Val Gly Ile Gly Ser Tyr Leu Val Arg Leu Gly | | |
| 1625 | 1630 | 1635 |
| Gln Arg Thr Ile Gln Val Glu Asp Lys Pro Ile Ile Leu Thr Gly | | |
| 1640 | 1645 | 1650 |

| | | | |
|---------|---------------------|---------------------|-------------|
| Ala Ser | Ala Ile Asn Lys Val | Leu Gly Thr Asp Ile | Tyr Thr Ser |
| 1655 | 1660 | 1665 | |
| Asn Leu | Gln Ile Gly Gly Thr | Gln Ile Met Tyr Lys | Asn Gly Ile |
| 1670 | 1675 | 1680 | |
| Ala His | Leu Thr Ala Ser Asn | Asp Met Lys Ala Ile | Glu Lys Ile |
| 1685 | 1690 | 1695 | |
| Met Thr | Trp Leu Ser Tyr Val | Pro Ala Lys Arg Asp | Met Ser Pro |
| 1700 | 1705 | 1710 | |
| Pro Leu | Leu Glu Thr Met Asp | Arg Trp Asp Arg Asp | Val Asp Phe |
| 1715 | 1720 | 1725 | |
| Lys Pro | Ala Lys Gln Val Pro | Tyr Glu Ala Arg Trp | Leu Ile Glu |
| 1730 | 1735 | 1740 | |
| Gly Lys | Trp Asp Ser Asn Asn | Asn Phe Gln Ser Gly | Leu Phe Asp |
| 1745 | 1750 | 1755 | |
| Lys Asp | Ser Phe Phe Glu Thr | Leu Ser Gly Trp Ala | Lys Gly Val |
| 1760 | 1765 | 1770 | |
| Ile Val | Gly Arg Ala Arg Leu | Gly Gly Ile Pro Val | Gly Val Ile |
| 1775 | 1780 | 1785 | |
| Ala Val | Glu Thr Lys Thr Ile | Glu Glu Ile Ile Pro | Ala Asp Pro |
| 1790 | 1795 | 1800 | |
| Ala Asn | Leu Asp Ser Ser Glu | Phe Ser Val Lys Glu | Ala Gly Gln |
| 1805 | 1810 | 1815 | |
| Val Trp | Tyr Pro Asn Ser Ala | Phe Lys Thr Ala Gln | Thr Ile Asn |
| 1820 | 1825 | 1830 | |
| Asp Phe | Asn Tyr Gly Glu Gln | Leu Pro Leu Ile Ile | Leu Ala Asn |
| 1835 | 1840 | 1845 | |
| Trp Arg | Gly Phe Ser Gly Gly | Gln Arg Asp Met Tyr | Asn Glu Val |
| 1850 | 1855 | 1860 | |
| Leu Lys | Tyr Gly Ser Phe Ile | Val Asp Ala Leu Val | Asp Tyr Lys |
| 1865 | 1870 | 1875 | |
| Gln Pro | Ile Leu Ile Tyr Ile | Pro Pro Phe Gly Glu | Leu Arg Gly |

| | | |
|---|------|------|
| 1880 | 1885 | 1890 |
| Gly Ser Trp Val Val Ile Asp Pro Thr Ile Asn Pro Glu Gln Met | | |
| 1895 | 1900 | 1905 |
| Glu Met Tyr Ala Asp Val Glu Ser Arg Gly Gly Val Leu Glu Pro | | |
| 1910 | 1915 | 1920 |
| Asp Gly Val Val Ser Ile Lys Tyr Arg Lys Glu Lys Met Ile Glu | | |
| 1925 | 1930 | 1935 |
| Thr Met Ile Arg Leu Asp Ser Thr Tyr Gly His Leu Arg Arg Thr | | |
| 1940 | 1945 | 1950 |
| Leu Thr Glu Lys Lys Leu Ser Leu Glu Lys Gln Asn Asp Leu Thr | | |
| 1955 | 1960 | 1965 |
| Lys Arg Leu Lys Ile Arg Glu Arg Gln Leu Ile Pro Ile Tyr Asn | | |
| 1970 | 1975 | 1980 |
| Gln Ile Ser Ile Gln Phe Ala Asp Leu His Asp Arg Ser Thr Arg | | |
| 1985 | 1990 | 1995 |
| Met Leu Val Lys Gly Val Ile Arg Asn Glu Leu Glu Trp Lys Lys | | |
| 2000 | 2005 | 2010 |
| Ser Arg Arg Phe Leu Tyr Trp Arg Leu Arg Arg Arg Leu Asn Glu | | |
| 2015 | 2020 | 2025 |
| Gly Gln Val Ile Lys Arg Leu Gln Lys Lys Thr Cys Asp Asn Lys | | |
| 2030 | 2035 | 2040 |
| Thr Lys Met Lys Tyr Asp Asp Leu Leu Lys Ile Val Gln Ser Trp | | |
| 2045 | 2050 | 2055 |
| Tyr Asn Asp Leu Asp Val Asn Asp Asp Arg Ala Val Val Glu Phe | | |
| 2060 | 2065 | 2070 |
| Ile Glu Arg Asn Ser Lys Lys Ile Asp Lys Asn Ile Glu Glu Phe | | |
| 2075 | 2080 | 2085 |
| Glu Ile Ser Leu Leu Ile Asp Glu Leu Lys Lys Lys Phe Glu Asp | | |
| 2090 | 2095 | 2100 |
| Arg Arg Gly Asn Ile Val Leu Glu Glu Leu Thr Arg Leu Val Asp | | |
| 2105 | 2110 | 2115 |

Ser Lys Arg Lys Arg

2120

<210> 75

<211> 319

<212> PRT

<213> Escherichia coli

<400> 75

Met Ser Leu Asn Phe Leu Asp Phe Glu Gln Pro Ile Ala Glu Leu Glu

1 5 10 15

Ala Lys Ile Asp Ser Leu Thr Ala Val Ser Arg Gln Asp Glu Lys Leu

20 25 30

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

35 40 45

Glu Leu Thr Arg Lys Ile Phe Ala Asp Leu Gly Ala Trp Gln Ile Ala

50 55 60

Gln Leu Ala Arg His Pro Gln Arg Pro Tyr Thr Leu Asp Tyr Val Arg

65 70 75 80

Leu Ala Phe Asp Glu Phe Asp Glu Leu Ala Gly Asp Arg Ala Tyr Ala

85 90 95

Asp Asp Lys Ala Ile Val Gly Gly Ile Ala Arg Leu Asp Gly Arg Pro

100 105 110

Val Met Ile Ile Gly His Gln Lys Gly Arg Glu Thr Lys Glu Lys Ile

115 120 125

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

130 135 140

Arg Leu Met Gln Met Ala Glu Arg Phe Lys Met Pro Ile Ile Thr Phe

145 150 155 160

Ile Asp Thr Pro Gly Ala Tyr Pro Gly Val Gly Ala Glu Glu Arg Gly

165 170 175

Gln Ser Glu Ala Ile Ala Arg Asn Leu Arg Glu Met Ser Arg Leu Gly

180 185 190

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

Arg Leu Lys Glu Leu Lys Leu Ile Asp Ser Ile Ile Pro Glu Pro Leu
 260 265 270
 Gly Gly Ala His Arg Asn Pro Glu Ala Met Ala Ala Ser Leu Lys Ala
 275 280 285
 Gln Leu Leu Ala Asp Leu Ala Asp Leu Asp Val Leu Ser Thr Glu Asp
 290 295 300
 Leu Lys Asn Arg Arg Tyr Gln Arg Leu Met Ser Tyr Gly Tyr Ala
 305 310 315

<210> 76

<211> 156

<212> PRT

<213> Escherichia coli

<400> 76

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

85 90 95
Pro Asp Ala Lys Ala Phe Ile Glu Val Gly Gln Lys Val Asn Val Gly
100 105 110
Asp Thr Leu Cys Ile Val Glu Ala Met Lys Met Met Asn Gln Ile Glu

115 120 125
Ala Asp Lys Ser Gly Thr Val Lys Ala Ile Leu Val Glu Ser Gly Gln
130 135 140
Pro Val Glu Phe Asp Glu Pro Leu Val Val Ile Glu
145 150 155

<210> 77

<211> 449

<212> PRT

<213> Escherichia coli

<400> 77

Met Leu Asp Lys Ile Val Ile Ala Asn Arg Gly Glu Ile Ala Leu Arg
1 5 10 15
Ile Leu Arg Ala Cys Lys Glu Leu Gly Ile Lys Thr Val Ala Val His

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

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

Asn Arg Ala Ile Ala Lys Arg Ile Gly Tyr Pro Val Ile Ile Lys Ala

145 150 155 160

Ser Gly Gly Gly Gly Gly Arg Gly Met Arg Val Val Arg Gly Asp Ala

165 170 175

Glu Leu Ala Gln Ser Ile Ser Met Thr Arg Ala Glu Ala Lys Ala Ala

180 185 190

Phe Ser Asn Asp Met Val Tyr Met Glu Lys Tyr Leu Glu Asn Pro Arg

195 200 205

His Val Glu Ile Gln Val Leu Ala Asp Gly Gln Gly Asn Ala Ile Tyr

210 215 220

Leu Ala Glu Arg Asp Cys Ser Met Gln Arg Arg His Gln Lys Val Val

225 230 235 240

Glu Glu Ala Pro Ala Pro Gly Ile Thr Pro Glu Leu Arg Arg Tyr Ile

245 250 255

Gly Glu Arg Cys Ala Lys Ala Cys Val Asp Ile Gly Tyr Arg Gly Ala

260 265 270

Gly Thr Phe Glu Phe Leu Phe Glu Asn Gly Glu Phe Tyr Phe Ile Glu

275 280 285

Met Asn Thr Arg Ile Gln Val Glu His Pro Val Thr Glu Met Ile Thr

290 295 300

Gly Val Asp Leu Ile Lys Glu Gln Leu Arg Ile Ala Ala Gly Gln Pro

305 310 315 320

Leu Ser Ile Lys Gln Glu Glu Val His Val Arg Gly His Ala Val Glu

325 330 335

Cys Arg Ile Asn Ala Glu Asp Pro Asn Thr Phe Leu Pro Ser Pro Gly

340 345 350

Lys Ile Thr Arg Phe His Ala Pro Gly Gly Phe Gly Val Arg Trp Glu

355 360 365

Ser His Ile Tyr Ala Gly Tyr Thr Val Pro Pro Tyr Tyr Asp Ser Met

370 375 380

Ile Gly Lys Leu Ile Cys Tyr Gly Glu Asn Arg Asp Val Ala Ile Ala

385 390 395 400

Arg Met Lys Asn Ala Leu Gln Glu Leu Ile Ile Asp Gly Ile Lys Thr

405 410 415

Asn Val Asp Leu Gln Ile Arg Ile Met Asn Asp Glu Asn Phe Gln His

420 425 430

Gly Gly Thr Asn Ile His Tyr Leu Glu Lys Lys Leu Gly Leu Gln Glu

435 440 445

Lys

<210> 78

<211> 304

<212> PRT

<213> Escherichia coli

<400> 78

Met Ser Trp Ile Glu Arg Ile Lys Ser Asn Ile Thr Pro Thr Arg Lys
1 5 10 15

| | | | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|----|--|
| Ala | Ser | Ile | Pro | Glu | Gly | Val | Trp | Thr | Lys | Cys | Asp | Ser | Cys | Gly | Gln | | |
| 20 | | | | | | 25 | | | | | 30 | | | | | | |
| Val | Leu | Tyr | Arg | Ala | Glu | Leu | Glu | Arg | Asn | Leu | Glu | Val | Cys | Pro | Lys | | |
| 35 | | | | | | 40 | | | | | 45 | | | | | | |
| Cys | Asp | His | His | Met | Arg | Met | Thr | Ala | Arg | Asn | Arg | Leu | His | Ser | Leu | | |
| 50 | | | | | | 55 | | | | | 60 | | | | | | |
| Leu | Asp | Glu | Gly | Ser | Leu | Val | Glu | Leu | Gly | Ser | Glu | Leu | Glu | Pro | Lys | | |
| 65 | | | | | | 70 | | | | | 75 | | | | | 80 | |

| | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Asp | Val | Leu | Lys | Phe | Arg | Asp | Ser | Lys | Lys | Tyr | Lys | Asp | Arg | Leu | Ala |
| | | | | 85 | | | | | | | 90 | | | | |
| Ser | Ala | Gln | Lys | Glu | Thr | Gly | Glu | Lys | Asp | Ala | Leu | Val | Val | Met | Lys |
| | | | | 100 | | | | | | | 105 | | | | |
| Gly | Thr | Leu | Tyr | Gly | Met | Pro | Val | Val | Ala | Ala | Ala | Phe | Glu | Phe | Ala |
| | | | | 115 | | | | | | | 120 | | | | |
| Phe | Met | Gly | Gly | Ser | Met | Gly | Ser | Val | Val | Gly | Ala | Arg | Phe | Val | Arg |
| | | | | 125 | | | | | | | 130 | | | | |

130

135

140

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

145 150 155 160

Ala Ser Gly Gly Ala Arg Met Gln Glu Ala Leu Met Ser Leu Met Gln

165 170 175

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

180 185 190

Pro Tyr Ile Ser Val Leu Thr Asp Pro Thr Met Gly Gly Val Ser Ala

195 200 205

Ser Phe Ala Met Leu Gly Asp Leu Asn Ile Ala Glu Pro Lys Ala Leu

210 215 220

Ile Gly Phe Ala Gly Pro Arg Val Ile Glu Gln Thr Val Arg Glu Lys

225 230 235 240

Leu Pro Pro Gly Phe Gln Arg Ser Glu Phe Leu Ile Glu Lys Gly Ala

245 250 255

Ile Asp Met Ile Val Arg Arg Pro Glu Met Arg Leu Lys Leu Ala Ser

260 265 270

Ile Leu Ala Lys Leu Met Asn Leu Pro Ala Pro Asn Pro Glu Ala Pro

275 280 285

Arg Glu Gly Val Val Val Pro Pro Val Pro Asp Gln Glu Pro Glu Ala

290 295 300

<210> 79

<211> 282

<212> PRT

<213> Chloroflexus aurantiacus

<400> 79

Met Glu Glu Thr Ala Ile Pro Gln Ser Leu Thr Pro Trp Asp Arg Val

1 5 10 15

Gln Leu Ala Arg His Pro Gln Arg Pro His Thr Leu Asp Tyr Ile Ala

20 25 30

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

35 40 45
 Asp Asp Pro Ala Met Val Gly Gly Met Ala Thr Phe Ala Gly Gln Thr
 50 55 60
 Val Met Val Ile Gly His Gln Lys Gly Asn Asp Thr Arg Glu Asn Met
 65 70 75 80
 Arg Arg Asn Phe Gly Met Pro His Pro Glu Gly Tyr Arg Lys Ala Gln

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

 145 150 155 160
 Leu Ala Ile Ser Val Ala Asp Arg Ile Leu Met Gln Glu Asn Ala Ile
 165 170 175
 Tyr Ser Val Ala Pro Pro Glu Ala Ala Ala Ser Ile Leu Trp Arg Asp
 180 185 190
 Ala Ala Lys Ala Pro Glu Ala Ala Arg Ala Leu Lys Leu Thr Ala Ala
 195 200 205
 Asp Leu Tyr Asp Leu Arg Ile Ile Asp Glu Val Ile Pro Glu Pro Pro

 210 215 220
 Gly Gly Ala His Ala Asp Arg Leu Thr Ala Ile Thr Thr Val Gly Glu
 225 230 235 240
 Arg Leu Arg Val His Leu Ala Asp Leu Gln Gln Arg Asp Ile Asp Thr
 245 250 255
 Leu Leu Arg Glu Arg Tyr Arg Lys Tyr Arg Ser Met Gly Gln Tyr Gln
 260 265 270
 Glu Gln Gln Met Asp Phe Phe Gly Arg Met
 275 280

<210> 80

<211> 180

<212> PRT

<213> Chloroflexus aurantiacus

<400> 80

Met Met Leu Trp Gly Ala Met Lys Asp Glu Thr Thr Glu Leu Pro Ala

1 5 10 15

Asp Gln Pro Asp Pro Phe Gly Leu Ala Ala Val Arg Val Leu Leu Gln

20 25 30

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

35 40 45

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

50 55 60

Ala Pro Leu Pro Thr Ala Pro Val Pro Ser Pro Ser Leu Pro Ala Thr

65 70 75 80

Pro Val Thr Pro Phe Val Gln Pro Pro Pro Ala Pro Glu Gly Pro Pro

85 90 95

Val Glu Met Pro Ala Gly His Thr Ile Thr Ala Pro Met Val Gly Thr

100 105 110

Phe Tyr Ala Ala Pro Ser Pro Arg Asp Arg Pro Phe Val Gln Glu Gly

115 120 125

Asp Glu Val Arg Val Gly Asp Thr Val Gly Ile Val Glu Ala Met Lys

130 135 140

Met Met Asn Glu Ile Glu Ser Asp Val Ala Gly Arg Val Ala Arg Ile

145 150 155 160

Leu Val Lys Asn Gly Gln Pro Val Glu Tyr Gly Gln Pro Leu Met Val

165 170 175

Ile Glu Pro Leu

180

<210> 81

<211> 455

<212> PRT

<213> Chloroflexus aurantiacus

<400> 81

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

1 5 10 15

Ile Ile Arg Ala Cys Gln Glu Leu Gly Ile Arg Thr Val Val Ala Tyr

20 25 30

Ser Thr Ala Asp Arg Asp Ser Leu Ala Val Arg Leu Ala Asp Glu Ala

35 40 45

Val Cys Ile Gly Pro Pro Pro Ala Ala Lys Ser Tyr Leu Asn Ala Pro

50 55 60

Ala Leu Ile Ser Ala Ala Leu Val Ser Gly Cys Asp Ala Ile His Pro

65 70 75 80

Gly Tyr Gly Phe Leu Ser Glu Asn Pro Tyr Phe Ala Glu Met Cys Ala

85 90 95

Asp Cys Lys Leu Thr Phe Ile Gly Pro Pro Pro Glu Pro Ile Arg Leu

100 105 110

Met Gly Asp Lys Ala Ile Gly Arg Glu Thr Met Arg Lys Ala Gly Val

115 120 125

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

130 135 140

Ile Asp Val Ala Arg Gln Ile Gly Tyr Pro Val Leu Leu Lys Pro Ser

145 150 155 160

Gly Gly Gly Gly Gly Arg Gly Met Arg Val Ala Tyr Asp Glu Ala Asp

165 170 175

Leu Gln Arg Ala Phe Pro Thr Ala Arg Ala Glu Ala Glu Ala Ala Phe

180 185 190

Gly Asn Gly Ala Leu Leu Leu Glu Lys Tyr Leu Thr Arg Val Arg His

195 200 205

Val Glu Ile Gln Val Leu Ala Asp Gln Tyr Gly His Ala Ile His Leu

210 215 220

Gly Glu Arg Asp Cys Ser Ala Gln Arg Arg His Gln Lys Ile Val Glu

225 230 235 240

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

245 250 255

Ala Asp Ala Val Arg Gly Ile Lys Ser Ile Gly Tyr Val Asn Ala Gly

260 265 270

Thr Leu Glu Phe Leu Leu Asp Gln Asp Gly Asn Tyr Tyr Phe Ile Glu

275 280 285

Met Asn Thr Arg Ile Gln Val Glu His Pro Val Thr Glu Gln Val Thr

290 295 300

Gly Ile Asp Leu Val Arg Trp Gln Leu Leu Ile Ala Ser Gly Glu Arg

305 310 315 320

Leu Thr Leu Arg Gln Glu Asp Ile Lys Ile Thr Arg His Ala Ile Glu

325 330 335

Cys Arg Ile Asn Ala Glu Asp Pro Glu Arg Asp Phe Leu Pro Ala Ser

340 345 350

Gly Glu Val Glu Phe Tyr Leu Pro Pro Gly Gly Pro Gly Val Arg Val

355 360 365

Asp Ser His Leu Tyr Ser Gly Tyr Thr Pro Pro Gly Thr Tyr Asp Ser

370 375 380

Leu Leu Ala Lys Ile Ile Thr Phe Gly Asp Thr Arg Asp Glu Ala Leu

385 390 395 400

Asn Arg Met Arg Arg Ala Leu Asn Glu Cys Val Ile Thr Gly Ile Lys

405 410 415

Thr Thr Ile Pro Phe Gln Leu Ala Leu Ile Asp Asp Pro Glu Phe Arg

420 425 430

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

435 440 445

Lys Glu Thr Leu Asn Pro Val

450 455

<210> 82

<211> 305

<212> PRT

<213> Chloroflexus aurantiacus

<400> 82

Met Lys Glu Phe Phe Arg Leu Ser Arg Lys Gly Phe Thr Gly Arg Glu
1 5 10 15

Asp Gln Asp Ser Ala Gln Ile Pro Asp Asp Leu Trp Val Lys Cys Ser
20 25 30

Ser Cys Arg Glu Leu Ile Tyr Lys Lys Gln Leu Asn Asp Asn Leu Lys

35 40 45

Val Cys Pro Lys Cys Gly His His Met Arg Leu Ser Ala His Glu Trp
50 55 60

Leu Gly Leu Leu Asp Val Gly Ser Phe Arg Glu Met Asp Ala Asn Leu
65 70 75 80

Leu Pro Thr Asp Pro Leu Gly Phe Val Thr Asp Glu Glu Ser Tyr Ala
85 90 95

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

100 105 110

Ile Ala Gly Ile Gly Ala Ile Ser Asn Met Gln Ile Cys Val Ala Val

115 120 125

Ala Asp Phe Ser Phe Met Gly Ala Ser Met Gly Ser Val Tyr Gly Glu
130 135 140

Lys Met Ala Arg Ser Ala Glu Arg Ala Ala Glu Leu Gly Val Pro Leu
145 150 155 160

Leu Thr Ile Asn Thr Ser Gly Gly Ala Arg Gln Gln Glu Gly Val Ile

165 170 175

Gly Leu Met Gln Met Ala Lys Val Thr Met Ala Leu Thr Arg Leu Ala
180 185 190

Asp Ala Gly Gln Pro His Ile Ala Leu Leu Val Asp Pro Cys Tyr Gly
195 200 205

Gly Val Thr Ala Ser Tyr Pro Ser Val Ala Asp Ile Ile Ile Ala Glu
210 215 220

Pro Gly Ala Asn Ile Gly Phe Ala Gly Lys Arg Leu Ile Glu Gln Ile

225 230 235 240
Met Arg Gln Lys Leu Pro Ala Gly Phe Gln Thr Ala Glu Phe Met Leu
 245 250 255
Glu His Gly Met Ile Asp Met Val Val Pro Arg Ser Glu Met Arg Asp
 260 265 270
Thr Leu Ala Arg Ile Leu Arg Leu Tyr Arg Gln Arg Ser Thr Ser Pro
 275 280 285
Ala Lys Ala Glu Leu Ala Gly Arg Arg Ala Thr Leu Pro Gln Pro Ile

290 295 300

Met

305

<210> 83

<211> 378

<212> PRT

<213> Bacillus subtilis

<400> 83

Met Ile Ile Gly Val Pro Lys Glu Ile Lys Asn Asn Glu Asn Arg Val
1 5 10 15
Ala Leu Thr Pro Gly Gly Val Ser Gln Leu Ile Ser Asn Gly His Arg
 20 25 30
Val Leu Val Glu Thr Gly Ala Gly Leu Gly Ser Gly Phe Glu Asn Glu
 35 40 45

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

Ala Ile Ala Tyr Glu Thr Val Ser Glu Gly Arg Thr Leu Pro Leu Leu
115 120 125

Thr Pro Met Ser Glu Val Ala Gly Arg Met Ala Ala Gln Ile Gly Ala
130 135 140

Gln Phe Leu Glu Lys Pro Lys Gly Gly Lys Gly Ile Leu Leu Ala Gly
145 150 155 160

Val Pro Gly Val Ser Arg Gly Lys Val Thr Ile Ile Gly Gly Gly Val
165 170 175

Val Gly Thr Asn Ala Ala Lys Met Ala Val Gly Leu Gly Ala Asp Val
180 185 190

Thr Ile Ile Asp Leu Asn Ala Asp Arg Leu Arg Gln Leu Asp Asp Ile
195 200 205

Phe Gly His Gln Ile Lys Thr Leu Ile Ser Asn Pro Val Asn Ile Ala
210 215 220

Asp Ala Val Ala Glu Ala Asp Leu Leu Ile Cys Ala Val Leu Ile Pro
225 230 235 240

Gly Ala Lys Ala Pro Thr Leu Val Thr Glu Glu Met Val Lys Gln Met
245 250 255

Lys Pro Gly Ser Val Ile Val Asp Val Ala Ile Asp Gln Gly Gly Ile
260 265 270

Val Glu Thr Val Asp His Ile Thr Thr His Asp Gln Pro Thr Tyr Glu
275 280 285

Lys His Gly Val Val His Tyr Ala Val Ala Asn Met Pro Gly Ala Val
290 295 300

Pro Arg Thr Ser Thr Ile Ala Leu Thr Asn Val Thr Val Pro Tyr Ala
305 310 315 320

Leu Gln Ile Ala Asn Lys Gly Ala Val Lys Ala Leu Ala Asp Asn Thr
325 330 335

Ala Leu Arg Ala Gly Leu Asn Thr Ala Asn Gly His Val Thr Tyr Glu
340 345 350

Ala Val Ala Arg Asp Leu Gly Tyr Glu Tyr Val Pro Ala Glu Lys Ala
355 360 365

Leu Gln Asp Glu Ser Ser Val Ala Gly Ala

370

375

<210> 84

<211> 505

<212> PRT

<213> Schizosaccharomyces pombe

<400> 84

Met Phe Thr Asp Tyr Pro Asn Asp Ile Asn Cys Glu Ser Pro Arg Met

1 5 10 15

Ser Asp Leu Asp Gly Phe Cys Gln Asn Ala Phe Ser Asp Leu Asn Ser

20 25 30

Leu Asn Gln Gln Val Phe Lys Ala Asn Tyr Ala Val Arg Gly Ala Leu

35 40 45

Ala Ile Leu Ala Asp Glu Ile Gln Asp Asp Leu Leu Glu Asn Pro Ser

50 55 60

Ser Tyr Pro Phe Ser Glu Ile Val Tyr Ala Asn Ile Gly Asn Pro Gln

65 70 75 80

Gln Met Gly Gln Ser Pro Ile Thr Phe Val Arg Gln Val Leu Ser Leu

85 90 95

Cys Gln Tyr Pro Thr Leu Leu Asp His Ala Glu Glu Lys Trp Phe Gln

100 105 110

Asn Leu Phe Pro Thr Asp Val Val Gln Arg Ser Lys Met Leu Leu Lys

115 120 125

Glu Ser Gly Ser Leu Gly Ala Tyr Ser Ala Ser Gln Gly Ile Pro Leu

130 135 140

Val Arg Arg His Val Ala Asp Phe Ile Arg Ala Arg Asp Gly Phe Asp

145 150 155 160

Cys Glu Pro Ser Asp Ile Tyr Leu Thr Ser Gly Ala Ser His Ala Ala

165 170 175

Arg Leu Ile Met Thr Leu Ile Ile Ala Arg Pro Thr Asp Gly Val Met

180 185 190

Val Pro Ala Pro Gln Tyr Pro Leu Tyr Gly Ala Gln Ile Asp Leu Met

195 200 205

Ser Gly Ser Met Val Ser Tyr Ser Leu Ser Glu Glu Asn Asn Trp Asp
210 215 220

Ile Asp Phe Asp Gln Phe Lys Lys Ser Phe Asp Glu Ala Ser Lys Lys
225 230 235 240

Gly Ile Asn Val Arg Leu Cys Val Val Ile Asn Pro Gly Asn Pro Thr
245 250 255

Gly Ala Cys Ile Ser Glu Asn Ser Met Glu Lys Val Leu Arg Phe Ala
260 265 270

Lys Ala Lys Gly Ile Val Leu Leu Ala Asp Glu Val Tyr Gln Asn Asn
275 280 285

Ile Tyr Gln Asn Lys Phe His Ser Phe Arg Arg Lys Leu Gly Glu Leu
290 295 300

Arg Glu Lys Glu Pro Asp Asn His Trp Asp Gln Val Ser Leu Ile Ser
305 310 315 320

Val Asn Ser Val Ser Lys Gly Gln Phe Gly Glu Cys Gly Gln Arg Gly
325 330 335

Gly Tyr Leu Asp Val Val Asn Ile Pro Glu Pro Ala Lys Asp Gln Ile
340 345 350

Leu Lys Leu Ala Thr Ile Asp Ile Cys Pro Pro Val Ala Gly Gln Leu
355 360 365

Leu Val Asp Met Leu Val Asn Pro Pro Lys Pro Gly Asp Pro Ser Tyr
370 375 380

Asp Leu Phe Ile Lys Glu Val Asp Glu Ile His Glu Ala Leu Arg Leu
385 390 395 400

Gln Cys Arg Gln Leu Tyr Glu Gly Thr Lys Arg Met Lys Arg Val Ser
405 410 415

Cys Leu Glu Pro His Gly Ala Met Tyr Leu His Pro Ser Val Ser Leu
420 425 430

Pro Glu Lys Leu Ile Thr Thr Ala Lys Ala Gln Lys Ile Gln Pro Asp
435 440 445

Glu Phe Tyr Ala Ile Glu Leu Leu Lys Arg Ser Gly Ile Cys Val Val

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

His Asn Glu Ile Met Asp Leu Tyr Glu
 500 505

<210> 85

<211> 507

<212> PRT

<213> *Saccharomyces cerevisiae*

<400> 85

Met Thr Met Thr His Gln Gln Asp Leu Lys Gly Val Phe Thr Ala Lys
 1 5 10 15
 Asp Leu Asp Phe Lys Pro Ala Gly Lys Ile Thr Lys Lys Asp Leu Asn
 20 25 30
 Thr Gly Val Thr Lys Ala Glu Tyr Ala Val Arg Gly Ala Ile Pro Thr
 35 40 45

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

Leu Phe Ser Arg Asp Ala Leu Glu Arg Ala Glu Arg Leu Leu Asn Asp
 115 120 125
 Ile Gly Gly Ser Ile Gly Ala Tyr Ser His Ser Gln Gly Val Pro Gly
 130 135 140
 Ile Arg Gln Thr Val Ala Asp Phe Ile Thr Arg Arg Asp Gly Gly Glu
 145 150 155 160

Pro Ala Thr Pro Glu Asp Ile Tyr Leu Thr Thr Gly Ala Ser Ser Ala
165 170 175

Ala Thr Ser Leu Leu Ser Leu Leu Cys Lys Asp Ser Gln Thr Gly Leu
180 185 190

Leu Ile Pro Ile Pro Gln Tyr Pro Leu Tyr Thr Ala Ser Ala Ser Leu
195 200 205

Phe Asn Ala Gln Val Leu Pro Tyr Tyr Leu Asp Glu Glu Ser Asn Trp
210 215 220

Ser Thr Asn Ser Asp Glu Ile Glu Lys Val Val Gln Asp Ala Leu Lys
225 230 235 240

Lys Gln Ile Arg Pro Ser Val Leu Ile Val Ile Asn Pro Gly Asn Pro
245 250 255

Thr Gly Ala Val Leu Ser Glu Glu Thr Ile Ala Arg Ile Cys Leu Ile
260 265 270

Ala Ala Lys Tyr Gly Ile Thr Ile Ile Ser Asp Glu Val Tyr Gln Glu
275 280 285

Asn Ile Phe Asn Asp Val Lys Phe His Ser Met Lys Lys Val Leu Arg
290 295 300

Lys Leu Gln His Leu Tyr Pro Gly Lys Phe Asp Asn Val Gln Leu Ala
305 310 315 320

Ser Leu His Ser Ile Ser Lys Gly Phe Met Asp Glu Cys Gly Gln Arg
325 330 335

Gly Gly Tyr Met Glu Ile Ile Gly Phe Ser Gln Glu Ile Arg Asp Ala
340 345 350

Leu Phe Lys Leu Met Ser Ile Ser Ile Cys Ser Val Val Thr Gly Gln
355 360 365

Ala Val Val Asp Leu Met Val Lys Pro Pro Gln Pro Gly Asp Glu Ser
370 375 380

Tyr Glu Gln Asp His Asp Glu Arg Leu Lys Ile Phe His Glu Met Arg
385 390 395 400

Thr Arg Ala Asn Leu Leu Tyr Glu Thr Phe Lys Glu Leu Glu Gly Ile

405 410 415
 Glu Cys Gln Lys Pro Gln Gly Ala Met Tyr Leu Phe Pro Arg Leu Val
 420 425 430

 Leu Pro Lys Lys Ala Leu Cys Glu Ser Glu Arg Leu Gly Ile Glu Pro
 435 440 445
 Asp Glu Phe Tyr Cys Thr Ser Leu Leu Glu Ser Thr Gly Ile Cys Thr
 450 455 460
 Val Pro Gly Ser Gly Phe Gly Gln Arg Pro Gly Thr Tyr His Val Arg
 465 470 475 480
 Thr Thr Phe Leu Ala Pro Gly Thr Lys Trp Ile Gln Asp Trp Lys Glu
 485 490 495

 Phe His Gln Asp Phe Phe Ser Lys Tyr Arg Asn
 500 505

 <210> 86
 <211> 471
 <212> PRT
 <213> Bacillus subtilis
 <400> 86
 Met Lys Asn Lys Trp Tyr Lys Pro Lys Arg His Trp Lys Glu Ile Glu
 1 5 10 15
 Leu Trp Lys Asp Val Pro Glu Glu Lys Trp Asn Asp Trp Leu Trp Gln
 20 25 30
 Leu Thr His Thr Val Arg Thr Leu Asp Asp Leu Lys Lys Val Ile Asn
 35 40 45

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

| | | | |
|---|-----|-----|-----|
| 100 | 105 | 110 | |
| Ser Pro Val Pro Gly Leu Thr His Arg Tyr Pro Asp Arg Val Leu Phe | | | |
| 115 | 120 | 125 | |
| Leu Val Thr Asn Gln Cys Ser Met Tyr Cys Arg Tyr Cys Thr Arg Arg | | | |
| 130 | 135 | 140 | |
| Arg Phe Ser Gly Gln Ile Gly Met Gly Val Pro Lys Lys Gln Leu Asp | | | |
| 145 | 150 | 155 | 160 |
| Ala Ala Ile Ala Tyr Ile Arg Glu Thr Pro Glu Ile Arg Asp Cys Leu | | | |
| 165 | 170 | 175 | |
| Ile Ser Gly Gly Asp Gly Leu Leu Ile Asn Asp Gln Ile Leu Glu Tyr | | | |
| 180 | 185 | 190 | |
| Ile Leu Lys Glu Leu Arg Ser Ile Pro His Leu Glu Val Ile Arg Ile | | | |
| 195 | 200 | 205 | |
| Gly Thr Arg Ala Pro Val Val Phe Pro Gln Arg Ile Thr Asp His Leu | | | |
| 210 | 215 | 220 | |
| Cys Glu Ile Leu Lys Lys Tyr His Pro Val Trp Leu Asn Thr His Phe | | | |
| 225 | 230 | 235 | 240 |
| Asn Thr Ser Ile Glu Met Thr Glu Glu Ser Val Glu Ala Cys Glu Lys | | | |
| 245 | 250 | 255 | |
| Leu Val Asn Ala Gly Val Pro Val Gly Asn Gln Ala Val Val Leu Ala | | | |
| 260 | 265 | 270 | |
| Gly Ile Asn Asp Ser Val Pro Ile Met Lys Lys Leu Met His Asp Leu | | | |
| 275 | 280 | 285 | |
| Val Lys Ile Arg Val Arg Pro Tyr Tyr Ile Tyr Gln Cys Asp Leu Ser | | | |
| 290 | 295 | 300 | |
| Glu Gly Ile Gly His Phe Arg Ala Pro Val Ser Lys Gly Leu Glu Ile | | | |
| 305 | 310 | 315 | 320 |
| Ile Glu Gly Leu Arg Gly His Thr Ser Gly Tyr Ala Val Pro Thr Phe | | | |
| 325 | 330 | 335 | |
| Val Val Asp Ala Pro Gly Gly Gly Gly Lys Ile Ala Leu Gln Pro Asn | | | |
| 340 | 345 | 350 | |

Tyr Val Leu Ser Gln Ser Pro Asp Lys Val Ile Leu Arg Asn Phe Glu
355 360 365

Gly Val Ile Thr Ser Tyr Pro Glu Pro Glu Asn Tyr Ile Pro Asn Gln
370 375 380

Ala Asp Ala Tyr Phe Glu Ser Val Phe Pro Glu Thr Ala Asp Lys Lys
385 390 395 400

Glu Pro Ile Gly Leu Ser Ala Ile Phe Ala Asp Lys Glu Val Ser Phe
405 410 415

Thr Pro Glu Asn Val Asp Arg Ile Lys Arg Arg Glu Ala Tyr Ile Ala
420 425 430

Asn Pro Glu His Glu Thr Leu Lys Asp Arg Arg Glu Lys Arg Asp Gln
435 440 445

Leu Lys Glu Lys Lys Phe Leu Ala Gln Gln Lys Lys Gln Lys Glu Thr
450 455 460

Glu Cys Gly Gly Asp Ser Ser
465 470

<210> 87

<211> 416

<212> PRT

<213> Porphyromonas gingivalis

<400> 87

Met Ala Glu Ser Arg Arg Lys Tyr Tyr Phe Pro Asp Val Thr Asp Glu
1 5 10 15

Gln Trp Asn Asp Trp His Trp Gln Val Leu Asn Arg Ile Glu Thr Leu
20 25 30

Asp Gln Leu Lys Lys Tyr Val Thr Leu Thr Ala Glu Glu Glu Gly
35 40 45

Val Lys Glu Ser Leu Lys Val Leu Arg Met Ala Ile Thr Pro Tyr Tyr
50 55 60

Leu Ser Leu Ile Asp Pro Glu Asn Pro Asn Cys Pro Ile Arg Lys Gln
65 70 75 80

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

Ser Ser Pro Ser Glu Arg Ile Asp Arg Cys Ile Asp Tyr Ile Ala Asn
145 150 155 160
Thr Pro Thr Val Arg Asp Val Leu Leu Ser Gly Gly Asp Ala Leu Leu
165 170 175
Val Ser Asp Glu Arg Leu Glu Tyr Ile Leu Lys Arg Leu Arg Glu Ile
180 185 190
Pro His Val Glu Ile Val Arg Ile Gly Ser Arg Thr Pro Val Val Leu
195 200 205

Pro Gln Arg Ile Thr Pro Gln Leu Val Asp Met Leu Lys Lys Tyr His
210 215 220
Pro Val Trp Leu Asn Thr His Phe Asn His Pro Asn Glu Val Thr Glu
225 230 235 240
Glu Ala Val Glu Ala Cys Glu Arg Met Ala Asn Ala Gly Ile Pro Leu
245 250 255
Gly Asn Gln Thr Val Leu Leu Arg Gly Ile Asn Asp Cys Thr His Val
260 265 270

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

[illegible]

| | | | |
|---|-----|-----|-----|
| 115 | 120 | 125 | |
| Met Tyr Cys Arg His Cys Thr Arg Arg Arg Phe Ala Gly Ser Ser Asp | | | |
| 130 | 135 | 140 | |
| Asp Ala Met Pro Met Asp Arg Ile Asp Lys Ala Ile Glu Tyr Ile Ala | | | |
| 145 | 150 | 155 | 160 |
| Lys Thr Pro Gln Val Arg Asp Val Leu Leu Ser Gly Gly Asp Ala Leu | | | |
| | 165 | 170 | 175 |
| Leu Val Ser Asp Lys Lys Leu Glu Ser Ile Ile Gln Lys Leu Arg Ala | | | |
| 180 | 185 | 190 | |
| Ile Pro His Val Glu Ile Ile Arg Ile Gly Ser Arg Thr Pro Val Val | | | |
| 195 | 200 | 205 | |
| Leu Pro Gln Arg Ile Thr Pro Glu Leu Cys Asn Met Leu Lys Lys Tyr | | | |
| 210 | 215 | 220 | |
| His Pro Ile Trp Leu Asn Thr His Phe Asn His Pro Gln Glu Val Thr | | | |
| 225 | 230 | 235 | 240 |
| Pro Glu Ala Lys Lys Ala Cys Glu Met Leu Ala Asp Ala Gly Val Pro | | | |
| 245 | 250 | 255 | |
| Leu Gly Asn Gln Thr Val Leu Leu Arg Gly Ile Asn Asp Ser Val Pro | | | |
| 260 | 265 | 270 | |
| Val Met Lys Arg Leu Val His Asp Leu Val Met Met Arg Val Arg Pro | | | |
| 275 | 280 | 285 | |
| Tyr Tyr Ile Tyr Gln Cys Asp Leu Ser Met Gly Leu Glu His Phe Arg | | | |
| 290 | 295 | 300 | |
| Thr Pro Val Ser Lys Gly Ile Glu Ile Ile Glu Gly Leu Arg Gly His | | | |
| 305 | 310 | 315 | 320 |
| Thr Ser Gly Tyr Ala Val Pro Thr Phe Val Val Asp Ala Pro Gly Gly | | | |
| 325 | 330 | 335 | |
| Gly Gly Lys Thr Pro Val Met Pro Gln Tyr Val Ile Ser Gln Ser Pro | | | |
| 340 | 345 | 350 | |
| His Arg Val Val Leu Arg Asn Phe Glu Gly Val Ile Thr Thr Tyr Thr | | | |
| 355 | 360 | 365 | |

Glu Pro Glu Asn Tyr Thr His Glu Pro Cys Tyr Asp Glu Glu Lys Phe
 370 375 380
 Glu Lys Met Tyr Glu Ile Ser Gly Val Tyr Met Leu Asp Glu Gly Leu
 385 390 395 400
 Lys Met Ser Leu Glu Pro Ser His Leu Ala Arg His Glu Arg Asn Lys
 405 410 415
 Lys Arg Ala Glu Ala Glu Gly Lys Lys
 420 425

<210> 89

<211> 517

<212> PRT

<213> Megasphaera elsdenii

<400> 89

Met Arg Lys Val Glu Ile Ile Thr Ala Glu Gln Ala Ala Gln Leu Val
 1 5 10 15
 Lys Asp Asn Asp Thr Ile Thr Ser Ile Gly Phe Val Ser Ser Ala His
 20 25 30
 Pro Glu Ala Leu Thr Lys Ala Leu Glu Lys Arg Phe Leu Asp Thr Asn
 35 40 45
 Thr Pro Gln Asn Leu Thr Tyr Ile Tyr Ala Gly Ser Gln Gly Lys Arg

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

115 120 125
 Ile Gly Leu Glu Thr Phe Leu Asp Pro Arg Gln Leu Gly Gly Lys Leu
 130 135 140
 Asn Asp Val Thr Lys Glu Asp Leu Val Lys Leu Ile Glu Val Asp Gly

145 150 155 160
 His Glu Gln Leu Phe Tyr Pro Thr Phe Pro Val Asn Val Ala Phe Leu
 165 170 175
 Arg Gly Thr Tyr Ala Asp Glu Ser Gly Asn Ile Thr Met Asp Glu Glu

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

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

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

 370 375 380
 Gly Thr Asn Val Ala Gly Cys Gly Gly Phe Pro Asn Ile Ser Gln Gln
 385 390 395 400

Thr Pro Asn Val Tyr Phe Cys Gly Thr Phe Thr Ala Gly Gly Leu Lys
 405 410 415
 Ile Ala Val Glu Asp Gly Lys Val Lys Ile Leu Gln Glu Gly Lys Ala
 420 425 430
 Lys Lys Phe Ile Lys Ala Val Asp Gln Ile Thr Phe Asn Gly Ser Tyr

435 440 445
 Ala Ala Arg Asn Gly Lys His Val Leu Tyr Ile Thr Glu Arg Cys Val
 450 455 460
 Phe Glu Leu Thr Lys Glu Gly Leu Lys Leu Ile Glu Val Ala Pro Gly
 465 470 475 480
 Ile Asp Ile Glu Lys Asp Ile Leu Ala His Met Asp Phe Lys Pro Ile
 485 490 495
 Ile Asp Asn Pro Lys Leu Met Asp Ala Arg Leu Phe Gln Asp Gly Pro

500 505 510
 Met Gly Leu Lys Lys
 515

<210> 90

<211> 145

<212> PRT

<213> Clostridium propionicum

<400> 90

Met Val Gly Lys Lys Val Val His His Leu Met Met Ser Ala Lys Asp
 1 5 10 15
 Ala His Tyr Thr Gly Asn Leu Val Asn Gly Ala Arg Ile Val Asn Gln
 20 25 30
 Trp Gly Asp Val Gly Thr Glu Leu Met Val Tyr Val Asp Gly Asp Ile

35 40 45
 Ser Leu Phe Leu Gly Tyr Lys Asp Ile Glu Phe Thr Ala Pro Val Tyr
 50 55 60
 Val Gly Asp Phe Met Glu Tyr His Gly Trp Ile Glu Lys Val Gly Asn
 65 70 75 80
 Gln Ser Tyr Thr Cys Lys Phe Glu Ala Trp Lys Val Ala Thr Met Val

85 90 95
 Asp Ile Thr Asn Pro Gln Asp Thr Arg Ala Thr Ala Cys Glu Pro Pro
 100 105 110
 Val Leu Cys Gly Arg Ala Thr Gly Ser Leu Phe Ile Ala Lys Lys Asp
 115 120 125
 Gln Arg Gly Pro Gln Glu Ser Ser Phe Lys Glu Arg Lys His Pro Gly
 130 135 140
 Glu
 145
 <210> 91
 <211> 260
 <212> PRT
 <213> Phytophthora sojae
 <400> 91
 Met Ala Ala Glu Tyr Glu Ser Ile Leu Thr Glu Val Arg Gly Lys Val
 1 5 10 15
 Ala Ile Ile Thr Leu Asn Arg Pro Lys Ala Leu Asn Ala Leu Cys Ser
 20 25 30
 Pro Leu Ile Glu Glu Leu Asn Gly Ala Ala His Ala Phe Asp Ala Asp
 35 40 45
 Pro Ser Ile Gly Ala Ile Val Ile Thr Gly Ser Gly Ser Lys Ala Phe
 50 55 60
 Ala Ala Gly Ala Asp Ile Lys Glu Met Ala Thr Lys Thr Phe Val Asp
 65 70 75 80
 Ala Tyr Lys Ser Asn Met Phe Ala Asn Trp Gly Asp Ile Thr Lys Val
 85 90 95
 Ser Lys Pro Val Ile Ala Ala Val Asn Gly Tyr Ala Leu Gly Gly Gly
 100 105 110
 Cys Glu Leu Ala Met Leu Cys Asp Leu Ile Ile Ala Gly Asp Ser Ala
 115 120 125
 Lys Phe Gly Gln Pro Glu Ile Thr Leu Gly Thr Ile Pro Gly Cys Gly

130

135

140

Gly Thr Gln Arg Leu Ile Arg Ala Val Gly Lys Ser Lys Ala Met Glu

145 150 155 160

Met Ile Leu Thr Gly Asn Met Ile Asp Ala Gln Gln Ala Glu Arg Asp

165 170 175

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

180 185 190

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

195 200 205

Met Ala Lys Glu Ala Val Asn Ala Ala Tyr Glu Gln Ser Leu Gln Glu

210 215 220

Gly Leu Lys Tyr Glu Ser Arg Leu Phe Trp Ser Ser Phe Ala Thr Lys

225 230 235 240

Asp Gln Lys Glu Gly Met Ala Ala Phe Val Glu Lys Arg Lys Ala Asp

245 250 255

Phe Lys Asp Glu

260

<210> 92

<211> 258

<212> PRT

<213> Chloroflexus aurantiacus

<400> 92

Met Ser Glu Glu Ser Leu Val Leu Ser Thr Ile Glu Gly Pro Ile Ala

1 5 10 15

Ile Leu Thr Leu Asn Arg Pro Gln Ala Leu Asn Ala Leu Ser Pro Ala

20 25 30

Leu Ile Asp Asp Leu Ile Arg His Leu Glu Ala Cys Asp Ala Asp Asp

35 40 45

Thr Ile Arg Val Ile Ile Ile Thr Gly Ala Gly Arg Ala Phe Ala Ala

50 55 60

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

65 70 75 80
 Thr Ser Gly Met Ile Ala Arg Trp Ala Arg Ile Ala Ala Val Arg Lys
 85 90 95
 Pro Val Ile Ala Ala Val Asn Gly Tyr Ala Leu Gly Gly Gly Cys Glu
 100 105 110
 Leu Ala Met Met Cys Asp Ile Ile Ile Ala Ser Glu Asn Ala Gln Phe
 115 120 125

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

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

Gly Arg

<210> 93

<211> 288

<212> PRT

<213> Rhodospirillum rubrum

<400> 93

Met Ala Ala Ala Pro Gly Pro Pro Ala Ala Arg Pro Arg Ala Ala Arg
 1 5 10 15

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

260 265 270
 Gly Met Asn Ala Phe Ile Glu Lys Arg Gln Pro Ser Phe Thr Asp Arg
 275 280 285
 <210> 94
 <211> 258
 <212> PRT
 <213> Rhodobacter capsulatus
 <400> 94
 Met Ser Tyr Gln Thr Leu Ile Val Glu Ile Ala Asp Gly Val Ala Leu
 1 5 10 15

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

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

 Gln Arg Leu Thr Arg Phe Val Gly Lys Ser Lys Ser Met Glu Met His
 145 150 155 160
 Leu Thr Gly Arg Phe Met Asp Ala Ala Glu Ala Glu Arg Ser Gly Leu
 165 170 175
 Val Ser Arg Val Leu Pro Leu Ala Asp Leu Leu Pro Glu Ala Leu Ala
 180 185 190

Thr Ala Arg Lys Ile Ala Glu Lys Ser Ala Ile Ala Thr Met Val Ala
195 200 205

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

Leu Tyr Glu Arg Arg Val Phe His Ala Leu Phe Ala Thr Glu Asp Gln
225 230 235 240

Lys Glu Gly Met Ala Ala Phe Ala Glu Lys Arg Pro Ala Lys Phe Ala
245 250 255

Asp Lys

<210> 95

<211> 290

<212> PRT

<213> Homo sapiens

<400> 95

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

1 5 10 15

Pro Pro Val Arg Cys Pro Ala Trp Arg Pro Phe Ala Ser Gly Ala Asn

20 25 30

Phe Glu Tyr Ile Ile Ala Glu Lys Arg Gly Lys Asn Asn Thr Val Gly

35 40 45

Leu Ile Gln Leu Asn Arg Pro Lys Ala Leu Asn Ala Leu Cys Asp Gly

50 55 60

Leu Ile Asp Glu Leu Asn Gln Ala Leu Lys Ile Phe Glu Glu Asp Pro

65 70 75 80

Ala Val Gly Ala Ile Val Leu Thr Gly Gly Asp Lys Ala Phe Ala Ala

85 90 95

Gly Ala Asp Ile Lys Glu Met Gln Asn Leu Ser Phe Gln Asp Cys Tyr

100 105 110

Ser Ser Lys Phe Leu Lys His Trp Asp His Leu Thr Gln Val Lys Lys

115 120 125

Pro Val Ile Ala Ala Val Asn Gly Tyr Ala Phe Gly Gly Gly Cys Glu

130 135 140

Leu Ala Met Met Cys Asp Ile Ile Tyr Ala Gly Glu Lys Ala Gln Phe

145 150 155 160

Ala Gln Pro Glu Ile Leu Ile Gly Thr Ile Pro Gly Ala Gly Gly Thr

165 170 175

Gln Arg Leu Thr Arg Ala Val Gly Lys Ser Leu Ala Met Glu Met Val

180 185 190

Leu Thr Gly Asp Arg Ile Ser Ala Gln Asp Ala Lys Gln Ala Gly Leu

195 200 205

Val Ser Lys Ile Cys Pro Val Glu Thr Leu Val Glu Glu Ala Ile Gln

210 215 220

Cys Ala Glu Lys Ile Ala Ser Asn Ser Lys Ile Val Val Ala Met Ala

225 230 235 240

Lys Glu Ser Val Asn Ala Ala Phe Glu Met Thr Leu Thr Glu Gly Ser

245 250 255

Lys Leu Glu Lys Lys Leu Phe Tyr Ser Thr Phe Ala Thr Asp Asp Arg

260 265 270

Lys Glu Gly Met Thr Ala Phe Val Glu Lys Arg Lys Ala Asn Phe Lys

275 280 285

Asp Gln

290

<210> 96

<211> 399

<212> PRT

<213> Pseudomonas fluorescens

<400> 96

Met Arg Val Val Leu Cys Lys Phe Ala Leu Leu Arg Ser His Asp Gly

1 5 10 15

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

20 25 30

Met Thr Ala Gln Val Ser Thr Glu Ala Ser His Ala Ala Ile Leu Gln
35 40 45

Asp Glu Val Leu Ala Glu Val Arg Asn His Ile Gly His Leu Thr Leu
50 55 60

Asn Arg Pro Ala Gly Leu Asn Ala Leu Thr Leu Gln Met Val Arg Ser
65 70 75 80

Leu Thr Ser Gln Leu Gln Ala Trp Ser Asp Asp Pro Gln Val Tyr Ala
85 90 95

Val Val Leu Arg Gly Ala Gly Glu Lys Ala Phe Cys Ala Gly Gly Asp
100 105 110

Ile Arg Ser Leu Tyr Asp Ser Phe Lys Asn Gly Asp Thr Leu His Gln
115 120 125

Asp Phe Phe Val Glu Glu Tyr Ala Leu Asp Leu Ala Ile His His Tyr
130 135 140

Arg Lys Pro Val Leu Ala Leu Met Asp Gly Phe Val Leu Gly Gly Gly
145 150 155 160

Met Gly Leu Val Gln Gly Ala Asp Leu Arg Val Val Thr Glu Arg Ser
165 170 175

Arg Leu Ala Met Pro Glu Val Ala Ile Gly Tyr Phe Pro Asp Val Gly
180 185 190

Gly Ser Tyr Phe Leu Pro Arg Ile Pro Gly Glu Leu Gly Ile Tyr Leu
195 200 205

Gly Val Thr Gly Val Gln Ile Arg Ala Ala Asp Ala Leu Tyr Cys Gly
210 215 220

Leu Ala Asp Trp Tyr Leu Asp Ser His Lys Leu Ala Asp Leu Asp Gln
225 230 235 240

Lys Leu Asp Asn Leu Arg Trp His Asp Ser Pro Leu Lys Asp Leu Gln
245 250 255

Gly Ala Leu Ala Arg Leu Ala Val Gln Gln Leu Pro Asp Ala Pro Leu
260 265 270

Ala Ala Leu Arg Pro Ala Ile Asp His Phe Phe Ala Leu Pro Asp Val

275

280

285

Pro Ser Ile Val Glu Gln Leu Gln Gln Val Thr Val Ala Asp Ser His

290

295

300

Glu Trp Ala Leu Asn Thr Val Ser Leu Met Gln Thr Arg Ser Pro Leu

305

310

315

320

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

325

330

335

Leu Glu Gln Cys Phe Ala Leu Glu Leu His Leu Asp Arg Gln Trp Phe

340

345

350

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

355

360

365

Asp Lys Ser Pro Arg Trp Asn Pro Pro Thr Leu His Gly Leu Ala Leu

370

375

380

Asn His Val Glu Ser Phe Phe His His Phe Glu Lys Val Val Lys

385

390

395

<210> 97

<211> 351

<212> PRT

<213> Bacillus cereus

<400> 97

Met Thr Glu Gln Val Leu Phe Ser Val Ser Glu Asn Gly Val Ala Thr

1

5

10

15

Ile Thr Leu Asn Arg Pro Lys Ala Leu Asn Ser Leu Ser Tyr Asp Met

20

25

30

Leu Gln Pro Ile Gly Gln Lys Leu Lys Glu Trp Glu His Asp Glu Arg

35

40

45

Ile Ala Leu Ile Val Leu Lys Gly Ala Gly Thr Lys Gly Phe Cys Ala

50

55

60

Gly Gly Asp Ile Lys Thr Leu Tyr Glu Ala Arg Ser Asn Glu Val Ala

65

70

75

80

Leu Gln His Ala Glu Arg Phe Phe Glu Glu Glu Tyr Glu Ile Asp Thr

85 90 95
 Tyr Ile Tyr Gln Tyr Thr Lys Pro Ile Ile Ala Cys Leu Asp Gly Ile
 100 105 110
 Val Met Gly Gly Gly Val Gly Leu Thr Asn Gly Ala Lys Tyr Arg Ile
 115 120 125
 Val Thr Glu Arg Thr Lys Trp Ala Met Pro Glu Met Asn Ile Gly Phe

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

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

 260 265 270
 Ser Pro Ile Ser Leu Lys Val Thr Leu Lys Gln Phe Ile Asp Gly Gln
 275 280 285
 Asp Lys Ser Val Glu Glu Cys Phe Ala Thr Asp Leu Ile Leu Ala Lys
 290 295 300
 Asn Phe Met Arg His Glu Asp Phe Phe Glu Gly Val Arg Ser Val Val
 305 310 315 320
 Val Asp Lys Asp Gln Asn Pro Asn Tyr Lys Tyr Lys Gln Leu Ser Asp

 325 330 335

Val Ser Glu Glu Asp Val Asn Arg Phe Phe Asn Leu Leu Asn Ala

340 345 350

<210> 98

<211> 381

<212> PRT

<213> Homo sapiens

<400> 98

Met Trp Arg Leu Met Ser Arg Phe Asn Ala Phe Lys Arg Thr Asn Thr

1 5 10 15

Ile Leu His His Leu Arg Met Ser Lys His Thr Asp Ala Ala Glu Glu

20 25 30

Val Leu Leu Glu Lys Lys Gly Cys Ala Gly Val Ile Thr Leu Asn Arg

35 40 45

Pro Lys Phe Leu Asn Ala Leu Thr Leu Asn Met Ile Arg Gln Ile Tyr

50 55 60

Pro Gln Leu Lys Lys Trp Glu Gln Asp Pro Glu Thr Phe Val Ile Ile

65 70 75 80

Ile Lys Gly Ala Gly Gly Lys Ala Phe Cys Ala Gly Gly Asp Ile Arg

85 90 95

Val Ile Ser Glu Ala Glu Lys Ala Lys Gln Lys Ile Ala Pro Val Phe

100 105 110

Phe Arg Glu Glu Tyr Met Leu Asn Asn Ala Val Gly Ser Cys Gln Lys

115 120 125

Pro Tyr Val Ala Leu Ile His Gly Ile Thr Met Gly Gly Gly Val Gly

130 135 140

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

145 150 155 160

Ala Met Pro Glu Thr Ala Ile Gly Leu Phe Pro Asp Val Gly Gly Gly

165 170 175

Tyr Phe Phe Ala Thr Thr Pro Arg Lys Thr Trp Leu Leu Pro Cys Ile

180 185 190

Asn Gly Phe Arg Leu Lys Gly Arg Asp Val Tyr Arg Ala Gly Ile Ala

| | | | |
|---|----------------------|-----|-----|
| 195 | 200 | 205 | |
| Thr His Phe Val Asp Ser Glu Lys Leu Ala Met Leu Glu Glu Asp Leu | | | |
| 210 | 215 | 220 | |
| Leu Ala Leu Lys Ser Pro Ser Lys Glu Asn Ile Ala Ser Val Leu Glu | | | |
| 225 | 230 | 235 | 240 |
| Asn Tyr His Thr Glu Ser Lys Ile Asp Arg Asp Lys Ser Phe Ile Leu | | | |
| | 245 | 250 | 255 |
| Glu Glu His Met Asp Lys Ile Asn Ser Cys Phe Ser Ala Asn Thr Val | | | |
| | 260 | 265 | 270 |
| Glu Glu Ile Ile Glu Asn Leu Gln Gln Asp Gly Ser Ser Phe Ala Leu | | | |
| | 275 | 280 | 285 |
| Glu Gln Leu Lys Val Ile Asn Lys Met Ser Pro Thr Ser Leu Lys Ile | | | |
| | 290 | 295 | 300 |
| Thr Leu Arg Gln Leu Met Glu Gly Ser Ser Lys Thr Leu Gln Glu Val | | | |
| 305 | 310 | 315 | 320 |
| Leu Thr Met Glu Tyr Arg Leu Ser Gln Ala Cys Met Arg Gly His Asp | | | |
| | 325 | 330 | 335 |
| Phe His Glu Gly Val Arg Ala Val Leu Ile Asp Lys Asp Gln Ser Pro | | | |
| | 340 | 345 | 350 |
| Lys Trp Lys Pro Ala Asp Leu Lys Glu Val Thr Glu Glu Asp Leu Asn | | | |
| | 355 | 360 | 365 |
| Asn His Phe Lys Ser Leu Gly Ser Ser Asp Leu Lys Phe | | | |
| | 370 | 375 | 380 |
| <210> | 99 | | |
| <211> | 261 | | |
| <212> | PRT | | |
| <213> | Megasphaera elsdenii | | |
| <400> | 99 | | |
| Val Lys Thr Val Tyr Thr Leu Gly Ile Asp Val Gly Ser Ser Ser Ser | | | |
| 1 | 5 | 10 | 15 |
| Lys Ala Val Ile Leu Glu Asp Gly Lys Lys Ile Val Ala His Ala Val | | | |

20 25 30
 Val Glu Ile Gly Thr Gly Ser Thr Gly Pro Glu Arg Val Leu Asp Glu
 35 40 45
 Val Phe Lys Asp Thr Asn Leu Lys Ile Glu Asp Met Ala Asn Ile Ile
 50 55 60
 Ala Thr Gly Tyr Gly Arg Phe Asn Val Asp Cys Ala Lys Gly Glu Val
 65 70 75 80
 Ser Glu Ile Thr Cys His Ala Lys Gly Ala Leu Phe Glu Cys Pro Gly
 85 90 95
 Thr Thr Thr Ile Leu Asp Ile Gly Gly Gln Asp Val Lys Ser Ile Lys
 100 105 110
 Leu Asn Gly Gln Gly Leu Val Met Gln Phe Ala Met Asn Asp Lys Cys
 115 120 125
 Ala Ala Gly Thr Gly Arg Phe Leu Asp Val Met Ser Lys Val Leu Glu
 130 135 140
 Ile Pro Met Ser Glu Met Gly Asp Trp Tyr Phe Lys Ser Lys His Pro
 145 150 155 160
 Ala Ala Val Ser Ser Thr Cys Thr Val Phe Ala Glu Ser Glu Val Ile
 165 170 175
 Ser Leu Leu Ser Lys Asn Val Pro Lys Glu Asp Ile Val Ala Gly Val
 180 185 190
 His Gln Ser Ile Ala Ala Lys Ala Cys Ala Leu Val Arg Arg Val Gly
 195 200 205
 Val Gly Glu Asp Leu Thr Met Thr Gly Gly Gly Ser Arg Asp Pro Gly
 210 215 220
 Val Val Asp Ala Val Ser Lys Glu Leu Gly Ile Pro Val Arg Val Ala
 225 230 235 240
 Leu His Pro Gln Ala Val Gly Ala Leu Gly Ala Ala Leu Ile Ala Tyr
 245 250 255
 Asp Lys Ile Lys Lys
 260
 <210> 100

<211> 428

<212> PRT

<213> *Megasphaera elsdenii*

<400> 100

Met Ser Glu Glu Lys Thr Val Asp Ile Glu Ser Met Ser Ser Lys Glu

1 5 10 15
Ala Leu Gly Tyr Phe Leu Pro Lys Val Asp Glu Asp Ala Arg Lys Ala
20 25 30
Lys Lys Glu Gly Arg Leu Val Cys Trp Ser Ala Ser Val Ala Pro Pro
35 40 45
Glu Phe Cys Thr Ala Met Asp Ile Ala Ile Val Tyr Pro Glu Thr His
50 55 60
Ala Ala Gly Ile Gly Ala Arg His Gly Ala Pro Ala Met Leu Glu Val

65 70 75 80
Ala Glu Asn Lys Gly Tyr Asn Gln Asp Ile Cys Ser Tyr Cys Arg Val
85 90 95
Asn Met Gly Tyr Met Glu Leu Leu Lys Gln Gln Ala Leu Thr Gly Glu
100 105 110
Thr Pro Glu Val Leu Lys Asn Ser Pro Ala Ser Pro Ile Pro Leu Pro
115 120 125
Asp Val Val Leu Thr Cys Asn Asn Ile Cys Asn Thr Leu Leu Lys Trp

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

195 200 205
Gln Arg Ser Ile Ala Ala Trp Asn Lys Ile Ala Thr Tyr Phe Gln Tyr

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

 260 265 270
 Gly Glu Asn Glu Lys Ser Arg Val Thr Trp Glu Gly Ile Ala Val Trp
 275 280 285
 Ile Ala Leu Gly His Thr Phe Lys Glu Leu Lys Gly Gln Gly Ala Leu
 290 295 300
 Met Thr Gly Ser Ala Tyr Pro Gly Met Trp Asp Val Ser Tyr Glu Pro
 305 310 315 320
 Gly Asp Leu Glu Ser Met Ala Glu Ala Tyr Ser Arg Thr Tyr Ile Asn

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

 385 390 395 400
 Phe Ser Glu Ala Gln Phe Asp Thr Arg Val Glu Ala Leu Ala Glu Met
 405 410 415
 Met Ala Asp Lys Lys Ala Asn Glu Gly Gly Asn His
 420 425

 <210> 101
 <211> 372
 <212> PRT
 <213> Megasphaera elsdenii
 <400> 101

Met Ser Gln Ile Asp Glu Leu Ile Ser Lys Leu Gln Glu Val Ser Asn
1 5 10 15

His Pro Gln Lys Thr Val Leu Asn Tyr Lys Lys Gln Gly Lys Gly Leu
20 25 30

Val Gly Met Met Pro Tyr Tyr Ala Pro Glu Glu Ile Val Tyr Ala Ala
35 40 45

Gly Tyr Leu Pro Val Gly Met Phe Gly Ser Gln Asn Pro Gln Ile Ser
50 55 60

Ala Ala Arg Thr Tyr Leu Pro Pro Phe Ala Cys Ser Leu Met Gln Ala
65 70 75 80

Asp Met Glu Leu Gln Leu Asn Gly Thr Tyr Asp Cys Leu Asp Ala Val
85 90 95

Ile Phe Ser Val Pro Cys Asp Thr Leu Arg Cys Met Ser Gln Lys Trp
100 105 110

His Gly Lys Ala Pro Val Ile Val Phe Thr Gln Pro Gln Asn Arg Lys
115 120 125

Ile Arg Pro Ala Val Asp Phe Leu Lys Ala Glu Tyr Glu His Val Arg
130 135 140

Thr Glu Leu Gly Arg Ile Leu Asn Val Lys Ile Ser Asp Leu Ala Ile
145 150 155 160

Gln Glu Ala Ile Lys Val Tyr Asn Glu Asn Arg Gln Val Met Arg Glu
165 170 175

Phe Cys Asp Val Ala Ala Gln Tyr Pro Gln Ile Phe Thr Pro Ile Lys
180 185 190

Arg His Asp Val Ile Lys Ala Arg Trp Phe Met Asp Lys Ala Glu His
195 200 205

Thr Ala Leu Val Arg Glu Leu Ile Asp Ala Val Lys Lys Glu Pro Val
210 215 220

Gln Pro Trp Asn Gly Lys Lys Val Ile Leu Ser Gly Ile Met Ala Glu
225 230 235 240

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

245 250 255
 Ala Asp Asp Leu Ala Gln Glu Ser Arg Gln Phe Arg Thr Asp Val Pro
 260 265 270

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

 Pro Glu Phe Glu Ala Ala Gly Val Arg Tyr Thr Val Leu Asp Leu Asp
 340 345 350
 Ile Glu Ser Pro Ser Leu Glu Gln Leu Arg Thr Arg Ile Gln Ala Phe
 355 360 365
 Ser Glu Ile Leu
 370
 <210> 102
 <211> 519
 <212> PRT
 <213> *Saccharomyces cerevisiae*
 <400> 102
 Met Phe Ser Arg Ser Thr Leu Cys Leu Lys Thr Ser Ala Ser Ser Ile
 1 5 10 15

 Gly Arg Leu Gln Leu Arg Tyr Phe Ser His Leu Pro Met Thr Val Pro
 20 25 30
 Ile Lys Leu Pro Asn Gly Leu Glu Tyr Glu Gln Pro Thr Gly Leu Phe
 35 40 45
 Ile Asn Asn Lys Phe Val Pro Ser Lys Gln Asn Lys Thr Phe Glu Val
 50 55 60
 Ile Asn Pro Ser Thr Glu Glu Glu Ile Cys His Ile Tyr Glu Gly Arg

| | | | |
|---|-----|-----|-----|
| 65 | 70 | 75 | 80 |
| Glu Asp Asp Val Glu Glu Ala Val Gln Ala Ala Asp Arg Ala Phe Ser | | | |
| | 85 | 90 | 95 |
| Asn Gly Ser Trp Asn Gly Ile Asp Pro Ile Asp Arg Gly Lys Ala Leu | | | |
| | 100 | 105 | 110 |
| Tyr Arg Leu Ala Glu Leu Ile Glu Gln Asp Lys Asp Val Ile Ala Ser | | | |
| | 115 | 120 | 125 |
| Ile Glu Thr Leu Asp Asn Gly Lys Ala Ile Ser Ser Ser Arg Gly Asp | | | |
| | 130 | 135 | 140 |
| Val Asp Leu Val Ile Asn Tyr Leu Lys Ser Ser Ala Gly Phe Ala Asp | | | |
| 145 | 150 | 155 | 160 |
| Lys Ile Asp Gly Arg Met Ile Asp Thr Gly Arg Thr His Phe Ser Tyr | | | |
| | 165 | 170 | 175 |
| Thr Lys Arg Gln Pro Leu Gly Val Cys Gly Gln Ile Ile Pro Trp Asn | | | |
| | 180 | 185 | 190 |
| Phe Pro Leu Leu Met Trp Ala Trp Lys Ile Ala Pro Ala Leu Val Thr | | | |
| | 195 | 200 | 205 |
| Gly Asn Thr Val Val Leu Lys Thr Ala Glu Ser Thr Pro Leu Ser Ala | | | |
| | 210 | 215 | 220 |
| Leu Tyr Val Ser Lys Tyr Ile Pro Gln Ala Gly Ile Pro Pro Gly Val | | | |
| 225 | 230 | 235 | 240 |
| Ile Asn Ile Val Ser Gly Phe Gly Lys Ile Val Gly Glu Ala Ile Thr | | | |
| | 245 | 250 | 255 |
| Asn His Pro Lys Ile Lys Lys Val Ala Phe Thr Gly Ser Thr Ala Thr | | | |
| | 260 | 265 | 270 |
| Gly Arg His Ile Tyr Gln Ser Ala Ala Ala Gly Leu Lys Lys Val Thr | | | |
| | 275 | 280 | 285 |
| Leu Glu Leu Gly Gly Lys Ser Pro Asn Ile Val Phe Ala Asp Ala Glu | | | |
| | 290 | 295 | 300 |
| Leu Lys Lys Ala Val Gln Asn Ile Ile Leu Gly Ile Tyr Tyr Asn Ser | | | |
| 305 | 310 | 315 | 320 |

Gly Glu Val Cys Cys Ala Gly Ser Arg Val Tyr Val Glu Glu Ser Ile

325 330 335

Tyr Asp Lys Phe Ile Glu Glu Phe Lys Ala Ala Ser Glu Ser Ile Lys

340 345 350

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

355 360 365

Gln Met Gln Leu Asn Lys Ile Leu Lys Tyr Val Asp Ile Gly Lys Asn

370 375 380

Glu Gly Ala Thr Leu Ile Thr Gly Gly Glu Arg Leu Gly Ser Lys Gly

385 390 395 400

Tyr Phe Ile Lys Pro Thr Val Phe Gly Asp Val Lys Glu Asp Met Arg

405 410 415

Ile Val Lys Glu Glu Ile Phe Gly Pro Val Val Thr Val Thr Lys Phe

420 425 430

Lys Ser Ala Asp Glu Val Ile Asn Met Ala Asn Asp Ser Glu Tyr Gly

435 440 445

Leu Ala Ala Gly Ile His Thr Ser Asn Ile Asn Thr Ala Leu Lys Val

450 455 460

Ala Asp Arg Val Asn Ala Gly Thr Val Trp Ile Asn Thr Tyr Asn Asp

465 470 475 480

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

485 490 495

Arg Glu Met Ser Val Asp Ala Leu Gln Asn Tyr Leu Gln Val Lys Ala

500 505 510

Val Arg Ala Lys Leu Asp Glu

515

<210> 103

<211> 495

<212> PRT

<213> Escherichia coli

<400> 103

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

Ala Ala Arg Gly Val Phe Glu Arg Gly Asp Trp Ser Leu Ser Ser Pro
 65 70 75 80
 Ala Lys Arg Lys Ala Val Leu Asn Lys Leu Ala Asp Leu Met Glu Ala
 85 90 95
 His Ala Glu Glu Leu Ala Leu Leu Glu Thr Leu Asp Thr Gly Lys Pro
 100 105 110
 Ile Arg His Ser Leu Arg Asp Asp Ile Pro Gly Ala Ala Arg Ala Ile
 115 120 125

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

Ser Glu Lys Ser Pro Leu Ser Ala Ile Arg Leu Ala Gly Leu Ala Lys
 195 200 205
 Glu Ala Gly Leu Pro Asp Gly Val Leu Asn Val Val Thr Gly Phe Gly
 210 215 220
 His Glu Ala Gly Gln Ala Leu Ser Arg His Asn Asp Ile Asp Ala Ile
 225 230 235 240
 Ala Phe Thr Gly Ser Thr Arg Thr Gly Lys Gln Leu Leu Lys Asp Ala

| | | |
|---|-----|-----|
| 245 | 250 | 255 |
| Gly Asp Ser Asn Met Lys Arg Val Trp Leu Glu Ala Gly Gly Lys Ser | | |
| 260 | 265 | 270 |
| Ala Asn Ile Val Phe Ala Asp Cys Pro Asp Leu Gln Gln Ala Ala Ser | | |
| 275 | 280 | 285 |
| Ala Thr Ala Ala Gly Ile Phe Tyr Asn Gln Gly Gln Val Cys Ile Ala | | |
| 290 | 295 | 300 |
| Gly Thr Arg Leu Leu Leu Glu Glu Ser Ile Ala Asp Glu Phe Leu Ala | | |
| 305 | 310 | 315 |
| | | 320 |
| Leu Leu Lys Gln Gln Ala Gln Asn Trp Gln Pro Gly His Pro Leu Asp | | |
| 325 | 330 | 335 |
| Pro Ala Thr Thr Met Gly Thr Leu Ile Asp Cys Ala His Ala Asp Ser | | |
| 340 | 345 | 350 |
| Val His Ser Phe Ile Arg Glu Gly Glu Ser Lys Gly Gln Leu Leu Leu | | |
| 355 | 360 | 365 |
| Asp Gly Arg Asn Ala Gly Leu Ala Ala Ala Ile Gly Pro Thr Ile Phe | | |
| 370 | 375 | 380 |
| Val Asp Val Asp Pro Asn Ala Ser Leu Ser Arg Glu Glu Ile Phe Gly | | |
| 385 | 390 | 395 |
| Pro Val Leu Val Val Thr Arg Phe Thr Ser Glu Glu Gln Ala Leu Gln | | |
| 405 | 410 | 415 |
| Leu Ala Asn Asp Ser Gln Tyr Gly Leu Gly Ala Ala Val Trp Thr Arg | | |
| 420 | 425 | 430 |
| Asp Leu Ser Arg Ala His Arg Met Ser Arg Arg Leu Lys Ala Gly Ser | | |
| 435 | 440 | 445 |
| Val Phe Val Asn Asn Tyr Asn Asp Gly Asp Met Thr Val Pro Phe Gly | | |
| 450 | 455 | 460 |
| Gly Tyr Lys Gln Ser Gly Asn Gly Arg Asp Lys Ser Leu His Ala Leu | | |
| 465 | 470 | 475 |
| Glu Lys Phe Thr Glu Leu Lys Thr Ile Trp Ile Ser Leu Glu Ala | | |
| 485 | 490 | 495 |

<210> 104

<211> 495

<212> PRT

<213> *Klebsiella pneumoniae*

<400> 104

Met Asn Phe Gln His Leu Ala Tyr Trp Gln Glu Lys Ala Lys Asn Leu

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

65 70 75 80
Ala Gln Arg Lys Ala Ile Leu Thr Arg Phe Ala Asp Leu Met Glu Ala
85 90 95
His Arg Glu Glu Leu Ala Leu Leu Glu Thr Leu Asp Thr Gly Lys Pro
100 105 110
Ile Arg His Ser Leu Arg Asp Asp Ile Pro Gly Ala Ala Arg Ala Ile
115 120 125
Arg Trp Tyr Ala Glu Ala Leu Asp Lys Val Tyr Gly Glu Val Ala Pro

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

195 200 205

Glu Ala Gly Leu Pro Asp Gly Val Leu Asn Val Val Ser Gly Phe Gly
 210 215 220
 His Glu Ala Gly Gln Ala Leu Ala Leu His Pro Asp Val Glu Val Ile
 225 230 235 240
 Thr Phe Thr Gly Ser Thr Arg Thr Gly Lys Gln Leu Leu Lys Asp Ala
 245 250 255
 Gly Asp Ser Asn Met Lys Arg Val Trp Leu Glu Ala Gly Gly Lys Ser
 260 265 270
 Ala Asn Ile Val Phe Ala Asp Cys Pro Asp Leu Gln Gln Ala Val Arg
 275 280 285
 Ala Thr Ala Gly Gly Ile Phe Tyr Asn Gln Gly Gln Val Cys Ile Ala
 290 295 300
 Gly Thr Arg Leu Leu Leu Glu Glu Ser Ile Ala Asp Glu Phe Leu Ala
 305 310 315 320
 Arg Leu Lys Ala Glu Ala Gln His Trp Gln Pro Gly Asn Pro Leu Asp
 325 330 335
 Pro Asp Thr Thr Met Gly Met Leu Ile Asp Asn Thr His Ala Asp Asn
 340 345 350
 Val His Ser Phe Ile Arg Gly Gly Glu Ser Gln Ser Thr Leu Phe Leu
 355 360 365
 Asp Gly Arg Lys Asn Pro Trp Pro Ala Ala Val Gly Pro Thr Ile Phe
 370 375 380
 Val Asp Val Asp Pro Ala Ser Thr Leu Ser Arg Glu Glu Ile Phe Gly
 385 390 395 400
 Pro Val Leu Val Val Thr Arg Phe Lys Ser Glu Glu Glu Ala Leu Lys
 405 410 415
 Leu Ala Asn Asp Ser Asp Tyr Gly Leu Gly Ala Ala Val Trp Thr Arg
 420 425 430
 Asp Leu Ser Arg Ala His Arg Met Ser Arg Arg Leu Lys Ala Gly Ser
 435 440 445
 Val Phe Val Asn Asn Tyr Asn Asp Gly Asp Met Thr Val Pro Phe Gly

| | | | |
|---|------|-----|-----|
| 450 | 455 | 460 | |
| Gly Tyr Lys Gln Ser Gly Asn Gly Arg Asp Lys Ser Leu His Ala Leu | | | |
| 465 | 470 | 475 | 480 |
| Glu Lys Phe Thr Glu Leu Lys Thr Ile Trp Ile Ala Leu Glu Ser | | | |
| | 485 | 490 | 495 |
| <210> 105 | | | |
| <211> 1053 | | | |
| <212> DNA | | | |
| <213> Issatchenkia orientalis | | | |
| <400> 105 | | | |
| atgtcttacg aaatcccaca aacacaaaag gcctgtgtct tttagcaaaa cggcggccca | 60 | | |
| atcacataca aggacattcc agttccaaag ccaaaaccta ctgagatttt agtcaaggtt | 120 | | |
| ctgtactctg gtgtctgccca caccgacttg cacgcatgga aggggtgactg gcctctagct | 180 | | |
| accaagttgc cattgggttg tggtcacgaa ggtgccggtg ttgttgttc caagggtgaa | 240 | | |
| aacgtcacct cttttgagat tggtgattac gcaggtatca agtgggtgaa tggttcatgt | 300 | | |
| atgggtttgt aattctgtga acaaggtgct gaaccaaact gtcctaaggc cgacttgagt | 360 | | |
| ggttacaccc acgacggttc ctccaacag tatgtactg ctgacgtat tcaagctgca | 420 | | |
| cacatctcca aggaaaccga cttggctggt gtgtctccaa tcttgtgtgc aggtgtcact | 480 | | |
| gtctacaagg ctttaaagac tgcagacctt agagcaggtg aatgggtttg tatttcggt | 540 | | |
| gcagctgggtg gtttaggttc tcttctatt caatatgcaa aggctatggg tctgagagtt | 600 | | |
| gttggtattg acggtggtga cgaagaag gaattgtga aatcccttg tgctgaagca | 660 | | |
| tttattgatt tcacaagac caaggatatc gtcaaggctg tccaagaggc aaccaatggt | 720 | | |
| ggtccacatg gtgtcatcaa tgtctctgtc tctgaagctg caatttctca atcttgtgaa | 780 | | |
| tacgttagac ctctaggtaa ggtgttctt gtgtgtttac cagcaggcgc acaagtcaaa | 840 | | |
| actggtgtct ttgaagccgt tgtcaagtct attgaaatta agggttctta tgtcggtaac | 900 | | |
| agaaaggata cggccgaagc acttgacttc tacaactagag gcttgggtcaa gtctccattc | 960 | | |
| aagattgtcg gtttatccga attgccaaaa gtctttgaac tcatggaaca gggtgaagatt | 1020 | | |
| ttaggtagaa tggctctaga cacctccaaa taa | 1053 | | |
| <210> 106 | | | |
| <211> 350 | | | |
| <212> PRT | | | |
| <213> Issatchenkia orientalis | | | |

<400> 106

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

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

290 295 300
 Ala Glu Ala Leu Asp Phe Tyr Thr Arg Gly Leu Val Lys Ser Pro Phe
 305 310 315 320
 Lys Ile Val Gly Leu Ser Glu Leu Pro Lys Val Phe Glu Leu Met Glu
 325 330 335
 Gln Gly Lys Ile Leu Gly Arg Met Val Leu Asp Thr Ser Lys
 340 345 350

<210> 107

<211> 1131

<212> DNA

<213> Issatchenkia orientalis

<400> 107

| | |
|--|-----|
| atgtttgcat caaccttcag aagtcaagct gtaagagctg caagatttac tagattccaa | 60 |
| tccacttttg ccattcctga gaagcaaatg ggtgttatct ttgaaactca tgggtggtcct | 120 |
| ttacaataca aggaaattcc agttccaaaa ccaaaaccaa ctgaaatttt aatcaatgtt | 180 |
| aaatactctg gtgtctgcc aaccgattta cagcatgga aaggtgactg gccattacca | 240 |
| gcaaagttac ccctagtgtg tggtcacgaa ggtgcgggca ttgttgttgc gaaaggttct | 300 |
| gcagttacca actttgagat tggcgattat gctggtatta agtgggttaa cggttcatgt | 360 |
| atgtcatgtg aattctgtga acaaggtgat gaatctaact gtgaacatgc cgatttgagt | 420 |
| | |
| ggttatactc atgatggttc tttccaacaa tatgccactg ctgacgctat tcaagctgca | 480 |
| aagatcccaa agggtagcga cttatctgaa gttgcgcaa ttttatgtgc tgggtgttact | 540 |
| gtctataaag ctttgaaaac tgctgattta agagcaggtc aatgggttgc gatttctggt | 600 |
| gccgtggtg gtctaggttc tcttctgtgc caatatgcaa aggcaatggg tctaagagtt | 660 |
| ttaggtatcg atggtggtga aggtaaaaag gaactttttg aacaatgtgg tggatgatgtg | 720 |
| tttatcgatt tcaccagata cccaagagat gcacctgaaa agatgggttgc tgatattaag | 780 |

gctgcaacta acggtttggg tccacacggt gttatcaatg tctctgtctc cccagctgct 840

atctctcaat catgtgacta tgttagagca actggtaagg ttgtccttgt cggtatgccca 900

tctgggtgctg tctgtaagtc tgatgtcttc actcatgttg ttaaatcctt acaaattaaa 960

ggttcttatg ttggtaacag agcagatacc agagaagctt tggaattctt taatgaaggt 1020

aaggtcagat ctccaatcaa gggtgtccca ttatctactt tacctgaaat ttacgaattg 1080

atggagcaag gtaagatttt aggtagatac gttgttgata cttctaaata a 1131

<210> 108

<211> 376

<212> PRT

<213> Issatchenkia orientalis

<400> 108

Met Phe Ala Ser Thr Phe Arg Ser Gln Ala Val Arg Ala Ala Arg Phe

1 5 10 15
Thr Arg Phe Gln Ser Thr Phe Ala Ile Pro Glu Lys Gln Met Gly Val

Ile Phe Glu Thr His Gly Gly Pro Leu Gln Tyr Lys Glu Ile Pro Val

35 40 45

Pro Lys Pro Lys Pro Thr Glu Ile Leu Ile Asn Val Lys Tyr Ser Gly

50 55 60

Val Cys His Thr Asp Leu His Ala Trp Lys Gly Asp Trp Pro Leu Pro

65 70 75 80

Ala Lys Leu Pro Leu Val Gly Gly His Glu Gly Ala Gly Ile Val Val

Ala Lys Gly Ser Ala Val Thr Asn Phe Glu Ile Gly Asp Tyr Ala Gly

Ile Lys Trp Leu Asn Gly Ser Cys Met Ser Cys Glu Phe Cys Glu Gln

115 120 125

Gly Asp Glu Ser Asn Cys Glu His Ala Asp Leu Ser Gly Tyr Thr His

130 135 140

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

145 150 155 160

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

<210> 109

<211> 1134

<212> DNA

<213> Issatchenkia orientalis

<400> 109

| | |
|---|------|
| atgttatcca agaccatcac tgctgcattg aggggcaata caactcgtac tgcattcaga | 60 |
| | |
| atcaatgccca ttagaagttt agcgatccca gctattccag agacacaaaa ggggtgttatc | 120 |
| ttttatgaga acggagggtga actattttac aaggacattc cagttccaaa gccaaagcca | 180 |
| aatgagattt tgggtgaatgt caagtattct ggtgtttgtc ataccgattt acacgcatgg | 240 |
| aaaggtgact ggcctttggc gaccaagtgt ccattgggtg gtggacatga aggtgccgga | 300 |
| gttgtgttg ctaaggggga caatgtcacc aactttgaaa ttggcgatta tgccggtatc | 360 |
| aagtgggtga atggttcatg tatgggggtg gaattttgcc aacaaggtgc agagccaaac | 420 |
| tgtccacagg ccgacttgag tggttacacc catgacgggt cctttcaaca atatgccact | 480 |
| | |
| gccgatgctg ttcaggcagc caagattcct cagggcactg atttggctca agttgcgcca | 540 |
| attttatgtg caggtattac tgtctataag gctttaaaga ctgcagaatt aagaccaggt | 600 |
| caatgggttg ccatttctgg tgctgctgga ggtttaggtt ctcttgctgt tcaatatgcc | 660 |
| aaggccatgg gtttgagagt tttgggtatt gatgggtgtg aggagaaggg caagtttgca | 720 |
| aagtctcttg gagctgaagt tttcattgat ttcaccaaat ccaaggacat tgtcaaggat | 780 |
| atccaagagg ccaccaatgg tgggccacat ggtgtcatta atgtttctgt tttccagct | 840 |
| gctatttctc aaagtaccca gtatgtcaga accttgggta aggttgtcct tgttggatta | 900 |
| | |
| ccagcgcattg ctgtatgcga gtcttcggtt ttcgaccatg ttgtcaagtc gattcaaatt | 960 |
| agaggctctt atgttggtaa caggaagat actagtgagg ctattgattt tttaccagg | 1020 |
| ggtttagtga agtcaccaat taagattgtt ggtttgagtg agttgccaaa gatctatgaa | 1080 |
| ttgatggagc aaggtgaagt tttaggcaga tatgttggtg acacttcgaa atga | 1134 |

<210> 110

<211> 377

<212> PRT

<213> Issatchenkia orientalis

<400> 110

Met Leu Ser Lys Thr Ile Thr Ala Ala Leu Arg Gly Asn Thr Thr Arg

1 5 10 15

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

20 25 30

Pro Glu Thr Gln Lys Gly Val Ile Phe Tyr Glu Asn Gly Gly Glu Leu

35 40 45

Phe Tyr Lys Asp Ile Pro Val Pro Lys Pro Lys Pro Asn Glu Ile Leu
 50 55 60
 Val Asn Val Lys Tyr Ser Gly Val Cys His Thr Asp Leu His Ala Trp
 65 70 75 80

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

 Asp Leu Ser Gly Tyr Thr His Asp Gly Ser Phe Gln Gln Tyr Ala Thr
 145 150 155 160
 Ala Asp Ala Val Gln Ala Ala Lys Ile Pro Gln Gly Thr Asp Leu Ala
 165 170 175
 Gln Val Ala Pro Ile Leu Cys Ala Gly Ile Thr Val Tyr Lys Ala Leu
 180 185 190
 Lys Thr Ala Glu Leu Arg Pro Gly Gln Trp Val Ala Ile Ser Gly Ala
 195 200 205

 Ala Gly Gly Leu Gly Ser Leu Ala Val Gln Tyr Ala Lys Ala Met Gly
 210 215 220
 Leu Arg Val Leu Gly Ile Asp Gly Gly Glu Glu Lys Gly Lys Phe Ala
 225 230 235 240
 Lys Ser Leu Gly Ala Glu Val Phe Ile Asp Phe Thr Lys Ser Lys Asp
 245 250 255
 Ile Val Lys Asp Ile Gln Glu Ala Thr Asn Gly Gly Pro His Gly Val
 260 265 270

 Ile Asn Val Ser Val Ser Pro Ala Ala Ile Ser Gln Ser Thr Gln Tyr
 275 280 285
 Val Arg Thr Leu Gly Lys Val Val Leu Val Gly Leu Pro Ala His Ala

290 295 300
Val Cys Glu Ser Ser Val Phe Asp His Val Val Lys Ser Ile Gln Ile
305 310 315 320
Arg Gly Ser Tyr Val Gly Asn Arg Glu Asp Thr Ser Glu Ala Ile Asp
325 330 335

Phe Phe Thr Arg Gly Leu Val Lys Ser Pro Ile Lys Ile Val Gly Leu
340 345 350
Ser Glu Leu Pro Lys Ile Tyr Glu Leu Met Glu Gln Gly Lys Ile Leu
355 360 365
Gly Arg Tyr Val Val Asp Thr Ser Lys
370 375

<210> 111

<211> 1347

<212> DNA

<213> Issatchenkia orientalis

<400> 111

atgtctcctt cacaaattaa cgttgacaac ttatctaatt ggactgaaga attcaaattct 60

gacccaaga ctcaaactgg gggttctgta ttgcaacatt ccaacattga tgaggtcttg 120

attaacagag atgcagaaat cgccaacaag catactttca accacaagat tgaaattgaa 180

ggtctacctg tcatggatca gaaggcttct ggtagatgtt ggttgtttgc atcgactaac 240

ttgatgcgtg ttactgcaat gaagaaatac aatttgaagg aaatcaagct ttcccatcg 300

tatttgtttt tctatgacaa attggaaaga gcaactatt tccttgaaca aatcatcgac 360

actcataagg aaccaatcga ttcaagattg gttcaatatt tcctgaccaa tccagttgaa 420

gatggtggtc aattcacat gatggcaca attgctacca aatacgggtg tgttcctgat 480

caagtctacc cagattcttt caacacaacc acttcgagga ttatgaacag attagtcaac 540

cacagattac gttcttatgc aatgacttta cgtaacgctc tagatgaagg taaagatgta 600

atgtccttga agaagagat gcaaaaagaa atttatcggt tgctaacaat gttccttggt 660

aaccaccaa agccaaacga agagtttgc tgggaattca ccgataaaga tggtaaatat 720

gaatctatta aaactacacc attaaaatat gcaactgaag ttttggattt ccatgctcca 780

gaatatgttt ccttggttaa tgaccaaga aataagtata acaagatggt tcaagttgaa 840

agattaggtg atgttgctgg tggcgaacca gttgcatact taaacttaga aattgaaaag 900

ttatctcaag ctgttgtaa cagaatcaaa aataacaaac cagttttctt tggtagcgat 960

acacctaaat ttatggataa aagtagaggt attatggata tcaatttatg ggactatgag 1020

ttattaggtt atgatgtccg taccatgtca aagaaggaaa gagttgtttt tggtagattct 1080

ttaatgaccc acgtatgtt gattactgca gtgcacgttg atgaaaatgg caaacctgtc 1140

agatacagag tcgaaaacag ttgggggtacc aagagtgggc aagaaggtaa ttacacaatg 1200

acccaagaat attttgaaga gtacgtttat caagtagtca ttgaaaagag tgaatttgct 1260

gccctaaacc tcgatgtttc cattctggag gataaagaac cagtcgtctt gccacattat 1320

gaccctatgg gtgcacttgc ttataa 1347

<210> 112

<211> 448

<212> PRT

<213> Issatchenkia orientalis

<400> 112

Met Ser Pro Ser Gln Ile Asn Val Asp Asn Leu Ser Asn Trp Thr Glu

1 5 10 15

Glu Phe Lys Ser Asp Ala Lys Thr Gln Ile Gly Gly Ser Val Leu Gln

20 25 30

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

35 40 45

Asn Lys His Ile Phe Asn His Lys Ile Glu Ile Glu Gly Leu Pro Val

50 55 60

Met Asp Gln Lys Ala Ser Gly Arg Cys Trp Leu Phe Ala Ser Thr Asn

65 70 75 80

Leu Met Arg Val Thr Ala Met Lys Lys Tyr Asn Leu Lys Glu Ile Lys

85 90 95

Leu Ser Pro Ser Tyr Leu Phe Phe Tyr Asp Lys Leu Glu Arg Ala Asn

100 105 110

Tyr Phe Leu Glu Gln Ile Ile Asp Thr His Lys Glu Pro Ile Asp Ser

115 120 125

Arg Leu Val Gln Tyr Phe Leu Thr Asn Pro Val Glu Asp Gly Gly Gln

130 135 140

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

385 390 395 400
 Thr Gln Glu Tyr Phe Glu Glu Tyr Val Tyr Gln Val Val Ile Glu Lys
 405 410 415
 Ser Glu Phe Ala Ala Leu Asn Leu Asp Val Ser Ile Leu Glu Asp Lys
 420 425 430
 Glu Pro Val Val Leu Pro Pro Tyr Asp Pro Met Gly Ala Leu Ala Leu

 435 440 445
 <210> 113
 <211> 1737
 <212> DNA
 <213> Issatchenkia orientalis
 <400> 113

atgttactca gatcactaaa ctcttctgct cgtttgtgtca aacaaacaac cagaacaaag 60
 gtttaggtatc tcagccacgt cagtgggtgca agcatggcga aacctacatt gaagaacaac 120
 tcgagagaat ccaacaaatc cagaaactat ctaattgctg ctgtgacagc attggctgta 180
 tcaacctcaa ttggagttgc cgtacatgtg aaggaccct tgtataacga tgctaccggc 240
 agtgattctc cgagaagtat atctgttgac gagtttgtca agcataattc acaaaacgac 300

 tgttggattg caatcaatgg caaggtttat gatttactg attttattcc aaaccatcca 360
 ggtgggggtac ctccattagt taatcatgct ggttatgatg gtactaaact ttatgagaaa 420
 ttgcattcaa aaggtacaat tgagaaattc ttgccaaagg ataagtttct gggtgtgtta 480
 gatggtgaag cgccaaaatt ggaagcagac tatttgggtgg acgatgatga acaagagaga 540
 ctggattatt tgaacaactt acctcctttg tcatctattc agaatgttta tgatttcgaa 600
 tacttggcca agaagatfff acctaaagat gcctgggcat attattcttg tgggtccgat 660
 gatgaaatca caatgagaga aaaccattat gcttatcaaa gagtttattt cagaccaaga 720

 atttgtgttg atgtcaagga agttgatact tcttatgaaa tgtaggcac taaaacctct 780
 gttccttttt atgtatctgc caccgctttg gctaaattag gccatcctga tggatgaatgc 840
 tcaattgcta gaggcgctgg taaggaaggt gtcgttcaaa tgatttcgac cctttcctca 900
 atgtcattag atgaaattgc cgctgctaga attccaggtg caaccaatg gttccaatta 960
 tacattaatg aggatagaaa tgcgctaaa ggtctgggtc aacatgcaga agacttgggt 1020
 atgaaggcta tctttataac tgttgatgct ccttctctag gtaacagaga aaaggataaa 1080
 agattaaagt ttgttaatga caccgatgtc gatttgggtg attccgcaga tcgaaacagt 1140

ggtagcttcaa aggcactatc ttcgttcatt gatgcttctg tctcttggaa tgacgtcaaa 1200
gcggtcaagt cgtggactaa attgcctgtc ttagttaaag gtgttcaaac agttgaagac 1260
gttattgaag cttacgatgc tggttgtcaa ggtgtgtgtt tgtcaaacca cggtagtagg 1320
caactagata ctgctcctcc tccaatcgaa ttattagctg aaactgttcc aactttgaag 1380
agattgggta aattaagacc agattttgaa attttaattg acggtgggtg caaaagaggt 1440
accgatattt tgaaagcagt cgcaatcggg ggccaagatg tcagagtctc agttggtatg 1500
ggtagacctt tcttatatgc caactcttgc tatggtgaag caggtgttag aaaattaatt 1560

caaaatctaa aggatgaatt agaatggat atgagattgt tgggtgtcac taaaatggac 1620
cagctatctt cgaaacatgt cgatactaaa cgtttgattg gtagagatgc gatcaactat 1680
ttgtatgata atgtatacag cccaatcgaa accgttaaatt tcaacaatga agattga 1737

<210> 114

<211> 578

<212> PRT

<213> Issatchenkia orientalis

<400> 114

Met Leu Leu Arg Ser Leu Asn Ser Ser Ala Arg Cys Val Lys Gln Thr

1 5 10 15

Thr Arg Thr Lys Val Arg Tyr Leu Ser His Val Ser Gly Ala Ser Met

20 25 30

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

35 40 45

Asn Tyr Leu Ile Ala Ala Val Thr Ala Leu Ala Val Ser Thr Ser Ile

50 55 60

Gly Val Ala Val His Val Lys Asp Pro Leu Tyr Asn Asp Ala Thr Gly

65 70 75 80

Ser Asp Ser Pro Arg Ser Ile Ser Val Asp Glu Phe Val Lys His Asn

85 90 95

Ser Gln Asn Asp Cys Trp Ile Ala Ile Asn Gly Lys Val Tyr Asp Phe

100 105 110

Thr Asp Phe Ile Pro Asn His Pro Gly Gly Val Pro Pro Leu Val Asn

115 120 125

His Ala Gly Tyr Asp Gly Thr Lys Leu Tyr Glu Lys Leu His Pro Lys

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

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

<210> 115

<211> 1698

<212> DNA

<213> Issatchenkia orientalis

<220><221> misc_feature

<222> (941)..(941)

<223> N= A, C, T, or G

<220><221> misc_feature

<222> (943)..(943)

<223> N= A, C, T, or G

<400> 115

| | |
|--|------|
| atgttaagat cccagttcaa aaacattttg aaaaatgtta acaagaacca ttctctaagg | 60 |
| agaactttta ctccagcac ctcaaaggct ggaaaaaatg cttcatacaa tgccaagatt | 120 |
| atatctgcaa ccgtggcctc gattgttgca gcagctggct cttatatgtt ggtccagcct | 180 |
| tcactagcta atgatgaggc acagtctgct aatccaacta ggaagatctc tgttgacgaa | 240 |
| tttgttaaac acaaccatgc cgatgattgt tggatcactg ttaacggtaa cgtctatgac | 300 |
| ttgactgatt tcatttcaat gcatccaggt ggtactaccc cattgattca aaatgcaggt | 360 |
| | |
| cacgacgcaa ctgaaattta caacaagatt catecaaagg gtacaatcga gaacttctta | 420 |
| ccaaaggaaa agcaattggg tgttttgat ggtgaagctc ctaaaatcga agttgtgctt | 480 |
| gacgaaaagg agaaacacag attggagttg ttgaatcatc tccctgctct ttccagaatt | 540 |
| caaaacattt atgatttcga acatattgct tctagagttt tgagcgacca agcatggaac | 600 |
| tactattcat gtggtgccga agatgaaatc accttgaggg aaaatcatta tgcttaccaa | 660 |
| agaatctact ttaagccaaa atgttgtgtc aatgttcag aagttgatac ctctcatgaa | 720 |
| attttaggta caaaagcttc tgttccttc tacgtttccg cagccgcttc tgcaaagttg | 780 |
| | |
| gggcacgagg atggtgaatg ttccattgct agaggcag gtaaggaagg cgttattcaa | 840 |
| atgatttctt ctttctcttc caactctttg gaggaattg cagaatccag aattcctggt | 900 |
| gcaacacaat ggtttcaatt atacgttaat gaagacaagg ntnttgtgaa gaagacttta | 960 |
| aaaagggccg aaaacttggg tatgaaggcc atctttgtca ctgtggacgc tgctagtaga | 1020 |
| ggtaatatag aaaaagacat tacaatgaga attaccgaag atacagatga gttaatagac | 1080 |
| gattcttctg ttagagctgg ttctacctc ggtgcattgc cagctttcat tgacaagagg | 1140 |
| ctgacttggg atgaagttaa ggatattcatt tcatggacca agttaccagt tttgctgaag | 1200 |
| | |
| ggtgttcaaa gaactgatga tattgagaag gcaattgata ttggttgtaa gggtgtgtc | 1260 |
| ttgtccaatc atggtgtag acaattagat acttctctc ctccaataga agttatggct | 1320 |
| gaatctgttc caatcctaaa gcaaaaggg aaactggatc caaatttcag tattttcgtt | 1380 |
| gatggtgggt ttagaagagg tacagatatt ttgaaagctt tggctattgg tggcagagac | 1440 |
| tgtaaagttg ctgttggtct gggtagacct ttctttatg caaatactgg ttatggtgaa | 1500 |
| aagggtgtca gaaagccgt gcaaatctca agagaagaat taaaggctga catgagaatg | 1560 |

ttgggcgtta cctctttgaa cgagctagac gactcttaca ttgacaccag aagattacta 1620

ggtagagatg ctgttaacca catatacaac aacaactact acccaatgtc taagattcaa 1680

ttcaaaaacg aaaaataa 1698

<210> 116

<211> 565

<212> PRT

<213> Issatchenkia orientalis

<220><221> MISC_FEATURE

<222> (314)..(314)

<223> Xaa = Asp, Gly, Ala, or Val

<220><221> MISC_FEATURE

<222> (315)..(315)

<223> Xaa = Ile, Val, Leu, or Phe

<400> 116

Met Leu Arg Ser Gln Phe Lys Asn Ile Leu Lys Asn Val Asn Lys Asn

1 5 10 15

His Ser Leu Arg Arg Thr Phe Thr Ser Ser Thr Ser Lys Ala Gly Lys

20 25 30

Asn Ala Ser Tyr Asn Ala Lys Ile Ile Ser Ala Thr Val Ala Ser Ile

35 40 45

Val Ala Ala Ala Gly Ser Tyr Met Leu Val Gln Pro Ser Leu Ala Asn

50 55 60

Asp Glu Ala Gln Ser Ala Asn Pro Thr Arg Lys Ile Ser Val Asp Glu

65 70 75 80

Phe Val Lys His Asn His Ala Asp Asp Cys Trp Ile Thr Val Asn Gly

85 90 95

Asn Val Tyr Asp Leu Thr Asp Phe Ile Ser Met His Pro Gly Gly Thr

100 105 110

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

115 120 125

Lys Ile His Pro Lys Gly Thr Ile Glu Asn Phe Leu Pro Lys Glu Lys

130

135

140

Gln Leu Gly Val Leu Asp Gly Glu Ala Pro Lys Ile Glu Val Val Leu

145

150

155

160

Asp Glu Lys Glu Lys His Arg Leu Glu Leu Leu Asn His Leu Pro Ala

165

170

175

Leu Ser Arg Ile Gln Asn Ile Tyr Asp Phe Glu His Ile Ala Ser Arg

180

185

190

Val Leu Ser Asp Gln Ala Trp Asn Tyr Tyr Ser Cys Gly Ala Glu Asp

195

200

205

Glu Ile Thr Leu Arg Glu Asn His Tyr Ala Tyr Gln Arg Ile Tyr Phe

210

215

220

Lys Pro Lys Cys Cys Val Asn Val Ala Glu Val Asp Thr Ser His Glu

225

230

235

240

Ile Leu Gly Thr Lys Ala Ser Val Pro Phe Tyr Val Ser Ala Ala Ala

245

250

255

Ser Ala Lys Leu Gly His Glu Asp Gly Glu Cys Ser Ile Ala Arg Gly

260

265

270

Ala Gly Lys Glu Gly Val Ile Gln Met Ile Ser Ser Phe Ser Ser Asn

275

280

285

Ser Leu Glu Glu Ile Ala Glu Ser Arg Ile Pro Gly Ala Thr Gln Trp

290

295

300

Phe Gln Leu Tyr Val Asn Glu Asp Lys Xaa Xaa Val Lys Lys Thr Leu

305

310

315

320

Lys Arg Ala Glu Asn Leu Gly Met Lys Ala Ile Phe Val Thr Val Asp

325

330

335

Ala Ala Ser Arg Gly Asn Arg Glu Lys Asp Ile Thr Met Arg Ile Thr

340

345

350

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

355

360

365

Thr Ser Gly Ala Leu Pro Ala Phe Ile Asp Lys Arg Leu Thr Trp Asp

370

375

380

Glu Val Lys Asp Ile Ile Ser Trp Thr Lys Leu Pro Val Leu Leu Lys
385 390 395 400

Gly Val Gln Arg Thr Asp Asp Ile Glu Lys Ala Ile Asp Ile Gly Cys
405 410 415

Lys Gly Val Val Leu Ser Asn His Gly Gly Arg Gln Leu Asp Thr Ser
420 425 430

Pro Pro Pro Ile Glu Val Met Ala Glu Ser Val Pro Ile Leu Lys Gln
435 440 445

Lys Gly Lys Leu Asp Pro Asn Phe Ser Ile Phe Val Asp Gly Gly Val
450 455 460

Arg Arg Gly Thr Asp Ile Leu Lys Ala Leu Ala Ile Gly Gly Arg Asp
465 470 475 480

Cys Lys Val Ala Val Gly Leu Gly Arg Pro Phe Leu Tyr Ala Asn Thr
485 490 495

Gly Tyr Gly Glu Lys Gly Val Arg Lys Ala Val Gln Ile Leu Arg Glu
500 505 510

Glu Leu Lys Ala Asp Met Arg Met Leu Gly Val Thr Ser Leu Asn Glu
515 520 525

Leu Asp Asp Ser Tyr Ile Asp Thr Arg Arg Leu Leu Gly Arg Asp Ala
530 535 540

Val Asn His Ile Tyr Asn Asn Asn Tyr Tyr Pro Met Ser Lys Ile Gln
545 550 555 560

Phe Lys Asn Glu Lys
565

<210> 117

<211> 1167

<212> DNA

<213> Issatchenkia orientalis

<400> 117

atggtgtccc ctgctgaaag attatctact attgcgtcca caatcaagcc aaacagaaaa 60

gattctacat cattacaacc agaagactat ccggaacatc cgttcaaggt gacggttggt 120

ggttcggta actgggggtg tacaattgcc aaggttatag cggaaaacac cgttgagaga 180
 cctcgtcaat ttcaaagaga tgtaaatatg tgggtctatg aagaattgat tgaaggcgaa 240
 aagttgactg aaatcataaa taccaaacac gaaaacgtca agtacttgcc aggtatcaag 300
 ttgccagtta acgtttgtgc agttccagac attgttgagg ctgtgacagg ctgagacttg 360
 attgtcttta atattcctca ccaattttta ccaagaattt tatcccaatt aaagggttaag 420
 gtgaatcaa aggctagagc aatttcttgt ttgaaagggt tggatgtcaa tcctaatgga 480
 tgtaagttgc tctccactgt tattactgaa gagttgggta tttattgtgg tgccttatca 540

gggtgctaatt tagctcctga agttgcacaa tgtaaatggt cggaaacaac tgttgcata 600
 acaattccgg acgatttcag aggtaaaggc aaggatattg accatcaaat tctaaagagt 660
 ttgttccata gaccttattt ccatgttcgt gttattagt atgttgacagg tatttccatt 720
 gccggtgcac tcaagaatgt cgttgcctatg gctgctggat ttgtcgaagg tttagggttg 780
 ggtgataatg caaaggctgc agtcatgaga ataggtttgg tggaaacat tcaatttgcc 840
 aagacttttt tcgatggctg tcatgtgca acctttactc atgaatctgc aggtgttgcc 900
 gacctaatca ctacctgtgc cggcggccgt aacgttagag ttggtagata tatggcacia 960

cattctgtct ctgcaacgga ggctgaagaa aagttgttga atggccaatc ctgtcaaggt 1020
 atccacacia ctagggaagt ttacgagttc ctctccaaca tgggcaggac agatgagttc 1080
 ccactattta ccaccaccta ccgtatcatc tacgaaaact tccaattga gaagtgcca 1140
 gaatgccttg aacctgtgga agattaa 1167

<210> 118

<211> 388

<212> PRT

<213> Issatchenkia orientalis

<400> 118

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

1 5 10 15

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

20 25 30

His Pro Phe Lys Val Thr Val Val Gly Ser Gly Asn Trp Gly Cys Thr

35 40 45

Ile Ala Lys Val Ile Ala Glu Asn Thr Val Glu Arg Pro Arg Gln Phe

50 55 60

Gln Arg Asp Val Asn Met Trp Val Tyr Glu Glu Leu Ile Glu Gly Glu

| | | | |
|---|-----|-----|-----|
| 65 | 70 | 75 | 80 |
| Lys Leu Thr Glu Ile Ile Asn Thr Lys His Glu Asn Val Lys Tyr Leu | | | |
| | 85 | 90 | 95 |
| Pro Gly Ile Lys Leu Pro Val Asn Val Val Ala Val Pro Asp Ile Val | | | |
| | 100 | 105 | 110 |
| Glu Ala Cys Ala Gly Ser Asp Leu Ile Val Phe Asn Ile Pro His Gln | | | |
| | 115 | 120 | 125 |
| Phe Leu Pro Arg Ile Leu Ser Gln Leu Lys Gly Lys Val Asn Pro Lys | | | |
| | 130 | 135 | 140 |
| Ala Arg Ala Ile Ser Cys Leu Lys Gly Leu Asp Val Asn Pro Asn Gly | | | |
| 145 | 150 | 155 | 160 |
| Cys Lys Leu Leu Ser Thr Val Ile Thr Glu Glu Leu Gly Ile Tyr Cys | | | |
| | 165 | 170 | 175 |
| Gly Ala Leu Ser Gly Ala Asn Leu Ala Pro Glu Val Ala Gln Cys Lys | | | |
| | 180 | 185 | 190 |
| Trp Ser Glu Thr Thr Val Ala Tyr Thr Ile Pro Asp Asp Phe Arg Gly | | | |
| | 195 | 200 | 205 |
| Lys Gly Lys Asp Ile Asp His Gln Ile Leu Lys Ser Leu Phe His Arg | | | |
| | 210 | 215 | 220 |
| Pro Tyr Phe His Val Arg Val Ile Ser Asp Val Ala Gly Ile Ser Ile | | | |
| 225 | 230 | 235 | 240 |
| Ala Gly Ala Leu Lys Asn Val Val Ala Met Ala Ala Gly Phe Val Glu | | | |
| | 245 | 250 | 255 |
| Gly Leu Gly Trp Gly Asp Asn Ala Lys Ala Ala Val Met Arg Ile Gly | | | |
| | 260 | 265 | 270 |
| Leu Val Glu Thr Ile Gln Phe Ala Lys Thr Phe Phe Asp Gly Cys His | | | |
| | 275 | 280 | 285 |
| Ala Ala Thr Phe Thr His Glu Ser Ala Gly Val Ala Asp Leu Ile Thr | | | |
| | 290 | 295 | 300 |
| Thr Cys Ala Gly Gly Arg Asn Val Arg Val Gly Arg Tyr Met Ala Gln | | | |
| 305 | 310 | 315 | 320 |

His Ser Val Ser Ala Thr Glu Ala Glu Glu Lys Leu Leu Asn Gly Gln
325 330 335

Ser Cys Gln Gly Ile His Thr Thr Arg Glu Val Tyr Glu Phe Leu Ser
340 345 350

Asn Met Gly Arg Thr Asp Glu Phe Pro Leu Phe Thr Thr Thr Tyr Arg
355 360 365

Ile Ile Tyr Glu Asn Phe Pro Ile Glu Lys Leu Pro Glu Cys Leu Glu
370 375 380

Pro Val Glu Asp

385

<210> 119

<211> 1566

<212> DNA

<213> Issatchenkia orientalis

<400

> 119

atgttgtccc tctctaaaca gtcaagaaac tttttcaaat tgaactatit ttcagtcacc 60
caaataagcaa aatgtctgc aacttcgcgc actttcccaa ttatcaacga aacttaccaa 120
cagccaaccg ggcttttcat caacaatgaa ttgttagtg caaagtcagg taagactttt 180
gatgttaaca cccaattga tgagtctctc atttgtaaag tccaacaggc cgatgctgaa 240
gatgttgaaa ttgccgttca agcagcatct aaagcttaca agacttggag atttacaccg 300
ccaaatgaaa gaggcagata ctgaacaaa ttggccgatt tgatggacga aaagagagac 360
ttacttgcca aaattgaatc cttgataat ggtaaggcct tacattgtgc aaaattcgat 420

gtcaatcttg tcattgaata tttcagatac tgtgcagggt actgtgataa aatcgatggt 480
agaacaatta caaccgatgt agaacatttt acctacacta gaaaggaacc tttaggtgtc 540
tgtggtgcaa ttacaccttg gaacttccca ttgctgatgt ttgcttggaa aatcggcccg 600
gcttttagcaa ccggtataac cattatcttg aagcctgcca gtgcaacacc tctatcaaac 660
ctctttactt gtacctgat caaggaggcg ggcatccag ccggtgttgt taatgttgtt 720
ccaggttccg gtagaggctg tggttaactcc attttacaac atcctaaaat taagaaggtt 780
gcgtttaccg gatctacaga agttggtaaa actgttatga aggaatgtgc taattccatc 840

aaaaaggtta ctctgaatt ggggtgtaag tctccaaaca ttgttttcaa agactgtaac 900
gttgaacaaa ccattcaaaa ttgattact ggtattttct tcaatggtgg tgaagtctgt 960

tgtgctgggt ctagaattta cattgaagca accgatgaga aatgggtatac tgaattcttg 1020
 accaaattca aggagactgt tgaaaaatta aagattggta acccatttga agagggtgtt 1080
 ttccaagggt cacaaaccac tccagatcaa ttccaaactg tcttggacta catcaccgct 1140
 gctaacgaat ccagcttgaa actattaact ggtgggtaaaa gaattggcaa taagggtatac 1200
 tttgttgagc caactatctt ctacgatgtt cctcaaaatt ccaagttaac tcaagaagaa 1260

atctttggtc cagttgctgt tgttttacct ttcaagtcca ctgaagaatt gattgaaaag 1320
 gcaaatgatt ccgattttgg cttagggtcc ggtattcaca ctgaagattt caacaaggca 1380
 atttgggttt ccgaaaggct tgaagcaggt tctgtttgga tcaacactta caatgatttc 1440
 caccagctg ctccattcgg tggttacaag gaatccggtta ttggcagaga aatgggtatt 1500
 gaagctttcg acaactatac tcaaaccaag ttagtttagag cttagagttaa caagccagct 1560
 ttttag 1566

<210> 120

<211> 521

<212> PRT

<213> Issatchenkia orientalis

<400> 120

Met Leu Ser Leu Ser Lys Gln Ser Arg Asn Phe Phe Lys Leu Asn Tyr

1 5 10 15

Phe Ser Val Thr Gln Ile Ala Lys Met Ser Ala Thr Ser Val Thr Phe

20 25 30

Pro Ile Ile Asn Glu Thr Tyr Gln Gln Pro Thr Gly Leu Phe Ile Asn

35 40 45

Asn Glu Phe Val Ser Ala Lys Ser Gly Lys Thr Phe Asp Val Asn Thr

50 55 60

Pro Ile Asp Glu Ser Leu Ile Cys Lys Val Gln Gln Ala Asp Ala Glu

65 70 75 80

Asp Val Glu Ile Ala Val Gln Ala Ala Ser Lys Ala Tyr Lys Thr Trp

85 90 95

Arg Phe Thr Pro Pro Asn Glu Arg Gly Arg Tyr Leu Asn Lys Leu Ala

100 105 110

Asp Leu Met Asp Glu Lys Arg Asp Leu Leu Ala Lys Ile Glu Ser Leu

| | | |
|---|-----|-----|
| 115 | 120 | 125 |
| Asp Asn Gly Lys Ala Leu His Cys Ala Lys Phe Asp Val Asn Leu Val | | |
| 130 | 135 | 140 |
| Ile Glu Tyr Phe Arg Tyr Cys Ala Gly Tyr Cys Asp Lys Ile Asp Gly | | |
| 145 | 150 | 155 |
| Arg Thr Ile Thr Thr Asp Val Glu His Phe Thr Tyr Thr Arg Lys Glu | | |
| 165 | 170 | 175 |
| Pro Leu Gly Val Cys Gly Ala Ile Thr Pro Trp Asn Phe Pro Leu Leu | | |
| 180 | 185 | 190 |
| Met Phe Ala Trp Lys Ile Gly Pro Ala Leu Ala Thr Gly Asn Thr Ile | | |
| 195 | 200 | 205 |
| Ile Leu Lys Pro Ala Ser Ala Thr Pro Leu Ser Asn Leu Phe Thr Cys | | |
| 210 | 215 | 220 |
| Thr Leu Ile Lys Glu Ala Gly Ile Pro Ala Gly Val Val Asn Val Val | | |
| 225 | 230 | 235 |
| Pro Gly Ser Gly Arg Gly Cys Gly Asn Ser Ile Leu Gln His Pro Lys | | |
| 245 | 250 | 255 |
| Ile Lys Lys Val Ala Phe Thr Gly Ser Thr Glu Val Gly Lys Thr Val | | |
| 260 | 265 | 270 |
| Met Lys Glu Cys Ala Asn Ser Ile Lys Lys Val Thr Leu Glu Leu Gly | | |
| 275 | 280 | 285 |
| Gly Lys Ser Pro Asn Ile Val Phe Lys Asp Cys Asn Val Glu Gln Thr | | |
| 290 | 295 | 300 |
| Ile Gln Asn Leu Ile Thr Gly Ile Phe Phe Asn Gly Gly Glu Val Cys | | |
| 305 | 310 | 315 |
| Cys Ala Gly Ser Arg Ile Tyr Ile Glu Ala Thr Asp Glu Lys Trp Tyr | | |
| 325 | 330 | 335 |
| Thr Glu Phe Leu Thr Lys Phe Lys Glu Thr Val Glu Lys Leu Lys Ile | | |
| 340 | 345 | 350 |
| Gly Asn Pro Phe Glu Glu Gly Val Phe Gln Gly Ala Gln Thr Thr Pro | | |
| 355 | 360 | 365 |

Asp Gln Phe Gln Thr Val Leu Asp Tyr Ile Thr Ala Ala Asn Glu Ser
370 375 380

Ser Leu Lys Leu Leu Thr Gly Gly Lys Arg Ile Gly Asn Lys Gly Tyr
385 390 395 400

Phe Val Glu Pro Thr Ile Phe Tyr Asp Val Pro Gln Asn Ser Lys Leu
405 410 415

Thr Gln Glu Glu Ile Phe Gly Pro Val Ala Val Val Leu Pro Phe Lys
420 425 430

Ser Thr Glu Glu Leu Ile Glu Lys Ala Asn Asp Ser Asp Phe Gly Leu
435 440 445

Gly Ser Gly Ile His Thr Glu Asp Phe Asn Lys Ala Ile Trp Val Ser
450 455 460

Glu Arg Leu Glu Ala Gly Ser Val Trp Ile Asn Thr Tyr Asn Asp Phe
465 470 475 480

His Pro Ala Ala Pro Phe Gly Gly Tyr Lys Glu Ser Gly Ile Gly Arg
485 490 495

Glu Met Gly Ile Glu Ala Phe Asp Asn Tyr Thr Gln Thr Lys Leu Val
500 505 510

Arg Ala Arg Val Asn Lys Pro Ala Phe
515 520

<210> 121

<211> 1506

<212> DNA

<213> Issatchenkia orientalis

<400> 121

atgtcagcac tgttcagaac cattgagact ccaaacggta aaacctgga acaaccactg 60

ggtctcttca tcgacaatga gtgggtgaaa acaaaccgta cttttgagac cattaatccg 120

tccacaggtg aggcgatctg tcatgtttac cgtgctgggg tccaggaggt gaacgacgct 180

gtcgaagctg caaatagagc atttagaaac gaatcttggt cagggtctaac tggttctcaa 240

cgtggcgatt tactgtatcg catgtaccaa gttatcaaaa gagacgccga gagcattgca 300

tcgattgagt ccatggataa tggtaaaccg tatgctgcag aatgcctaga tggagattta 360

ggtgaagctg ctgacgtttt caaatattat gccggttggg ccgacaagat caccggtgaa 420
 ctcatgtgct cgagtgtatt gggtaagaat aagatgtgtt atgtcgagcc tacaccactg 480
 ggtgccgttg gcggtatagt cccttggaa ttcccgttta ccatgatggc atggaaaatt 540
 gccccggcac tggcgacggg ttgtacagt gttatgaagt caagtgaagt cacaccgttg 600
 acggcattat ggtatggcaa gattgcactt gaagtgggtc tacctaaagg tgtacttaac 660
 atcctctccg gttttggatc ggaatgttga tcggccatgg cttcacatcc aaagtggct 720

 aagatagcgt tcaactggctc aactgcaact ggtaaaaaaa tcatggaagc agcaggtggt 780
 tccaacttga aaaaggttac actagagtgt ggtggtaaat ctccttacat tgtttttgat 840
 gatgctgact tagaattggc agtagaatgg gcatattggg gtatttggta taacaaaggt 900
 gaggttttga ctcaacttc gagatttttg attcaggaag acattttacga taagtttgtt 960
 gagagttttg ttgagttgac caagacgaga gcaatcactg ctgatccgtt tgatgataga 1020
 tgcactatcg ggccttttgg ttctagctca cagtacgaaa aagtcaaaaa gtacgttgaa 1080
 ataggtaaaa atgaaggagc aaagctacta actggcaaat tcatcgacgg gccaggctat 1140

 ttctgtgagc catttatctt cagtgaatgc actgacgata tgacaatcat gaaagaggaa 1200
 atctttggcc ctgtttggg gattactaaa ttctcaacgg ttaaagaggc gatcgagaga 1260
 gccaatgcta cgacttacgg tttaggagct gcgttgtttt cctctaacat aacaaaggca 1320
 cattctgtgg ctgccaagtt ggaggtgga atggtgtgga tcaattctaa tgggtattct 1380
 gatatccaca ttccatttgg tggttccaaa atgagtggtg taggtaggga gttggggcca 1440
 tacgcactag acttgtttac tgagaaaaag gcagttcatg tcaacttacc gcttccggtc 1500
 aagtga 1506

<210> 122

<211> 501

<212> PRT

<213> Issatchenkia orientalis

<400> 122

Met Ser Ala Leu Phe Arg Thr Ile Glu Thr Pro Asn Gly Lys Thr Leu

1 5 10 15

Glu Gln Pro Leu Gly Leu Phe Ile Asp Asn Glu Trp Val Lys Thr Asn

20 25 30

Arg Thr Phe Glu Thr Ile Asn Pro Ser Thr Gly Glu Ala Ile Cys His

35 40 45

Val Tyr Arg Ala Gly Val Gln Glu Val Asn Asp Ala Val Glu Ala Ala

50 55 60

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

65 70 75 80

Arg Gly Asp Leu Leu Tyr Arg Met Tyr Gln Val Ile Lys Arg Asp Ala

85 90 95

Glu Ser Ile Ala Ser Ile Glu Ser Met Asp Asn Gly Lys Pro Tyr Ala

100 105 110

Ala Glu Cys Leu Asp Gly Asp Leu Gly Glu Ala Ala Asp Val Phe Lys

115 120 125

Tyr Tyr Ala Gly Trp Ala Asp Lys Ile Thr Gly Glu Leu Ile Gly Ser

130 135 140

Ser Val Leu Gly Lys Asn Lys Met Cys Tyr Val Glu Pro Thr Pro Leu

145 150 155 160

Gly Ala Val Gly Gly Ile Val Pro Trp Asn Phe Pro Phe Thr Met Met

165 170 175

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

180 185 190

Lys Ser Ser Glu Val Thr Pro Leu Thr Ala Leu Trp Tyr Gly Lys Ile

195 200 205

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

210 215 220

Phe Gly Ser Asp Val Gly Ser Ala Met Ala Ser His Pro Lys Leu Ala

225 230 235 240

Lys Ile Ala Phe Thr Gly Ser Thr Ala Thr Gly Lys Lys Ile Met Glu

245 250 255

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

260 265 270

Lys Ser Pro Tyr Ile Val Phe Asp Asp Ala Asp Leu Glu Leu Ala Val

275 280 285

Glu Trp Ala Tyr Trp Gly Ile Trp Tyr Asn Lys Gly Glu Val Cys Thr

290 295 300
 Ser Thr Ser Arg Phe Leu Ile Gln Glu Asp Ile Tyr Asp Lys Phe Val

 305 310 315 320
 Glu Ser Phe Val Glu Leu Thr Lys Thr Arg Ala Ile Thr Ala Asp Pro

 325 330 335
 Phe Asp Asp Arg Cys Thr Ile Gly Pro Leu Val Ser Ser Ser Gln Tyr

 340 345 350
 Glu Lys Val Lys Lys Tyr Val Glu Ile Gly Lys Asn Glu Gly Ala Lys

 355 360 365
 Leu Leu Thr Gly Lys Phe Ile Asp Gly Pro Gly Tyr Phe Cys Glu Pro

 370 375 380
 Phe Ile Phe Ser Glu Cys Thr Asp Asp Met Thr Ile Met Lys Glu Glu
 385 390 395 400
 Ile Phe Gly Pro Val Val Gly Ile Thr Lys Phe Ser Thr Val Lys Glu

 405 410 415
 Ala Ile Glu Arg Ala Asn Ala Thr Thr Tyr Gly Leu Gly Ala Ala Leu

 420 425 430
 Phe Ser Ser Asn Ile Thr Lys Ala His Ser Val Ala Ala Lys Leu Glu

 435 440 445
 Ala Gly Met Val Trp Ile Asn Ser Asn Gly Asp Ser Asp Ile His Ile

 450 455 460
 Pro Phe Gly Gly Ser Lys Met Ser Gly Ile Gly Arg Glu Leu Gly Pro
 465 470 475 480
 Tyr Ala Leu Asp Leu Phe Thr Glu Lys Lys Ala Val His Val Asn Leu

 485 490 495
 Ser Leu Pro Val Lys

 500

 <210
 > 123
 <211> 1506
 <212> DNA
 <213> Issatchenkia orientalis

<400> 123

| | |
|---|------|
| atggctcttc cacttgcaac tacaatctcc ttatcaagcg gcaaaacatt agaacagcca | 60 |
| attggtttat ttattgataa tgaatttgtc aatccaattt ctgtttctaa tgcaagaaca | 120 |
| ctaacaacct tcaaccecaag cacaggtagag ccaataaccg atgttcattg tgcctcagct | 180 |
| gcagatgttg atgttcgggt aaatgctgca aacaaggcaa tggaaacatg gaaagacatt | 240 |
| gacacctagc ttcgtgtcga acttttacta aaattggcca gcttagttga cgagcattcc | 300 |
| caagcaattg ctgaaattga agcactagac tcgggtaaac cattgtactc gaatgcactg | 360 |
| gcggatgttc aatcggttgc tgagtactta aggtactgtg ccggttgggc ggataaatta | 420 |
| cacggtagcg aaattccaat aaactctaag gtaatggcta ttacaaaacg tgtaccctta | 480 |
| gttgtcggct gcatcattcc atggaactac ccaatttcaa tggcctcctg gaagtctctg | 540 |
| ccagcattgg ctgccggatg tactattgta atgaagtaa gtgagataac cccgttatcg | 600 |
| ttactttatt ttgcgaattt ggtcaaatta gcaggtttcc ctaagggtgt ttttaatgtc | 660 |
| gtctctggat ttggtagatg tgttggctca gcgctttcaa atcacccaaa gttgggtaag | 720 |
| attgcattta caggctcgac cttgaccggg caaaagggtga tggcggatgc tgccagatca | 780 |
| aatttgaaaa gcgtatcttt ggaatgtggt ggtaaacttc cacttattgt cttcgaagat | 840 |
| gcagaattgg atgaatgcgt taaatgggca agttttggtg tcatgtataa caccggacaa | 900 |
| aattgtactg ccaattctcg tattattgtg catgataagg tttatgatca atttatcgaa | 960 |
| aagttcctgt ctcaactcaa ggaagattgg aaaatgggag atgtcatgaa tgaaaagact | 1020 |
| acattgggac cacttgtagc ccaacaacaa tatgagcgtg ttcagtcgta tattgatata | 1080 |
| ggtgtcaaag aaggggctac actgattcaa ccgcttaagg agagcactcc atcaaatgga | 1140 |
| ttctacatct ctctactgt ttttactaac gttaaggaag atatgagaat tgttaaggag | 1200 |
| gaaatatttg gtctgtcgt aactatctcc aaattctcaa ctgaggaaga ggcaatttca | 1260 |
| aaggcgaatg atacaattta tggcttagct gcaatgttat ttactactaa ttttgaacgt | 1320 |
| gccaacagag ttgctgataa gctggaagct ggcagtggtg acattaatag ctctaacaac | 1380 |
| gagagtacca aagttccatt tggaggaatg aagatgagtg gtattggaag agagttgggg | 1440 |
| caagaagcat ttaatttgta cactgttaca aagagtatct attatagtta tgggtgctaag | 1500 |
| ctttaa | 1506 |

<210> 124

<211> 501

<212> PRT

<213> Issatchenkia orientalis

<400> 124

```

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

Val Ala Val Asn Ala Ala Asn Lys Ala Met Glu Thr Trp Lys Asp Ile
65           70           75           80
Asp Pro Thr Val Arg Val Glu Leu Leu Leu Lys Leu Ala Ser Leu Val
           85           90           95
Asp Glu His Ser Gln Ala Ile Ala Glu Ile Glu Ala Leu Asp Ser Gly
           100          105          110
Lys Pro Leu Tyr Ser Asn Ala Leu Ala Asp Val Gln Ser Val Ala Glu
           115          120          125

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

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

```

| | | |
|---|-----|-----|
| 245 | 250 | 255 |
| Ala Ala Arg Ser Asn Leu Lys Ser Val Ser Leu Glu Cys Gly Gly Lys | | |
| 260 | 265 | 270 |
| Ser Pro Leu Ile Val Phe Glu Asp Ala Glu Leu Asp Glu Cys Val Lys | | |
| 275 | 280 | 285 |
| Trp Ala Ser Phe Gly Val Met Tyr Asn Thr Gly Gln Asn Cys Thr Ala | | |
| 290 | 295 | 300 |
| Asn Ser Arg Ile Ile Val His Asp Lys Val Tyr Asp Gln Phe Ile Glu | | |
| 305 | 310 | 315 |
| | | 320 |
| Lys Phe Leu Ser Gln Leu Lys Glu Asp Trp Lys Met Gly Asp Val Met | | |
| 325 | 330 | 335 |
| Asn Glu Lys Thr Thr Leu Gly Pro Leu Val Ser Gln Gln Gln Tyr Glu | | |
| 340 | 345 | 350 |
| Arg Val Gln Ser Tyr Ile Asp Ile Gly Val Lys Glu Gly Ala Thr Leu | | |
| 355 | 360 | 365 |
| Ile Gln Pro Leu Lys Glu Ser Thr Pro Ser Asn Gly Phe Tyr Ile Ser | | |
| 370 | 375 | 380 |
| Pro Thr Val Phe Thr Asn Val Lys Glu Asp Met Arg Ile Val Lys Glu | | |
| 385 | 390 | 395 |
| Glu Ile Phe Gly Pro Val Val Thr Ile Ser Lys Phe Ser Thr Glu Glu | | |
| 405 | 410 | 415 |
| Glu Ala Ile Ser Lys Ala Asn Asp Thr Ile Tyr Gly Leu Ala Ala Met | | |
| 420 | 425 | 430 |
| Leu Phe Thr Thr Asn Phe Glu Arg Ala Asn Arg Val Ala Asp Lys Leu | | |
| 435 | 440 | 445 |
| Glu Ala Gly Ser Val Tyr Ile Asn Ser Ser Asn Asn Glu Ser Thr Lys | | |
| 450 | 455 | 460 |
| Val Pro Phe Gly Gly Met Lys Met Ser Gly Ile Gly Arg Glu Leu Gly | | |
| 465 | 470 | 475 |
| Gln Glu Ala Phe Asn Leu Tyr Thr Val Thr Lys Ser Ile Tyr Tyr Ser | | |
| 485 | 490 | 495 |

Tyr Gly Ala Lys Leu

500

<210> 125

<211> 1544

<212> DNA

<213> Issatchenkia orientalis

<220><221> misc_feature

<222> (1319)..(1319)

<223> N= A, C, G, OR T

<400> 125

| | |
|--|------|
| atgatggcg caactacagc aaaaattagt attccaaatg gtaacaaata cgagcaacct | 60 |
| acaggtttgt tcatcaatgg tgagtttgtt gcttcaagt atggtaaaac tgcagaagtt | 120 |
| gagaatccag gcaatggaaa cattgtatgt tctgtccact tagcttctat tgaggatatt | 180 |
| aataccgccg tagaagctgc tgaagatgca tttttcaaaa ggtgggccac catcagtgg | 240 |
| aaagccaagg gagaatactt gagtaagatt gccgatctaa tcgttaaata ttctgatcaa | 300 |
| ttggcagatc tagaggctat tgaatcaggt aagccaaagg acaccaatgc aatctttgat | 360 |
| | |
| gttttacatt cggtgatgt tttcagatac tatgtctggca aggtgtcac tgcacaaagc | 420 |
| ggcaagacta tcgagtcga actctccaaa ttacataca cagtttacga gccctatggt | 480 |
| gtttgtgccg ctatcatcgc atggaacttc ccaatgagca catttgctg gaaagttgcg | 540 |
| gcattgttag ctgctggtaa tacaatggtt gtcaaaactt ccgagctgac tccgttatct | 600 |
| gcattgttca tgtgtaagat tttccaagaa gcagatctac ctgctggagt tataaacgtc | 660 |
| acatgtggtt taggttctgt tgcaggtgtt cgattgagt aacatgaaa ggttcagaaa | 720 |
| atttcgttta ctggctccac tggcgttgggt aagttgatcc aagaatccgc agcaaagtct | 780 |
| | |
| aacttaaagt attgtacgt tgaatgtggt ggtaagtctc cgtagtgat ttacaggat | 840 |
| gcagatcttg agcaagcagt gaagtgggt gcctttggta ttttttcaa caaaggtgaa | 900 |
| atttgcacag cctcttcag aatatatgtt caagaatcag tctatgaaa attttgact | 960 |
| atgtacaagg atcatgtgga agaagcctat gtccaaggag aacagtttgc cactggtgtt | 1020 |
| aacgttgggc ctactgtctg caaagcccaa caagagaaaa tactggccta cattgaaagt | 1080 |
| gccaagcaag aaggtggtag aattatcact ggtggtaaaa taccatctta cacgaacaaa | 1140 |
| aatggttact atctcgaacc aacaattatt gcagattgta accaggatat gaagtagtc | 1200 |
| | |
| agggaagaga ttttcggacc agtcgttact gtatccaaat tcactagtga tgaagaagcc | 1260 |

atcaaattaa gcaatgattc cgaatatggc ttggcagcat atttattcac tystsrasna 1320
 ssrgttyrgy taaaatyrrh thraaggacc tcgttagatc tcagaattat atcagaaaag 1380
 tgcaaagcgg acaggctctt gtcaacttca cctttgcggc tgatttcagg ttgccatttg 1440
 gcggatataa gatgagtggg aacggaagag agcttgggtga tgaaggactg agtgctttcc 1500
 agcaagtcaa agcagtiacac attaattcca ctgggaagtt gtaa 1544

<210> 126

<211> 503

<212> PRT

<213> Issatchenkia orientalis

<400> 126

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

Glu Ala Ala Glu Asp Ala Phe Phe Lys Arg Trp Ala Thr Ile Ser Gly
 65 70 75 80
 Lys Ala Lys Gly Glu Tyr Leu Ser Lys Ile Ala Asp Leu Ile Val Lys
 85 90 95
 Tyr Ser Asp Gln Leu Ala Asp Leu Glu Ala Ile Glu Ser Gly Lys Pro
 100 105 110
 Lys Asp Thr Asn Ala Ile Phe Asp Val Leu His Ser Ala Asp Val Phe
 115 120 125

Arg Tyr Tyr Ala Gly Lys Ala Val Thr Ala Gln Ser Gly Lys Thr Ile
 130 135 140
 Glu Ser Glu Leu Ser Lys Phe Thr Tyr Thr Val Tyr Glu Pro Tyr Gly
 145 150 155 160
 Val Cys Ala Ala Ile Ile Ala Trp Asn Phe Pro Met Ser Thr Phe Ala
 165 170 175

Trp Lys Val Ala Ala Cys Leu Ala Ala Gly Asn Thr Met Val Val Lys
180 185 190

Thr Ser Glu Leu Thr Pro Leu Ser Ala Leu Phe Met Cys Lys Ile Phe
195 200 205

Gln Glu Ala Asp Leu Pro Ala Gly Val Ile Asn Val Thr Cys Gly Leu
210 215 220

Gly Ser Val Ala Gly Val Arg Leu Ser Glu His Glu Lys Val Gln Lys
225 230 235 240

Ile Ser Phe Thr Gly Ser Thr Gly Val Gly Lys Leu Ile Gln Glu Ser
245 250 255

Ala Ala Lys Ser Asn Leu Lys Tyr Cys Thr Leu Glu Cys Gly Gly Lys
260 265 270

Ser Pro Leu Val Ile Tyr Glu Asp Ala Asp Leu Glu Gln Ala Val Lys
275 280 285

Trp Ala Ala Phe Gly Ile Phe Phe Asn Lys Gly Glu Ile Cys Thr Ala
290 295 300

Ser Ser Arg Ile Tyr Val Gln Glu Ser Val Tyr Asp Lys Phe Leu Thr
305 310 315 320

Met Tyr Lys Asp His Val Glu Glu Ala Tyr Val Gln Gly Glu Gln Phe
325 330 335

Ala Thr Gly Val Asn Val Gly Pro Thr Val Cys Lys Ala Gln Gln Glu
340 345 350

Lys Ile Leu Ala Tyr Ile Glu Ser Ala Lys Gln Glu Gly Gly Arg Ile
355 360 365

Ile Thr Gly Gly Lys Ile Pro Ser Tyr Thr Asn Lys Asn Gly Tyr Tyr
370 375 380

Leu Glu Pro Thr Ile Ile Ala Asp Cys Asn Gln Asp Met Lys Val Val
385 390 395 400

Arg Glu Glu Ile Phe Gly Pro Val Val Thr Val Ser Lys Phe Thr Ser
405 410 415

Asp Glu Glu Ala Ile Lys Leu Ser Asn Asp Ser Glu Tyr Gly Leu Ala

420 425 430
Ala Tyr Leu Phe Thr Lys Asp Leu Val Arg Ser Gln Asn Tyr Ile Arg
435 440 445
Lys Val Gln Ser Gly Gln Val Phe Val Asn Phe Thr Phe Ala Ala Asp
450 455 460
Phe Arg Leu Pro Phe Gly Gly Tyr Lys Met Ser Gly Asn Gly Arg Glu
465 470 475 480
Leu Gly Asp Glu Gly Leu Ser Ala Phe Gln Gln Val Lys Ala Val His
485 490 495
Ile Asn Leu Thr Gly Lys Leu

500

<210> 127

<211> 1716

<212> DNA

<213> Issatchenkia orientalis

<220><221> misc_feature

<222> (655)..(655)

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

<400> 127

| | |
|---|-----|
| atggctccaa ctgctgttga tatccataac gagtacaaac agaatgtttc caacgaacag | 60 |
| gaaattcctt tcaacaaaac tgaagaaag tcatcgattg catctaaatt aggactgaat | 120 |
| ccagacgcta agattcacta caattctgct gttcctatat tatacgaaga tggtttaaag | 180 |
| gaaaaaggta caaccatttc ctcttctggt gcattgattg cattctctgg ttccaaaaca | 240 |
| ggtagatctc caaaggacaa aagaattgtc gatgaagaga ctcaacaga caacatctgg | 300 |
| tggggtccag tcaataagaa ggttgatgaa aacacttgga atatctcgaa atctagagcg | 360 |
| | |
| attgattatt tgagaacaag agagaaggtt tacattatcg atgcttttgc tggttgggat | 420 |
| ccaagataca gaattaaggt tagaattgtc tgtgctagag cttaccatgc tttgttcag | 480 |
| aagaatatgt taattagacc aacaacggaa gaattaaaga actttggtga gcctgatttc | 540 |
| accatttgga atgcaggtea attccctgct aatgittaca ctaagggtat gacttcttca | 600 |
| acttctgttg aaataaatth caagtctthr ysgymthrs rsrthrsrva gtasnhysr | 660 |
| atggaaatgg ttatcctagg tactgaatac gcaggtgaaa tgaagaaagg tatctttacc | 720 |

gttatgttct acttgatgcc aatcagacac aaggtttta ctttacactc ttctgcaa 780

caaggtaaaa aggatggatga tgtcacatta ttctttgggt tatctgggtac aggtaaaaca 840

accttgtctg cagatcctca tagagaattg attgggatg atgaacattg ctggctctgat 900

catgggtgtt tcaacattga aggtggatgt tatgctaagt gtttggactt atctgctgaa 960

agagaacctg agattttcaa tgcaattagg ttggatctg tcttggagaa tgtgtctat 1020

gatccagtgt atagaactgt tgactattcc gctgctaagt tcaactgaaa tactagatgt 1080

gcttatecta tcgactttat tcttctgct aagatcccat gtctggcaga ttctcatcca 1140

aagaatatgt ttcttttaac ttgtgatgca agagggtgtt tgccacctgt ctccaagcta 1200

actaatgcac aagtcatgta tcactttatc tctggttaca cctccaagat ggcaggtacc 1260

gaagtgggtg tcaactgaacc agaagcaacc ttctctgcat gttttgggtca acctttctta 1320

gttttacatc caatgaata cgcacaacaa ctctctgata aaatggctga acattcttcc 1380

accgcttgggt tattgaatac cggttggact ggtcaatctt atgttaaagg tggttaagaga 1440

tgtccattga agtatactag agcaatttta gatgctattc actctgggtga gcttgcaaaa 1500

caggaattcg aaacataccc tactttcgggt ttacaagttc caaaaacttg tccaggtgtc 1560

ccagaaagtg ttctgaaccc atctaaacac tgggctactg gtgaagctga tttcaaggt 1620

gaagtcacta acttggctaa attatttctg gagaactttg aaaagtattc tgcagaatgt 1680

actgcagaag ttgttctgctc tggctctgct ttataa 1716

<210> 128

<211> 560

<212> PRT

<213> Issatchenkia orientalis

<400> 128

Met Ala Pro Thr Ala Val Asp Ile His Asn Glu Tyr Lys Gln Asn Val

1 5 10 15

Ser Asn Glu Gln Glu Ile Pro Phe Asn Lys Thr Glu Arg Lys Ser Ser

20 25 30

Ile Ala Ser Lys Leu Gly Leu Asn Pro Asp Ala Lys Ile His Tyr Asn

35 40 45

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

50 55 60

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

| | | | | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|--|-----|-----|
| 65 | | | | | | 70 | | | | | | 75 | | | | | | 80 |
| Gly | Arg | Ser | Pro | Lys | Asp | Lys | Arg | Ile | Val | Asp | Glu | Glu | Thr | Ser | Thr | | | |
| | | | | | 85 | | | | | | 90 | | | | | | 95 | |
| Asp | Asn | Ile | Trp | Trp | Gly | Pro | Val | Asn | Lys | Lys | Val | Asp | Glu | Asn | Thr | | | |
| | | | | | 100 | | | | | | 105 | | | | | | 110 | |
| Trp | Asn | Ile | Ser | Lys | Ser | Arg | Ala | Ile | Asp | Tyr | Leu | Arg | Thr | Arg | Glu | | | |
| | | | | | 115 | | | | | | 120 | | | | | | 125 | |
| Lys | Val | Tyr | Ile | Ile | Asp | Ala | Phe | Ala | Gly | Trp | Asp | Pro | Arg | Tyr | Arg | | | |
| | | | | | 130 | | | | | | 135 | | | | | | 140 | |
| Ile | Lys | Val | Arg | Ile | Val | Cys | Ala | Arg | Ala | Tyr | His | Ala | Leu | Phe | Met | | | |
| | | | | | 145 | | | | | | 150 | | | | | | 155 | 160 |
| Lys | Asn | Met | Leu | Ile | Arg | Pro | Thr | Thr | Glu | Glu | Leu | Lys | Asn | Phe | Gly | | | |
| | | | | | 165 | | | | | | 170 | | | | | | 175 | |
| Glu | Pro | Asp | Phe | Thr | Ile | Trp | Asn | Ala | Gly | Gln | Phe | Pro | Ala | Asn | Val | | | |
| | | | | | 180 | | | | | | 185 | | | | | | 190 | |
| Tyr | Thr | Lys | Gly | Met | Thr | Ser | Ser | Thr | Ser | Val | Glu | Ile | Asn | Phe | Lys | | | |
| | | | | | 195 | | | | | | 200 | | | | | | 205 | |
| Ser | Met | Glu | Met | Val | Ile | Leu | Gly | Thr | Glu | Tyr | Ala | Gly | Glu | Met | Lys | | | |
| | | | | | 210 | | | | | | 215 | | | | | | 220 | |
| Lys | Gly | Ile | Phe | Thr | Val | Met | Phe | Tyr | Leu | Met | Pro | Ile | Arg | His | Lys | | | |
| | | | | | 225 | | | | | | 230 | | | | | | 235 | 240 |
| Val | Leu | Thr | Leu | His | Ser | Ser | Ala | Asn | Gln | Gly | Lys | Lys | Asp | Gly | Asp | | | |
| | | | | | 245 | | | | | | 250 | | | | | | 255 | |
| Val | Thr | Leu | Phe | Phe | Gly | Leu | Ser | Gly | Thr | Gly | Lys | Thr | Thr | Leu | Ser | | | |
| | | | | | 260 | | | | | | 265 | | | | | | 270 | |
| Ala | Asp | Pro | His | Arg | Glu | Leu | Ile | Gly | Asp | Asp | Glu | His | Cys | Trp | Ser | | | |
| | | | | | 275 | | | | | | 280 | | | | | | 285 | |
| Asp | His | Gly | Val | Phe | Asn | Ile | Glu | Gly | Gly | Cys | Tyr | Ala | Lys | Cys | Leu | | | |
| | | | | | 290 | | | | | | 295 | | | | | | 300 | |
| Asp | Leu | Ser | Ala | Glu | Arg | Glu | Pro | Glu | Ile | Phe | Asn | Ala | Ile | Arg | Phe | | | |
| | | | | | 305 | | | | | | 310 | | | | | | 315 | 320 |

Gly Ser Val Leu Glu Asn Val Val Tyr Asp Pro Val Asp Arg Thr Val
325 330 335
Asp Tyr Ser Ala Ala Asn Val Thr Glu Asn Thr Arg Cys Ala Tyr Pro
340 345 350
Ile Asp Phe Ile Pro Ser Ala Lys Ile Pro Cys Leu Ala Asp Ser His
355 360 365
Pro Lys Asn Ile Val Leu Leu Thr Cys Asp Ala Arg Gly Val Leu Pro
370 375 380
Pro Val Ser Lys Leu Thr Asn Ala Gln Val Met Tyr His Phe Ile Ser
385 390 395 400
Gly Tyr Thr Ser Lys Met Ala Gly Thr Glu Val Gly Val Thr Glu Pro
405 410 415
Glu Ala Thr Phe Ser Ala Cys Phe Gly Gln Pro Phe Leu Val Leu His
420 425 430
Pro Met Lys Tyr Ala Gln Gln Leu Ser Asp Lys Met Ala Glu His Ser
435 440 445
Ser Thr Ala Trp Leu Leu Asn Thr Gly Trp Thr Gly Gln Ser Tyr Val
450 455 460
Lys Gly Gly Lys Arg Cys Pro Leu Lys Tyr Thr Arg Ala Ile Leu Asp
465 470 475 480
Ala Ile His Ser Gly Glu Leu Ala Lys Gln Glu Phe Glu Thr Tyr Pro
485 490 495
Thr Phe Gly Leu Gln Val Pro Lys Thr Cys Pro Gly Val Pro Glu Ser
500 505 510
Val Leu Asn Pro Ser Lys His Trp Ala Thr Gly Glu Ala Asp Phe Lys
515 520 525
Ala Glu Val Thr Asn Leu Ala Lys Leu Phe Ala Glu Asn Phe Glu Lys
530 535 540
Tyr Ser Ala Glu Cys Thr Ala Glu Val Val Ala Ala Gly Pro Ala Leu
545 550 555 560

<210> 129

<211> 267

<212> PRT

<213> *Saccharomyces cerevisiae*

<400> 129

Met Ser Gln Gly Arg Lys Ala Ala Glu Arg Leu Ala Lys Lys Thr Val

1 5 10 15

Leu Ile Thr Gly Ala Ser Ala Gly Ile Gly Lys Ala Thr Ala Leu Glu

20 25 30

Tyr Leu Glu Ala Ser Asn Gly Asp Met Lys Leu Ile Leu Ala Ala Arg

35 40 45

Arg Leu Glu Lys Leu Glu Glu Leu Lys Lys Thr Ile Asp Gln Glu Phe

50 55 60

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

65 70 75 80

Lys Ile Lys Pro Phe Ile Glu Asn Leu Pro Gln Glu Phe Lys Asp Ile

85 90 95

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

100 105 110

Gly Gln Ile Ala Thr Glu Asp Ile Gln Asp Val Phe Asp Thr Asn Val

115 120 125

Thr Ala Leu Ile Asn Ile Thr Gln Ala Val Leu Pro Ile Phe Gln Ala

130 135 140

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

145 150 155 160

Ala Tyr Pro Thr Gly Ser Ile Tyr Cys Ala Ser Lys Phe Ala Val Gly

165 170 175

Ala Phe Thr Asp Ser Leu Arg Lys Glu Leu Ile Asn Thr Lys Ile Arg

180 185 190

Val Ile Leu Ile Ala Pro Gly Leu Val Glu Thr Glu Phe Ser Leu Val

195 200 205

Arg Tyr Arg Gly Asn Glu Glu Gln Ala Lys Asn Val Tyr Lys Asp Thr

210 215 220

Thr Pro Leu Met Ala Asp Asp Val Ala Asp Leu Ile Val Tyr Ala Thr
225 230 235 240

Ser Arg Lys Gln Asn Thr Val Ile Ala Asp Thr Leu Ile Phe Pro Thr
245 250 255

Asn Gln Ala Ser Pro His His Ile Phe Arg Gly
260 265

<210> 130

<211> 420

<212> DNA

<213> Streptomyces avermitilis

<400> 130

atgttaagaa ccatgttcaa atctaagatt cacagagcaa ctgttactca agcagatctc 60
cattatgttg gticcgttac tattgatgca gacttgtag acgcagcaga cttgttgcca 120
ggtgaattgg ttcacatcgt tgacattacg aacggtgcta gattggaaac ttacgtcatt 180

gaaggatgaac gtggttccgg tgttgttggg atcaatgggtg ctgccgtca tttagtccat 240
cctggatgac ttgttatcat tatctcctat gcacaagttt cagatgcaga agcacgtgca 300
ttgcgtccaa gaggttgtca cgttgacaga gacaatagag ttgttgcgt tgggtcggat 360
ccagccgaac cagtccagg ttccgaccaa gctagatccc cacaagctgt tactgcataa 420

<210> 131

<211> 384

<212> DNA

<213> Clostridium acetobutylicum

<400> 131

atgcacttga acatgttgaa gtccaagatc cacagagcta ccgtcgttca agcagacttg 60
aactacgtcg gttccatcac catcgacaga aacttgatgg acaaggcaaa catcttgaa 120

tacgaaaagg tcgagatcgc aaacatcaac aacggtgcaa gattcgaaac ctacgtcatc 180
gctggtgagg ctggttccgg tatcatctgt ttgaacgggtg ctgctgcaag atgtgcacaa 240
gctgggtgaca aggttatcat catgtgttac tgttccttga cccagagaaga agcttccgag 300
cacagaccaa aggtcgtttt cgtaacgac gacaactcca tctccaacgt caccgaatac 360
gagaagcacg gcaccatcgg ttaa 384

<210> 132

<211> 354

<212> DNA

<213> Helicobacter pylori

<400> 132

atgaccttcg agatgttgta ctccaagatc cacagagcaa ccatcaccga cgcaaacttg 60

aactacatcg gctccatcac catcgacgag gacttggcta agttggctaa gttgagagag 120

ggtatgaagg tcgaaatcgt cgacgtcaac aacggcgaga gattctccac ctacgtcatc 180

ttgggtaaga agagaggtga aatctgcgtc aacgggtgcag cagccagaaa ggtcgctatc 240

ggtgacgtcg tcatcatctt ggcttacgca tccatgaacg aggacgagat caacgctcac 300

aagccatcca tcgtcttggg cgacgaaaag aacgaaatct tggaaaaggg ttaa 354

<210> 133

<211> 117

<212> PRT

<213> Helicobacter pylori

<400> 133

Met Thr Phe Glu Met Leu Tyr Ser Lys Ile His Arg Ala Thr Ile Thr

| | | | |
|-----|-----|-----|-----|
| 1 | 5 | 10 | 15 |
| Asp | Ala | Asn | Leu |
| Asn | Tyr | Ile | Gly |
| Ser | Ile | Thr | Ile |
| Asp | Glu | Asp | Leu |
| 20 | 25 | 30 | |
| Ala | Lys | Leu | Ala |
| Lys | Leu | Arg | Glu |
| Gly | Met | Lys | Val |
| Glu | Ile | Val | Asp |
| 35 | 40 | 45 | |
| Val | Asn | Asn | Gly |
| Glu | Arg | Phe | Ser |
| Thr | Tyr | Val | Ile |
| Leu | Gly | Lys | Lys |
| 50 | 55 | 60 | |
| Arg | Gly | Glu | Ile |
| Cys | Val | Asn | Gly |
| Ala | Ala | Ala | Arg |
| Lys | Val | Ala | Ile |

| | | | |
|-----|-----|-----|-----|
| 65 | 70 | 75 | 80 |
| Gly | Asp | Val | Val |
| Ile | Ile | Leu | Ala |
| Tyr | Ala | Ser | Met |
| Asn | Glu | Asp | Glu |
| 85 | 90 | 95 | |
| Ile | Asn | Ala | His |
| Lys | Pro | Ser | Ile |
| Val | Leu | Val | Asp |
| Glu | Lys | Asn | Glu |
| 100 | 105 | 110 | |
| Ile | Leu | Glu | Lys |
| Gly | | | |
| 115 | | | |

<210> 134

<211> 387

<212> DNA

<213> Bacillus sp. TS25

<400> 134

atgtacagaa ccatgatgaa gtccaagttg cacagagcga ccgtcaccga agcaaacttg 60

aactacgtcg gttccatcac catcgaccaa gacttgatgg aagctgcaga catcttggaa 120

aacgagaagg tccaaatcgt caacaacaac aacgggtgcta gattcgaaac ctacgtcatc 180

gctgggtccaa gaggtccgg caccatctgt ttgaacgggtg cagcagcaag attggtccaa 240

ccaggtgaca ccgttatcat catctcctac gcaatgttgg aagaagccga ggctagaaag 300

caccaacctg tcgtcgtttt gttgaaccca gacaacacca tccaagaatt gatcagagaa 360

acccacgggtg ctaccgctac cgtctaa 387

<210> 135

<211> 128

<212> PRT

<213> Bacillus sp. TS25

<400> 135

Met Tyr Arg Thr Met Met Lys Ser Lys Leu His Arg Ala Thr Val Thr

1 5 10 15

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

20 25 30

Met Glu Ala Ala Asp Ile Leu Glu Asn Glu Lys Val Gln Ile Val Asn

35 40 45

Asn Asn Asn Gly Ala Arg Phe Glu Thr Tyr Val Ile Ala Gly Pro Arg

50 55 60

Gly Ser Gly Thr Ile Cys Leu Asn Gly Ala Ala Ala Arg Leu Val Gln

65 70 75 80

Pro Gly Asp Thr Val Ile Ile Ile Ser Tyr Ala Met Leu Glu Glu Ala

85 90 95

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

100 105 110

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

115 120 125

<210> 136

<211> 411

<212> DNA

<213> *Corynebacterium glutamicum*

<400> 136

```
atgttgagaa ccattctggg ttccaaaatc cacagagcta ccgttaccca agcagacttg      60
gactacgttg gttcagtcac catcgatgca gacttgggtc atgccgcagg tttgatcgaa      120
gggtgaaaagg ttgccatcgt cgacatcacc aatggcgcta gattggaaac ctacgttatc      180
gttgggtgatg ctggcaccgg taacatctgt atcaacggtg cagcagcaca cttgatcaac      240
ccaggtgatt tggttatcat catgtcctac ttgcaagcta ccgatgcaga ggccaaggct      300
tacgaaccaa agatcgttca cgtcgacgca gacaacagaa tcgttgcttt gggtaacgac      360
```

```
ttggcagaag ctttgcttgg ttccggtttg ttgacctcaa gatctatcta a              411
```

<210> 137

<211> 136

<212> PRT

<213> *Corynebacterium glutamicum*

<400> 137

Met Leu Arg Thr Ile Leu Gly Ser Lys Ile His Arg Ala Thr Val Thr

1 5 10 15

Gln Ala Asp Leu Asp Tyr Val Gly Ser Val Thr Ile Asp Ala Asp Leu

20 25 30

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

35 40 45

Ile Thr Asn Gly Ala Arg Leu Glu Thr Tyr Val Ile Val Gly Asp Ala

50 55 60

Gly Thr Gly Asn Ile Cys Ile Asn Gly Ala Ala Ala His Leu Ile Asn

65 70 75 80

Pro Gly Asp Leu Val Ile Ile Met Ser Tyr Leu Gln Ala Thr Asp Ala

85 90 95

Glu Ala Lys Ala Tyr Glu Pro Lys Ile Val His Val Asp Ala Asp Asn

100 105 110

Arg Ile Val Ala Leu Gly Asn Asp Leu Ala Glu Ala Leu Pro Gly Ser

| | | | |
|---|------------------------|-----|----|
| 115 | 120 | 125 | |
| Gly Leu Leu Thr Ser Arg Ser Ile | | | |
| 130 | 135 | | |
| <210> | 138 | | |
| <211> | 384 | | |
| <212> | DNA | | |
| <213> | Bacillus licheniformis | | |
| <400> | 138 | | |
| atgtacagaa ccttgatgtc cgctaagttg cacagagcta gagtcaccga agcaaacttg | 60 | | |
| aactacgttg gtccgtcac catcgacgag gacttggttg acgcagtcgg tatgatggca | 120 | | |
| aacgaaaagg tccaaatcgt caacaacaac aacggtgcta gattggaaac ctacatcatc | 180 | | |
| | | | |
| cctggtgaga gaggttccgg tgcgtctgt ttgaacggtg ctgcagctag attggtccaa | 240 | | |
| gttggtgacg tcgttatcat cgtttcctac gccatgatgt ccgaagaaga agcaaagacc | 300 | | |
| cacaagccaa aggtcgctgt ttgaacgaa agaaacgaaa tcgaagagat gttgggtcaa | 360 | | |
| gaaccagcta gaaccatctt gtaa | 384 | | |
| <210> | 139 | | |
| <211> | 127 | | |
| <212> | PRT | | |
| <213> | Bacillus licheniformis | | |
| <400> | 139 | | |
| Met Tyr Arg Thr Leu Met Ser Ala Lys Leu His Arg Ala Arg Val Thr | | | |
| 1 | 5 | 10 | 15 |
| | | | |
| Glu Ala Asn Leu Asn Tyr Val Gly Ser Val Thr Ile Asp Glu Asp Leu | | | |
| 20 | 25 | 30 | |
| Leu Asp Ala Val Gly Met Met Ala Asn Glu Lys Val Gln Ile Val Asn | | | |
| 35 | 40 | 45 | |
| Asn Asn Asn Gly Ala Arg Leu Glu Thr Tyr Ile Ile Pro Gly Glu Arg | | | |
| 50 | 55 | 60 | |
| Gly Ser Gly Val Val Cys Leu Asn Gly Ala Ala Ala Arg Leu Val Gln | | | |
| 65 | 70 | 75 | 80 |
| | | | |
| Val Gly Asp Val Val Ile Ile Val Ser Tyr Ala Met Met Ser Glu Glu | | | |
| 85 | 90 | 95 | |

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

100

105

110

Glu Ile Glu Glu Met Leu Gly Gln Glu Pro Ala Arg Thr Ile Leu

115

120

125

<210> 140

<211> 1356

<212> DNA

<213> Streptomyces avermitilis

<400> 140

atgacaccac agccaaaccc acaagtcggt gcagcagtc aagctgcaga tagagcacac 60

gtcttcact cttggctgc acaagaattg atcgatccat tggctgttc tggcgcagag 120

ggttcctact tctgggatta cgatggtaga cgttaccttg acttcacctc cggttagtc 180

ttaccaaca tcggttacca acacccaaag gttgtcgcag ctatccaaga acaagccgca 240

tctttgacta cttttgcccc agctttcgca gttgaagcaa gatccgaagc tgcaagattg 300

atcgctgagc gtactccagg tgatttagac aaaatcttct tcaccaacgg tggcgcagat 360

gctatcgagc acgctgttcg tatggcaaga atccacgtg gtagacaaa ggtcttatcc 420

gcatacagat cataccagg tggtagacaa caggcagtc acatcactgg tgatccaagg 480

agatgggcat ccgattccgc ttctgtggc gttgtccact tctgggctcc atacttatac 540

agatccagat tctacccga aactgagcaa caagaatgtg agcgtgctct tgagcacttg 600

gaaactacta tcgccttcga aggtccagg actattgccg ctatcgtttt ggaaactgtc 660

ccaggctactg ctggatcat ggttcctcca ccaggttact tagcagggtg tagagaattg 720

tgtgacaaat acggtatcgt cttcgtcttg gatgaagtca tggctggttt cggcagaact 780

ggcgaatggt tcgctgcaga cttattcgat gttaccccag acttgatgac cttcgtctag 840

ggtgtcaact caggttacgt tccattgggt ggtgttgcta tctccggcaa aatcgagag 900

actttcggta agagagctta cccagggtgt ttgacgtact ccggtcaccc tcttgcttgc 960

gcagccgctg ttgtactat caacgttatg gcagaagaag gtgtcgttga aaacgctgca 1020

aacttgggtg cttagattat cgaaccagg ttgagagaac ttgcagagag acaccatca 1080

gttgggtgaag ttagagggtg tggatgttc tgggctttgg aattgggttaa ggatagagaa 1140

accagagaac ctttgggtccc atacaatgcc gctgggtgaag cgaacgcacc aatggctgct 1200

ttcgggtgcag ctgcaaaggc aaacggtttg tggccattca tcaacatgaa cagaaccac 1260

gttgttctc cttgtaacgt taccgaagcc gaagctaaag aaggcttggc agcattggat 1320

gcagctttat ctgttcaga ttagtacact gtctaa 1356

<210> 141

<211> 1416

<212> DNA

<213> *Saccharomyces cerevisiae*

<400> 141

atgtccatct gtgagcaata ctaccagaa gaaccaacca agccaaccgt caagaccgaa 60

tccatccctg gtccagaatc ccaaaagcaa ttgaaggaat tgggtgaagt tttcgacacc 120

agaccagctt acttcttggc agactacgaa aagtccttgg gtaactacat caccgacgtc 180

gacggtaaca cctacttggc cttgtacgt ccaatctcct ccatcgctt gggttacaac 240

aaccagcat tgatcaaagc cgctcaatcc ccagaaatga tcagagcatt ggttgacaga 300

ccagccttgg gtaacttccc ttccaaggac ttggacaaga tcttgaagca aatcttgaag 360

tccgtccaa agggtcagga ccacgtctgg tccggtttgt ccggtgcaga cgaaacgaa 420

ttggctttca aggtgcatt catctactac agagctaagc aaagaggta cgacgcagac 480

tttccgaaa aggaaaactt gtccgttatg gacaacgacg caccaggtgc tccacacttg 540

gcagttttgt cttcaagag agctttccac ggtagattgt tcgcatccgg ttccactacc 600

tgttccaagc caatccaaa gttggacttc cctgctttcc actggccaca cgccaatac 660

ccatcctacc agtacccttt ggacgaaaac tccgacgcaa acagaaagga ggacgaccac 720

tgcttggcaa tcgttgaaga attgatcaag acctgggtcca tcctgttgc tgcgttgatc 780

atcgaaccaa tccaatccga ggggtgtgac aaccacgcct ccaagtactt cttgcaaaag 840

ttgagagaca tcaccttgaa gtacaacgtc gtctacatca tcgacgaagt ccaaactggc 900

gttggcgcaa ccggtgaagt gtggtgtcac gaatacgagc acatccagcc tccagtcgac 960

ttggtcacct tctccaagaa gttccaatcc gctggtact tcttcacga ccctaagttc 1020

atcccaaaca agccatacag acaattcaac acctggtgtg gtgaaccagc aagaatgatc 1080

atcgcagggt caatcgtca agagatctcc gacaagaagt tgaccgaaca atgctccaga 1140

gtcggtgact acttgttcaa gaagttggaa ggtttgcaaa agaagtaccc agagaacttc 1200

caaaacttga gaggttaagg tagaggcacc ttcacgtt gggacttgcc aactggcgag 1260

aagagagact tgttgttgaa gaagttgaag ttgaacggct gtaacgttgg tggctgtgct 1320

gttcacgcag tcagattgag accttccttg accttcgaag agaagcacgc agacatcttc 1380

atcgaagctt tggctaagtc cgtcaacgaa ttgtaa 1416

<210> 142

<211> 1428

<212> DNA

<213> *Saccharomyces kluyveri*

<400> 142

atgccatcct actccgttgc agaattgtac taccagacg aacctaccga acctaagatc 60

tccacctcct cctaccagg tccaaaggca aagcaagaat tggaaaagt gtccaacgtc 120

ttcgacacca gagcagctta cttgttggca gactactaca agtcccgtgg taactacatc 180

gttgaccagg acggtaacgt cttgttggac gtttacgtc aaatctctc catcgcttg 240

ggttacaaca acccagaaat ctgaaggtt gcaaagtccg acgcaatgtc cgttgcatg 300

gccaaccgtc cagcattggc ttgtttccca tccaacgact acggtcaatt gttggaagac 360

ggtttgttga aggcagcacc acaaggtcaa gacaagatct ggaccgcttt gtccggttcc 420

gacgcaaacg aaaccgcctt caaagcctgc ttcatgtacc aagctgcgaa gaagagaaac 480

ggtagatcct tctccaccga agaattggaa tccgttatgg acaaccaatt gccaggcacc 540

tccgaaatgg ttatctgttc cttcgaaaag ggtttccacg gtcgtttgtt cggttccttg 600

tccactacca gatccaagcc tatccacaag ttggacatcc cagctttcaa ctggcctaaa 660

gccccattcc cagacttgaa gtaccattg gaagaaaaca aggaagccaa caaggctgaa 720

gaatcctcct gtatcgaaaa gttctccaa atcgttcaag agtggcaagg taagatcgct 780

gcagttatca tcgaaccaat ccagtcagag ggtggcgaca accacgttc ctccgacttc 840

ttccaaaagt tgagagaaat caccatcgaa aacggtatct tgatgatcgt cgacgaagtt 900

caaaccggig tcggtgctac cggcaagatg tgggcacacg aacctggaa cttgtccaac 960

cctccagact tggttacctt ctccaagaag ttccaagcag caggtttcta ctaccacgac 1020

ccaaagtigc aaccagacca gccattcaga cagtccaaca cctgggtgtg tgacccatcc 1080

aaggctttga tcgccaaggt tatctacgaa gaaatcgta agcacgactt ggtcaccaga 1140

actgccgaag tcggtaacta cttgttcaac agattggaaa agttgttcga aggttaagaac 1200

tacatccaga acttgagagg taagggtcaa ggcacctaca tcgctttcga cttcggcacc 1260

tcctccgaga gagactcctt cttgtccaga ttgagatgta acggtgcaaa cgctcgtggt 1320

tgcggtgact ccgctgtcag attgagacca tccttgacct tcgaagagaa gcacgcagac 1380

gtcttggttt ccatcttcga caagaccttg agacaattgt acggctaa 1428

<210> 143

<211> 747

<212> DNA

<213> *Escherichia coli*

<400> 143

atgatcgitt tgggtaccgg tgcaaccgca ggtttcggcg aatgtatcac cagaagattc 60
atccagcagg gtcacaaggt tatcgctacc ggtagaagac aagagagatt gcaagaattg 120
aaggacgagt tgggtgacaa cttgtacatc gctcaattgg acgttagaaa cagagcagct 180
atcgaagaaa tgttggcatc cttgccagct gaatggtgca acatcgacat cttggtcaac 240

aacgttggtt tggcattggg tatggaacca gtcacaagg ctagtgttga ggactgggag 300
accatgatcg acaccaacaa caagggtttg gtctacatga ccagagcagt tttgcctggt 360
atggttgaaa gaaaccacgg tcacatcatc aacatcggtt ccaccgtgg ttcctggcca 420
tacgttggcg gtaacgtcta cgggtgtacc aaggctttcg ttagacagtt ctccttgaac 480
ttgagaaccg acttgcacgg caccgtgtt agagttaccg acatcgaacc aggtttggtt 540
ggtggcaccg aattctccaa cgtcagattc aaggcgacg acggttaaggc tgaaaagacc 600
tacaaaaaca ccgtcgcttt gaccccagaa gacgtttcag aggtctgttg gtgggtcagt 660

accttgccag cacacgtcaa catcaacacc ttggaaatga tgccagtcac ccaatcctac 720
gcaggtttga acgttcacag acaataa 747

<210> 144

<211> 804

<212> DNA

<213> *Saccharomyces cerevisiae*

<400> 144

atgtcccaag gtagaaaggc agcagaaaga ttggcaaaga agaccgtctt gatcaccggt 60
gcgtccgctg gtatcggtaa ggctaccgag ttggagtact tggaagcatc caacggtgac 120
atgaagttag tcttggcagc aagaagattg gagaagttgg aagaattgaa gaagaccatc 180
gaccaagaat tcccaaacgc taaggtccac gttgcacaat tggacatcac ccaagcagag 240

aagatcaagc cattcatcga aaacttgcca caagaattca aggacatcga catcttggtc 300
aacaacgctg gtaaggcgtt gggttccgac agagttggtc aaatcgcaac cgaagacatc 360
caagcgtct tcgacaccaa cgtcaccgct ttgatcaaca tcaccaagc tgttttgcca 420
atcttccaag cgaagaactc cggtagacatc gtcaacttgg gttccatcgc tggtagagac 480
gcatacccaa ccggctccat ctactgcgcc tccaagttcg ctgtcgggtgc tttcaccgac 540
tccttgagaa aggaattgat caacaccaag atcagagtca tcttgattgc ccctgggttg 600
gtcgaaccg aattctcctt ggtagatagc agaggtaacg aagaacaagc aaagaacgtt 660

tacaaggaca ctacccatt gatggccgac gacgttgcag acttgatcgt ttacgtacc 720

| | |
|--|-----|
| tccagaaagc aaaacaccgt tatcgagac accttgatct tcccaacca ccaagcatcc | 780 |
| ccacaccaca tcttcagagg ttaa | 804 |
| <210> 145 | |
| <211> 420 | |
| <212> DNA | |
| <213> Streptomyces avermitilis | |
| <400> 145 | |
| atgcttagaa ccatgttcaa atccaagatc cacagagcaa ccgtcactca agcagatttg | 60 |
| cattacgttg gttctgttac tattgacgca gacttacttg atgcagccga tttacttcct | 120 |
| ggtgagcttg ttcataatgt tgatatcacc aatgggtgcgc gtcttgaaac ctatgtcatt | 180 |
| gagggtagac gtgggtccgg tgtcgtcggc atcaatggcg cagctgcaca cctcgtccat | 240 |
| ccaggtgacc tggatcatcat catttcttat gcacaggtct ccgatgtga agcccggtgcc | 300 |
| ttaagaccaa gagtcgttca cgtcgacaga gataacagag ttgtcgcttt aggtgcagac | 360 |
| ccagcagagc cagttccagg ttccgatcaa gctagatcac cacaggctgt taccgcctaa | 420 |
| <210> 146 | |
| <211> 420 | |
| <212> DNA | |
| <213> Streptomyces avermitilis | |
| <400> 146 | |
| atgttaagaa ctatgtttaa gtccaagatt cacagagcta ccgtcaccca agcagatttg | 60 |
| cattacgtcg gtccggttac cattgatgca gatttactcg atgcggcaga cttgttacct | 120 |
| ggtgaactag ttcacattgt tgacattacc aatgggtgcta gattggaac ttacgtcatt | 180 |
| gaaggtgaaa gaggtagtagg tgtcgttggc atcaatggcg cagctgtcga cttagttcac | 240 |
| ccaggtgact tagtcatcat catttcatat gcacaagttt ccgatgcgga agctagagca | 300 |
| ttaagaccaa gagttgttca tgttgataga gacaacagag tcgttgactc tgggtgcagat | 360 |
| cctgctgaac cagttccagg ttcatatcaa gctaggtctc cacaagcagt tactgcataa | 420 |
| <210> 147 | |
| <211> 420 | |
| <212> DNA | |
| <213> Streptomyces avermitilis | |
| <400> 147 | |
| atgttaagaa ctatgttcaa gagtaagatt catagggcta ccgttaccca ggcagatcta | 60 |

cattatgttg gttctgttac cattgacgca gatttgttgg atgcagcaga cttgttacca 120
 ggtgaactcg ttcacattgt cgatatacc aacggtgcc aactggaaac ttatgtcatt 180
 gaaggtaga gaggcagtgg tgtgtcggc attaacggtg ccgcagcaca cttggttcat 240
 ccaggtgact tggatcatcat catttcttac gcacaagtct ccgatgccga agctagagca 300
 ttgagaccta gactcgttca cgttgacaga gacaatagag ttgttgctct tgggtcagat 360
 ccagctgagc cagttccagg ttccgatcaa gcgagatctc ctcaagctgt tactgcataa 420
 <210> 148
 <211> 384
 <212> DNA
 <213> Bacillus licheniformis

<400> 148
 atgtatagaa ccttgatgag tgcaaagctt cacagagcaa gagtactga agcaaacttg 60
 aactacgttg gttctgttac tattgacgaa gacttacttg atgcagtcgg tatgatggca 120
 aacgagaaaag ttcaaattgt caacaataac aatggtgcgc gtcttgaaac ctatatcatt 180
 cctggtgaac gtggttccgg tgtcgtctgc ttgaatggcg cagctgcaag gctcgtccaa 240
 gttggtgacg tcgtcatcat tgtttcttat gcaatgatgt ccgaagaaga agccaaaaca 300
 cataagccaa aggtcgtctgt cctcaatgaa agaaacgaaa ttgaggaaat gttaggtcag 360
 gaaccagcga gaaccatctt ataa 384

<210> 149
 <211> 384
 <212> DNA
 <213> Bacillus licheniformis
 <400> 149
 atgtatagaa ccttgatgag tgcaaagctt cacagagcaa gagtactga agcaaacttg 60
 aactacgttg gttctgttac tattgacgaa gacttacttg atgcagtcgg tatgatggca 120
 aacgagaaaag ttcaaattgt caacaataac aatggtgcgc gtcttgaaac ctatatcatt 180
 cctggtgaac gtggttccgg tgtcgtctgc ttgaatggcg cagctgcaag gctcgtccaa 240
 gttggtgacg tcgtcatcat tgtttcttat gcaatgatgt ccgaagaaga agccaaaaca 300
 cataagccaa aggtcgtctgt cctcaatgaa agaaacgaaa ttgaggaaat gttaggtcag 360
 gaaccagcga gaaccatctt ataa 384

<210> 150
 <211> 384

<212> DNA

<213> *Bacillus licheniformis*

<400> 150

| | |
|--|-----|
| atgtacagaa cgtaaagtc tgccaagtta cacagagcga gagttactga agcaaattctc | 60 |
| aactatgttg gttcagttac cattgatgag gatctattgg atgctgttgg tatgatggca | 120 |
| aatgaaaagg ttcaaactgt caataacaac aatgggtgcta gattagaaac ttacatcatt | 180 |
| cctggtgaaa gaggttcagg tgttgtttgc ttaaaccgtg ctgccgcaag attagttcaa | 240 |
| gtcggtagtg ttgttatcat cgtctcttat gctatgatga gtgaggaaga agctaagact | 300 |

| | |
|---|-----|
| cataagccta aggttgccgt tctcaatgag agaaacgaaa tcgaagaaat gcttggccaa | 360 |
| gaacctgcca gaaccatcct gtaa | 384 |

<210> 151

<211> 384

<212> DNA

<213> *Bacillus licheniformis*

<400> 151

| | |
|--|-----|
| atgtatagaa cattgatgtc tgcaaagttg catagggtc gtgttactga ggcaaattcta | 60 |
| aactacgtcg gttccgtcac aatcgatgaa gatctactcg atgccgttgg tatgatggcc | 120 |
| aatgaaaaag ttcaaattgt caacaacaac aacgggtgcaa gattggagac ctatatcatt | 180 |
| cctggtgaaa gaggttcagg tgttgtttgc ctgaacgggtg ctgctgcgag gttggtccaa | 240 |

| | |
|---|-----|
| gtcggtagtg ttgtcattat cgtttcttat gcaatgatgt cagaagaaga agctaagacc | 300 |
| cataagccaa aggttgctgt ttgaacgaa agaaatgaga ttgaggaaat gttaggtcaa | 360 |
| gaaccagcaa gaacaatctt ataa | 384 |

<210> 152

<211> 54

<212> DNA

<213> *Issatchenkia orientalis*

<400> 152

| | |
|--|----|
| taaaacgacg gccagtgaat tccgcggcgg ccgcgagtc atcggttcct gtca | 54 |
|--|----|

<210> 153

<211> 59

<212> DNA

<213> *Issatchenkia orientalis*

<400> 153

cataagaaaa tcaaagacag aaggcgcgcc ttgctagca tttttgtgtt ttgctgtgt 59

<210> 154

<211> 59

<212> DNA

<213> Artificial Sequence

<220><223> Artificial primer

<400> 154

acacagcaaa acacaaaaat gctagcaaag gcgcgccttc tgtctttgat tttcttatg 59

<210> 155

<211> 45

<212> DNA

<213> Artificial Sequence

<220><223> ARTIFICIAL PRIMER

<400> 155

gggggaaaga actaccaata ggcctccttt aattcggaga aaatc 45

<210> 156

<211> 45

<212> DNA

<213> Artificial Sequence

<220><223> ARTIFICIAL PRIMER

<400> 156

gattttctcc gaattaaagg aggcctattg gtagttcttt ccccc 45

<210> 157

<211> 51

<212> DNA

<213> Artificial Sequence

<220><223> ARTIFICIAL PRIMER

<400> 157

aaaataaact agtaaaataa attaattaat tatctagaga gggggttata t 51

<210> 158

<211> 51

<212> DNA

<213> Artificial Sequence

<220><223> ARTIFICIAL PRIMER

<400> 158

atataacccc ctctctagat aattaattaa tttatitttac tagttttatt t 51

<210> 159

<211> 53

<212> DNA

<213> Artificial Sequence

<220><223> ARTIFICIAL PRIMER

<400> 159

gaccatgatt acgccaagct ccgcggcggc cgcccagtca aaaccttctt etc 53

<210> 160

<211> 52

<212> DNA

<213> Artificial Sequence

<220><223> ARTIFICIAL PRIMER

<400> 160

taaaacgacg gccagtgaat tccgcggcgg ccgcctttga aggagcttgc ca 52

<210> 161

<211> 49

<212> DNA

<213> Artificial Sequence

<220><223> ARTIFICIAL PRIMER

<400> 161

ctattccttc ctcaaattgc tgtttaaagc cgttgaagat ctattctcc 49

<210> 162

<211> 49

<212> DNA

<213> Artificial Sequence

<220><223> ARTIFICIAL PRIMER

<400> 162

ggagaataga tcttcaacgc gtttaaagc caatttgagg aaggaatag 49

<210> 163

<211> 55

<212> DNA

<213> Artificial Sequence

<220><223> ARTIFICIAL PRIMER

<400> 163

aatgttcatt ttacattcag atgttaatta aggtctagat gtgtttgttt gtgtg 55

<210> 164

<211> 55

<212> DNA

<213> Artificial Sequence

<220><223> ARTIFICIAL PRIMER

<400> 164

cacacaaaca aacacatcta gaccttaatt aacatctgaa tgtaaaatga acatt 55

<210> 165

<211> 53

<212> DNA

<213> Artificial Sequence

<220><223> ARTIFICIAL PRIMER

<400> 165

gaccatgatt acgccaagct ccgcggcggc cgcaatgcca agagttatgg ggc 53

<210> 166

<211> 23

<212> DNA

<213> Artificial Sequence

<220><223> ARTIFICIAL PRIMER

<400> 166

ggctacccta tatatggtga gcg 23

<210> 167

<211> 21

<212> DNA

<213> Artificial Sequence

<220><223> ARTIFICIAL PRIMER

<400> 167

gggtccaagt tatccaagca g 21

<210> 168

<211> 29
 <212> DNA
 <213> Artificial Sequence
 <220><223> ARTIFICIAL PRIMER
 <400> 168
 ccttaattaa ccgtaaagtt gtctcaatg 29

<210> 169
 <211> 55
 <212> DNA
 <213> Artificial Sequence
 <220><223> ARTIFICIAL PRIMER
 <400> 169
 cagcaaaaaca caaaaattct agaaaattta attaacatct gaatgtaaaa tgaac 55

<210> 170
 <211> 55
 <212> DNA
 <213> Artificial Sequence
 <220><223> ARTIFICIAL PRIMER
 <400> 170
 gttcatttta cattcagatg ttaattaaat ttcttagaat ttttgtgttt tgctg 55

<210> 171
 <211> 32
 <212> DNA
 <213> Artificial Sequence
 <220><223> ARTIFICIAL PRIMER
 <400> 171
 gctctagata aaatgccctc ttattctgtc gc 32

<210> 172
 <211> 20
 <212> DNA
 <213> Artificial Sequence
 <220><223> ARTIFICIAL PRIMER
 <400> 172

| | |
|---|----|
| gactggatca ttatgactcc | 20 |
| <210> 173 | |
| <211> 44 | |
| <212> DNA | |
| <213> Artificial Sequence | |
| <220><223> ARTIFICIAL PRIMER | |
| <400> 173 | |
| acgccttgcc aaatgcggcc gcgagtcctat cggttcctgt caga | 44 |
| <210> 174 | |
| <211> 49 | |
| <212> DNA | |
| <213> Artificial Sequence | |
| <220><223> ARTIFICIAL PRIMER | |
| <400> 174 | |
| tcaaagacag aattaattaa gctctagaat ttttggtgtt tgctgtgtt | 49 |
| <210> 175 | |
| <211> 49 | |
| <212> DNA | |
| <213> Artificial Sequence | |
| <220><223> ARTIFICIAL PRIMER | |
| <400> 175 | |
| aacacaaaaa ttctagagct taattaattc tgtctttgat tttcttatg | 49 |
| <210> 176 | |
| <211> 32 | |
| <212> DNA | |
| <213> Artificial Sequence | |
| <220><223> ARTIFICIAL PRIMER | |
| <400> 176 | |
| gttttaaacct ttaattcgga gaaaatctga tc | 32 |
| <210> 177 | |
| <211> 30 | |
| <212> DNA | |
| <213> Artificial Sequence | |
| <220><223> ARTIFICIAL PRIMER | |

<400> 177

gttaacggta ccgagctcta agtagtggtg

30

<210> 178

<211> 41

<212> DNA

<213> Artificial Sequence

<220><223> ARTIFICIAL PRIMER

<400> 178

aaccgatgga ctgcggcgcg catttggtcaa ggcgtatcta t

41

<210> 179

<211> 30

<212> DNA

<213> Artificial Sequence

<220><223> ARTIFICIAL PRIMER

<400> 179

ggcgccagca atttgaggaa ggaataggag

30

<210> 180

<211> 52

<212> DNA

<213> Artificial Sequence

<220><223> ARTIFICIAL PRIMER

<400> 180

aaactattag attaattaag ctgcgatgt gtttgtttgt gtgttttgtg tg

52

<210> 181

<211> 52

<212> DNA

<213> Artificial Sequence

<220><223> ARTIFICIAL PRIMER

<400> 181

caaacaaca catcgcgagc ttaattaatc taatagttta atcacagctt at

52

<210> 182

<211> 42

<212> DNA

<213> Artificial Sequence

<220><223> ARTIFICIAL PRIMER

<400> 182

tacattatgg tagcgccgc gtgtgacatt ttgatccact cg 42

<210> 183

<211> 47

<212> DNA

<213> Artificial Sequence

<220><223> ARTIFICIAL PRIMER

<400> 183

tggatcaaaa tgtcacacgc ggccgctacc ataatgtatg cgttgag 47

<210> 184

<211> 27

<212> DNA

<213> Artificial Sequence

<220><223> ARTIFICIAL PRIMER

<400> 184

gggccctaaa agtgttggtg tattaga 27

<210> 185

<211> 33

<212> DNA

<213> Artificial Sequence

<220><223> ARTIFICIAL PRIMER

<400> 185

gctagctcaa caaactcttt atcagattta gca 33

<210> 186

<211> 30

<212> DNA

<213> Artificial Sequence

<220><223> ARTIFICIAL PRIMER

<400> 186

gctagcgagg aaaagaagtc taacctttgt 30

<210> 187

<211> 35
 <212> DNA
 <213> Artificial Sequence
 <220><223> ARTIFICIAL PRIMER
 <400> 187
 tcgcgataaa atgtcaactg tggaagatca ctcct 35
 <210> 188
 <211> 27
 <212> DNA
 <213> Artificial Sequence
 <220><223> ARTIFICIAL PRIMER
 <400> 188
 ttaattaagc tgctggcgct tcctctt 27
 <210> 189
 <211> 32
 <212> DNA
 <213> Artificial Sequence
 <220><223> ARTIFICIAL PRIMER
 <400> 189
 tcgcgataaa atgtccagag gcttctttac tg 32
 <210> 190
 <211> 30
 <212> DNA
 <213> Artificial Sequence
 <220><223> ARTIFICIAL PRIMER
 <400> 190
 ttaattaact aaaggtctct cagcacagag 30
 <210> 191
 <211> 40
 <212> DNA
 <213> Artificial Sequence
 <220><223> ARTIFICIAL PRIMER
 <400> 191
 gttaaccggg tttaacata gcctcatgaa atcagccatt 40

<210> 192

<211> 44

<212> DNA

<213> Artificial Sequence

<220><223> ARTIFICIAL PRIMER

<400> 192

gggcccatat ggcgcccggg gcgttgaaga tctattctcc agca 44

<210> 193

<211> 30

<212> DNA

<213> Artificial Sequence

<220><223> ARTIFICIAL PRIMER

<400> 193

gtttaaacga ttggtagttc tttccccctc 30

<210> 194

<211> 46

<212> DNA

<213> Artificial Sequence

<220><223> ARTIFICIAL PRIMER

<400> 194

aataaattaa gggcccttta tcgcgagagg gggttatatg tgtaaa 46

<210> 195

<211> 50

<212> DNA

<213> Artificial Sequence

<220><223> ARTIFICIAL PRIMER

<400> 195

tataaccccc tctcgcgata aaggccctt aatttatttt actagtttat 50

<210> 196

<211> 27

<212> DNA

<213> Artificial Sequence

<220><223> ARTIFICIAL PRIMER

| | |
|--|----|
| <400> 196 | |
| gtttaaactt ttgtagcacc tcctggt | 27 |
| <210> 197 | |
| <211> 48 | |
| <212> DNA | |
| <213> Artificial Sequence | |
| <220><223> ARTIFICIAL PRIMER | |
| <400> 197 | |
| cagcaaaaca caaaaagcta gctaaaatgt tacgtacat gttcaaaa | 48 |
| <210> 198 | |
| <211> 46 | |
| <212> DNA | |
| <213> Artificial Sequence | |
| <220><223> ARTIFICIAL PRIMER | |
| <400> 198 | |
| gaaaatcaaa gacagaaggc gcgccttatg ctgtaacagc ctgcgg | 46 |
| <210> 199 | |
| <211> 49 | |
| <212> DNA | |
| <213> Artificial Sequence | |
| <220><223> ARTIFICIAL PRIMER | |
| <400> 199 | |
| cagcaaaaca caaaaagcta gctaaaatgt taagaacat gttcaaatc | 49 |
| <210> 200 | |
| <211> 47 | |
| <212> DNA | |
| <213> Artificial Sequence | |
| <220><223> ARTIFICIAL PRIMER | |
| <400> 200 | |
| gaaaatcaaa gacagaaggc gcgccttatg cagtaacagc ttgtggg | 47 |
| <210> 201 | |
| <211> 20 | |
| <212> DNA | |
| <213> Artificial Sequence | |

<220><223> ARTIFICIAL PRIMER

<400> 201

gcatgggtgt gcaagcgacg

20

<210> 202

<211> 19

<212> DNA

<213> Artificial Sequence

<220><223> ARTIFICIAL PRIMER

<400> 202

ggtgctgcat ttgctgctg

19

<210> 203

<211> 25

<212> DNA

<213> Artificial Sequence

<220><223> ARTIFICIAL PRIMER

<400> 203

tgtatacagg atcgaagaat agaag

25

<210> 204

<211> 22

<212> DNA

<213> Artificial Sequence

<220><223> ARTIFICIAL PRIMER

<400> 204

gaacgtctac aacgaggtga ac

22

<210> 205

<211> 21

<212> DNA

<213> Artificial Sequence

<220><223> ARTIFICIAL PRIMER

<400> 205

ggcgcgcctc gcgataaaat g

21

<210> 206

<211> 32

| | |
|-------------------------------------|----|
| <212> DNA | |
| <213> Artificial Sequence | |
| <220><223> ARTIFICIAL PRIMER | |
| <400> 206 | |
| agggccctta attaattatg cagtaacagc tt | 32 |
| <210> 207 | |
| <211> 20 | |
| <212> DNA | |
| <213> Artificial Sequence | |
| <220><223> ARTIFICIAL PRIMER | |
| <400> 207 | |
| cgctacgata cgctacgata | 20 |
| <210> 208 | |
| <211> 20 | |
| <212> DNA | |
| <213> Artificial Sequence | |
| <220><223> ARTIFICIAL PRIMER | |
| <400> 208 | |
| ctcccttccc tgatagaagg | 20 |
| <210> 209 | |
| <211> 21 | |
| <212> DNA | |
| <213> Artificial Sequence | |
| <220><223> ARTIFICIAL PRIMER | |
| <400> 209 | |
| ggggagcaat ttgccaccag g | 21 |
| <210> 210 | |
| <211> 25 | |
| <212> DNA | |
| <213> Artificial Sequence | |
| <220><223> ARTIFICIAL PRIMER | |
| <400> 210 | |
| ctccttcatt taactatacc agacg | 25 |

<210> 211
 <211> 26
 <212> DNA
 <213> Artificial Sequence
 <220><223> ARTIFICIAL PRIMER
 <400> 211
 gacagatgta aggaccaata gtgtcc 26
 <210> 212
 <211> 23
 <212> DNA
 <213> Artificial Sequence
 <220><223> ARTIFICIAL PRIMER
 <400> 212
 ccatatcgaa atctagcccg tcc 23
 <210> 213
 <211> 58
 <212> DNA
 <213> Artificial Sequence
 <220><223> ARTIFICIAL PRIMER
 <400> 213
 cataagaaaa tcaaagacag aaggcgcgcc ttgctagct tttgtgttt tgctgtgt 58
 <210> 214
 <211> 19
 <212> DNA
 <213> Artificial Sequence
 <220><223> ARTIFICIAL PRIMER
 <400> 214
 attggacaca acattatat 19
 <210> 215
 <211> 46
 <212> DNA
 <213> Artificial Sequence
 <220><223> ARTIFICIAL PRIMER
 <400> 215

gtaaaacgac ggccagtgaat ttgttaacat aggctccaac atctcg 46

<210> 216

<211> 38

<212> DNA

<213> Artificial Sequence

<220><223> ARTIFICIAL PRIMER

<400> 216

ggaaccgatg gactcgcggc cgcgtgggat attggaag 38

<210> 217

<211> 22

<212> DNA

<213> Artificial Sequence

<220><223> ARTIFICIAL PRIMER

<400> 217

ggaggtgcta caaaaggaat tc 22

<210> 218

<211> 48

<212> DNA

<213> Artificial Sequence

<220><223> ARTIFICIAL PRIMER

<400> 218

gaccatgatt acgccaagct ccgcggagtc aaaaccttct tctctacc 48

<210> 219

<211> 46

<212> DNA

<213> Artificial Sequence

<220><223> ARTIFICIAL PRIMER

<400> 219

gtaaaacgac ggccagtgaat ttctttgaag gagcttgcca agaaac 46

<210> 220

<211> 23

<212> DNA

<213> Artificial Sequence

<220><223> ARTIFICIAL PRIMER
 <400> 220
 cctattcctt cctcaaattg ctg 23
 <210> 221
 <211> 38
 <212> DNA
 <213> Artificial Sequence
 <220><223> ARTIFICIAL PRIMER
 <400> 221
 ataactcttg gcattgcggc cgccaagtta gttagagc 38
 <210> 222
 <211> 51
 <212> DNA
 <213> Artificial Sequence
 <220><223> ARTIFICIAL PRIMER
 <400> 222
 atgaccatga ttacgccaag ctccgcgga aagacggtgt attagtgtt g 51
 <210> 223
 <211> 64
 <212> DNA
 <213> Artificial Sequence
 <220><223> ARTIFICIAL PRIMER
 <400> 223
 cacaaaacac acaaacacac acagctagca aaggcgcgcc atctaatagt ttaatcacag 60
 cttg 64
 <210> 224
 <211> 47
 <212> DNA
 <213> Artificial Sequence
 <220><223> ARTIFICIAL PRIMER
 <400> 224
 gccgttgcag caaatgtcga ggcctgtgtg acattttgat ccactcg 47
 <210> 225
 <211> 47

<212> DNA

<213> Artificial Sequence

<220><223> ARTIFICIAL PRIMER

<400> 225

cgagtggatc aaaatgtcac acaggcctcg acatttgctg caacggc 47

<210> 226

<211> 51

<212> DNA

<213> Artificial Sequence

<220><223> ARTIFICIAL PRIMER

<400> 226

cattttacat tcagatgtta attaattatc tagatgttg tgttgttg g 51

<210> 227

<211> 51

<212> DNA

<213> Artificial Sequence

<220><223> ARTIFICIAL PRIMER

<400> 227

cgacaacaac aacaacatct agataattaa ttaacatctg aatgtaaat g 51

<210> 228

<211> 51

<212> DNA

<213> Artificial Sequence

<220><223> ARTIFICIAL PRIMER

<400> 228

gctctaacta acttggcggc cgcttttatt ataaaattat atattattct t 51

<210> 229

<211> 47

<212> DNA

<213> Artificial Sequence

<220><223> ARTIFICIAL PRIMER

<400> 229

aacacacaaa caaacacagc tagctaaaat gttaagaact atgttta 47

<210> 230

<211> 46

<212> DNA

<213> Artificial Sequence

<220><223> ARTIFICIAL PRIMER

<400> 230

gattaaacta ttagatggcg cgccttatgc agtaactgct tgtgga 46

<210> 231

<211> 42

<212> DNA

<213> Artificial Sequence

<220><223> ARTIFICIAL PRIMER

<400> 231

cgacggccag tgaattcggtt aaccggtttc gatgggattc cc 42

<210> 232

<211> 46

<212> DNA

<213> Artificial Sequence

<220><223> ARTIFICIAL PRIMER

<400> 232

caggaaccga tggactcgcg gccgctccct tctctaaatg gactgc 46

<210> 233

<211> 48

<212> DNA

<213> Artificial Sequence

<220><223> ARTIFICIAL PRIMER

<400> 233

tatataattt tataataaaa gcggccgcac caggggttta gtgaagtc 48

<210> 234

<211> 47

<212> DNA

<213> Artificial Sequence

<220><223> ARTIFICIAL PRIMER

| | |
|---|----|
| <400> 234 | |
| catgattacg ccaagctccg cggccataac tgacatttat ggtaagg | 47 |
| <210> 235 | |
| <211> 21 | |
| <212> DNA | |
| <213> Artificial Sequence | |
| <220><223> ARTIFICIAL PRIMER | |
| <400> 235 | |
| ccaacaatct taattggtga c | 21 |
| <210> 236 | |
| <211> 20 | |
| <212> DNA | |
| <213> Artificial Sequence | |
| <220><223> ARTIFICIAL PRIMER | |
| <400> 236 | |
| ggctgttacc gcctaattaa | 20 |
| <210> 237 | |
| <211> 23 | |
| <212> DNA | |
| <213> Artificial Sequence | |
| <220><223> ARTIFICIAL PRIMER | |
| <400> 237 | |
| gttcttaaca ttttagctag ctg | 23 |
| <210> 238 | |
| <211> 30 | |
| <212> DNA | |
| <213> Artificial Sequence | |
| <220><223> ARTIFICIAL PRIMER | |
| <400> 238 | |
| acgcgtcgac tcgacatttg ctgcaacggc | 30 |
| <210> 239 | |
| <211> 29 | |
| <212> DNA | |
| <213> Artificial Sequence | |

<220><223> ARTIFICIAL PRIMER

<400> 239

ctagtctaga tgttggtgtt gttgtcggt

29

<210> 240

<211> 49

<212> DNA

<213> Artificial Sequence

<220><223> ARTIFICIAL PRIMER

<400> 240

caaaacacag caaaacacaa aaagctagca tgtatagaac cttgatgag

49

<210> 241

<211> 40

<212> DNA

<213> Artificial Sequence

<220><223> ARTIFICIAL PRIMER

<400> 241

caaagacaga aggcgcgcct tataagatgg ttctcgctgg

40

<210> 242

<211> 50

<212> DNA

<213> Artificial Sequence

<220><223> ARTIFICIAL PRIMER

<400> 242

caaaacacac aaacaaacac agctagcatg tacagaacgt taatgtctgc

50

<210> 243

<211> 47

<212> DNA

<213> Artificial Sequence

<220><223> ARTIFICIAL PRIMER

<400> 243

gtgattaaac tattagatgg cgcgccttac aggatggttc tggcagg

47

<210> 244

<211> 703

<212> DNA

<213> *Issatchenkia orientalis*

<400> 244

| | |
|--|-----|
| gagtcacatcg gtccctgtca gatgggatac tcttgacgtg gaaaattcaa acagaaaaaa | 60 |
| aacccaata atgaaaaata acactacgtt atatccgtgg taccctctat cgtatcgtat | 120 |
| cgtagcgtat cgtagcgtac cgtatcacag tatagtctaa tattccgtat cttattgtat | 180 |
| cctatcctat tcgatcctat tgtatttcag tgcaccattt taatttctat tgctataatg | 240 |
| | |
| tccttattag ttgccactgt gaggtgacca atggacgagg gcgagccgtt cagaagccgc | 300 |
| gaagggtgtt ctcccatga atttcttaag gagggcggct cagctccgag agtgaggcga | 360 |
| gacgtctcgg tcagcgtatc ccccttcctc ggcttttaca aatgatgcgc tcttaatagt | 420 |
| gtgtcgttat ccttttgga ttgacggggg agggaaattg attgagcgca tccatatttt | 480 |
| tgcggactgc tgaggacaat ggtggttttt ccgggtggcg tgggctacaa atgatacgat | 540 |
| ggtttttttc ttttcggaga aggcgtataa aaaggacacg gagaacccat ttattctaaa | 600 |
| aacagttgag cttctttaat tattttttga tataatattc tattattata tattttcttc | 660 |
| | |
| ccaataaaac aaaataaaac aaaacacagc aaaacacaaa aat | 703 |

<210> 245

<211> 961

<212> DNA

<213> *Issatchenkia orientalis*

<400> 245

| | |
|--|-----|
| agcaatttga ggaaggaata ggagaaggag aagcaatttc taggaaagag caaggtgtgc | 60 |
| aacagcatgc tctgaatgat attttcagca atagttcagt tgaagaacct gttggcgtat | 120 |
| ctacatcact tcctacaaac aacaccacga attgcgtccg tggtagcgca actacgaatg | 180 |
| gcattgtcaa tgccaatgcc agtgacata cacgtgcaag tcccaccggt tccctgcccg | 240 |
| gctatggtag agacaagaag gacgataccg gcatacgacat caacagtttc aacagcaatg | 300 |
| | |
| cgtttggcgt cgacgcgtcg atggggctgc cgtatttga tttggacggg ctagatttcg | 360 |
| atatggatat ggatatggat atggatatgg agatgaattt gaatttagat ttgggtcttg | 420 |
| atttgggggtt ggaattaaaa ggggataaca atgagggttt tcctgttgat ttaaacaatg | 480 |
| gacgtgggag gtgattgatt taacctgac caaaaggggt atgtctattt tttagagtgt | 540 |
| gtctttgtgt caaattatgg tagaatgtgt aaagtagtat aaactttcct ctcaatgac | 600 |
| gaggttttaa acacccccg ggtgagccga gccgagaatg gggcaattgt tcaatgtgaa | 660 |

atagaagtat cgagtgagaa acttgggtgt tggccagcca agggggaagg aaaatggcgc 720

gaatgctcag gtgagattgt ttiggaattg ggtgaagcga ggaaatgagc gacccggagg 780

ttgtgacttt agtggcggag gaggacggag gaaaagccaa gagggaagtg tatataaggg 840

gagcaatttg ccaccaggat agaattggat gagttataat tctactgtat ttattgtata 900

atttatttct cttttatata caaacacatt acaaaacaca caaacacac aaacaaacac 960

a 961

<210> 246

<211> 992

<212> DNA

<213> Issatchenkia orientalis

<400> 246

tattggtagt tctttcccc tctcaagctg gcgtgaaatg caaccttacg gcgtctacgt 60

tactacaagg tccagaaagt gtaggtattg ctactatfff tattttttat tggttctgga 120

gaaatgcaga cagtcaatga acacaactgt ctcaatatgc atctatgcac atgcacacac 180

acacacatca caggtacccc tacaagaga ggtctcttga taatgtttca ttaccacgtg 240

gcatcccccc ccccccccc aataaacaag tggccgagtt ccctgttgc agaggaggac 300

aaaaaaaccg ctgggtgttg taccattatg cagcaactag cacaacaaac aaccgacca 360

gacatacaaa tcaacaacac ttgcctaaag acaccctttc cagggaggat ccactcccaa 420

cgtctctcca taatgtctct gttggcccat gtctctgtcg ttgacaccgt aaccacacca 480

accaaccctg ccattgtact gggatggctg tccatagaca cctctccaac ggggaacacc 540

tcattcgtaa accgccaagg ttaccgttcc tctgactcg ccccggtgtt gatgctgcgc 600

acctgtggtt gcccaacatg gttgtatatc gtgtaaccac accaacacat gtgcagcaca 660

tgtgtttaaa agagtgtcat ggaggtggat catgatggaa gtggacttta ccacttgga 720

actgtctcca ctcccgggaa gaaaagaccg ggcgtatcac gcggttgcct caatggggca 780

atttgaagg agaaatatag ggaaatcac gtcgtctctg gacggggaag agttccagac 840

tatgaggggg gggggtggta tataagaca ggagatgtcc accccagag agaggaagaa 900

gttgaactt tagaagagag agataacttt cccagtgct catcaatata caaccaaaca 960

caaactctat atttacacat ataaccct ct 992

<210> 247

<211> 613

<212> DNA

<213> Issatchenkia orientalis

<400> 247

| | |
|--|-----|
| ttgctgcaac ggcaacatca atgtccacgt ttacacacct acatttatat ctatatttat | 60 |
| atttatattt atttatattt gctacttagc ttctatagtt agttaatgca ctcacgatat | 120 |
| tcaaaattga cacccttcaa ctactcccta ctattgtcta ctactgtcta ctactcctct | 180 |
| ttactatagc tgctcccaat aggtccacc aataggctct gtcaatacat tttgcgccgc | 240 |
| | |
| cacctttcag gttgtgtcac tctgaagga ccatattggg taatcgtgca atttctggaa | 300 |
| gagagtccgc gagaagttag gccccactg taaatcctcg agggggcatg gagtatgggg | 360 |
| catggaggat ggaggatggg gggggggggg gggaaaatag gtagcgaaag gaccgcgtat | 420 |
| cacccacccc ggagaactcg ttgccgggaa gtcatatattc gacactccgg ggagtctata | 480 |
| aaaggcgggt tttgtctttt gccagttgat gttgctgaga ggacttgttt gccgtttctt | 540 |
| ccgatttaac agtatagaat caaccactgt taattataca cgttatacta acacaacaaa | 600 |
| aacaaaaaca acg | 613 |

<210> 248

<211> 399

<212> DNA

<213> Issatchenkia orientalis

<400> 248

| | |
|--|-----|
| tctgtctttg attttcttat gttattcaaa acatctgccc caaaatctaa cgattatata | 60 |
| tattcctacg tataactgta tagctaatta ttgatttatt tgtacataaa aaccacataa | 120 |
| atgtaaaagc aagaaaaaaa ataactaagg agaaggatca atatctcatt tataatgctc | 180 |
| gccaaagcag cgtacgtgaa ttttaatcaa gacatcaaca aatcttgcaa cttgggttata | 240 |
| tcgtcttctc acccactcac ccgttttctt acattgttga acacaaatat atacaggggt | 300 |
| atgtctcaag gtcaagtgca gtttcaacag agactacctc aaggtaacctc ttcagaaatg | 360 |
| | |
| cagaacttca ctcttgatca gattttctcc gaattaaag | 399 |

<210> 249

<211> 422

<212> DNA

<213> Issatchenkia orientalis

<400> 249

| | |
|---|-----|
| atctaatagt ttaatcacag cttatagtct attatagttt tcttttttaa acattgttgt | 60 |
| atttgtccc cccccctcta attgatgatg attatcctat aagaatccaa taaaacgatg | 120 |

gaaactaata ctttctcctt tgtcatgtgg tcttttagtat ttcttgaaca ttggctctga 180
 tttctcgact ttatagtcct attaaaaatct ctgttagttc tcgatcgttg tatctcgttt 240
 cttgtctctt tgggtgatga ttttgcgtgc gaacatgttt ttttcccttt ctctcaccat 300

catcgtgtag ttcttgtcac catcccccca ccccttcctt ctctcattga ttctataaga 360
 gcttatccac agagggtcag taacgaggta gttaaacctt cgagtggatc aaaatgtcac 420
 ac 422

<210> 250
 <211> 421
 <212> DNA
 <213> Issatchenkia orientalis
 <400> 250

tttatattac tagtttattt ttgctcctga gaataggatt acaaacactt aaagtcttta 60
 attacaacta tatataatat tctgttgggt ttcttgaatt gggtcgtgc gattcatgcc 120
 tcccattcac caaagggtga gtgggaaata acggttttac tgcggttaatt agcagaggca 180

agaacaggat acactttttg atgataaatc tgtattatag tcgagcctat ttaggaaatc 240
 aaattttctt gtgtttactt ttcaaataaa taatgttcga aaatttttac tttactcctt 300
 catttaacta taccagagct tatatcatca acaccttctg accatataca gctcaagatg 360
 tttaagagtc tgttaaattt tttcaatcca ttcatggag taccaggagg tgctacaaaa 420
 g 421

<210> 251
 <211> 281
 <212> DNA
 <213> Issatchenkia orientalis
 <400> 251

catctgaatg taaaatgaac attaaaatga attactaaac tttacgtcta ctttacaatc 60

tataaacttt gttaaatcat ataacgaaat acactaatac acaatcctgt acgtatgtaa 120
 tacttttacc catcaaggat tgagaaaaaa aagtaatgat tccttgggcc attaaaactt 180
 agacccccaa gcttggatag gtcactctct attttcgttt ctcccttccc tgatagaagg 240
 gtgatatgta attaagaata atatataatt ttataataaa a 281

<210> 252
 <211> 1394
 <212> DNA

<213> Issatchenkia orientalis

<400> 252

catagcctca tgaatcagc catttgcttt tgttcaacga tcttttgaaa ttgttgttgt 60
tcttggtagt taagttgac catcttggt tatgttgtgt gtatgttgta gttattctta 120

gtatattcct gtcctgagtt tagtgaaaca taatatcgcc ttgaaatgaa aatgctgaaa 180
ttcgtcgaca tacaattttt caaacttttt ttttttcttg gtgcacggac atgtttttta 240
aggaagtact ctataccagt tattcttcac aaatttaatt gctggagaat agatcttcaa 300
cgctttaata aagtagtttg tttgtcaagg atggcgcat acaaagaaag atcagaatca 360
cacacttccc ctgttgctag gagacttttc tccatcatgg aggaaaagaa gtctaacctt 420
tgtgcatcat tggatattac tgaactgaa aagcttctct ctattttgga cactattggt 480
ccttacatct gtctagttaa aacacacatc gatattgttt ctgattttac gtatgaagga 540

actgtgttgc ctttgaagga gcttgccaag aaacataatt ttatgatttt tgaagataga 600
aaatttgctg atattggtaa cactgttaaa aatcaatata aatctgggtgt cttccgtatt 660
gccgaatggg ctgacatcac taatgcacat ggtgtaacgg gtgcaggtat tgtttctggc 720
ttgaaggagg cagcccaaga aacaaccagt gaacctagag gtttgctaata gcttgctgag 780
ttatcatcaa agggttcttt agcatatggt gaatatacag aaaaaacagt agaaattgct 840
aaatctgata aagagtttgt cattggtttt attgcgcaac acgatatggg cggtagagaa 900
gaaggttttg actggatcat tatgactcca ggggttgggt tagatgacaa aggtgatgca 960

cttgggtcaac aatatagaac tgttgatgaa gttgtaaaga ctggaacgga taccataatt 1020
gttggttagag gtttgtacgg tcaaggaaga gatcctatag agcaagctaa aagataccaa 1080
caagctgggtt ggaatgctta tttaaacaga tttaaatgat tcttacacaa agatttgata 1140
catgtacact agtttaata agcatgaaaa gaattacaca agcaaaaaaa aaaaaataaa 1200
tgaggtaact tacgttcacc tacaacaaa aaaactagat agagtaaaat cttaagattt 1260
agaaaaagtt gtttaacaaa ggcttttagta tgtgaatttt taatgtagca aagcgataac 1320
taataaacat aaacaaaagt atggttttct ttatcagtca aatcattatc gattgattgt 1380

tccgcgtatc tgca 1394

<210> 253

<211> 523

<212> DNA

<213> Issatchenkia orientalis

<400> 253

gtataaactg tatgtcatta taaacaggga aggttgacat tgtctagcgg caatcattgt 60
ctcatttggg tcattaactt tggttctgtt ctgggaaacg ggtaccaact ctctcagagt 120
gcttcaaaaa tttttcagca catttggtta gacatgaact ttctctgctg gttaaggatt 180
cagaggtgaa gtcttgaaca caatcgttga aacatctgtc cacaagagat gtgtatagcc 240
tcatgaaatc agccatttgc ttttgttcaa cgatcttttg aaattgttgt tgttcttggg 300

agttaagttg atccatcttg gcttatgttg tgtgtatgtt gtagttattc ttagtatatt 360
cctgtcctga gtttagtgaa acataatata gccttgaaat gaaaatgctg aaattcgtcg 420
acatacaatt tttcaaactt ttttttttct ttgggtgcacg gacatgtttt taaaggaagt 480
actctatacc agttattctt caccctgcag ggtacgtagc atg 523

<210> 254

<211> 626

<212> DNA

<213> Issatchenkia orientalis

<220><221> misc_feature

<222> (28)..(28)

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

<400> 254

gggcaacaaa gcctccaga tttgatanat tticaatttg tgcittgaat catgacttcc 60

acctgtttgg tccgcaagaa cagtaaatg cgcaatttgt ttctcccttc tgcttaaaaa 120
ccatgcacct ttaatatatt ctggaaagat aaagaacaga attgttgcgt agaaacaagt 180
agcagagccg taaatgagaa aaatatactt ccaagctggt aatttccctt ttattagtcc 240
aatacagtgt ccgaagacc caccaagaat accagcaagg gtgttgaaat ataatgtaga 300
tcttagtggt tgttctgatt tcttccacca cattccgcta ataatacata aagacggtaa 360
tattccggct tcaaatacgc caagaaaaaa cctcacggta accaaaccac caaagctatg 420
acatgcagcc atgcacataa gtaagccgcc ccaaatgaac aaacaaatag acacaaattt 480

gccaatctta actcgtggca acaaaaaaaaa ggatatgaac tcaccttaata aataaccgaa 540
ataaaaagta gaagcaactg tggaaaattg agaaccatgt aaatttgtgt cttctttcaa 600
tgtataaaca gccgcaatac ctaggg 626

<210> 255

<211> 1416

<212> DNA

<213> Saccharomyces cerevisiae

<400> 255

atgtttgctt tctactttct caccgcatgc accactttga aggggtgtttt cggagtttct 60
ccgagttaca atggtcttgg tctcacccca cagatgggtt gggacagctg gaatacgttt 120
gcctgcgatg tcagtgaaca gctacttcta gacactgctg atagaatttc tgacttgggg 180

ctaaaggata tgggttacaa gtatgtcatc ctagatgact gttggtctag cggcagggat 240
tccgacgggt tctctgttgc agacaagcac aaatttccca acggtatggg ccatgttgca 300
gaccacctgc ataataacag ctttcttttc ggtatgtatt cgtctgctgg tgagtacacc 360
tgtgctgggt accctgggtc tctggggcgt gaggaagaag atgctcaatt ctttgcaaat 420
aaccgcgttg actacttgaa gtatgataat tgttacaata aaggtaatt tggtacacca 480
gacgtttctt accaccgtta caaggccatg tcagatgctt tgaataaaac tggtaggcct 540
attttctatt ctctatgtaa ctggggtcag gatttgacat ttactgggg ctctggtatc 600

gccaatctt ggagaatgag cggagatatt actgctgagt tcacccgtcc agatagcaga 660
tgtccctgtg acggtgacga atatgattgc aagtacgccg gtttccattg ttctattatg 720
aatattctta acaaggcagc tccaatgggg caaaatgcag gtgttgggtg ttggaacgat 780
ctggacaatc tagaggtcgg agtcggtaat ttgactgacg atgaggaaaa ggccatttc 840
tctatgtggg caatggtaaa gtccccactt atcattgggtg ccgacgtgaa tcaactaaag 900
gcattcttct actcgatcta cagtcaagcc tctgtcatcg caattaatca agatccaaag 960
ggtatccag ccacaagagt ctggagatat tatgtttcag acaccgatga atatggacaa 1020

ggtgaaattc aaatgtggag tgggtccgtt gacaatggtg accaagtgtg tgctttattg 1080
aatggaggaa gcgtagcaag accaatgaac acgaccttgg aagagatttt ctttgacagc 1140
aatttgggtt caaaggaact gacatcgact tgggatattt acgacttatg ggccaacaga 1200
gttgacaact ctacggcgtc tgctatcctt gaacagaata aggcagccac cggatatttc 1260
tacaatgcta cagagcagtc ttataaagac ggtttgtcta agaataatg aagactgttt 1320
ggccagaaaa ttggtagtct ttctccaaat gctatactta acacaactgt tccagctcat 1380
ggtatcgctt tctataggtt gagaccctcg gcttaa 1416

<210> 256

<211> 707

<212> DNA

<213> Kluyveromyces thermotolerans

<400> 256

cactcgcaag ctgtgccatc gcccaacggt taattataag aaatcaacat cagccaacaa 60

ctattttcgt cccctctttt tcagtggtaa cgagcaatta cattagtaag agactatttt 120
 cttcagtgat ttgtaatfff ttttcagtga ttgtgaattc tttctcgaaa tatgcgggct 180
 taacttatcc ggacattcac tacatgcaag gaaaaacgag aaccgcggag atttcctcag 240
 taagtaacaa tgatgatctt tttacgttc atcatcactt tccaaagtcc taagctataa 300
 gttcaagcct agatacgctg aaaaactcct gaccaacaat gtaaagaaaa caattacaat 360

tgttaaggttg aaaacatcta aaaatgaaat attttattgt acatgcacac cctgatagtc 420
 attctcttac ttcacccctg aaagacgtgg ctgtacaaga gttggaatcg caaggtcatg 480
 aggttaaagt tagtgatctt tatgctcaaa agtggaaggc ctcaatagac cgtgacgact 540
 tcgagcagct tttcgcaaga agagaggtta aaaatacccc aagcttctta tgaagcgtat 600
 gccagaggag cattaacaaa agacgtaaat caggaacagg aaaaacttat ttgggcggac 660
 tttgtcattt tgcgtttcc tatatgggtg tcttctatgc cggctag 707

<210> 257

<211> 655

<212> DNA

<213> Kluyveromyces thermotolerans

<400> 257

ggttaattat aagaaatcaa catcagccaa caactatfff cgtcccccctc ttttcagtgg 60
 taacgagcaa ttacattagt aagagactat tttcttcagt gatttgtaat tttttttcag 120
 tgatttgtaa ttttttctcg aaatatgcgg gcttaactta tccggacatt cactacatgc 180
 aaggaaaaac gagaaccgag gagatttcct cagtaagtaa caatgatgat ctttttacgc 240
 ttcacatca ctttccaaag ttctaagcta taagttcaag cctagatagc ctgaaaaact 300
 cctgaccaac aatgtaaaga aaacaattac aattgtaagg ttgaaaacat ctaaaaatga 360
 aatatfffat tgtacatgca caccctgata gtcattctct tacttcatcc ctgaaagacg 420

tggctgtaca agagttggaa tcgcaaggtc atgaggttaa agttagtat ctttatgctc 480
 aaaagtggaa ggccctcaata gaccgtgacg acttcgagca gcttttcgca agaagagagg 540
 ttaaaaaatc cccaagcttc ttatgaagcg tatgccagag gagcattaac aaaagacgta 600
 aatcaggaac aggaaaaact tatttgggag gactttgtca tttgtcgtt tccta 655

<210> 258

<211> 523

<212> DNA

<213> Issatchenkia orientalis

<400> 258

gtataaactg tatgtcatta taaacaggga aggttgacat tgtctagcgg caatcattgt 60
ctcatttggg tcattaactt tggttctgtt ctgggaaacg ggtaccaact ctctcagagt 120

gcttcaaaaa tttttcagca catttggtta gacatgaact ttctctgctg gttaaggatt 180
cagaggtgaa gtcttgaaca caatcgttga aacatctgtc cacaagagat gtgtatagcc 240
tcatgaaatc agccatttgc ttttgttcaa cgatcttttg aaattgttgt tgttcttggg 300
agttaagttg atccatcttg gcttatgttg tgtgtatgtt gtagttattc ttagtatatt 360
cctgtcctga gtttagtgaa acataatata gccttgaaat gaaaatgctg aaattcgtcg 420
acatacaatt tttcaaactt ttttttttct ttgggtgcacg gacatgtttt taaaggaagt 480
actctatacc agttattctt caccctgcag ggtacgtagc atg 523

<210> 259

<211> 707

<212> DNA

<213> Issatchenkia orientalis

<400> 259

tgataacagc aatctgcatg tggattcata atattagggtg tttattgtgg gtattaacgt 60
gtctacctgc tttagctggg gctattatgg taaataagat tgatccacac gaaaatcgac 120
atgctgcatt agctggaatt tatttgatgg gattctacaa tgtgcatgg acattaatgt 180
tggcgttggg atcttcaaac acctcaggat ctaccaagaa gacgttcatg tatgtttctg 240
ttgcagtttg gtatgcagtt ggtaacatta tcgggtccata ttttttcaaa ggttctcaag 300
caccaaagta tccgatgggg atacatgcca tggaagcttc cttttcaatt atggcgttca 360

ccggtatcct ttactatctt atgctgagaa tacagaataa acaaaagcaa tcgttaacta 420
tgaatgatat aaaacactta gaatacgagg aacaatcaat cgataatgat ttgactgata 480
aagaaaacca tacttttgtt tatgtttatt agttatcgct ttgctacatt aaaaattcac 540
atactaaagc ctttgtttaa caacttttct taaatcataa gattttactc tatctagttt 600
ttttggttgt aggtgaacgt aaagtacctc atttaatttt tttttgcttg tgtaattctt 660
ttcatgttta tttaaactag tgtacatgta tcaaatcttt gtgtaag 707

<210> 260

<211> 22

<212> DNA

<213> Artificial Sequence

<220><223> ARTIFICIAL PRIMER

<400> 260
gcaactgatg ttcacgaatg cg 22
<210> 261
<211> 20
<212> DNA
<213> Artificial Sequence
<220><223> ARTIFICIAL PRIMER
<400> 261
ttgccgttgc agcaaattctc 20
<210> 262
<211> 20
<212> DNA
<213> Artificial Sequence
<220><223> ARTIFICIAL PRIMER
<400> 262
aatgatccat ggtccgagag 20
<210> 263
<211> 21
<212> DNA
<213> Artificial Sequence
<220><223> ARTIFICIAL PRIMER
<400> 263
acggcagtat accctatcag g 21
<210> 264
<211> 28
<212> DNA
<213> Artificial Sequence
<220><223> ARTIFICIAL PRIMER
<400> 264
gcaccaagag cagttttccc atctattg 28
<210> 265
<211> 36
<212> DNA
<213> Artificial Sequence

<220><223> ARTIFICIAL PRIMER

<400> 265

ccatatagtt cttttctaac atcaacatca cacttc

36

<210> 266

<211> 37

<212> DNA

<213> Artificial Sequence

<220><223> ARTIFICIAL PRIMER

<400> 266

ggagaataga tcttcaacgc ttttaataaag tagtttg

37

<210> 267

<211> 25

<212> DNA

<213> Artificial Sequence

<220><223> ARTIFICIAL PRIMER

<400> 267

cgtgttgcg c aataaaacca atgac

25

<210> 268

<211> 20

<212> DNA

<213> Artificial Sequence

<220><223> ARTIFICIAL PRIMER

<400> 268

cctcgaagag ctgaatttg

20

<210> 269

<211> 20

<212> DNA

<213> Artificial Sequence

<220><223> ARTIFICIAL PRIMER

<400> 269

gtagtgaatg tccggataag

20

<210> 270

<211> 20

<212> DNA
 <213> Artificial Sequence
 <220><223> ARTIFICIAL PRIMER
 <400> 270
 gcaaggtcat gaggttaaag 20
 <210> 271
 <211> 20
 <212> DNA
 <213> Artificial Sequence
 <220><223> ARTIFICIAL PRIMER
 <400> 271
 aacacttatg gcgtctcctc 20

 <210> 272
 <211> 23
 <212> DNA
 <213> Artificial Sequence
 <220><223> ARTIFICIAL DNA PRIMER
 <400> 272
 cgaaccaatt caagaaaacc aac 23
 <210> 273
 <211> 18
 <212> DNA
 <213> Artificial Sequence
 <220><223> ARTIFICIAL DNA PRIMER
 <400> 273
 tctgtccctt ggcgcgcg 18
 <210> 274
 <211> 21
 <212> DNA
 <213> Artificial Sequence
 <220><223> ARTIFICIAL DNA PRIMER
 <400> 274
 cttttcaaac agataagtac c 21

<210> 275
 <211> 20
 <212> DNA
 <213> Artificial Sequence
 <220><223> ARTIFICIAL DNA PRIMER
 <400> 275
 atgggctgac ctgaaaattc 20
 <210> 276
 <211> 23
 <212> DNA
 <213> Artificial Sequence
 <220><223> ARTIFICIAL DNA PRIMER
 <400> 276
 gctgaaaata tcattcagag cat 23
 <210> 277
 <211> 18
 <212> DNA
 <213> Artificial Sequence
 <220><223> ARTIFICIAL DNA PRIMER
 <400> 277
 actgttgatg tcgatgcc 18

 <210> 278
 <211> 20
 <212> DNA
 <213> Artificial Sequence
 <220><223> ARTIFICIAL DNA PRIMER
 <400> 278
 tctgaatgca gtacgagttg 20
 <210> 279
 <211> 20
 <212> DNA
 <213> Artificial Sequence
 <220><223> ARTIFICIAL DNA PRIMER
 <400> 279

gaaggggggtc caagttatcc 20

<210> 280

<211> 20

<212> DNA

<213> Artificial Sequence

<220><223> ARTIFICIAL DNA PRIMER

<400> 280

gatatgggcg gtagagaaga 20

<210> 281

<211> 20

<212> DNA

<213> Artificial Sequence

<220><223> ARTIFICIAL DNA PRIMER

<400> 281

gctccttcaa aggcaacaca 20

<210> 282

<211> 20

<212> DNA

<213> Artificial Sequence

<220><223> ARTIFICIAL DNA PRIMER

<400> 282

tgaactatca catgaacgta 20

<210> 283

<211> 20

<212> DNA

<213> Artificial Sequence

<220><223> ARTIFICIAL DNA PRIMER

<400> 283

tcaaggtagg gtcacttaac 20

<210> 284

<211> 20

<212> DNA

<213> Artificial Sequence

<220><223> ARTIFICIAL DNA PRIMER
 <400> 284
 tgattccttc aatcacaggt 20
 <210> 285
 <211> 20
 <212> DNA
 <213> Artificial Sequence
 <220><223> ARTIFICIAL DNA PRIMER
 <400> 285
 aaccgacctt tcgaatgcct 20
 <210> 286
 <211> 51
 <212> DNA
 <213> Artificial Sequence
 <220><223> ARTIFICIAL DNA PRIMER
 <400> 286
 accatgatta cgccaagctt ggtaccttgg gggtttacgct tacagcgtac t 51
 <210> 287
 <211> 61
 <212> DNA
 <213> Artificial Sequence
 <220><223> ARTIFICIAL DNA PRIMER
 <400> 287
 tcaacgcccc gggggatctg gatccgcggc cgcaagaaat tcttttcttt tcccctttat 60
 a 61
 <210> 288
 <211> 58
 <212> DNA
 <213> Artificial Sequence
 <220><223> ARTIFICIAL DNA PRIMER
 <400> 288
 aaccgatgga ctctcgagg gatccgcggc cgcgcataac aaaattgtgc ctaaccca 58
 <210> 289
 <211> 53

<212> DNA

<213> Artificial Sequence

<220><223> ARTIFICIAL DNA PRIMER

<400> 289

acgactcact ataggcgcaa ttgggccccg ggaaaaggag agagaaaagg aga 53

<210> 290

<211> 55

<212> DNA

<213> Artificial Sequence

<220><223> ARTIFICIAL DNA PRIMER

<400> 290

tatgaccatg attacgcaa gcttggtacc tctacagaa gtagcaccag cacca 55

<210> 291

<211> 60

<212> DNA

<213> Artificial Sequence

<220><223> ARTIFICIAL DNA PRIMER

<400> 291

ggaaccgatg gactcctcga gggatccgcg gccgcagact accgtgtaaa taagagtacc 60

<

210> 292

<211> 61

<212> DNA

<213> Artificial Sequence

<220><223> ARTIFICIAL DNA PRIMER

<400> 292

tcttcaacgc cccgggggat ctggatccgc ggccgcattt gatataaacg cttctataat 60

a 61

<210> 293

<211> 58

<212> DNA

<213> Artificial Sequence

<220><223> ARTIFICIAL DNA PRIMER

<400> 293

gtaatacgac tcactatagg gcgaattggg cccaacatct gctgctgtaa tatattca 58

<210> 294

<211> 25

<212> DNA

<213> Artificial Sequence

<220><223> ARTIFICIAL DNA PRIMER

<400> 294

gcatgtctgt taactctcaa accat 25

<210> 295

<211> 23

<212> DNA

<213> Artificial Sequence

<220><223> ARTIFICIAL DNA PRIMER

<400> 295

tccataatcc aatagatcat ccc 23

<210> 296

<211> 22

<212> DNA

<213> Artificial Sequence

<220><223> ARTIFICIAL DNA PRIMER

<400> 296

aacacaatgg aaccaaccta gt 22

<

210> 297

<211> 45

<212> DNA

<213> Artificial Sequence

<220><223> ARTIFICIAL DNA PRIMER

<400> 297

atttacacgg tagtctgcgg ccgccatagc ctcatgaaat cagcc 45

<210> 298

<211> 40

<212> DNA

<213> Artificial Sequence

<220><223> ARTIFICIAL DNA PRIMER

<400> 298

ctctaggttc actggttggt tcttgggctg cctccttcaa 40

<210> 299

<211> 40

<212> DNA

<213> Artificial Sequence

<220><223> ARTIFICIAL DNA PRIMER

<400> 299

ttgaaggagg cagccaaga aacaaccagt gaacctagag 40

<210> 300

<211> 56

<212> DNA

<213> Artificial Sequence

<220><223> ARTIFICIAL DNA PRIMER

<400> 300

tattatagaa gcgtttatat caaatgcggc cgcggatcca gatcccccg ggcgtt 56

<210> 301

<211> 20

<212> DNA

<213> Artificial Sequence

<220><223> ARTIFICIAL DNA PRIMER

<400> 301

aatgatcaac ttgagaggta 20

<210> 302

<211> 20

<212> DNA

<213> Artificial Sequence

<220><223> ARTIFICIAL DNA PRIMER

<400> 302

caggtctgtt acataaagca 20

<210> 303

<211> 21

<212> DNA
 <213> Artificial Sequence
 <220><223> ARTIFICIAL DNA PRIMER
 <400> 303
 gtttacgctc aaatctcctc c 21
 <210> 304
 <211> 24
 <212> DNA
 <213> Artificial Sequence
 <220><223> ARTIFICIAL DNA PRIMER
 <400> 304
 ggtacatgaa gcaggctttg aagg 24
 <210> 305
 <211> 24
 <212> DNA
 <213> Artificial Sequence
 <220><223> ARTIFICIAL DNA PRIMER
 <400> 305
 gattgtgtcc gtcagccttt gctc 24

 <210> 306
 <211> 21
 <212> DNA
 <213> Artificial Sequence
 <220><223> ARTIFICIAL DNA PRIMER
 <400> 306
 taccatattt tcagaggatc a 21
 <210> 307
 <211> 21
 <212> DNA
 <213> Artificial Sequence
 <220><223> ARTIFICIAL DNA PRIMER
 <400> 307
 aggatgttct tgcctgcaag t 21
 <210> 308

<211> 26
 <212> DNA
 <213> Artificial Sequence
 <220><223> ARTIFICIAL DNA PRIMER
 <400> 308
 gatgatatag ttgatgcttt ccaaag 26

<210> 309
 <211> 33
 <212> DNA
 <213> Artificial Sequence
 <220><223> ARTIFICIAL DNA PRIMER
 <400> 309
 gggccccttc atttacgaaa taaagtgcg cg 33

<210> 310
 <211> 40
 <212> DNA
 <213> Artificial Sequence
 <220><223> ARTIFICIAL DNA PRIMER
 <400> 310
 gcggccgcaa ataaatttaa aataaacgat atcaaaattc 40

<210> 311
 <211> 30
 <212> DNA
 <213> Artificial Sequence
 <220><223> ARTIFICIAL DNA PRIMER
 <400> 311
 cgacgcaaaa gaagggtca gccgaaaaag 30

<210> 312
 <211> 32
 <212> DNA
 <213> Artificial Sequence
 <220><223> ARTIFICIAL DNA PRIMER
 <400> 312

ccatttcttt ttcggtgag cccttcttg gc 32

<210> 313

<211> 32

<212> DNA

<213> Artificial Sequence

<220><223> ARTIFICIAL DNA PRIMER

<400> 313

gcggccgcaa taacctcagg gagaactttg gc 32

<210> 314

<211> 30

<212> DNA

<213> Artificial Sequence

<220><223> ARTIFICIAL DNA PRIMER

<400> 314

gagctcccaa acaaggtctc agtaggatcg 30

<210> 315

<211> 28

<212> DNA

<213> Artificial Sequence

<220><223> ARTIFICIAL DNA PRIMER

<400> 315

cgccataagg agaggctgta gatttgtc 28

<210> 316

<211> 25

<212> DNA

<213> Artificial Sequence

<220><223> ARTIFICIAL DNA PRIMER

<400> 316

ccaggacatc ttgcttgcaa tgtcg 25

<210> 317

<211> 34

<212> DNA

<213> Artificial Sequence

<220><223> ARTIFICIAL DNA PRIMER

<400> 317

gttccatcgg gccctaaag gtctctcacg acag

34

<210> 318

<211> 33

<212> DNA

<213> Artificial Sequence

<220><223> ARTIFICIAL DNA PRIMER

<400> 318

caaacacatc gcgataaaat gtccagaggc ttc

33

<210> 319

<211> 22

<212> DNA

<213> Artificial Sequence

<220><223> ARTIFICIAL DNA PRIMER

<400> 319

gcaagacctt ggatctgaag gg

22

<210> 320

<211> 25

<212> DNA

<213> Artificial Sequence

<220><223> ARTIFICIAL DNA PRIMER

<400> 320

cgaaccaatt caagaaaacc aacag

25

<210> 321

<211> 28

<212> DNA

<213> Artificial Sequence

<220><223> ARTIFICIAL DNA PRIMER

<400> 321

gggcccgtcc cttggcgacg ccctgatc

28

<210> 322

<211> 33

<212> DNA

<213> Artificial Sequence

<220><223> ARTIFICIAL DNA PRIMER

<400> 322

gcggccgcta tttttgtgtt ttgctgtgtt ttg 33

<210> 323

<211> 37

<212> DNA

<213> Artificial Sequence

<220><223> ARTIFICIAL DNA PRIMER

<400> 323

gcggccgccca tctgaatgta aaatgaacat taaaatg 37

<210> 324

<211> 29

<212> DNA

<213> Artificial Sequence

<220><223> ARTIFICIAL DNA PRIMER

<400> 324

gagctccccc agttgtgtgtt gcaattaac 29

<210> 325

<211> 22

<212> DNA

<213> Artificial Sequence

<220><223> ARTIFICIAL DNA PRIMER

<400> 325

gaagagacgt acaagatccg cc 22

<210> 326

<211> 22

<212> DNA

<213> Artificial Sequence

<220><223> ARTIFICIAL DNA PRIMER

<400> 326

taggaatggt gcatcatcca ac 22

<210> 327

<211> 40
 <212> DNA
 <213> Artificial Sequence
 <220><223> ARTIFICIAL DNA PRIMER
 <400> 327
 ttcttatctg aaaactccga gttcgcaaag aaggttgaag 40
 <210> 328
 <211> 45
 <212> DNA
 <213> Artificial Sequence
 <220><223> ARTIFICIAL DNA PRIMER
 <400> 328
 ttattgaaat taatccaagg attcaagtg aacatacaat tactg 45
 <210> 329
 <211> 42
 <212> DNA
 <213> Artificial Sequence
 <220><223> ARTIFICIAL DNA PRIMER
 <400> 329
 tttccaaaa agttcgtgag ttcgatggtt gtatgattat gg 42
 <210> 330
 <211> 22
 <212> DNA
 <213> Artificial Sequence
 <220><223> ARTIFICIAL DNA PRIMER
 <400> 330
 cctctaccac caccacaaa tg 22
 <210> 331
 <211> 24
 <212> DNA
 <213> Artificial Sequence
 <220><223> ARTIFICIAL DNA PRIMER
 <400> 331
 gggaagaac taagaagaag tatg 24

<210> 332
 <211> 44
 <212> DNA
 <213> Artificial Sequence
 <220><223> ARTIFICIAL DNA PRIMER
 <400> 332
 caatccttct agagaggggg ttatatgtgt aaatatagag ttg 44

<210> 333
 <211> 36
 <212> DNA
 <213> Artificial Sequence
 <220><223> ARTIFICIAL DNA PRIMER
 <400> 333
 ttctaccctc gagattgggtt cttccccct ctcaag 36

<210> 334
 <211> 44
 <212> DNA
 <213> Artificial Sequence
 <220><223> ARTIFICIAL DNA PRIMER
 <400> 334
 gtggtcaatc tagaaaatat gactgacaaa atctccctag gtac 44

<210> 335
 <211> 30
 <212> DNA
 <213> Artificial Sequence
 <220><223> ARTIFICIAL DNA PRIMER
 <400> 335
 caattttgga gctgattcca aatcgtaaac 30

<210> 336
 <211> 31
 <212> DNA
 <213> Artificial Sequence
 <220><223> ARTIFICIAL DNA PRIMER

| | |
|---|----|
| <400> 336 | |
| caagagtatc ccattctgaca ggaaccgatg g | 31 |
| <210> 337 | |
| <211> 27 | |
| <212> DNA | |
| <213> Artificial Sequence | |
| <220><223> ARTIFICIAL DNA PRIMER | |
| <400> 337 | |
| gctggagaat agatcttcaa cgccccg | 27 |
| <210> 338 | |
| <211> 29 | |
| <212> DNA | |
| <213> Artificial Sequence | |
| <220><223> ARTIFICIAL DNA PRIMER | |
| <400> 338 | |
| cggagaaggc gtataaaaag gacacggag | 29 |
| <210> 339 | |
| <211> 47 | |
| <212> DNA | |
| <213> Artificial Sequence | |
| <220><223> ARTIFICIAL DNA PRIMER | |
| <400> 339 | |
| ggataaaagt attacatacg tacaggattg tgtattagtg tatttcg | 47 |
| <210> 340 | |
| <211> 40 | |
| <212> DNA | |
| <213> Artificial Sequence | |
| <220><223> ARTIFICIAL DNA PRIMER | |
| <400> 340 | |
| cctccagtgt ttttctctct gtctctttgt ttttttttc | 40 |
| <210> 341 | |
| <211> 46 | |
| <212> DNA | |
| <213> Artificial Sequence | |

<220><223> ARTIFICIAL DNA PRIMER

<400> 341

ggttaattaa tttatttgta cataaaaacc acataaatgt aaaagc 46

<210> 342

<211> 35

<212> DNA

<213> Artificial Sequence

<220><223> ARTIFICIAL DNA PRIMER

<400> 342

gaattccttt aattcggaga aaatctgac aagag 35

<210> 343

<211> 945

<212> DNA

<213> Metallosphaera sedula

<400> 343

atgaccgaaa aggtctccgt cgtcgggtgca ggcgtcatcg gtgtcggttg ggctaccttg 60

ttcgttcca agggttactc cgtctccttg tacaccgaaa agaaggaaac cttggacaag 120

ggatcgtgaaa agttgcgtaa ctacgtccaa gtcatgaaga acaactccca aatcaccgaa 180

gacgtcaaca ccgtcatctc cagagtttcc cctactacca acttggacga agccgttaga 240

ggtgcaaact tcgtcatcga ggctgtcatc gaagactacg acgctaagaa gaagatcttc 300

ggttacttgg actccgtctt ggacaaggaa gttatcttgg catcctccac ctccggtttg 360

ttgatcaccg aagtccaaaa ggctatgtcc aagcaccag aaagagctgt catgcccac 420

ccatggaacc caccacactt gttgccttgg gtcgaaatcg ttccaggtga gaagacctcc 480

atggaagtcg ttgagagaac caagtccttg atggaaaagt tggacagaat cgtcgtcgtt 540

ttgaagaagg aaatcccagg ttctcatcgg aacagattgg cattcgcttt gttcagagaa 600

gctgtctact tggttgacga ggggtgtcga accgtcgagg acatcgacaa gggttatgact 660

gccgctatcg gcttgagatg ggccttcatg ggtccattct tgacctacca cttgggtggt 720

ggtgagggtg gtttgaata cttcttcaac agaggtttcg gttacggtgc caacgaatgg 780

atgcacacct tggctaagta cgacaagttc ccatacaccg gtgtcaccaa ggccatccag 840

caaatgaagg aatactcctt catcaagggt aagaccttcc aagagatctc caagtggaga 900

gacgaaaagt tgttgaagg ctacaagttg gtctgggaaa agtaa 945

<210> 344

<211> 20
 <212> DNA
 <213> Artificial Sequence
 <220><223> ARTIFICIAL DNA PRIMER
 <400> 344
 aggttacctt agtacgaagg 20

<210> 345
 <211> 20
 <212> DNA
 <213> Artificial Sequence
 <220><223> ARTIFICIAL DNA PRIMER
 <400> 345
 ctattcttac gatgaaggcg 20

<210> 346
 <211> 31
 <212> DNA
 <213> Artificial Sequence
 <220><223> ARTIFICIAL DNA PRIMER
 <400> 346
 ccttaattaa ttatccacgg aagatatgat g 31

<210> 347
 <211> 32
 <212> DNA
 <213> Artificial Sequence
 <220><223> ARTIFICIAL DNA PRIMER
 <400> 347
 gctctagata aaatgtccca aggtagaaaa gc 32

<210> 348
 <211> 30
 <212> DNA
 <213> Artificial Sequence
 <220><223> ARTIFICIAL DNA PRIMER
 <400> 348

| | |
|---|----|
| gctagctaaa atgtttggta atatttccca | 30 |
| <210> 349 | |
| <211> 28 | |
| <212> DNA | |
| <213> Artificial Sequence | |
| <220><223> ARTIFICIAL DNA PRIMER | |
| <400> 349 | |
| ttaattaact atttatctaa tgatcctc | 28 |
| <210> 350 | |
| <211> 41 | |
| <212> DNA | |
| <213> Artificial Sequence | |
| <220><223> ARTIFICIAL DNA PRIMER | |
| <400> 350 | |
| tctagataaa atgtctatta gtgaaaaata ttttcctcaa g | 41 |
| <210> 351 | |
| <211> 43 | |
| <212> DNA | |
| <213> Artificial Sequence | |
| <220><223> ARTIFICIAL DNA PRIMER | |
| <400> 351 | |
| ttaattaact tttaaatttt ggaaaaagct tgatcaataa tgg | 43 |