



(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2014-0092759
(43) 공개일자 2014년07월24일

(51) 국제특허분류(Int. Cl.)
C12N 15/61 (2006.01) C12N 15/53 (2006.01)
C12N 15/63 (2006.01) C12P 7/16 (2006.01)
C12P 7/52 (2006.01)
(21) 출원번호 10-2013-7027906
(22) 출원일자(국제) 2012년03월23일
심사청구일자 없음
(85) 번역문제출일자 2013년10월23일
(86) 국제출원번호 PCT/US2012/030479
(87) 국제공개번호 WO 2012/129555
국제공개일자 2012년09월27일
(30) 우선권주장
61/467,249 2011년03월24일 미국(US)
(뒷면에 계속)

(71) 출원인
부타맥스 어드밴스드 바이오퓨얼스 엘엘씨
미국 19880 델라웨어주 윌밍톤 빌딩 356 듀폰 엑스페리멘탈 스테이션 루트 141 & 헨리 클레이
(72) 발명자
엔토니, 래리, 카메론
미국 19014 펜실베이니아주 아스톤 코리 레인 501
김슨, 캐서린, 제이.
미국 19809 델라웨어주 윌밍톤 메이플 애비뉴 302
(뒷면에 계속)
(74) 대리인
김영, 양영준, 양영환

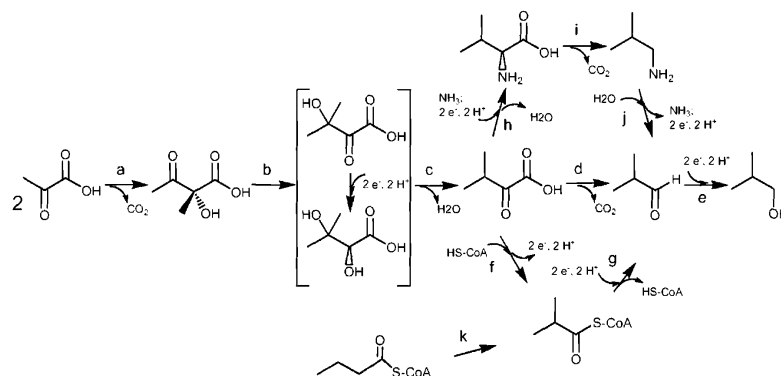
전체 청구항 수 : 총 74 항

(54) 발명의 명칭 숙주 세포 및 아이소부탄올의 제조 방법

(57) 요약

제조합 효모 숙주 세포 및 아이소부탄올의 제조에 있어서의 이의 사용 방법이 본 발명에서 제공된다. 제공된 효모 숙주 세포는 아이소부탄올 생합성 경로와, 감소되거나 또는 제거된 알데히드 탈하이드로게나아제 활성, 감소되거나 또는 제거된 아세토라테이트 리덕타아제 활성 중 적어도 하나; 또는 케톤-산 리덕토아이스모머라아제 활성을 갖는 폴리펩티드를 코딩하는 이중성 폴리뉴클레오티드를 포함한다.

대표도 - 도1



(72) 발명자

해, 홍시양

미국 19803 델라웨어주 월밍톤 턱어웨이 드라이브 120

후양, 리사, 엘.

미국 19707 델라웨어주 호케신 올드 월밍톤 로드 1423

오키프, 다니엘, 피.

미국 19078-2731 펜실베이니아주 리들리 파크 쇼우 로드 328

크릭케버그, 아서 레오

미국 19806 델라웨어주 월밍톤 펜실베이니아 애비뉴 1404

리, 유젠

미국 08534 뉴저지주 페닝톤 켄츠데일 드라이브 4

마지오-홀, 로리, 앤

미국 19810 델라웨어주 월밍톤 노쓰 록필드 드라이브 3413

맥엘바인, 제시카

미국 19807 델라웨어주 월밍톤 마운트 에어리 드라이브 7

넬슨, 마크, 제이.

미국 19711 델라웨어주 뉴어크 피스크 레인 707

파트나익, 란잔

미국 19702 델라웨어주 뉴어크 도날드 프레스톤 드라이브 5

로스만, 스티븐 캐리

미국 19810 델라웨어주 월밍톤 룽우드 드라이브 2621

(30) 우선권주장

61/467,261 2011년03월24일 미국(US)

61/467,271 2011년03월24일 미국(US)

61/472,474 2011년04월06일 미국(US)

61/472,484 2011년04월06일 미국(US)

61/472,487 2011년04월06일 미국(US)

61/472,497 2011년04월06일 미국(US)

61/570,513 2011년12월14일 미국(US)

특허청구의 범위

청구항 1

엔지니어링된(engineered) 아이소부탄을 생성 경로와,

a. i. 1. 비피도박테리움 안굴라툼(*Bifidobacterium angulatum*), 비피도박테리움 덴티움(*Bifidobacterium dentium*), 자이모모나스 모빌리스(*Zymomonas mobilis*), 클로스트리듐 베이제린키이(*Clostridium beijerinckii*) 또는 엔에어로스티페스 카카에(*Anaerostipes caccae*)로부터 유래된 케톨-산 리덕토아이스머라아제(ketol-acid reductoisomerase; KARI) 효소에 대하여 동일성이 약 90% 이상인 폴리펩티드 또는 이의 활성 단편;

2. 서열 번호 27, 29, 31, 33, 35, 37, 39, 41, 43, 45, 47, 49, 51, 53, 55, 57, 59, 61, 63 또는 65의 서열에 대하여 동일성이 약 90% 이상이거나 또는 동일성이 약 95% 이상인 폴리펩티드로 이루어진 군으로부터 선택되는, KARI 활성을 갖는 이중성 폴리펩티드; 또는

ii. a)의 KARI 활성을 갖는 이중성 폴리펩티드를 코딩하는 이중성 폴리뉴클레오티드 중 적어도 하나와;

b. 엔지니어링된 아이소부탄을 생성 경로의 성능을 향상시키는 적어도 하나의 숙주 세포 변형을 포함하는 재조합 숙주 세포.

청구항 2

제1항에 있어서, a)와 b)의 조합은 아이소부탄을 생성 경로의 성능의 상승적 증가로 이어지는 재조합 숙주 세포.

청구항 3

제1항 또는 제2항에 있어서, 숙주 세포 변형은 감소되거나 또는 제거된 알데히드 데하이드로게나아제 발현 또는 활성 및 감소되거나 또는 제거된 아세토락테이트 리덕타아제 발현 또는 활성을 포함하는 재조합 숙주 세포.

청구항 4

제3항에 있어서, 감소되거나 또는 제거된 알데히드 데하이드로게나아제 발현 또는 활성 및 감소되거나 또는 제거된 아세토락테이트 리덕타아제 발현 또는 활성을 포함하며, KARI 활성을 갖는 이중성 폴리펩티드는 NADH에 대한 K_M 이 300 μM 미만인 재조합 숙주 세포.

청구항 5

제1항에 있어서, KARI 활성을 갖는 이중성 폴리펩티드는 서열 번호 27, 29, 141, 143, 275 또는 277의 서열에 대한 동일성이 약 90% 이상이거나 또는 약 95% 이상인 재조합 숙주 세포.

청구항 6

제1항 내지 제5항 중 어느 한 항에 있어서, KARI 활성을 갖는 이중성 폴리펩티드는 서열 번호 27의 S56 및 S58에 상응하는 아미노산의 치환을 포함하는 재조합 숙주 세포.

청구항 7

제6항에 있어서, KARI 활성을 갖는 이중성 폴리펩티드는 서열 번호 27의 I86, N87, T131 또는 T191에 상응하는 아미노산들 중 하나 이상의 치환을 추가로 포함하는 재조합 숙주 세포.

청구항 8

제1항 내지 제7항 중 어느 한 항에 있어서, 약 48시간 후 세포 1 g당 약 3 g 이상, 약 4 g 이상 또는 약 5 g 이상의 유효 아이소부탄을 생산성을 가지며, 상기 48시간 중 적어도 마지막의 약 24시간은 혐기성 조건 하에 있는 재조합 숙주 세포.

청구항 9

제1항 내지 제8항 중 어느 한 항에 있어서, KARI 활성을 갖는 이중성 폴리펩티드는 pH 6.8에서 NADH에 대한 K_M 이 약 350 μM 미만, 약 100 μM 미만, 약 50 μM 미만 또는 약 10 μM 미만인 재조합 숙주 세포.

청구항 10

제1항에 있어서, KARI 활성을 갖는 이중성 폴리펩티드는 서열 번호 27의 아미노산 A41, S56, S58, I87, T131, T191, R227 또는 Q246에 상응하는 위치들 중 하나 이상에서 아미노산 치환을 포함하는 재조합 숙주 세포.

청구항 11

제1항 내지 제10항 중 어느 한 항에 있어서, 상기 숙주 세포 변형은 알데히드 데하이드로게나아제 활성을 갖는 폴리펩티드를 코딩하는 내인성 폴리뉴클레오티드 내에 결실, 돌연변이 또는 치환 중 적어도 하나를 포함하는 재조합 숙주 세포.

청구항 12

제11항에 있어서, 알데히드 데하이드로게나아제 활성을 갖는 상기 폴리펩티드는 효소 위원회 번호(Enzyme Commission Number) EC 1.21.3, EC 1.2.1.4, EC1.2.1.5 또는 이의 조합에 상응하는 재조합 숙주 세포.

청구항 13

제11항 또는 제12항에 있어서, 상기 숙주 세포는 에스. 세레비지애(*S. cerevisiae*)이며, 알데히드 데하이드로게나아제 활성을 갖는 상기 폴리펩티드는 ALD2, ALD3, ALD4, ALD5, ALD6 또는 이의 상동체인 재조합 숙주 세포.

청구항 14

제11항 또는 제12항에 있어서, 상기 숙주 세포는 케이. 락티스(*K. lactis*)이며, 알데히드 데하이드로게나아제 활성을 갖는 상기 폴리펩티드는 KLLA0F00440, KLLA0E23057, KLLA0D10021 또는 KLLA0D09999G인 재조합 숙주 세포.

청구항 15

제11항 또는 제12항에 있어서, 상기 숙주 세포는 피. 스티피티스(*P. stipitis*)이며, 알데히드 데하이드로게나아제 활성을 갖는 상기 폴리펩티드는 ALD2, ALD3, ALD4, ALD5 또는 ALD7인 재조합 숙주 세포.

청구항 16

제11항 또는 제12항에 있어서, 상기 숙주 세포는 락토바실러스 플란타룸(*Lactobacillus plantarum*)이며, 알데히드 데하이드로게나아제 활성을 갖는 상기 폴리펩티드는 AldH인 재조합 숙주 세포.

청구항 17

제11항 또는 제12항에 있어서, 상기 숙주 세포는 이. 콜라이(*E. coli*)이며, 알데히드 데하이드로게나아제 활성을 갖는 상기 폴리펩티드는 aldA, aldB 또는 aldH인 재조합 숙주 세포.

청구항 18

제1항 내지 제17항 중 어느 한 항에 있어서, 알데히드 옥시다아제 활성을 갖는 폴리펩티드를 코딩하는 내인성 폴리뉴클레오티드 내에 결실, 돌연변이 또는 치환 중 적어도 하나 이상을 포함하는 재조합 숙주 세포.

청구항 19

제18항에 있어서, 알데히드 옥시다아제 활성을 갖는 상기 폴리펩티드는 효소 위원회 번호 EC 1.2.3.1에 상응하는 재조합 숙주 세포.

청구항 20

제19항에 있어서, 알데히드 옥시다아제 활성을 갖는 상기 폴리펩티드는 AOX1, AOX2 또는 이의 조합인 재조합 숙주 세포.

청구항 21

제1항 내지 제20항 중 어느 한 항에 있어서, 아세토락테이트 리덕타아제 활성을 갖는 폴리펩티드를 코딩하는 내인성 폴리뉴클레오티드 내에 적어도 하나 이상의 결실, 돌연변이 또는 치환을 포함하는 재조합 숙주 세포.

청구항 22

제21항에 있어서, 내인성 폴리뉴클레오티드는 서열 번호 676, 서열 번호 678, 서열 번호 680, 서열 번호 682, 서열 번호 684, 서열 번호 686, 서열 번호 688, 서열 번호 690, 서열 번호 692, 서열 번호 694, 서열 번호 696, 서열 번호 702, 서열 번호 704, 서열 번호 706, 서열 번호 708, 서열 번호 710, 서열 번호 712, 서열 번호 714, 서열 번호 716, 서열 번호 718, 서열 번호 720, 서열 번호 722, 서열 번호 724, 서열 번호 726, 서열 번호 728 및 서열 번호 730의 서열로 이루어진 군으로부터 선택되는 재조합 숙주 세포.

청구항 23

제21항에 있어서, 아세토락테이트 리덕타아제 활성을 갖는 폴리펩티드는 YMR226C인 재조합 숙주 세포.

청구항 24

제21항 내지 제23항 중 어느 한 항에 있어서, 효모 숙주 세포인 재조합 숙주 세포.

청구항 25

제24항에 있어서, 효모는 사카로마이세스(*Saccharomyces*), 쉬조사카로마이세스(*Schizosaccharomyces*), 한세놀라(*Hansenula*), 칸디다(*Candida*), 클루이베로마이세스(*Kluyveromyces*), 야로위아(*Yarrowia*), 이사첸키아(*Issatchenkia*) 또는 피키아(*Pichia*)인 재조합 숙주 세포.

청구항 26

제1항 또는 제25항에 있어서, 사카로마이세스 세레비지애인 재조합 숙주 세포.

청구항 27

제1항 내지 제26항 중 어느 한 항에 있어서, 엔지니어링된 아이소부탄올 생성 경로는 하기의 기질의 생성물로의 전환을 포함하며, 각각의 기질의 생성물로의 전환은 숙주 세포에 의해 재조합적으로 발현되는 효소에 의해 촉매되는 재조합 숙주 세포:

- a. 피루베이트의 아세토락테이트로의 전환;
- b. 아세토락테이트의 2,3-다이하이드록시아이소발레레이트로의 전환;
- c. 2,3-다이하이드록시아이소발레레이트의 2-케토아이소발레레이트로의 전환;
- d. 2-케토아이소발레레이트의 아이소부티르알데히드로의 전환; 및
- e. 아이소부티르알데히드의 아이소부탄올로의 전환.

청구항 28

제27항에 있어서, 기질의 생성물로의 전환들 전부는 숙주 세포에 대하여 이중성인 효소에 의해 촉매되는 재조합 숙주 세포.

청구항 29

제28항에 있어서, 숙주 세포에 대하여 이중성인 효소들 중 적어도 하나를 코딩하는 적어도 하나의 이중성 폴리뉴클레오티드는 숙주 세포에 염색체 통합되는 재조합 숙주 세포.

청구항 30

제27항 내지 제29항 중 어느 한 항에 있어서, 기질의 생성물로의 전환은 사실상 세포질에 국소화된 효소에 의해 촉매되는 재조합 숙주 세포.

청구항 31

제27항 내지 제30항 중 어느 한 항에 있어서, 아이소부티르알데히드의 아이소부탄올로의 전환에 있어서의 기질의 생성물로의 전환은 NADH를 보조 인자로 이용하는 알코올 데하이드로게나아제 효소에 의해 촉매되는 제조합 숙주 세포.

청구항 32

제1항 내지 제31항 중 어느 한 항에 있어서, 아세토락테이트의 2,3-다이하이드록시-2-메틸부티레이트로의 전환 능력의 감소, 파괴 또는 제거를 추가로 포함하는 제조합 숙주 세포.

청구항 33

제1항 내지 제32항 중 어느 한 항에 있어서, 효모이며, 감소되거나, 파괴되거나 또는 제거된 피루베이트 데카르복실라아제 발현 또는 활성을 갖는 제조합 숙주 세포.

청구항 34

제33항에 있어서, 감소되거나, 파괴되거나 또는 제거된 PDC1, PDC5 또는 PDC6 활성 또는 이의 조합을 갖는 제조합 숙주 세포.

청구항 35

제1항 내지 제34항 중 어느 한 항에 있어서, 감소되거나 또는 제거된 NAD-의존성 글리세롤-3-포스페이트 데하이드로게나아제 발현 또는 활성을 갖는 제조합 숙주 세포.

청구항 36

제35항에 있어서, 감소된 GPD2 발현 또는 활성을 갖는 제조합 숙주 세포.

청구항 37

제1항 내지 제36항 중 어느 한 항에 있어서, 감소되거나 또는 제거된 FRA2 발현 또는 활성을 갖는 제조합 숙주 세포.

청구항 38

제1항 내지 제37항 중 어느 한 항에 있어서, 숙주 세포는 혐기성 조건 하에서 아이소부탄올을 생성하며, 글리세롤에 대한 아이소부탄올의 몰비는 1 초과인 제조합 숙주 세포.

청구항 39

제1항 내지 제38항 중 어느 한 항에 있어서, KARI 활성을 갖는 폴리펩티드는 표 Z에 제공된 프로파일 HMM을 $<10^{-3}$ 의 프로파일 HMM E 값과 매칭시키는 제조합 숙주 세포.

청구항 40

제1항 내지 제39항 중 어느 한 항에 있어서, 이론적 수율의 약 25% 초과, 약 50% 초과, 약 75% 초과 또는 약 90% 초과 수율로 아이소부탄올을 생성하는 제조합 숙주 세포.

청구항 41

제1항 내지 제40항 중 어느 한 항에 있어서, 아이소부탄올이 생성되며, 에탄올이 또한 생성되는 제조합 숙주 세포.

청구항 42

아이소부탄올을 제조하는 방법으로서,

a. 제1항 내지 제41항 중 어느 한 항의 제조합 숙주 세포를 제공하는 단계, 및

b. a)의 숙주 세포를 아이소부탄올을 생성하는 조건 하에서 탄소 기질과 접촉시키는 단계를 포함하는 방법.

청구항 43

제42항에 있어서, 접촉의 적어도 일부분은 혐기성 조건 하에서 일어나는 방법.

청구항 44

제42항 또는 제43항에 있어서, 접촉은 추출제의 존재 하에 일어나는 방법.

청구항 45

제44항에 있어서, 접촉은 수성 상 및 유기 상을 포함하는 2상(two-phase) 시스템을 형성하기에 충분한 양의 유기 추출제의 존재 하에 일어나는 방법.

청구항 46

제42항 내지 제45항 중 어느 한 항에 있어서, 아이소부탄올의 유효 비율, 유효 역가 또는 유효 수율 중 하나 이상은 (i) KARI 활성을 갖는 이중성 폴리펩티드 또는 KARI 활성을 갖는 폴리펩티드를 코딩하는 이중성 폴리뉴클레오티드 중 적어도 하나 및 (ii) 엔지니어링된 아이소부탄올 생성 경로의 성능을 향상시키는 적어도 하나의 변형을 포함하지 않는 재조합 숙주 세포와 비교하여 증가되는 방법.

청구항 47

제42항 내지 제46항 중 어느 한 항에 있어서, 2,3-다이하이드록시-2-메틸 부티레이트(DHMB) 생성, 아이소부티르산 생성 또는 이들 둘 모두는 (i) KARI 활성을 갖는 이중성 폴리펩티드 또는 KARI 활성을 갖는 폴리펩티드를 코딩하는 이중성 폴리뉴클레오티드 중 적어도 하나 및 (ii) 엔지니어링된 아이소부탄올 생성 경로의 성능을 향상시키는 적어도 하나의 변형을 포함하지 않는 재조합 숙주 세포와 비교하여 감소되는 방법.

청구항 48

제42항 내지 제47항 중 어느 한 항에 있어서, 글리세롤에 대한 아이소부탄올의 몰비는 1 초과인 방법.

청구항 49

숙주 세포와 제42항 내지 제28항 중 어느 한 항에 방법에 따라 제조된 아이소부탄올을 포함하는 발효 조성물.

청구항 50

아이소부탄올을 제조하는 방법으로서,

a. 아이소부탄올을 생성하는 재조합 숙주 세포를 제공하는 단계와,

b. a)의 숙주 세포를 아이소부탄올을 생성하는 조건 하에서 탄소 기질과 접촉시키는 단계를 포함하며;

접촉의 적어도 일부분은 혐기성 조건 하에서 일어나고;

생성된 글리세롤에 대한 아이소부탄올의 비는 1 초과인 방법.

청구항 51

부탄올을 제조하는 방법으로서,

a. 부탄올을 생성하는 조건 하에서 피루베이트를 아세트락테이트로 전환시킬 수 있는 생합성 경로를 포함하는 재조합 효모를 성장시키는 단계; 및

b. DHMB를 배양물로부터 제거하는 단계를 포함하는 방법.

청구항 52

부탄올 및 약 0.5 mM 이하의 DHMB를 포함하는, 제51항의 방법에 의해 제조된 조성물.

청구항 53

i) 부탄올을 생성할 수 있는 재조합 효모, ii) 부탄올, 및 iii) 약 0.5 mM 이하의 DHMB를 포함하는 조성물.

청구항 54

재조합 숙주 세포를 제조하는 방법으로서,

- a. 알데히드 탈하이드로게나아제 활성 또는 알데히드 옥시다아제 활성을 갖는 폴리펩티드를 코딩하는 폴리뉴클레오티드 내에 변형을 포함하는 재조합 숙주 세포를 제공하는 단계; 및
- b. 아이소부탄올 생합성 경로의 폴리펩티드를 코딩하는 폴리뉴클레오티드로 숙주 세포를 형질전환시키는 단계를 포함하는 방법.

청구항 55

아이소부티르알데히드의 아이소부티르산으로의 전환을 감소시키거나 또는 제거하는 방법으로서,

- a. 제1항 내지 제41항 중 어느 한 항의 재조합 숙주 세포를 제공하는 단계; 및
- b. 감소되거나 또는 제거된 알데히드 탈하이드로게나아제 및/또는 알데히드 옥시다아제 활성을 포함하지 않는 재조합 숙주 세포와 비교하여 아이소부티르알데히드의 아이소부티르산으로의 전환을 감소시키거나 또는 제거하는 조건을 숙주 세포에 가하는 단계를 포함하는 방법.

청구항 56

아이소부탄올과 제1항 내지 제41항 중 어느 한 항의 재조합 숙주 세포를 포함하는 조성물.

청구항 57

케툴-산 리덕토아이스머라아제 활성을 갖는, 서열 번호 29, 31, 33, 35, 37, 39, 41, 43, 45, 47, 49, 51, 53, 55, 57, 59, 61, 63, 65, 417, 419, 421, 423, 425, 427, 429, 431, 433, 434, 435, 436, 437, 438, 439, 440, 441, 442, 443, 444, 445, 446, 447, 448, 449, 450, 451, 452, 453, 454, 455, 456, 457, 458, 459, 460, 461, 462, 463, 464, 465, 466, 467, 468, 469, 470, 471, 472, 473, 474, 475, 476, 477, 478, 479, 480, 481, 482, 483, 484, 485, 486, 487, 488, 489, 490, 491, 492, 493, 494, 495, 496, 497, 498, 499, 500, 501, 502, 503, 504, 505, 506, 507, 508, 509, 510, 511, 512, 513, 514, 515, 516, 517, 518, 519, 520, 521, 522, 523, 524, 525, 526, 527, 528, 529, 530, 531, 532, 533, 534, 535, 536, 537, 624, 626, 628, 630, 632의 서열에 대하여 약 90% 이상의 동일성 또는 약 95% 이상의 동일성 또는 약 99% 이상의 동일성을 포함하는 폴리펩티드 또는 이의 활성 단편.

청구항 58

케툴-산 리덕토아이스머라아제 활성을 갖는, 서열 번호 417, 419, 421, 423, 425 또는 427의 서열에 대하여 약 90% 이상의 동일성 또는 약 95% 이상의 동일성 또는 약 99% 이상의 동일성을 포함하는 폴리펩티드 또는 이의 활성 단편.

청구항 59

제57항 또는 제58항의 폴리펩티드를 코딩하는 폴리뉴클레오티드.

청구항 60

아세트락테이트를 2,3-다이하이드록시아이소발레레이트로 전환시키는 방법으로서,

- a. 제57항 또는 제58항의 폴리펩티드를 제공하는 단계와;
- b. 2,3-다이하이드록시아이소발레레이트를 생성하는 조건 하에서 a)의 폴리펩티드를 아세트락테이트와 접촉시키는 단계를 포함하는 방법.

청구항 61

제57항 또는 제58항의 이중성 폴리펩티드 또는 제59항의 폴리뉴클레오티드를 포함하는 재조합 숙주 세포.

청구항 62

소비되는 당 1 몰당 0.01 몰 미만의 2,3-다이하이드록시-2-메틸 부티레이트 (DHMB)를 생성하는, 피루베이트를 아세토락테이트로 전환시킬 수 있는 생합성 경로를 포함하는 재조합 효모.

청구항 63

약 1.0 mM/시간 미만의 속도로 DHMB를 생성하는, 피루베이트를 아세토락테이트로 전환시킬 수 있는 생합성 경로를 포함하는 재조합 효모.

청구항 64

DHMB 생성량의 약 1.5배 이상인 양의 2,3-다이하이드록시-3-아이소발레레이트 (DHIV)를 생성하는, 피루베이트를 아세토락테이트로 전환시킬 수 있는 생합성 경로를 포함하는 재조합 효모.

청구항 65

부탄올을 제조하는 방법으로서,

a) 부탄올을 생성하는 조건 하에서 피루베이트를 아세토락테이트로 전환시킬 수 있는 생합성 경로를 포함하는 재조합 효모를 성장시키는 단계; 및

b) DHIV 및/또는 DHMB의 농도를 측정하는 단계를 포함하며;

단계 a) 및 b)를 동시에 또는 순차적으로 그리고 임의의 순서로 수행할 수 있는 방법.

청구항 66

제65항에 있어서, 측정은 액체 크로마토그래피-질량 분광법을 포함하는 방법.

청구항 67

케톨-산 리덕토아이스머라아제 (KARI) 활성을 증가시키는 방법으로서, a) 아세토락테이트, KARI 효소 및 아세토락테이트 리덕타아제 효소를 포함하는 조성물을 제공하는 단계, 및 b) DHMB의 수준을 감소시키는 단계를 포함하는 방법.

청구항 68

제67항에 있어서, DHMB 수준을 감소시키는 상기 단계는 아세토락테이트 리덕타아제 효소 활성을 감소시킴으로써 성취되는 방법.

청구항 69

제67항에 있어서, DHMB 수준을 감소시키는 단계는 상기 조성물로부터 DHMB를 제거함으로써 성취되는 방법.

청구항 70

다이하이드록시산 데하이드라타아제 (DHAD) 활성을 증가시키는 방법으로서, a) 다이하이드록시아이소발레레이트 (DHIV) 및 DHAD 효소를 포함하는 조성물을 제공하는 단계, 및 b) DHMB 수준을 감소시키는 단계를 포함하는 방법.

청구항 71

숙주 세포에서 KARI 효소의 생산성을 증가시키는 방법으로서, 숙주 세포는 이중성 KARI 효소와, 아세토락테이트 리덕타아제 발현 또는 활성을 감소시키거나, 파괴하거나 또는 제거하는 적어도 하나의 유전자 변형을 포함하는 엔지니어링된 아이소부탄올 생성 경로를 포함하며, KARI 효소 활성은 DHMB의 존재 하에서 감소되는 방법.

청구항 72

제71항에 있어서, KARI는 이. 콜라이 또는 엘. 락티스(*L. lactis*) KARI에 대하여 동일성이 약 90% 이상, 약 95% 이상 또는 약 99% 이상인 방법.

청구항 73

제71항 또는 제72항에 있어서, 감소되거나, 파괴되거나 또는 제거된 아세토락테이트 리덕타아제 발현 또는 활성은 DHMB의 존재를 사실상 감소시키는 방법.

청구항 74

제1항에 있어서, 서열 번호 33 또는 35의 서열을 포함하는, KARI 활성을 갖는 폴리펩티드 또는 이의 활성 단편을 코딩하는 이중성 폴리뉴클레오티드를 포함하는 제조합 숙주 세포.

명세서**기술 분야**

[0001] 미국 정부의 실시권

[0002] 본 발명은 미국 에너지부에 의해 수여된 협정 DE-AR0000006 하의 미국 정부 지원으로 이루어졌다. 미국 정부는 본 발명에 있어서 소정의 권리를 갖는다.

[0003] 본 발명은 아이소부탄올의 발효 생산을 위한 제조합 숙주 세포 및 방법에 관한 것이다.

[0004] 관련-출원과의 상호 참조

[0005] 본 출원은 35 U.S.C. § 119(e) 하에서 미국 특허 출원 제61/472,484 (2011년 4월 6일자로 출원), 미국 특허 출원 제61/467,261호 (2011년 3월 24일자로 출원), 미국 특허 출원 제61/472,487호 (2011년 4월 6일자로 출원), 미국 특허 출원 제61/467,271호 (2011년 3월 24일자로 출원), 미국 특허 출원 제61/570,513호 (2011년 12월 14일자로 출원), 미국 특허 출원 제61/467,249호 (2011년 3월 24일자로 출원), 미국 특허 출원 제61/472,497호 (2011년 4월 6일자로 출원), 및 미국 특허 출원 제61/472,474호 (2011년 4월 6일자로 출원)의 우선권을 주장하는데, 상기 미국 특허 출원 각각은 전체적으로 본 명세서에 참고로 포함된다.

[0006] EFS-웹을 통하여 전자 제출된 서열 목록의 언급

[0007] 전자 제출된 서열 목록 (명칭: 20120323_CL5367USNA_SEQLIST.txt; 크기: 2,003,893 바이트; 생성일: 2012년 3월 23일)의 내용은 전체적으로 본 명세서에 참고로 포함된다.

배경 기술

[0008] 부탄올은 연료 첨가제로서, 플라스틱 산업에서 공급 원료 화학물질로서 및 식품 및 조미료 산업에서 식품 등급의 추출제로서 유용한, 중요한 산업적 화학물질이다. 매해, 4,535,923,700 내지 5,443,108,440 kg(100억 내지 120억 파운드)의 부탄올이 석유화학적 수단으로 생산되며, 이러한 원자재 화학물질에 대한 요구는 미래에는 아마도 증가할 것이다.

[0009] 아이소부탄올의 화학적 합성 방법, 예를 들어 옥소 합성, 일산화탄소의 촉매적 수소화 (문헌[Ullmann's Encyclopedia of Industrial Chemistry, 6th edition, 2003, Wiley-VCH Verlag GmbH and Co., Weinheim, Germany, Vol. 5, pp. 716-719]) 및 메탄올과 n-프로판올의 구에르베트(Guerbet) 축합 (문헌[Carlini et al., J. Molec. Catal. A. Chem. 220:215-220, 2004])이 공지되어 있다. 이들 방법은 석유화학물질 유래의 출발 물질을 사용하며, 일반적으로 비용이 많이 들고, 환경 친화적이지 못하다. 식물 유래된 원료로부터의 아이소부탄올의 생성은 온실 가스 방출을 최소화할 것이며, 본 기술 분야에서 진보를 나타낼 것이다.

[0010] 효모 발효의 부산물로서 아이소부탄올이 생물학적으로 생성된다. 이것은 진균류에 의한 불완전한 아미노산 대사의 결과로서 형성되는 "퓨젤유(fusel oil)"의 성분이다. 아이소부탄올은 L-발린의 이화작용으로부터 특이적으로 생성된다. L-발린의 아민기가 질소 공급원으로서 수확된 후에, 생성된 α--케토산은 소위 에를리히 경로(Ehrlich pathway)의 효소들에 의해 탈카르복실화되고 환원되어 아이소부탄올이 된다 (문헌[Dickinson et al., J. Biol. Chem. 273:25752-25756, 1998]).

[0011] 직접적으로 당류로부터 부탄올을 생합성하는 것에 대한 개선 및 대안은 경제적 실행 가능성을 개선시킬 것이며, 본 기술 분야에서의 진보를 나타낼 것이다.

발명의 내용

[0012] 재조합 효모 숙주 세포 및 아이소부탄올의 제조 방법이 본 발명에서 제공된다.

[0013] 일부 실시 형태에서, 재조합 숙주 세포는 엔지니어링된(engineered) 아이소부탄올 생성 경로와, (a) (i) (1) 비피도박테리움 안굴라툼(*Bifidobacterium angulatum*), 비피도박테리움 덴티움(*Bifidobacterium dentium*), 자이모모나스 모빌리스(*Zymomonas mobilis*), 클로스트리듐 베이제린키이(*Clostridium beijerinckii*) 또는 언에어로스티페스 카카에(*Anaerostipes caccae*)로부터 유래된 케톨-산 리덕토아이스머라아제(ketol-acid reductoisomerase; KARI) 효소에 대하여 동일성이 약 90% 이상인 폴리펩티드 또는 이의 활성 단편; (2) 서열 번호 27, 29, 31, 33, 35, 37, 39, 41, 43, 45, 47, 49, 51, 53, 55, 57, 59, 61, 63 또는 65의 서열에 대하여 동일성이 약 90% 이상이거나 또는 동일성이 약 95% 이상인 폴리펩티드로 이루어진 군으로부터 선택되는, 케톨-산 리덕토아이스머라아제 (KARI) 활성을 갖는 이중성 폴리펩티드; 또는 (ii) (a)의 KARI 활성을 갖는 이중성 폴리펩티드를 코딩하는 이중성 폴리뉴클레오티드 중 적어도 하나와; (b) 엔지니어링된 아이소부탄올 생성 경로의 성능을 향상시키는 적어도 하나의 숙주 세포 변형을 포함한다. 일부 실시 형태에서, (a)와 (b)의 조합의 결과는 아이소부탄올 생성 경로의 성능의 상승적 증가로 이어진다.

[0014] 일부 실시 형태에서, 재조합 숙주 세포는 아이소부탄올 생합성 경로와, (a) (i) 비피도박테리움 안굴라툼, 비피도박테리움 덴티움, 자이모모나스 모빌리스, 클로스트리듐 베이제린키이 또는 언에어로스티페스 카카에로부터 유래된 KARI 효소에 대하여 동일성이 약 90% 이상인 폴리펩티드, 또는 이의 활성 단편, (ii) 서열 번호 27, 29, 31, 33, 35, 37, 39, 41, 43, 45, 47, 49, 51, 53, 55, 57, 59, 61, 63 또는 65에 대하여 동일성이 약 90% 이상이거나 또는 동일성이 약 95% 이상인 폴리펩티드, (iii) 비피도박테리움 안굴라툼, 비피도박테리움 덴티움, 엔테로코커스 갈리나룸(*Enterococcus gallinarum*), 스트렙토코커스 서모필레스(*Streptococcus thermophiles*), 자이모모나스 모빌리스, 클로스트리듐 베이제린키이, 언에어로스티페스 카카에 또는 락토코커스 락티스(*Lactococcus lactis*) 아종 크레모리스(*cremoris*) MG1363으로부터 유래된 KARI 효소에 대하여 동일성이 약 90% 이상이거나 또는 동일성이 약 95% 이상인 폴리펩티드 또는 이의 활성 단편 (여기서, 이 폴리펩티드는 NADH에 대한 K_M 이 약 50 미만임), (iv) 스탕필로코커스 카피티스(*Staphylococcus capitis*) SK14, 스탕필로코커스 에피데르미디스(*Staphylococcus epidermidis*) M23864-#1, 스탕필로코커스 호미니스(*Staphylococcus hominis*) SK119, 스탕필로코커스 아우레우스(*Staphylococcus aureus*) 아종 아우레우스 TCH130, 스탕필로코커스 와르네리(*Staphylococcus warneri*) L37603, 스탕필로코커스 에피데르미디스 #23144, 스탕필로코커스 사프로피티쿠스(*Staphylococcus saprophyticus*) 아종 사프로피티쿠스 ATCC15305, 스탕필로코커스 카르노수스(*Staphylococcus carnosus*) 아종 카르노수스 TM300, 리스테리아 모노사이토게네스(*Listeria monocytogenes*) EGD-e, 리스테리아 그라이이(*Listeria grayi*) DSM 20601, 엔테로코커스 카셀리플라부스(*Enterococcus casseliflavus*) EC30, 엔테로코커스 갈리나룸 EG2, 마크로코커스 카세오라이트쿠스(*Macrocooccus caseolyticus*) JCSC5402, 스트렙토코커스 베스티블라리스(*Streptococcus vestibularis*), 스트렙토코커스 뮤탄스(*Streptococcus mutans*) UA159, 스트렙토코커스 고르도니이(*Streptococcus gordonii*) str, 크가쿠스(*cgakkus*) sybstr. CH1, 스트렙토코커스 수이스(*Streptococcus suis*) 89/1591, 스트렙토코커스 인판타리우스(*Streptococcus infantarius*) 아종 인판타리우스 ATCC BAA-102, 락토코커스 락티스 아종 크레모리스 MG1363, 락토코커스 락티스, 류코노스톡 메센테로이데스(*Leuconostoc mesenteroides*) 아종 메센테로이데스 ATCC8293, 락토바실러스 부크네리(*Lactobacillus buchneri*) ATCC 11577, 스탕필로코커스 헤모라이티쿠스(*Staphylococcus haemolyticus*) JCSC1435, 스탕필로코커스 에피데르미디스 ATCC12228, 스트렙토코커스 뉴모니아(*Streptococcus pneumoniae*) CGSP14, 스트렙토코커스 뉴모니아 TIGR4, 스트렙토코커스 상귀니스(*Streptococcus sanguinis*) SK36, 스트렙토코커스 살리바리우스(*Streptococcus salivarius*) SK126, 스트렙토코커스 서모필루스(*Streptococcus thermophilus*) LMD-9, 스트렙토코커스 뉴모니아 CCRI 1974M2, 락토코커스 락티스 아종 락티스 I11403, 류코노스톡 메센테로이데스 아종 크레모리스 ATCC19254, 류코노스톡 메센테로이데스 아종 크레모리스, 락토바실러스 브레비스(*Lactobacillus brevis*) 아종 그라베센시스(*gravesensis*) ATCC27305 또는 락토코커스 락티스 아종 락티스 NCD02118로부터 유래된 KARI 효소에 대하여 동일성이 약 90% 이상이거나 또는 동일성이 약 95% 이상인 폴리펩티드 또는 이의 활성 단편 (상기 이중성 폴리펩티드는 NADH에 대한 K_M 이 약 50 미만임), (v) 서열 번호 67, 69, 71, 73, 75, 77, 79, 81, 83, 85, 87, 89, 91, 93, 95, 97, 99, 101, 103, 105, 107, 109, 111, 113, 115, 117, 119, 121, 123, 125, 127, 129, 131, 133 또는 135의 서열에 대하여 동일성이 약 90% 이상이거나 또는 동일성이 약 95% 이상인, KARI 활성을 갖는 이중성 폴리펩티드 (여기서, 상기 이중성 폴리펩티드는 NADH에 대한 K_M 이 약 50 미만임)로 이루어진 군으로부터 선택되는, 케톨-산 리덕토아이스머라아제 (KARI) 활성을 갖는 이중성 폴리펩티드; (b) (a)의 KARI 활성을 갖는 이중성 폴리펩티드를 코딩하는 이중성 폴리뉴클레오티드; (c) 감소되거나 또는 제거된 알데히드 탈하이드로게나아제 활성; (d) 감소되거나 또는 제거된 알데히드 옥시다아제 활성; (e) 감소되거나 또는 제거된 아세토락테이

트 리덕타아제 활성; 또는 (f) 이들의 조합을 포함한다.

- [0015] 일 실시 형태에서, 재조합 숙주 세포는 감소되거나 또는 제거된 알데히드 데하이드로게나아제 발현 활성 및 감소되거나 또는 제거된 아세토락테이트 리덕타아제 발현 또는 활성을 포함한다.
- [0016] 다른 실시 형태에서, 재조합 숙주 세포는 (i) 감소되거나 또는 제거된 알데히드 데하이드로게나아제 발현 또는 활성 또는 감소되거나 또는 제거된 아세토락테이트 리덕타아제 발현 또는 활성 및 (ii) KARI 활성을 갖고 NADH에 대한 K_M 이 300 μ M 미만인 폴리펩티드를 코딩하는 이중성 폴리뉴클레오티드를 포함한다.
- [0017] 다른 실시 형태에서, 재조합 숙주는 서열 번호 27, 29, 141, 143, 275 또는 277의 서열에 대한 동일성이 약 90% 이상이거나 또는 약 95% 이상인, KARI 활성을 갖는 이중성 폴리펩티드를 포함한다.
- [0018] 일부 실시 형태에서, 재조합 숙주 세포는 서열 번호 27의 S56 및 S58에 상응하는 아미노산의 치환을 포함하는, KARI 활성을 갖는 이중성 폴리펩티드를 포함한다. 일부 실시 형태에서, KARI 활성을 갖는 폴리펩티드는 서열 번호 27의 I86, N87, T131, 또는 T191에 상응하는 아미노산들 중 하나 이상의 치환을 추가로 포함한다. 일부 실시 형태에서, KARI 활성을 갖는 폴리펩티드는 서열 번호 31, 33, 35, 37, 39, 41, 43, 45, 47, 49, 51, 53, 55, 57, 59, 61, 63 또는 65의 서열에 대하여 동일성이 90% 이상이거나 또는 동일성이 95% 이상이다.
- [0019] 일부 실시 형태에서, 재조합 숙주 세포는 약 48시간 후 세포 1 g당 약 3 g 이상, 약 4 g 이상 또는 약 5 g 이상의 유효 아이소부탄올 생산성을 가지며, 여기서, 상기 48시간 중 적어도 마지막의 약 24시간은 혐기성 조건 하에 있다.
- [0020] 일부 실시 형태에서, 재조합 숙주 세포는 pH 6.8에서 NADH에 대한 K_M 이 약 350 μ M 미만, 약 100 μ M 미만, 약 50 μ M 미만, 또는 약 10 μ M 미만인, KARI 활성을 갖는 이중성 폴리펩티드를 포함한다.
- [0021] 일부 실시 형태에서, 재조합 숙주 세포는 서열 번호 376, 382, 378 또는 275의 서열에 대하여 동일성이 약 90% 이상이거나 또는 동일성이 약 95% 이상인, KARI 활성을 갖는 이중성 폴리펩티드를 포함한다.
- [0022] 일부 실시 형태에서, 재조합 숙주 세포는 언에어로스티페스 카카에로부터 유래된 KARI 효소 (서열 번호 27)의 아미노산 A41, S56, S58, I87, T131, T191, R227 또는 Q246에 상응하는 위치들 중 하나 이상에서 아미노산 치환을 포함하는, KARI 활성을 갖는 이중성 폴리펩티드를 포함한다.
- [0023] 일부 실시 형태에서, 재조합 숙주 세포는 서열 번호 33 또는 서열 번호 35의 서열을 포함하는, KARI 활성을 갖는 이중성 폴리펩티드 또는 이의 활성 단편을 포함한다.
- [0024] 다른 실시 형태에서, 재조합 숙주 세포는 알데히드 데하이드로게나아제 활성을 갖는 폴리펩티드를 코딩하는 내인성 폴리뉴클레오티드 내에 결실, 돌연변이 및/또는 치환을 포함한다. 일부 실시 형태에서, 알데히드 데하이드로게나아제 활성을 갖는 폴리펩티드는 아이소부티르알데히드의 아이소부티르산으로의 전환을 촉매한다. 일부 실시 형태에서, 알데히드 데하이드로게나아제 활성을 갖는 폴리펩티드는 효소 위원회 번호(Enzyme Commission Number) EC 1.2.1.3, EC 1.2.1.4, 및/또는 EC1.2.1.5에 상응한다. 일부 실시 형태에서, 숙주 세포는 에스. 세레비시아(*S. cerevisiae*)이며, 알데히드 데하이드로게나아제 활성을 갖는 폴리펩티드는 ALD2, ALD3, ALD4, ALD5, ALD6 또는 이의 상동체이다. 일부 실시 형태에서, 숙주 세포는 케이. 락티스이며, 알데히드 데하이드로게나아제 활성을 갖는 폴리펩티드는 KLLA0F00440, KLLA0E23057, KLLA0D10021, 또는 KLLA0D09999G이다. 일부 실시 형태에서, 숙주 세포는 피. 스티피티스(*P. stipitis*)이며, 알데히드 데하이드로게나아제 활성을 갖는 폴리펩티드는 ALD2, ALD3, ALD4, ALD5, 또는 ALD7이다. 일부 실시 형태에서, 숙주 세포는 락토바실러스 플란타룸(*Lactobacillus plantarum*)이며, 알데히드 데하이드로게나아제 활성을 갖는 상기 폴리펩티드는 AldH이다. 일부 실시 형태에서, 숙주 세포는 이. 콜라이(*E. coli*)이며, 알데히드 데하이드로게나아제 활성을 갖는 폴리펩티드는 aldA, aldB 또는 aldH이다.
- [0025] 다른 실시 형태에서, 숙주 세포는 알데히드 옥시다아제 활성을 갖는 폴리펩티드를 코딩하는 내인성 폴리뉴클레오티드 내에 결실, 돌연변이 및/또는 치환을 포함한다. 일부 실시 형태에서, 알데히드 옥시다아제 활성을 갖는 폴리펩티드는 아이소부티르알데히드의 아이소부티르산으로의 전환을 촉매한다. 일부 실시 형태에서, 알데히드 옥시다아제 활성을 갖는 폴리펩티드는 효소 위원회 번호 EC 1.2.3.1에 상응한다. 일부 실시 형태에서, 알데히드 옥시다아제 활성을 갖는 폴리펩티드는 AOX1 및/또는 AOX2이다.
- [0026] 다른 실시 형태에서, 숙주 세포는 아세토락테이트 리덕타아제 활성을 갖는 폴리펩티드를 코딩하는 내인성 폴리뉴클레오티드 내에 결실, 돌연변이 및/또는 치환을 포함한다. 일부 실시 형태에서, 아세토락테이트 리덕타아제

활성을 갖는 폴리펩티드는 서열 번호 676, 서열 번호 678, 서열 번호 680, 서열 번호 682, 서열 번호 684, 서열 번호 686, 서열 번호 688, 서열 번호 690, 서열 번호 692, 서열 번호 694, 서열 번호 696, 서열 번호 702, 서열 번호 704, 서열 번호 706, 서열 번호 708, 서열 번호 710, 서열 번호 712, 서열 번호 714, 서열 번호 716, 서열 번호 718, 서열 번호 720, 서열 번호 722, 서열 번호 724, 서열 번호 726, 서열 번호 728 및 서열 번호 730의 서열로 이루어진 군으로부터 선택되는 폴리뉴클레오티드에 의해 코딩되는 폴리펩티드를 포함한다. 일부 실시 형태에서, 아세트락테이트 리덕타아제 활성을 갖는 폴리펩티드는 YMR226C이다.

[0027] 다른 실시 형태에서, 재조합 숙주 세포는 효모 숙주 세포이다. 일부 실시 형태에서, 효모는 사카로마이세스 (*Saccharomyces*), 쉬조사카로마이세스(*Schizosaccharomyces*), 한센울라(*Hansenula*), 칸디다(*Candida*), 클루이 베로마이세스(*Kluyveromyces*), 야로위아(*Yarrowia*), 이사첸키아(*Issatchenkia*) 또는 피키아(*Pichia*)로 이루어진 군으로부터 선택된다. 일부 실시 형태에서, 숙주 세포는 사카로마이세스 세레비지애(*Saccharomyces cerevisiae*)이다.

[0028] 다른 실시 형태에서, 숙주 세포는 박테리아 세포이다. 일부 실시 형태에서, 박테리아 세포는 클로스트리듐, 자 이모모나스, 에스케리키아(*Escherichia*), 살모넬라(*Salmonella*), 로도코커스(*Rhodococcus*), 슈도모나스(*Pseudomonas*), 바실러스(*Bacillus*), 락토바실러스, 엔테로코커스, 페디오코커스(*Pediococcus*), 알칼리게네스(*Alcaligenes*), 클렙시엘라(*Klebsiella*), 파에니바실러스(*Paenibacillus*), 아트로박터(*Arthrobacter*), 코리네 박테리움(*Corynebacterium*), 브레비박테리움(*Brevibacterium*), 락토코커스, 류코노스톡, 오에노코커스(*Oenococcus*), 페디오코커스, 또는 스트렙토코커스 세포이다. 일부 실시 형태에서, 박테리아 세포는 이. 콜라 이가 아니다.

[0029] 일부 실시 형태에서, 재조합 숙주 세포의 엔지니어링된 아이소부탄올 생성 경로는 하기의 기질의 생성물로의 전환, 즉 (a) 피루베이트의 아세트락테이트로의 전환, (b) 아세트락테이트의 2,3-다이하이드록시아이소발레레이트의 전환, (c) 2,3-다이하이드록시아이소발레레이트의 2-케토아이소발레레이트로의 전환, (d) 2-케토아이소발레레이트의 아이소부티르알데히드로의 전환, 및 (e) 아이소부티르알데히드의 아이소부탄올로의 전환을 포함하며, 상기의 기질의 생성물로의 전환들 중 하나 초과는 숙주 세포에 대하여 이중성인 효소에 의해 촉매된다. 일부 실시 형태에서, 기질의 생성물로의 전환들 전부는 숙주 세포에 대하여 이중성인 효소에 의해 촉매된다. 일부 실시 형태에서, 숙주 세포에 대하여 이중성인 효소를 코딩하는 적어도 하나의 이중성 폴리뉴클레오티드는 숙주 세포에 염색체 통합된다. 일부 실시 형태에서, 기질의 생성물로의 전환은 사실상 세포질에 국소화된 효소에 의해 촉매된다. 일부 실시 형태에서, 아이소부티르알데히드의 아이소부탄올로의 전환에 있어서의 기질의 생성물로의 전환은 NADH를 보조 인자로 이용하는 알코올 데하이드로게나아제 효소에 의해 촉매된다. 일부 실시 형태에서, 아세트락테이트의 2,3-다이하이드록시아이소발레레이트로의 전환은 NADH를 보조 인자로서 이용할 수 있는 KARI에 의해 촉매된다.

[0030] 일부 실시 형태에서, 숙주 세포는 플라스미드 pLH702 또는 pLH701, 또는 동일 코딩 영역을 갖는 플라스미드를 포함한다. 일부 실시 형태에서, 숙주 세포는 플라스미드 pBP915 또는 동일 코딩 영역을 갖는 플라스미드를 포함한다. 일부 실시 형태에서, 숙주 세포는 플라스미드 pYZ067 Δ ki vD Δ hADH 또는 동일 코딩 영역을 갖는 플라스미드를 포함한다.

[0031] 일부 실시 형태에서, 숙주 세포는 아세트락테이트의 2,3-다이하이드록시-2-메틸부티레이트로의 전환 능력의 감소, 파괴 또는 제거를 포함한다.

[0032] 일부 실시 형태에서, 숙주 세포는 효모이며, 감소되거나 또는 제거된 피루베이트 데카르복실라아제 발현 또는 활성을 갖는다. 일부 실시 형태에서, 숙주 세포는 감소되거나 또는 제거된 PDC1, PDC5 또는 PDC6 활성 또는 이의 조합을 갖는다.

[0033] 일부 실시 형태에서, 숙주 세포는 감소되거나 또는 제거된 NAD-의존성 글리세롤-3-포스페이트 데하이드로게나아제 발현 또는 활성을 갖는다. 일부 실시 형태에서, 숙주 세포는 감소된 GPD2 활성을 갖는다.

[0034] 일부 실시 형태에서, 숙주 세포는 감소되거나 또는 제거된 FRA2 발현 또는 활성을 갖는다.

[0035] 일부 실시 형태에서, 숙주 세포는 혐기성 조건 하에서 아이소부탄올을 생성하며, 글리세롤에 대한 아이소부탄올의 몰비는 1 초과이다.

[0036] 일부 실시 형태에서, 케톨-산 리덕토아이소머라아제 활성을 갖는 폴리펩티드는 표 Z에 제공된 주어진 프로파일 HMM을 $<10^{-3}$ 의 프로파일 HMM E 값과 매칭시킨다.

- [0037] 일부 실시 형태에서, 숙주 세포는 이론적 수율의 약 25% 초과, 약 50% 초과, 약 75% 초과 또는 약 90% 초과 수율로 아이소부탄올을 생성한다.
- [0038] 일부 실시 형태에서, 아이소부탄올 및 에탄올이 생성된다.
- [0039] 아이소부탄올의 제조 방법은 상기에 기재된 재조합 숙주 세포를 제공하는 단계 및 아이소부탄올을 생성하는 조건 하에서 숙주 세포를 탄소 기질과 접촉시키는 단계를 포함하는 방법을 포함한다. 일부 실시 형태에서, 상기 접촉의 적어도 일부분은 혐기성 조건 하에서 일어난다. 일부 실시 형태에서, 상기 접촉은 추출제의 존재 하에 일어난다. 일부 실시 형태에서, 상기 접촉은 수성 상 및 유기 상을 포함하는 2상(two-phase) 시스템을 형성하기에 충분한 양의 유기 추출제의 존재 하에 일어난다. 일부 실시 형태에서, 아이소부탄올의 유효 비율, 유효 역가 또는 유효 수율 중 하나 이상은 KARI 활성을 갖는 이중성 폴리펩티드, KARI 활성, 감소되거나 또는 제거된 알데히드 데하이드로게나아제 활성, 감소되거나 또는 제거된 알데히드 옥시다아제 활성, 감소되거나 또는 제거된 아세토락테이트 리덕타아제 활성 또는 이의 조합을 갖는 폴리펩티드를 코딩하는 이중성 폴리뉴클레오티드를 포함하지 않는 재조합 숙주 세포를 이용한 방법과 비교하여 증가된다. 일부 실시 형태에서, 아이소부탄올의 유효 비율, 유효 역가 또는 유효 수율 중 하나 이상은 (i) KARI 활성을 갖는 이중성 폴리펩티드 또는 KARI 활성을 갖는 폴리펩티드를 코딩하는 이중성 폴리뉴클레오티드 및 (ii) 엔지니어링된 아이소부탄올 생성 경로의 성능을 향상시키는 적어도 하나의 변형을 포함하지 않는 재조합 숙주 세포를 이용한 방법과 비교하여 증가된다. 일부 실시 형태에서, DHMB 생성, 아이소부티르산 생성 또는 이들 둘 모두는 KARI 활성을 갖는 이중성 폴리펩티드, KARI 활성, 감소되거나 또는 제거된 알데히드 데하이드로게나아제 활성, 감소되거나 또는 제거된 알데히드 옥시다아제 활성, 감소되거나 또는 제거된 아세토락테이트 리덕타아제 활성 또는 이의 조합을 갖는 폴리펩티드를 코딩하는 이중성 폴리뉴클레오티드를 포함하지 않는 재조합 숙주 세포를 이용한 방법과 비교하여 감소된다. 일부 실시 형태에서, DHMB 생성, 아이소부티르산 생성 또는 이들 둘 모두는 (i) KARI 활성을 갖는 이중성 폴리펩티드 또는 KARI 활성을 갖는 폴리펩티드를 코딩하는 이중성 폴리뉴클레오티드 및 (ii) 엔지니어링된 아이소부탄올 생성 경로의 성능을 향상시키는 적어도 하나의 변형을 포함하지 않는 재조합 숙주 세포를 이용한 방법과 비교하여 감소된다. 일부 실시 형태에서, 글리세롤에 대한 아이소부탄올의 몰비는 1 초과이다.
- [0040] 아이소부탄올의 제조 방법은 또한 아이소부탄올을 생성하는 재조합 숙주 세포를 제공하는 단계 및 아이소부탄올을 생성하는 조건 하에서 숙주 세포를 탄소 기질과 접촉시키는 단계를 포함하며, 여기서, 상기 접촉의 적어도 일부분은 혐기성 조건 하에서 일어나고, 생성된 글리세롤에 대한 아이소부탄올의 비는 1 초과이다.
- [0041] 아이소부탄올의 제조 방법은 또한 부탄올을 생성하는 조건 하에서 피루베이트를 아세토락테이트로 전환시킬 수 있는 생합성 경로를 포함하는 재조합 효모를 성장시키는 단계 및 DHMB를 배양물로부터 제거하는 단계를 포함한다.
- [0042] 이러한 방법에 의해 생성된 조성물이 또한 본 발명에서 제공된다. 일부 실시 형태에서, 조성물은 상기에 제공된 재조합 숙주 세포 및 아이소부탄올을 포함한다. 일부 실시 형태에서, 조성물은 부탄올 및 약 0.5 mM 이하의 DHMB를 포함한다.
- [0043] 발효 조성물이 또한 본 발명에서 제공된다. 일부 실시 형태에서, 발효 조성물은 상기에 제공된 방법에 따라 제조된 아이소부탄올 및 숙주 세포를 포함한다.
- [0044] i) 부탄올을 생성할 수 있는 재조합 효모, ii) 부탄올, 및 iii) 약 0.5 mM 이하의 DHMB를 포함하는 조성물이 또한 제공된다.
- [0045] 재조합 숙주 세포를 생성하는 방법이 또한 제공된다. 이러한 방법은 (a) 알데히드 데하이드로게나아제 활성 또는 알데히드 옥시다아제 활성을 갖는 폴리펩티드를 코딩하는 폴리뉴클레오티드 내에 변형을 포함하는 재조합 숙주 세포를 제공하는 단계; 및 (b) 아이소부탄올 생합성 경로의 폴리펩티드를 코딩하는 폴리뉴클레오티드로 숙주 세포를 형질전환시키는 단계를 포함할 수 있다.
- [0046] 아이소부티르알데히드의 아이소부티르산으로의 전환을 감소시키거나 또는 제거하는 방법이 또한 제공된다. 이러한 방법은 (a) 제1항 내지 제52항 중 어느 한 항의 재조합 숙주 세포를 제공하는 단계; 및 (b) 감소되거나 또는 제거된 알데히드 데하이드로게나아제 및/또는 알데히드 옥시다아제 활성을 포함하지 않는 재조합 숙주 세포를 사용하는 방법과 비교하여 아이소부티르알데히드의 아이소부티르산으로의 전환을 감소시키거나 또는 제거하는 조건을 숙주 세포에 가하는 단계를 포함할 수 있다.
- [0047] 소정의 폴리펩티드가 또한 본 발명에서 제공된다. 일부 실시 형태에서, 폴리펩티드는 서열 번호 29, 31, 33,

35, 37, 39, 41, 43, 45, 47, 49, 51, 53, 55, 57, 59, 61, 63, 65, 417, 419, 421, 423, 425, 427, 429, 431, 433, 434, 435, 436, 437, 438, 439, 440, 441, 442, 443, 444, 445, 446, 447, 448, 449, 450, 451, 452, 453, 454, 455, 456, 457, 458, 459, 460, 461, 462, 463, 464, 465, 466, 467, 468, 469, 470, 471, 472, 473, 474, 475, 476, 477, 478, 479, 480, 481, 482, 483, 484, 485, 486, 487, 488, 489, 490, 491, 492, 493, 494, 495, 496, 497, 498, 499, 500, 501, 502, 503, 504, 505, 506, 507, 508, 509, 510, 511, 512, 513, 514, 515, 516, 517, 518, 519, 520, 521, 522, 523, 524, 525, 526, 527, 528, 529, 530, 531, 532, 533, 534, 535, 536, 537, 624, 626, 628, 630, 632의 서열 또는 이의 활성 단편에 대하여 약 90% 이상의 동일성 또는 약 95% 이상의 동일성 또는 약 99% 이상의 동일성을 포함하며, 케툴-산 리덕토아이스머라아제 활성을 갖는다. 일부 실시 형태에서, 폴리펩티드는 서열 번호 417, 419, 421, 423, 425, 또는 427의 서열 또는 이의 활성 단편에 대하여 약 90% 이상의 동일성 또는 약 95% 이상의 동일성 또는 약 99% 이상의 동일성을 포함하며, 케툴-산 리덕토아이스머라아제 활성을 갖는다. 일부 실시 형태에서, 폴리펩티드는 서열 번호 927, 928, 196, 266, 267, 389, 405, 637, 781, 782, 783, 835, 853, 854, 855, 856, 857, 또는 859의 서열에 대하여 동일성이 약 90% 이상이거나, 동일성이 약 95% 이상이거나, 또는 동일성이 약 99% 이상인 서열을 포함한다.

[0048] 이러한 폴리펩티드를 코딩하는 폴리뉴클레오티드 및 이러한 폴리뉴클레오티드 및 폴리펩티드를 포함하는 숙주 세포가 또한 제공된다.

[0049] 아세토락테이트를 2,3-다이하이드록시아이소발레레이트로 전환시키는 방법이 또한 제공된다. 예를 들어, 이러한 방법은 (a) 상기에 기재된 폴리펩티드를 제공하는 단계, 및 (b) 2,3-다이하이드록시아이소발레레이트가 생성되는 조건 하에서 폴리펩티드를 아세토락테이트와 접촉시키는 단계를 포함할 수 있다.

[0050] 재조합 효모 세포가 또한 본 발명에서 제공된다. 일부 실시 형태에서, 재조합 효모는 피루베이트를 아세토락테이트로 전환시킬 수 있는 생합성 경로를 포함하며, 효모는 소비되는 당 1몰당 0.01몰 미만의 2,3-다이하이드록시-2-메틸 부티레이트 (DHMB)를 생성한다. 일부 실시 형태에서, 재조합 효모는 피루베이트를 아세토락테이트로 전환시킬 수 있는 것을 포함하며, 효모는 약 1.0 mM/시간 미만의 속도로 DHMB를 생성한다. 일부 실시 형태에서, 재조합 효모는 피루베이트를 아세토락테이트로 전환시킬 수 있는 생합성 경로를 포함하며, 효모는 DHMB의 생성량의 약 1.5배 이상인 양의 2,3-다이하이드록시-3-아이소발레레이트 (DHIV)를 생성한다.

[0051] DHMB 생성에 연루된 유전자를 확인하는 방법이 또한 제공된다. 일부 실시 형태에서, 본 방법은 (i) 적어도 2개 이상의 유전자 결실을 포함하는 효모 주의 컬렉션(collection)을 제공하는 단계; (ii) 개별 효모 주에 의해 생성되는 DHMB의 양을 측정하는 단계; (iii) 시간당 약 1.0 mM 이하의 DHMB를 생성하는 효모 주를 선발하는 단계; 및 (iv) 선발된 효모 주에서 결실된 유전자를 확인하는 단계를 포함한다. 일부 실시 형태에서, 본 방법은 (i) 적어도 2가지 이상의 유전자를 과다발현하는 효모 주의 컬렉션을 제공하는 단계; (ii) 개별 효모 주에 의해 생성되는 DHMB의 양을 측정하는 단계; (iii) 약 1.0 mM 이상의 DHMB를 생성하는 효모 주를 선발하는 단계; 및 (iv) 선발된 효모 주에서 과다발현된 유전자를 확인하는 단계를 포함한다.

[0052] 부탄올의 제조 방법이 또한 제공된다. 일부 실시 형태에서, 본 방법은 (a) 부탄올을 생성하는 조건 하에서 피루베이트를 아세토락테이트로 전환시킬 수 있는 생합성 경로를 포함하는 재조합 효모를 성장시키는 단계; 및

[0053] b) DHIV 및/또는 DHMB 농도를 측정하는 단계를 포함한다. 일부 실시 형태에서, 상기 성장 및 측정은 동시에 또는 순차적으로 그리고 임의의 순서로 수행될 수 있다. 일부 실시 형태에서, 상기 측정은 액체 크로마토그래피-질량 분광법을 포함한다.

[0054] 케툴-산 리덕토아이스머라아제 (KARI) 활성을 증가시키는 방법이 또한 제공된다. 일부 실시 형태에서, 본 방법은 (a) 아세토락테이트, KARI 효소 및 아세토락테이트 리덕타아제 효소를 포함하는 조성물을 제공하는 단계, 및 (b) DHMB의 수준을 감소시키는 단계를 포함한다. 일부 실시 형태에서, DHMB 수준을 감소시키는 단계는 아세토락테이트 리덕타아제 효소 활성을 감소시킴으로써 성취된다. 일부 실시 형태에서, DHMB 수준을 감소시키는 단계는 상기 조성물로부터 DHMB를 제거함으로써 성취된다.

[0055] 일부 실시 형태에서, 숙주 세포에서 KARI 효소의 생산성을 증가시키는 것은 이중성 KARI 효소와, 아세토락테이트 리덕타아제 발현 또는 활성을 감소시키거나, 파괴하거나 또는 제거하는 적어도 하나의 유전자 변형을 포함하는 숙주 세포를 배양하는 것을 포함할 수 있으며, 여기서, KARI 효소 활성은 DHMB의 존재 하에서 감소된다. 일부 실시 형태에서, KARI는 이. 콜라이또는 엘. 락티스 KARI에 대하여 동일성이 약 90% 이상, 약 95% 이상 또는 약 99% 이상이다. 일부 실시 형태에서, 감소되거나, 파괴되거나 또는 제거된 아세토락테이트 리덕타아제 발현 또는 활성은 DHMB의 존재를 사실상 감소시킨다.

- [0056] 다이하이드록시산 데하이드라타아제 (DHAD) 활성을 증가시키는 방법이 또한 제공된다. 일부 실시 형태에서, 본 방법은 (a) 다이하이드록시아이소발레레이트 (DHIV) 및 DHAD 효소를 포함하는 조성물을 제공하는 단계 및 (b) DHMB 수준을 감소시키는 단계를 포함한다.
- [0057] 도면의 간단한 설명 및 본 출원과 함께 전자 출원된 표를 참고로 포함시킴
- [0058] 본 발명은 하기의 발명을 실시하기 위한 구체적인 내용, 도면 및 첨부된 서열 설명 - 이는 본 출원의 일부를 형성함 - 으로부터 더욱 완전히 이해될 수 있다.
- [0059] [도면의 간단한 설명]
- [0060] <도 1>
- [0061] 도 1은 4개의 상이한 아이소부탄올 생합성 경로를 나타낸다. "a", "b", "c", "d", "e", "f", "g", "h", "i", "j" 및 "k"로 표시된 단계는 하기에 기재된 기질의 생성물로의 전환을 나타낸다.
- [0062] <도 2>
- [0063] 도 2는 슈도모나스 플루오레센스(*Pseudomonas fluorescens*) 유래의 KARI의 아미노산 서열 ("PF5"; 서열 번호 5) 및 언에어로스티페스 카카에 유래의 KARI ("K9"; 서열 번호 27)의 아미노산 서열의 정렬을 도시한다. 볼드체의 위치들은 본 명세서에 기재된 바와 같이 돌연변이 유발용으로 표적화된다.
- [0064] <도 3a, 도 3b 및 도 3c>
- [0065] 도 3a, 도 3b 및 도 3c는 비피도박테리움 안굴라툼 DSM 20098 ("K1"; 서열 번호 141), 비피도박테리움 덴티움 ATCC 27678 ("K2"; 서열 번호 143), 클로스트리듐 베이제린키이 NCIMB 8052 ("K7"; 서열 번호 275), 언에어로스티페스 카카에 DSM 14662 ("K9"; 서열 번호 27), 엔테로코커스 갈리나룸 EG2 ("K25"서열 번호 376), 스트렙토코커스 서모필루스 LMD-9 ("K26" 서열 번호 121), 락토코커스 락티스 아종 크레모리스 MG1363 ("K29"; 서열 번호 382), 자이모모나스 모빌리스 ("S2"; 서열 번호 277), 및 락토코커스 락티스 ("LTS"; 서열 번호 380) 유래의 KARI 효소의 아미노산 서열의 정렬을 도시한다. 볼드체의 위치들은 본 명세서에 기재된 바와 같이 돌연변이 유발용으로 표적화된다.
- [0066] <도 4>
- [0067] 도 4는 pLH556 (pHR81-Pilv5-Pf5.KARI) 벡터 (서열 번호 138)의 플라스미드 지도이다.
- [0068] <도 5>
- [0069] 도 5는 2가지의 주, PNY1910 및 PNY2242의 아이소부탄올의 비생성률(specific rate of isobutanol production), Qp를 나타낸다.
- [0070] <도 6>
- [0071] 도 6은 PNY1910 (삼각형) 및 PNY2242 (마름모형)를 이용한 발효 시간 코스 동안의 배양 상청액 중의 DHIV + DHMB의 축적을 나타낸다. (DHMB 및 DHIV는 사용한 HPLC 방법에 의해서는 구별되지 않는다.)
- [0072] <도 7>
- [0073] 도 7은 글리세롤, 피루브산, 2,3-부탄다이올 (BDO), DHIV/DHMB, α-케토아이소발레레이트 (aKIV), 및 아이소부티르산 (iBuAc)의 수율을 나타낸다. DHIV 및 DHMB는 이들이 사용한 HPLC 방법에 의해 구별되지 않기 때문에 함께 나타낸다.
- [0074] <도 8>
- [0075] 도 8은 실시예 16에 기재된 K9G9 변이체의 V_{max}/K_M 값들의 요약을 나타낸다.
- [0076] <도 9a, 도 9b 및 도 9c>
- [0077] 도 9a, 도 9b 및 도 9c는 실시예 19에 기재된 K9 변이체의, 글리세롤에 대한 아이소부탄올의 몰 수율 비, 아이소부탄올 몰 수율 및 아이소부탄올의 역가를 나타낸다.
- [0078] <도 10>
- [0079] 도 10은 아이소부탄올 생합성 경로를 나타낸다. 단계 "a"는 피루베이트의 아세토락테이트로의 전환을

나타낸다. 단계 "b"는 아세토락테이트의 DHIV로의 전환을 나타낸다. 단계 "c"는 DHIV의 KIV로의 전환을 나타낸다. 단계 "d"는 KIV의 아이소부티르알데히드로의 전환을 나타낸다. 단계 "e"는 아이소부티르알데히드의 아이소부탄올로의 전환을 나타낸다. 단계 "f"는 아세토락테이트의 DHMB로의 전환을 나타낸다.

[0080] <도 11>

[0081] 도 11은 자낭균 효모의 중 유래의 YMR226c 상동체들의 계통수를 나타낸다. 사상균 (뉴로스포라 크라사 (*Neurospora crassa*)) 서열이 외집단(outgroup)으로서 포함된다.

[0082] <도 12>

[0083] 도 12는 YMR226C에 대하여 상동성을 갖는 ORF의 뉴클레오티드 서열의 다중 서열 정렬(multiple sequence alignment; MSF 형식)을 나타낸다. 나타낸 유전자명은 표 7에 주어진 등록 번호 및 서열 번호에 상응한다. 정렬은 얼라인엑스(AlignX; 벡터 엔티아이(Vector NTI))에 의해 생성되었다.

[0084] <도 13>

[0085] 도 13은 시간이 지남에 따른 DHMB의 몰 수율의 그래프를 나타낸다.

[0086] <도 14>

[0087] 도 14는 효모 주 NYLA84에서의 아이소부탄올 및 아이소부티르산의 생성을 도시한다.

[0088] 표 Z-는 전체적으로 본 명세서에 참고로 포함된 미국 특허 공개 제2010/0197519호 및 미국 특허 공개 제 2009/0163376호에 기재된 바와 같은 그리고 표 1에 열거된, 실험적으로 검증된 KARI 효소의 프로파일 HMM의 표 이다 (본 출원과 함께 전자 출원되고 참고로 포함됨).

[0089] [표 1]

실험적으로 검증된 KARI 효소

GI명	등록 번호	서열 번호	미생물
70732562	YP_262325.1	5	슈도모나스 플루오레센스 Pf-5
15897495	NP_342100.1	1	솔폴로부스 솔파타리쿠스(<i>Sulfolobus solfataricus</i>) P2
18313972	NP_560639.1	2	파이로바쿨룸 아에로필룸(<i>Pyrobaculum aerophilum</i>) str. IM2
76801743	YP_326751.1	7	나트로노모나스 파라오니스(<i>Natronomonas pharaonis</i>) DSM 2160
16079881	NP_390707.1	8	바실러스 서브틸리스(<i>Bacillus subtilis</i>) 아종 서브틸리스 str. 168
19552493	NP_600495.1	9	코리네박테리움 글루타미쿰(<i>Corynebacterium glutamicum</i>) ATCC 13032
6225553	O32414	10	파에오스프리릴룸 몰리쉬아눔(<i>Phaeosprillum molischianum</i>)
17546794	NP_520196.1	3	랄스토니아 솔라나세아룸(<i>Ralstonia solanacearum</i>) GMI1000
56552037	YP_162876.1	11	자이모모나스 모빌리스 아종 모빌리스 ZM4
114319705	YP_741388.1	12	알칼리림니콜라 에르리체이(<i>Alkalilimnicola ehrlichei</i>) MLHE-1
57240359	ZP_00368308.1	13	캠필로박터 라리(<i>Campylobacter lari</i>) RM2100
120553816	YP_958167.1	14	마리노박터 아쿠아에올레이(<i>Marinobacter aquaeolei</i>) VT8
71065099	YP_263826.1	15	사이크로박터 아르크티쿠스(<i>Psychrobacter arcticus</i>) 273-4
83648555	YP_436990.1	16	하헬라 제주엔시스(<i>Hahella chejuensis</i>) KCTC 2396

[0090]

74318007	YP_315747.1	17	티오바실러스 테니트리피칸스(<i>Thiobacillus denitrificans</i>) ATCC 25259
67159493	ZP_00420011.1	18	아조토박터 비넬란디아(<i>Azotobacter vinelandii</i>) AvOP
66044103	YP_233944.1	19	슈도모나스 시린가에(<i>Pseudomonas syringae</i>) pv. 시린가에 B728a
28868203	NP_790822.1	20	슈도모나스 시린가에 pv. 토마토(<i>Pseudomonas syringae</i> str. DC3000)
26991362	NP_746787.1	21	슈도모나스 푸티다(<i>Pseudomonas putida</i>) KT2440
104783656	YP_610154.1	22	슈도모나스 엔토모필라(<i>Pseudomonas entomophila</i>) L48
146306044	YP_001186509.1	23	슈도모나스 멘도시나(<i>Pseudomonas mendocina</i>) ymp
15599888	NP_253382.1	4	슈도모나스 아에루기노사(<i>Pseudomonas aeruginosa</i>) PAO1
42780593	NP_977840.1	24	바실러스 세레우스(<i>Bacillus cereus</i>) ATCC 10987
42781005	NP_978252.1	25	바실러스 세레우스 ATCC 10987
266346	Q01292	6	스피나시아 올레라세아(<i>Spinacia oleracea</i>)

[0091]

[0092]

슈도모나스 플루오레센스 Pf-5 KARI 내의 11개의 보조 인자 스위칭 위치에 상응하는, 정렬에 있어서 컬럼을 나타내는 프로파일 HMM 내의 11개의 위치는 위치 24, 33, 47, 50, 52, 53, 61, 80, 115, 156 및 170으로 확인된다. 표 Z는 본 출원과 함께 전자 제출되며, 본 명세서에 참고로 포함된다.

[0093]

본 출원과 함께 출원된 서열 목록에 제공된 서열 (명칭: 20120323_CL5367USNA_SEQLIST.txt; 크기: 2,003,893 바이트; 생성일: 2012년 3월 23일)은 본 명세서에 참고로 포함된다.

[0094]

세계 지적 재산권 기구(World Intellectual Property Organization; WIPO) 표준 ST.25 (2009)와 일치하여, 서열 목록 및 본 명세서에 주어진 소정의 프라이머는 뉴클레오티드 a 또는 g 또는 c 또는 t를 나타내기 위하여 N을 사용한다. K는 g 또는 t를 나타내는 데 사용된다. M은 a 또는 c를 나타내는 데 사용된다.

발명을 실시하기 위한 구체적인 내용

[0095]

달리 정의되지 않으면, 본 명세서에서 사용된 모든 기술 및 과학 용어는 본 발명이 속하는 기술 분야의 숙련자가 일반적으로 이해하는 것과 동일한 의미를 갖는다. 상충될 경우, 정의를 비롯한 본 명세서가 좌우할 것이다. 또한, 문맥에 의해 다르게 필요로 하지 않는 한, 단수의 용어는 복수를 포함할 것이며, 복수의 용어는 단수를 포함할 것이다. 특허 또는 특허 공개의 특정한 섹션만이 참고로 포함되는 것으로 지시되지 않으면, 본 명세서에 언급된 모든 간행물, 특허 및 다른 참고 문헌은 마치 각각의 개별 간행물 또는 특허 출원이 특정하게 그리고 개별적으로 참고로 포함되는 것으로 지시되는 것처럼 사실상 전체적으로 참고로 포함된다.

[0096]

본 명세서에 개시된 것과 유사하거나 또는 등가인 방법 및 재료가 본 발명의 실시 또는 시험에서 사용될 수 있지만, 적합한 방법 및 재료는 하기에 개시되어 있다. 재료, 방법 및 실시예는 단지 예시적인 것이며 제한하고자 하는 것은 아니다. 본 발명의 다른 특징 및 이점은 발명을 실시하기 위한 구체적인 내용 및 특허청구범위로부터 명백해질 것이다.

[0097]

피루베이트-이용 생합성 경로를 통한 아이소부탄올의 생합성에서의 마지막 단계는 아이소부티르알데히드의 아이소부탄올로의 전환이다 (도 1). 이 경로에서의 부반응은 아이소부티르알데히드의 아이소부티르산으로의 전환이며, 이는 아이소부탄올로의 전환에 이용가능한 아이소부티르알데히드의 감소된 양 및 감소된 아이소부탄올 수율로 이어진다. 효율적인 생합성 공정에 있어서, 증가된 양의 아이소부티르알데히드가 아이소부탄올로의 전환에 이용가능해지고 아이소부탄올 수율이 증가되도록, 아이소부티르알데히드의 아이소부티르산으로의 전환을 방지할 필요가 있다.

[0098]

알데히드 데하이드로게나아제는 알데히드의 산화 (탈수소화)를 촉매하는 효소 패밀리이다 (문헌[Wang et al.,

J. Bacteriol. 180:822-30, 1998]; 문헌[Navarro-Avino *et al.*, *Yeast* 15:829-42, 1999]; 및 문헌[Saint-Prix *et al.*, *Microbiology* 150:2209-20, 2004]). 알데히드 데하이드로게나아제 활성이 감소되거나 또는 제거되도록 변형될 수 있으며, 아이소부티르알데히드의 아이소부티르산의 전환을 감소시키거나 또는 제거할 수 있어서, 증가된 양의 아이소부티르알데히드가 아이소부탄올로의 전환에 이용가능하고 아이소부탄올 수율이 증가되도록 하는 적합한 알데히드 데하이드로게나아제를 확인할 필요가 있다.

[0099] 알데히드 옥시다아제는 알데히드로부터의 카르복실산의 생성을 촉매하는 효소 패밀리아다 (문헌[Nomura *et al.*, *Biosci. Biotechnol. Biochem.* 62:1134-7, 1998]; 및 문헌[Johnson *et al.*, *Genetics* 151:1379-1391, 1999]). 알데히드 옥시다아제 활성이 감소되거나 또는 제거되도록 변형될 수 있으며, 아이소부티르알데히드의 아이소부티르산의 전환을 감소시키거나 또는 제거할 수 있어서, 증가된 양의 아이소부티르알데히드가 아이소부탄올로의 전환에 이용가능하고 아이소부탄올 수율이 증가되도록 하는 적합한 알데히드 옥시다아제를 확인할 필요가 있다.

[0100] 유전자 엔지니어링된 효모에서의 부탄올의 생성을 위한 생합성 경로는 아세토락테이트의 2,3-다이하이드록시-3-아이소발레레이트 (DHIV)로의 전환을 포함하는데, 상기 DHIV는 후속적으로 부탄올로 전환된다. 도 10을 참조하라. 그러나, 전체 부탄올 생성을 감소시키는 이 경로에서의 부반응은 아세토락테이트의 2,3-다이하이드록시-2-메틸부티레이트 (DHMB)로의 전환이다. 실제, 본 출원인은 DHMB가 아이소부탄올 생성 경로의 효소(다이하이드록시산 데하이드라타아제 및 케톨-산 리덕토아이스머라아제)에 대하여 저해 효과를 가짐을 발견하였다. 효율적인 생합성 과정에 있어서, 아세토락테이트의 DHMB로의 전환을 방지할 필요가 있다.

[0101] 본 출원인은 아이소부탄올 생합성 경로와; i) 감소되거나 또는 제거된 알데히드 데하이드로게나아제 활성, ii) 감소되거나 또는 제거된 알데히드 옥시다아제 활성, iii) 감소되거나 또는 제거된 아세토락테이트 리덕타아제 활성, iv) 케톨-산 리덕토아이스머라아제 활성을 갖는 폴리펩티드를 코딩하는 이중성 폴리뉴클레오티드, 및 v) 케톨-산 리덕토아이스머라아제 활성을 갖는 이중성 폴리펩티드 중 적어도 하나를 포함하는 재조합 효모 숙주 세포를 제공함으로써 기술된 문제점들을 해결하였다. 또한, 본 출원인은 이러한 숙주 세포를 이용하여 부탄올을 제조하는 방법을 제공한다. 이러한 재조합 숙주 세포는 생합성 경로의 생성물 (예를 들어, 아이소부탄올, 1-부탄올, 또는 2-부탄올)의 생성을 증가시키고/시키거나 경로 중간체의 바람직하지 못한 부산물로의 전환을 감소시키거나 또는 제거하는 데 사용될 수 있다. 또한 본 출원인은 다양한 후보 효소를 평가하기에 적합한 스크리닝 전략을 제공하였다. 확인된 효소들은 생합성 경로의 생성물 (예를 들어, 아이소부탄올, 1-부탄올, 또는 2-부탄올)의 생성을 향상시키고/시키거나 경로 중간체의 바람직하지 못한 부산물로의 전환을 감소시키거나 또는 제거하기 위하여 변경될 수 있다.

[0102] 본 발명을 추가로 정의하기 위하여, 하기의 용어, 약어 및 정의가 명세서에 제공된다.

[0103] 본 명세서에 개시된 폴리펩티드와 관련하여 "유래된"은 지시된 유기체에 존재하는 KARI의 아미노산 서열과, 그 유기체의 유전 물질로부터 직접적으로 클로닝된 것을 기초로 하여 합성된 서열을 포함함이 이해될 것이다.

[0104] 본 명세서에 사용되는 바와 같이, 용어 "포함하다", "포함하는", "포함하고 있다", "포함하고 있는", "갖다", "갖는", "함유하다" 또는 "함유하는" 또는 이의 임의의 다른 변이형은 임의의 다른 정수 또는 정수의 군의 배제가 아닌, 기술된 정수 또는 정수의 군의 포함을 암시하는 것으로 이해될 것이며, 이는 비배제적이거나 또는 개방형(open-ended)인 것으로 의도된다. 예를 들어, 요소들의 목록을 포함하는 조성물, 혼합물, 공정, 방법, 용품, 또는 장치는 반드시 그러한 요소만으로 제한되지는 않고, 명확하게 열거되지 않거나 그러한 조성물, 혼합물, 공정, 방법, 용품, 또는 장치에 내재적인 다른 요소를 포함할 수 있다. 더욱이, 명백히 반대로 기술되지 않는다면, "또는"은 포괄적인 '또는'을 말하며 배타적인 '또는'을 말하는 것은 아니다. 예를 들어, 조건 A 또는 B는 하기 중 어느 하나에 의해 만족된다: A는 참 (또는 존재함)이고 B는 거짓 (또는 존재하지 않음), A는 거짓 (또는 존재하지 않음)이고 B는 참 (또는 존재함), A 및 B가 모두가 참 (또는 존재함).

[0105] 본 명세서에 사용되는 바와 같이, 명세서 및 특허청구범위 전체에 걸쳐 사용되는 용어 "이루어지다", 또는 "이루어진다" 또는 "이루어지는"과 같은 변이형은 임의의 기재된 정수 또는 정수의 군의 포함을 나타내지만, 추가의 정수 또는 정수의 군은 특정 방법, 구조 또는 조성물에 부가될 수 없음을 나타낸다.

[0106] 본 명세서에 사용되는 바와 같이, 명세서 및 특허청구범위 전체에 걸쳐 사용되는 용어 "본질적으로 이루어지다", 또는 "본질적으로 이루어진다" 또는 "본질적으로 이루어지는"과 같은 변이형은 임의의 기재된 정수 또는 정수의 군의 포함, 및 특정 방법, 구조 또는 조성물의 기본 또는 신규 특성을 실질적으로 변화시키지 않는 임의의 기재된 정수 또는 정수의 군의 선택적 포함을 나타낸다. M.P.E.P. § 2111.03을 참조하라.

[0107] 또한, 본 발명의 요소 또는 성분 앞의 부정 관사 "a" 및 "an"은 그 요소 또는 성분의 경우의 수, 즉, 출현의 수

에 관해서는 비제한적인 것으로 의도된다. 따라서, 부정 관사는 하나 또는 적어도 하나를 포함하는 것으로 판독되어야 하며, 요소 또는 성분의 단수형 단어는 그 수가 단수형을 명백하게 의미하는 것이 아니라면 복수형을 또한 포함한다.

- [0108] 본 명세서에 사용되는 바와 같이, 용어 "발명" 또는 "본 발명"은 비제한적인 용어이며, 특정 발명의 임의의 단일 실시 형태를 언급하는 것으로 의도되지 않고, 제시된 또는 이후에 보정되고 보충되는 특허청구범위에 또는 명세서에 기재된 모든 가능한 실시 형태들을 포함한다.
- [0109] 본 명세서에 사용되는 바와 같이, 이용된 발명의 성분 또는 반응물의 양을 수식하는 용어 "약"은, 예를 들어, 현실 세계에서 농축물 또는 용액을 제조하는 데 사용되는 전형적인 측정 및 액체 취급 절차; 이들 절차에서의 우연한 오류; 조성물을 제조하거나 또는 방법을 실시하기 위해 이용되는 성분의 제조, 공급원 또는 순도의 차이 등을 통하여 일어날 수 있는 수치적 양의 변동을 말한다. 용어 "약"은 또한 특정 초기 혼합물에서 생기는 조성물에 있어서 상이한 평형 조건으로 인해 달라지는 양을 포함한다. 용어 "약"으로 수식되든지 수식되지 않든지 간에, 특허청구범위는 그 양에 대한 등가의 것을 포함한다. 일 실시 형태에서, 용어 "약"은 보고된 수치의 10% 이내, 또는 보고된 수치의 5% 이내를 의미한다.
- [0110] 본 명세서에 사용되는 바와 같이, "상승적"은 조합에 의해 생성되는 상가적 효과 이상의 효과(greater-than-additive effect; 즉, 개별 효과들의 합보다 더 큰 효과) 또는 개별 효과들이 상가적인 것으로 예상되지 않을 때의 상가적 효과를 말한다. 또한 상기 용어는 필요한 다른 화합물이 덜 생성되게 하는 하나의 화합물의 첨가를 말한다.
- [0111] 본 명세서에 사용되는 바와 같이, 어구 "부탄올 생합성 경로"는 1-부탄올, 2-부탄올, 또는 아이소부탄올을 제조하는 효소적 경로를 말한다. 예를 들어, 아이소부탄올 생합성 경로는 본 명세서에 참고로 포함된 미국 특허 공개 제2007/0092957호에 개시되어 있다.
- [0112] 어구 "아이소부탄올 생합성 경로"는 아이소부탄올을 생성하는 효소적 경로를 말한다. 소정의 아이소부탄올 생합성 경로가 도 1에 예시되어 있고 본 명세서에 기재되어 있다. 때로, "아이소부탄올 생합성 경로"는 "아이소부탄올 생성 경로"와 동의어로 사용된다.
- [0113] 본 명세서에 사용되는 바와 같이, 용어 "부탄올"은 1-부탄올, 2-부탄올, 아이소부탄올 또는 이의 혼합물을 말한다. 아이소부탄올은 2-메틸-1-프로판올로도 공지되어 있다.
- [0114] "엔지니어링된 알코올 생성 경로" (예를 들어 엔지니어링된 부탄올 또는 아이소부탄올 생성 경로)를 포함하는 재조합 숙주 세포는 숙주 세포에 정상적으로 존재하는 것과는 상이한 방식으로 알코올을 생성하는 변형된 경로를 포함하는 숙주 세포를 말한다. 이러한 차이는 전형적으로 숙주 세포에 의해서는 생성되지 않는 알코올의 생성, 또는 증가된 또는 더욱 효율적인 생성을 포함한다.
- [0115] 본 명세서에 사용되는 바와 같이, 어구 "이종성 생합성 경로"는 효소들 중 적어도 하나가 그 생합성 경로를 포함하는 숙주 세포에 대하여 내인성이 아닌, 생성물을 생성하는 효소 경로를 말한다.
- [0116] 본 명세서에 사용되는 바와 같이, 용어 "추출제"는 발효 브로쓰로부터 부탄올을 추출하는 데 사용될 수 있는 하나 이상의 유기 용매를 말한다.
- [0117] 본 명세서에 사용되는 바와 같이, 어구 "유효 아이소부탄올 생산성"은 세포 1 g당 생성되는 아이소부탄올의 g 단위의 총량을 말한다.
- [0118] 본 명세서에 사용되는 바와 같이, 어구 "유효 역가"는 발효 배지 1 리터당 발효에 의해 생성되는 특정 알코올 (예를 들어 부탄올)의 총량을 말한다. 부탄올의 총량은 (i) 발효 배지 중 부탄올의 양; (ii) 유기 추출제로부터 회수된 부탄올의 양; 및 (iii) 가스 스트립핑(stripping)이 이용될 경우, 가스 상으로부터 회수되는 부탄올의 양을 포함한다.
- [0119] 본 명세서에 사용되는 바와 같이, 어구 "유효 속도(effective rate)"는 발효 시간당 발효 배지 1 리터당 발효에 의해 생성되는 부탄올의 총량을 말한다.
- [0120] 본 명세서에 사용되는 바와 같이, 어구 "유효 수율"은 생매에 의해 소비되는 발효가능한 탄소 기질의 단위당 생성되는 부탄올의 양을 말한다.
- [0121] 본 명세서에 사용되는 바와 같이, 용어 "분리"는 "회수"와 동의어이며, 초기 혼합물 중 화학적 화합물의 순도 또는 농도보다 더 큰 순도 또는 더 높은 농도의 화합물을 수득하기 위해 초기 혼합물로부터 상기 화합물을 제거

하는 것을 말한다.

- [0122] 본 명세서에 사용되는 바와 같이, 어구 "수성 상"은 발효 브로쓰를 수불산화성 유기 추출제와 접촉시켜 수득되는 2상 혼합물 중 수성 상을 말한다. 발효 추출을 포함하는 본 명세서에 기재된 방법의 실시 형태에서, 어구 "발효 브로쓰"는 또한 2상 발효 추출에 있어서 수성 상을 특징적으로 말한다.
- [0123] 본 명세서에 사용되는 바와 같이, 어구 "유기 상"은 발효 브로쓰와 수불산화성 유기 추출제를 접촉시켜 수득되는 2상 혼합물의 비수성 상을 말한다.
- [0124] 본 명세서에 사용되는 바와 같이, 용어 "PDC-", "PDC 녹아웃(knockout)" 또는 "PDC-KO"는 세포가 피루베이트 데카르복실라아제 효소 활성이 사실상 또는 완전히 결여되도록, 피루베이트 데카르복실라아제(PDC)를 코딩하는 유전자의 발현을 감소시키거나 또는 불활성화시키는 유전자 변형을 갖는 세포를 말한다. 세포가 하나 초과 발현(활성) PDC 유전자를 갖는다면, 활성 PDC 유전자 각각은 불활성화되거나 또는 최소 발현을 가지고, 이럼으로써 PDC- 세포를 생성할 수 있다.
- [0125] 용어 "폴리뉴클레오티드"는 단수의 핵산 뿐만 아니라 복수의 핵산도 포함하는 것으로 의도되며, 핵산 분자 또는 제작물, 예를 들어 메신저(messenger) RNA (mRNA) 또는 플라스미드 DNA(pDNA)를 말한다. 폴리뉴클레오티드는 미번역 5' 및 3' 서열 및 코딩 서열을 포함하는, 전장 cDNA 서열 또는 이의 단편의 뉴클레오티드 서열을 함유할 수 있다. 폴리뉴클레오티드는 미변형 RNA 또는 DNA, 또는 변형 RNA 또는 DNA일 수 있는 임의의 폴리리보뉴클레오티드 또는 폴리데옥시리보뉴클레오티드로 구성될 수 있다. 예를 들어, 폴리뉴클레오티드는 단일- 및 이중-가닥 DNA, 단일-가닥 영역과 이중-가닥 영역의 혼합물인 DNA, 단일- 및 이중-가닥 RNA 및 단일-가닥 영역과 이중-가닥 영역의 혼합물인 RNA, 단일-가닥 또는 더욱 전형적으로는 이중-가닥 또는 단일-가닥 영역과 이중-가닥 영역의 혼합물일 수 있는 DNA 및 RNA를 포함하는 하이브리드 분자로 구성될 수 있다. "폴리뉴클레오티드"는 화학적으로, 효소적으로 또는 대사적으로 변형된 형태를 포함한다.
- [0126] 폴리뉴클레오티드 서열은 "단리된"으로 지칭될 수 있으며, 여기서, 이는 그의 천연 환경으로부터 제거되었다. 예를 들어, 벡터에 함유된 다이하이드록시산 테하이드라타아제 활성을 갖는 폴리펩티드 또는 폴리펩티드 단편을 코딩하는 이중성 폴리뉴클레오티드는 본 발명의 목적에 있어서 단리된 것으로 간주된다. 단리된 폴리뉴클레오티드의 추가 예는 이중성 숙주 세포에서 유지된 재조합 폴리뉴클레오티드 또는 용액 중의, (부분적으로 또는 사실상) 정제된 폴리뉴클레오티드를 포함한다. 본 발명에 따른 단리된 폴리뉴클레오티드 또는 핵산은 합성에 의해 생성된 이러한 분자를 추가로 포함한다. DNA의 중합체 형태인 단리된 폴리뉴클레오티드 단편은 cDNA, 게놈 DNA 또는 합성 DNA의 하나 이상의 절편으로 이루어질 수 있다.
- [0127] 어구 "NAD(P)H 소비 분석"은 KARI 보조 인자, NAD(P)H가 효소 반응물로부터 사라지는 것의 측정을 포함하는, KARI 효소의 비활성의 측정을 위한 효소 분석을 말한다. 이러한 분석은 문헌[Aulabaugh and Schloss, Biochemistry 29: 2824-2830, 1990]에 기재되어 있으며, 이는 전체적으로 본 명세서에 참고로 포함된다.
- [0128] 용어 "NAD(P)H"는 NADH 또는 NADPH 중 어느 하나를 말한다.
- [0129] "KARI"는 효소 케톨-산 리덕토아이스머라아제의 약어이다.
- [0130] NADPH의 아데노실 2'-포스페이트와 관련하여 KARI 효소의 다양한 아미노산 잔기의 위치를 언급할 때 용어 "근접(close proximity)"은 상기 효소에 결합된 NADPH의 아데노실 2'-포스페이트의 인 원자의 약 4.5 Å 이내에 있는, 효소의 구조의 3차원 모델에서의 아미노산을 의미한다.
- [0131] 어구 "케톨-산 리덕토아이스머라아제" ("KARI"로 약칭됨), 및 "아세토하이드록시산 아이소머로리덕타아제"는 상호교환가능하게 사용되며, (S)-아세토라테이트의 2,3-다이하이드록시아이소발레레이트로의 반응을 촉매할 수 있는 효소로서 EC 번호 EC 1.1.1.86로 분류되는 것을 말한다 (문헌[Enzyme Nomenclature 1992, Academic Press, San Diego]). 본 명세서에 사용되는 바와 같이, 어구 "클래스 I 케톨-산 리덕토아이스머라아제 효소"는 전형적으로 330 내지 340개의 아미노산 잔기를 갖는 짧은 형태를 의미하며, 전형적으로 대략 490개의 잔기를 갖는, 클래스 II로 불리우는 긴 형태와는 다르다. 이들 효소는 이. 콜라이 (아미노산 서열 번호 942; 젠뱅크 등록 번호 NC_000913 REGION: 3955993.3957468), 비브리오 콜레라(*Vibrio cholerae*) (젠뱅크 등록 번호 NC_002505 REGION: 157441.158925), 슈도모나스 아에루기노사 (젠뱅크 등록 번호 NC_002516, REGION: 5272455.5273471), 및 슈도모나스 플루오레센스 (아미노산 서열 번호 943; 젠뱅크 등록 번호 NC_004129 REGION: 6017379.6018395)를 포함하지만 이에 한정되지 않는 다수의 공급원으로부터 입수가능하다. KARI 효소는 예를 들어 미국 특허 제7,910,342호 및 미국 특허 제8,129,162호와, 미국 특허 공개 제2010/0197519호에 기재되어 있으며, 상기 미

국 특허 및 미국 특허 공개 전부는 전체적으로 본 명세서에 참고로 포함된다.

- [0132] KARI는 다양한 유기체에서 발견되며, 아미노산 서열의 중간 비교에 의하면, 2가지 유형의 이 효소, 즉 진균류 및 대부분의 박테리아에서 발견되는 짧은 형태 (클래스 I) 및 식물에 전형적인 긴 형태 (클래스 II)가 있음이 나타났다. 클래스 I KARI는 전형적으로 330 내지 340개의 아미노산 잔기를 갖는다. 긴 형태의 KARI 효소는 약 490개의 아미노산 잔기를 갖는다. 그러나, 일부 박테리아, 예를 들어 에스케리키아 콜라이는 긴 형태를 보유하며, 여기서, 그 아미노산 서열은 식물에서 발견되는 것과 눈에 띄게 다르다. KARI는 *ilvC* 유전자에 의해 코딩되며, 최소 배지에서 이. 콜라이 및 다른 박테리아의 성장에 필수적인 효소이다. 클래스 II KARI는 일반적으로 225개 잔기의 N-말단 도메인 및 287개 잔기의 C-말단 도메인으로 이루어진다. NADPH-결합 부위를 함유하는 N-말단 도메인은 $\alpha\beta$ 구조를 가지며, 다른 피리딘 뉴클레오티드-의존성 옥시도리덕타아제에서 발견되는 도메인과 비슷하다. 상기 C-말단 도메인은 거의 전적으로 α -나선으로 이루어진다.
- [0133] 케톨-산 리덕토아이스머라아제 (KARI) 효소는 엔지니어링된 미생물을 이용하여 아이소부탄올을 생성하는 경로에서 유용하다 (본 명세서에 참고로 포함된 미국 특허 제7,851,188호 및 미국 특허 제7,993,889호).
- [0134] NADH를 이용할 수 있는 KARI는 가장 일반적으로 사용되는 미생물 세포에 있어서 기존의 당 분해 경로 및 다른 대사 경로에 의해 생성되는 NADH를 활용할 수 있으며, 향상된 아이소부탄올 생성으로 이어질 수 있다. 문헌 [Rane et al., Arch. Biochem. Biophys., 338: 83-89, 1997]에는 이. 콜라이로부터 단리된 케톨 산 리덕토아이스머라아제의 보조 인자 스위칭이 논의되어 있다. 미국 특허 공개 제2009/0163376호 및 미국 특허 공개 제2010/0197519호 (이들 각각은 전체적으로 본 명세서에 참고로 포함됨)에는 NADH를 이용할 수 있는 KARI 효소의 생성이 기재되어 있다. 미국 특허 공개 제2010/0143997호 (이는 전체적으로 본 명세서에 참고로 포함됨)에는 NADH에 대한 개선된 K_M 값을 갖는 이. 콜라이 변이체가 기재되어 있다.
- [0135] 어구 "케톨-산 리덕토아이스머라아제 활성" 및 "KARI 활성"은 기질의 생성물로의 전환, (S)-아세트락테이트의 2,3-다이하이드록시아이소발레레이트로의 전환을 촉매하는 능력을 말한다.
- [0136] 어구 "아세트락테이트 신타아제"는 피루베이트의 아세트락테이트 및 CO₂로의 전환을 촉매하는 효소를 말한다. 아세트락테이트는 2가지의 입체이성체 ((R) 및 (S))를 가지며; 상기 효소는 생물 시스템에 의해 만들어지는 (S)- 이성체를 선호한다. 소정의 아세트락테이트 신타아제가 EC 번호 2.2.1.6으로 공지되어 있다 (문헌[*Enzyme Nomenclature 1992*, Academic Press, San Diego]). 이들 효소는 바실러스 서브틸리스 (젠뱅크 번호: CAB15618, Z99122, 각각 미국 국립 생물공학 정보 센터(National Center for Biotechnology Information; NCBI)) 아미노산 서열, NCBI 뉴클레오티드 서열), 클렙시엘라 뉴모니아(*Klebsiella pneumoniae*) (젠뱅크 번호: AAA25079, M73842) 및 락토코커스 락티스 (젠뱅크 번호: AAA25161, L16975)를 포함하지만 이에 한정되지 않는 다수의 공급원으로부터 입수가 가능하다.
- [0137] 어구 "아세트하이드록시산 데하이드라타아제"는 2,3-다이하이드록시아이소발레레이트의 α -케토아이스발레레이트로의 전환을 촉매하는 효소를 말한다. 소정의 아세트하이드록시산 데하이드라타아제는 EC 번호 4.2.1.9로 공지되어 있다. 이들 효소는 이. 콜라이 (젠뱅크 번호: YP_026248, NC_000913, 에스. 세레비지에 (젠뱅크 번호: NP_012550, NC_001142), 엠. 마리팔루디스(*M. maripaludis*) (젠뱅크 번호: CAF29874, BX957219), 비. 서브틸리스 (젠뱅크 번호: CAB14105, Z99115), 락토코커스 락티스 (서열 번호 926), 및 스트렙토코커스 뮤탄스 (서열 번호 939)를 포함하지만 이에 한정되지 않는 수많은 미생물로부터 입수가 가능하다.
- [0138] 어구 "분지쇄 α -케토산 데카복실라제"는 α -케토아이스발레레이트의 아이소부티르알데히드 및 CO₂로의 전환을 촉매하는 효소를 말한다. 소정의 분지쇄 α -케토산 데카복실라아제는 EC 번호 4.1.1.72로 공지되어 있으며, 락토코커스 락티스 (젠뱅크 번호: AAS49166, AY548760; CAG34226, AJ746364, 살모넬라 타이피무륨(*Salmonella typhimurium*) (젠뱅크 번호: NP-461346, NC-003197), 클로스트리듐 아세트부틸리쿰(*Clostridium acetobutylicum*) (젠뱅크 번호: NP-149189, NC-001988), 마크로코커스 카세오라이티쿠스 (서열 번호 940), 및 리스테리아 그라이이 (서열 번호 941)를 포함하지만 이에 한정되지 않는 다수의 공급원으로부터 입수가 가능하다.
- [0139] 어구 "분지쇄 알코올 데하이드로게나아제"는 아이소부티르알데히드의 아이소부탄올로의 전환을 촉매하는 효소를 말한다. 소정의 분지쇄 알코올 데하이드로게나아제는 EC 번호 1.1.1.265로 공지되어 있지만, 다른 알코올 데하이드로게나아제 (구체적으로는, EC 1.1.1.1 또는 1.1.1.2)로 또한 분류될 수 있다. 이들 효소는 NADH(환원형 니코틴아미드 아데닌 다이뉴클레오티드) 및/또는 NADPH를 전자 공여체로 이용하며, 에스. 세레비지에 (젠뱅크 번호: NP_010656, NC_001136; NP_014051, NC_001145), 이. 콜라이 (젠뱅크 번호: NP_417484), 씨. 아세트부틸리쿰 (젠뱅크 번호: NP_349892, NC_003030), 비. 인디카(*B. indica*) (아미노산 서열 번호 945), 에이. 자일로

속시단스(*A. xylosoxidans*) (아미노산 서열 번호 944)를 포함하지만 이에 한정되지 않는 다수의 공급원으로부터 입수가 가능하다.

- [0140] 어구 "분지쇄 케토산 데하이드로게나아제"는 NAD^+ (니코틴아미드 아데닌 다이뉴클레오타이드)를 전자 수용체로서 사용하여 α -케토아이소발레레이트의 아이소부티릴-CoA (아이소부티릴-보조 인자 A)로의 전환을 촉매하는 효소를 말한다. 소정의 분지쇄 케토산 데하이드로게나아제는 EC 번호 1.2.4.4로 공지되어 있다. 이들 분지쇄 케토산 데하이드로게나아제는 4개의 서브유닛(subunit)을 포함하며, 모든 서브유닛으로부터의 서열은 비. 서브틸리스 (젠뱅크 번호: CAB14336, Z99116; CAB14335, Z99116; CAB14334, Z99116; 및 CAB14337, Z99116) 및 슈도모나스 퓨티다 (젠뱅크 번호: AAA65614, M57613; AAA65615, M57613; AAA65617, M57613; 및 AAA65618, M57613)를 포함하지만 이에 한정되지 않는 수많은 미생물로부터 입수가 가능하다.
- [0141] 본 명세서에 사용되는 바와 같이, "알데히드 데하이드로게나아제 활성"은 본 명세서에 제공된 예들을 비롯하여 알데히드 데하이드로게나아제의 생물학적 기능을 갖는 임의의 폴리펩티드를 말한다. 이러한 폴리펩티드는 알데히드의 산화(탈수소화)를 촉매하는 폴리펩티드를 포함한다. 이러한 폴리펩티드는 아이소부티르알데히드의 아이소부티르산으로의 전환을 촉매하는 폴리펩티드를 포함한다. 또한 이러한 폴리펩티드는 효소 위원회 번호 EC 1.2.1.3, EC 1.2.1.4 또는 EC 1.2.1.5에 상응하는 폴리펩티드를 포함한다. 이러한 폴리펩티드는 본 기술 분야에 잘 알려진 그리고 본 명세서에 개시된 방법에 의해 결정될 수 있다.
- [0142] 본 명세서에 사용되는 바와 같이, "알데히드 옥시다아제 활성"은 본 명세서에 제공된 예들을 비롯하여 알데히드 옥시다아제의 생물학적 기능을 갖는 임의의 폴리펩티드를 말한다. 이러한 폴리펩티드는 알데히드로부터의 카르복실산을 촉매하는 폴리펩티드를 포함한다. 이러한 폴리펩티드는 아이소부티르알데히드의 아이소부티르산으로의 전환을 촉매하는 폴리펩티드를 포함한다. 또한 이러한 폴리펩티드는 효소 위원회 번호 EC 1.2.3.1에 상응하는 폴리펩티드를 포함한다. 이러한 폴리펩티드는 본 기술 분야에 잘 알려진 그리고 본 명세서에 개시된 방법에 의해 결정될 수 있다.
- [0143] 본 명세서에 사용되는 바와 같이, "피루베이트 데카르복실라아제 활성"은 본 명세서에 제공된 예들을 비롯하여 피루베이트 데카르복실라아제 효소의 생물학적 기능을 갖는 임의의 폴리펩티드의 활성을 말한다. 이러한 폴리펩티드는 피루베이트의 아세트알데히드로의 전환을 촉매하는 폴리펩티드를 포함한다. 또한 이러한 폴리펩티드는 효소 위원회 번호 4.1.1.1에 상응하는 폴리펩티드를 포함한다. 이러한 폴리펩티드는 본 기술 분야에 잘 알려진 그리고 본 명세서에 개시된 방법에 의해 결정될 수 있다. 피루베이트 데카르복실라이트 활성을 갖는 폴리펩티드는 예로서 PDC1, PDC5, PDC6 또는 이의 임의의 조합일 수 있다.
- [0144] 본 명세서에 사용되는 바와 같이, "아세토락테이트 리덕타아제 활성"은 아세토락테이트의 DHMB로의 전환을 촉매하는 능력을 갖는 임의의 폴리펩티드의 활성을 말한다. 이러한 폴리펩티드는 본 기술 분야에 잘 알려진 그리고 본 명세서에 개시된 방법에 의해 결정될 수 있다.
- [0145] 본 명세서에 사용되는 바와 같이, "DHMB"는 2,3-다이하이드록시-2-메틸 부티레이트를 말한다. DHMB는 2S, 3S 배열을 갖는 "패스트(fast) DHMB", 및 2S, 3R 배열을 갖는 "슬로우(slow) DHMB"를 포함한다. 문헌[Kaneko et al., *Phytochemistry* 39: 115-120 (1995)]을 참조하는데, 상기 문헌은 전체적으로 본 명세서에 참고로 포함되며, 패스트 DHMB를 안글리세르산으로 지칭하고 슬로우 DHMB를 티글리세르산으로 지칭한다.
- [0146] 본 명세서에 사용되는 바와 같이, "활성 감소"는 폴리펩티드의 공지된 생물학적 활성이, 활성 감소로 이어지는 변화 이전의 그 폴리펩티드의 상기 생물학적 활성과 비교할 때 임의로 측정가능하게 감소되는 것을 말한다. 이러한 변화는 폴리펩티드 또는 폴리펩티드를 코딩하는 폴리뉴클레오타이드의 변형을 포함할 수 있으며, 이는 본 명세서에 기재된 바와 같다. 본 명세서에 개시된 폴리펩티드의 활성 감소는 본 기술 분야에 잘 알려져 있는 그리고 본 명세서에 개시된 방법에 의해 결정될 수 있다.
- [0147] 본 명세서에 사용되는 바와 같이, "활성 제거"는 폴리펩티드의 공지된 생물학적 활성이, 활성 제거로 이어지는 변화 이전의 그 폴리펩티드의 상기 생물학적 활성과 비교할 때 완전히 무효화됨을 말한다. 이러한 변화는 폴리펩티드 또는 폴리펩티드를 코딩하는 폴리뉴클레오타이드의 변형을 포함할 수 있으며, 이는 본 명세서에 기재된 바와 같다. 활성 제거는 폴리펩티드의 생물학적 활성이, 활성 제거로 이어지는 변화 이전의 그 폴리펩티드의 상기 생물학적 활성과 비교할 때 측정가능하지 않은 것을 포함한다. 본 명세서에 개시된 폴리펩티드의 활성 제거는 본 기술 분야에 잘 알려져 있는 그리고 본 명세서에 개시된 방법에 의해 결정될 수 있다.
- [0148] 용어 "탄소 기질" 또는 "발효가능한 탄소 기질"은 본 발명의 숙주 유기체에 의하여 대사될 수 있는 탄소 공급원 및 특히 단당류, 올리고당류, 다당류 및 일원자-탄소(one-carbon) 기질 또는 이들의 혼합물로 이루어진 군으로

부터 선택된 탄소 공급원을 말한다. 탄소 기질의 비제한적 예는 본 명세서에 제공되며, 단당류, 올리고당류, 다당류, 에탄올, 락테이트, 석시네이트, 글리세롤, 이산화탄소, 메탄올, 글루코스, 프룩토스, 수크로스, 자일로스, 아라비노스, 텍스트로스 또는 이의 혼합물을 포함하지만, 이에 한정되지 않는다. 다른 탄소 기질은 에탄올, 락테이트, 석시네이트 또는 글리세롤을 포함할 수 있다.

[0149] 본 명세서에 사용되는 바와 같이, "발효 브로쓰"는 물, 당류 (발효가능한 탄소 공급원), 용해된 고체 (존재할 경우), 알코올을 생성하는 미생물, 생성물 알코올 및 발효 용기 내에서 유지되는 물질의 모든 다른 구성 성분의 혼합물을 의미하고, 상기 발효 용기 내에서 생성물 알코올은 존재하는 미생물에 의해 당류의 알코올, 물 및 이산화탄소(CO_2)로의 반응에 의해 만들어지고 있다. 때로, 본 명세서에 사용되는 바와 같이, 어구 "발효 배지" 및 "발효 혼합물"은 "발효 브로쓰"와 동의어로 사용될 수 있다.

[0150] 본 명세서에 사용되는 바와 같이, "바이오매스"는 임의의 셀룰로오스 또는 리그노셀룰로오스 물질과, 셀룰로오스를 포함하며 헤미셀룰로오스, 리그닌, 전분, 올리고당류, 이당류 및/또는 단당류를 선택적으로 추가로 포함하는 물질을 비롯하여, 발효가능한 당을 제공하는 가수분해성 전분을 함유하는 천연 생성물을 말한다. 바이오매스는 또한 단백질 및/또는 지질과 같은 추가의 성분을 포함할 수 있다. 바이오매스는 단일 공급원으로부터 유래할 수 있거나, 또는 바이오매스는 1가지 초과 공급원으로부터 유래된 혼합물을 포함할 수 있다. 예를 들어, 바이오매스는 옥수수 속대 및 옥수수 대의 혼합물, 또는 풀과 잎의 혼합물을 포함할 수 있다. 바이오매스는 바이오에너지 작물, 농업적 잔사, 도시 고형 폐기물(municipal solid waste), 산업 고형 폐기물, 종이 제조로부터의 슬러지, 정원 폐기물(yard waste), 나무 및 삼림 폐기물을 포함하지만, 이에 한정되지 않는다. 바이오매스의 예는, 옥수수 낱알, 옥수수 속대, 옥수수 잔사, 예를 들어 옥수수 껍질, 옥수수 대, 풀, 밀, 호밀, 밀짚, 보리, 보릿짚, 건초, 벼짚, 스위치그래스(switchgrass), 폐지, 사탕수수 버가스(bagasse), 수수, 대두, 낱알의 제분으로부터 수득된 성분, 나무, 가지, 뿌리, 잎, 나뭇조각, 톱밥, 관목 및 떨기나무, 야채, 과일, 꽃, 동물 퇴비 및 이들의 혼합물을 포함하지만, 이에 한정되지 않는다.

[0151] 본 명세서에 사용되는 바와 같이, "공급 원료"는 발효가능한 탄소 공급원을 함유하는 생성물을 의미한다. 적합한 공급 원료는 호밀, 밀, 옥수수, 사탕수수 및 이의 혼합물을 포함하지만, 이에 한정되지 않는다.

[0152] 본 명세서에 사용되는 바와 같이, 어구 "호기성 조건"은 산소의 존재 하의 성장 조건을 의미한다.

[0153] 본 명세서에 사용되는 바와 같이, 어구 "미세호기성 조건"은 낮은 수준의 산소 (즉, 보통의 대기중 산소 수준 미만)가 있는 성장 조건을 의미한다.

[0154] 본 명세서에 사용되는 바와 같이, 어구 "혐기성 조건"은 산소의 부재 하의 성장 조건을 의미한다.

[0155] 본 명세서에 사용되는 바와 같이, 용어 "비활성"은 주어진 양의 단백질 중의 활성의 유닛으로 정의된다. 따라서, 비활성은 직접적으로 측정되지 않으며, 1) 효소 샘플 1 ml당 유닛 단위의 활성을 2) 그 샘플 중 단백질의 농도로 나눔으로써 계산되고, 따라서 비활성은 유닛/mg으로 표현되는데, 여기서, 효소 유닛은 형성된 생성물의 몰/분으로 정의된다. 순수한 완전 활성 효소의 샘플의 비활성은 그 효소의 특징이다. 단백질의 혼합물의 샘플의 비활성은 관심대상의 활성 효소로 구성된 그 샘플 중의 단백질의 상대 분율의 측정치이다.

[0156] 용어 " k_{cat} " 및 " K_M "은 당업자에게 공지되어 있으며, 문헌[Enzyme Structure and Mechanism, 2nd ed. (Ferst; W.H. Freeman Press, NY, 1985; pp 98-120)]에 기재되어 있다. 미하엘리스 상수(Michaelis constant), K_M 은 최대 속도의 절반에 이르게 되는 기질의 농도이다. 흔히 "턴오버수(turnover number)"로 불리는 용어 " k_{cat} "은 단위 시간당 활성 부위당 생성물로 전환되는 기질 분자의 최대 수, 또는 단위 시간당 효소가 턴오버하는 횟수로서 정의된다. $k_{cat} = V_{max}/[E]$ 이며, 여기서, $[E]$ 는 효소 농도이다 (페르스트(Ferst)의 상기 문헌). 본 명세서에서, 어구 "총 턴오버" 및 "총 턴오버수"는 KARI 효소와 기질의 반응에 의해 형성되는 생성물의 양을 말하는 데 사용된다.

[0157] 어구 "촉매 효율"은 효소의 k_{cat}/K_M 으로 정의된다. 촉매 효율을 이용하여 기질에 대한 효소의 특이성을 정량화한다.

[0158] 어구 "단리된 핵산 분자", "단리된 핵산 단편" 및 "유전자 제작물"은 상호교환가능하게 사용될 것이며, 합성, 비천연 또는 변경 뉴클레오타이드 염기를 선택적으로 함유하는, 단일 가닥 또는 이중 가닥인 RNA 또는 DNA의 중합체를 의미할 것이다. DNA의 중합체 형태인 단리된 핵산 단편은 cDNA, 게놈 DNA 또는 합성 DNA의 하나 이상의

절편으로 이루어질 수 있다.

[0159] 용어 "아미노산"은 "단백질" 또는 폴리펩티드의 기본 화학 구조 단위를 말한다. 하기의 약어는 특정 아미노산을 확인하기 위해 본 명세서에서 사용된다.

아미노산	3문자 약어	1문자 약어
알라닌	Ala	A
아르기닌	Arg	R
아스파라긴	Asn	N
아스파르트산	Asp	D
시스테인	Cys	C
글루타민	Gln	Q
글루탐산	Glu	E
글리신	Gly	G
히스티딘	His	H
류신	Leu	L
라이신	Lys	K
메티오닌	Met	M
페닐알라닌	Phe	F
프롤린	Pro	P
세린	Ser	S
트레오닌	Thr	T
트립토판	Trp	W
티로신	Tyr	Y
발린	Val	V

[0160]

[0161] 용어 "유전자"는 선택적으로, 코딩 서열에 앞서는 조절 서열(5' 비코딩 서열) 및 코딩 서열에 뒤따르는 조절 서열(3' 비코딩 서열)을 포함하는, 특정 단백질로 발현될 수 있는 핵산 단편을 말한다. "천연 유전자"는 그 자신의 조절 서열과 함께 자연에서 발견되는 유전자를 말한다. "키메라 유전자"는 자연에서 함께 발견되지 않는 조절 및 코딩 서열을 포함하는, 천연 유전자가 아닌 임의의 유전자를 말한다. 따라서, 키메라 유전자는 상이한 공급원으로부터 유래된 조절 서열과 코딩 서열, 또는 동일한 공급원으로부터 유래되었지만, 자연에서 발견되는 것과 상이한 방식으로 배열된 조절 서열과 코딩 서열을 포함할 수 있다. "내인성 유전자"는 미생물의 게놈에서 그의 자연적 위치로 있는 천연 유전자를 말한다. "외래" 유전자는 숙주 미생물 내에서 정상적으로 발견되는 것이 아니라 유전자 전달에 의해 숙주 미생물 내로 도입된 유전자를 말한다. 외래 유전자는 비천연 미생물 내로 삽입된 천연 유전자, 또는 키메라 유전자를 포함할 수 있다. "트랜스유전자(transgene)"는 형질전환 절차에 의해 게놈 내로 도입된 유전자이다.

[0162] 본 명세서에 사용되는 바와 같이, "천연"은 자연에서 발견되는, 존재할 경우 그 자신의 조절 서열을 포함하는 폴리뉴클레오타이드, 유전자 또는 폴리펩티드의 형태를 말한다.

[0163] 본 명세서에 사용되는 바와 같이, 어구 "코딩 서열" 또는 "코딩 영역"은 특정 아미노산 서열을 코딩하는 DNA 서열을 말한다. "적합한 조절 서열"은 코딩 서열의 상류(5' 비코딩 서열), 이의 내부 또는 이의 하류(3' 비코딩 서열)에 위치하고, 전사, RNA 프로세싱 또는 안정성, 또는 관련된 코딩 서열의 번역에 영향을 미치는 뉴클레오타이드 서열을 말한다. 조절 서열은 프로모터, 번역 리더 서열(translation leader sequence), 인트론, 폴리아데닐화 인식 서열, RNA 프로세싱 부위, 이펙터(effector) 결합 부위 및 스템-루프 구조(stem-loop structure)를 포함할 수 있다.

[0164] 본 명세서에 사용되는 바와 같이, "내인성"은 유기체의 게놈 내의 또는 유기체 내의 그의 자연적 위치의 폴리뉴클레오타이드, 유전자 또는 폴리펩티드의 천연 형태를 말한다. "내인성 폴리뉴클레오타이드"는 유기체의 게놈 내의 그의 자연적 위치의 천연 폴리뉴클레오타이드를 포함한다. "내인성 유전자"는 유기체의 게놈 내의 그의 자연적 위치의 천연 유전자를 포함한다. "내인성 폴리펩티드"는 유기체의 게놈 내의 그의 자연적 위치의 천연 폴리뉴클레오타이드 또는 유전자로부터 전사되고 번역되는, 유기체 내의 그의 자연적 위치의 천연 폴리펩티드를 포함한다.

[0165] 폴리뉴클레오타이드, 유전자 또는 폴리펩티드에 관하여 사용되는 경우, 용어 "이종성"은 숙주 유기체 내에서 보통

관찰되지 않는 폴리뉴클레오티드, 유전자 또는 폴리펩티드를 말한다. 또한, "이종성"은 상응하는 천연 유전자와 상이한 형태로, 예를 들어, 유기체의 게놈 내의 그의 자연적 위치에 있지 않게, 소스 유기체 내에 재도입되는 천연 코딩 영역 또는 이의 부분을 포함한다. 이종성 폴리뉴클레오티드 또는 유전자는 예를 들어, 유전자 전달에 의해 숙주 유기체 내로 도입될 수 있다. 이종성 유전자는 천연 숙주 내로 재도입되는 비천연 조절 영역을 갖는 천연 코딩 영역을 포함할 수 있다. 예를 들어, 이종성 유전자는 천연 숙주 내로 재도입되는 비천연 조절 영역을 포함하는 키메라 유전자의 일부분인 천연 코딩 영역을 포함할 수 있다. "이종성 폴리펩티드"는 상응하는 천연 폴리펩티드와 상이한 형태로 소스 유기체 내로 재도입되는 천연 폴리펩티드를 포함한다.

[0166] "트랜스진"은 형질전환 절차에 의해 게놈 내로 도입된 유전자이다.

[0167] 본 명세서에 사용되는 바와 같이, 용어 "변형"은 폴리뉴클레오티드에 의해 코딩되는 폴리펩티드의 감소되거나 또는 제거된 활성으로 이어지는, 본 명세서에 개시된 폴리뉴클레오티드의 변화와, 폴리펩티드의 감소되거나 또는 제거된 활성으로 이어지는, 본 명세서에 개시된 폴리펩티드의 변화를 말한다. 이러한 변화는 본 기술 분야에 잘 알려진 방법에 의해 이루어질 수 있으며, 이는 결실, 돌연변이 (예를 들어, 자연 돌연변이 유발, 랜덤 돌연변이 유발, 돌연변이 유발 유전자에 의해 야기되는 돌연변이 유발, 또는 트랜스포존(transposon) 돌연변이 유발), 치환, 삽입, 하향 조절, 세포 위치의 변경, 폴리뉴클레오티드 또는 폴리펩티드의 상태의 변경 (예를 들어, 메틸화, 포스포릴화 또는 유비퀴틴화), 보조 인자의 제거, 안티센스 RNA/DNA의 도입, 간섭 RNA/DNA의 도입, 화학적 변형, 공유적 변형, UV 또는 X선을 이용한 조사, 상동 재조합, 유사 분열 재조합, 프로모터 대체법 및/또는 이의 조합을 포함하지만, 이에 한정되지 않는다. 어떤 뉴클레오티드 또는 아미노산 잔기가 변형될 수 있는지를 결정함에 있어서의 지침은 특정 폴리뉴클레오티드 또는 폴리펩티드의 서열을 상동성 폴리뉴클레오티드 또는 폴리펩티드의 것, 예를 들어 효모 또는 박테리아의 서열과 비교하고, 높은 상동성의 영역 (보존된 영역) 또는 콘센서스(consensus) 서열에서 이루어지는 변형의 수를 최대화함으로써 찾아질 수 있다.

[0168] 어구 "재조합 유전자 발현 요소"는 단백질에 대한 선행하는 (5' 비코딩 서열) 및 뒤따르는 (3' 종결 서열) 코딩 서열을 포함하는 하나 이상의 특정 단백질을 발현하는 핵산 단편을 말한다. 키메라 유전자는 재조합 유전자 발현 요소이다. 오페론의 코딩 영역은 작동가능하게 연결된 프로모터 및 종결 영역과 함께 재조합 유전자 발현 요소를 형성할 수 있다.

[0169] "조절 서열"은 코딩 서열의 상류 (5' 비코딩 서열), 그의 내부 또는 그의 하류 (3' 비코딩 서열)에 위치하고, 전사, RNA 프로세싱 또는 안정성 또는 관련된 코딩 서열의 번역에 영향을 미치는 뉴클레오티드 서열을 말한다. 조절 서열은 프로모터, 인핸서, 오퍼레이터, 리프레서, 전사 종결 신호, 번역 리더 서열, 인트론, 폴리아데닐화 인식 서열, RNA 프로세싱 부위, 이펙터 결합 부위 및 스템-루프 구조를 포함할 수 있다.

[0170] 용어 "프로모터"는 코딩 서열 또는 기능성 RNA의 발현을 제어할 수 있는 핵산 서열을 말한다. 일반적으로, 코딩 서열은 프로모터 서열에 대해 3'에 위치한다. 프로모터는, 천연 유전자로부터 그 전체가 유래될 수 있거나, 자연에서 발견되는 상이한 프로모터로부터 유래된 상이한 요소로 구성될 수 있거나, 또는 심지어 합성 핵산 절편을 포함할 수 있다. 상이한 프로모터는 상이한 조직 또는 세포 유형에서, 또는 상이한 발생 단계에서, 또는 상이한 환경적 또는 생리학적 조건에 응답하여 유전자의 발현을 유도할 수 있음을 당업자는 이해할 것이다. 대부분의 시점에 대부분의 세포 유형에서 유전자의 발현을 야기하는 프로모터를 통상적으로 "구성성(constitutive) 프로모터"라고 지칭한다. 반면, "유도성 프로모터"는 프로모터가 프로모터-특이적 신호 또는 분자에 의하여 유도되거나 켜지는 경우 유전자가 발현되게 한다. 대부분의 경우에 조절 서열들의 정확한 경계는 완전히 정의된 것은 아니기 때문에, 상이한 길이의 DNA 단편들은 동일한 프로모터 활성을 가질 수 있음이 추가로 인지된다. 예를 들어, "FBA1 프로모터"는 FBA1 유전자의 프로모터 영역으로부터 유래된 단편을 지칭하는데 사용될 수 있음이 이해될 것이다.

[0171] 본 명세서에 사용되는 바와 같이, 어구 "종결 서열"은 코딩 서열의 하류에 위치하는 DNA 서열을 말한다. 이는 폴리아데닐화 인식 서열 및 mRNA 프로세싱 또는 유전자 발현에 영향을 미칠 수 있는 조절 신호를 코딩하는 다른 서열을 포함한다. 폴리아데닐화 신호는 통상적으로 mRNA 전구체의 3' 말단에 폴리아데닐산 트래кт(tract)를 추가하는 데에 영향을 미치는 것을 특징으로 한다. 3' 영역은 전사, RNA 프로세싱 또는 안정성, 또는 관련 코딩 서열의 번역에 영향을 미칠 수 있다. 또한, 대부분의 경우에 조절 서열들의 정확한 경계는 완전히 정의된 것은 아니기 때문에, 상이한 길이의 DNA 단편들이 동일한 종결 서열 활성을 가질 수 있음이 인지된다. 예를 들어, "CYC1 종결 서열"은 CYC1 유전자의 종결 영역으로부터 유래된 단편을 말하는 데 사용될 수 있다.

[0172] 어구 "작동가능하게 연결된"은, 하나의 기능이 다른 것에 의해 영향을 받도록 단일 핵산 단편 상에서 핵산 서열들이 회합됨을 말한다. 예를 들어, 프로모터가 코딩 서열의 발현에 영향을 미칠 수 있는 경우에, 프로모터는

그 코딩 서열에 작동가능하게 연결된다(즉, 코딩 서열은 프로모터의 전사 제어 하에 있다). 코딩 서열은 센스 또는 안티센스 배향으로 조절 서열에 작동가능하게 연결될 수 있다.

[0173] 본 명세서에 사용되는 바와 같이, 용어 "발현"은, 본 발명의 핵산 단편으로부터 유래된 센스(mRNA) 또는 안티센스 RNA의 전사 및 안정한 축적을 말한다. 발현은 또한 mRNA의 폴리펩티드로의 번역을 말할 수 있다.

[0174] 본 명세서에 사용되는 바와 같이, 용어 "과다발현"은 동일하거나 관련된 유전자의 내인성 발현보다 더 높은 발현을 말한다. 이중성 유전자는, 그의 발현이 비견되는 내인성 유전자의 것보다 더 높을 경우 과다발현되는 것이다.

[0175] 용어 과다발현은 숙주 세포에서 핵산 또는 단백질의 수준의 증가를 말한다. 따라서, 과다발현은 숙주 세포에서 내인성 서열의 전사 또는 번역 수준의 증가에서 생길 수 있거나 또는 숙주 세포 내로의 이중성 서열의 도입에서 생길 수 있다. 또한 과발현은 핵산 또는 단백질 서열의 안정성의 증가에서 생길 수 있다.

[0176] 본 명세서에 사용되는 바와 같이, 용어 "형질전환"은 유전적으로 안정적인 유전성을 야기하는, 숙주 미생물의 게놈 내로의 핵산 단편의 전달을 말한다. 형질전환된 핵산 단편을 함유하는 숙주 미생물은 "트랜스제닉(transgenic)" 또는 "재조합" 또는 "형질전환" 미생물로 지칭된다.

[0177] 용어 "플라스미드", "벡터" 및 "카세트"는, 세포의 중추적인 대사의 일부가 아닌 유전자를 흔히 운반하며 통상적으로 원형의 이중 가닥 DNA 단편의 형태인 염색체 이외의 요소(extra chromosomal element)를 말한다. 이러한 요소는 임의의 공급원으로부터 유래된, 단일 가닥 또는 이중 가닥 DNA 또는 RNA의 선형 또는 원형의, 자가 복제 서열, 게놈 통합 서열, 파지 또는 뉴클레오티드 서열일 수 있으며, 여기서 많은 뉴클레오티드 서열들은 독특한 제작물로 연결되거나 재조합되었는데, 상기 제작물은 프로모터 단편 및 선택된 유전자 생성물의 DNA 서열을 적절한 3' 미번역 서열과 함께 세포 내로 도입할 수 있다. "형질전환 카세트"는 외래 유전자를 함유하고, 그 외래 유전자에 더하여 특정 숙주 세포의 형질전환을 용이하게 하는 요소들을 갖는 특정 벡터를 말한다. "발현 카세트"는 외래 유전자를 함유하며, 외래 유전자에 더하여, 외래 숙주에서 외래 유전자의 발현이 향상되게 하는 요소를 갖는 특정 벡터를 말한다.

[0178] 어구 "부위 포화 라이브러리"는 20가지 이하의 모든 가능한 아미노산을 한꺼번에 갖는 특정 아미노산 위치에서의 랜덤 치환을 포함하는 라이브러리를 말한다.

[0179] 어구 "오류 발생이 쉬운(error-prone) PCR"은 특정 뉴클레오티드 서열에서의 돌연변이들의 랜덤화된 라이브러리를 생성하는 불완전한 PCR 또는 '슬로피(sloppy)' PCR 반응 조건을 부과함으로써 랜덤 복제 오류를 부과하는 것을 말한다.

[0180] 본 명세서에서 사용된 어구 "코돈 축퇴성(codon degeneracy)"은 코딩된 폴리펩티드의 아미노산 서열에 영향을 주지 않으면서 뉴클레오티드 서열의 변이를 허용하는 유전 암호의 성질을 말한다. 주어진 아미노산을 특정하기 위한 뉴클레오티드 코돈의 사용에 있어서 특정 숙주 세포에서 나타나는 "코돈-바이어스(codon-bias)"가 당업자에게 주지되어 있다. 그러므로, 숙주 세포 내의 발현을 개선하기 위한 유전자를 합성할 경우에는, 그의 코돈 사용 빈도가 숙주 세포가 선호하는 코돈의 사용 빈도에 근접하도록 유전자를 고안하는 것이 바람직하다.

[0181] 다양한 숙주의 형질전환을 위한 핵산 분자의 코딩 영역 또는 유전자에 관련된 용어 "코돈-최적화"는, DNA에 의해 코딩되는 폴리펩티드를 변경하지 않으면서 숙주 유기체의 통상적 코돈 이용을 반영하도록 핵산 분자의 코딩 영역 또는 유전자 내의 코돈을 변경함을 말한다. 이러한 최적화는 적어도 하나 또는 하나 초과 또는 상당한 수의 코돈을 유기체의 유전자에서 더욱 자주 사용되는 하나 이상의 코돈으로 대체하는 것을 포함한다.

[0182] 임의의 폴리펩티드 쇄의 아미노산을 코딩하는 코돈을 포함하는 뉴클레오티드 서열 내의 일탈은 유전자를 코딩하는 서열 내의 변이를 가능하게 한다. 각 코돈이 3개의 뉴클레오티드로 이루어지고, DNA를 포함하는 뉴클레오티드가 4개의 특정 염기에 제한되기 때문에, 뉴클레오티드의 64개의 가능한 조합이 존재하며, 이 중 61개는 아미노산을 코딩한다 (나머지 3개 코돈은 번역을 종결하는 신호를 코딩한다). 코돈이 어떠한 아미노산을 코딩하는지를 보여주는 "유전자 암호"는 본 명세서에서 표 2a에 나타내었다. 결과적으로, 많은 아미노산이 1개 초과 코돈에 의해 지정된다. 예를 들어, 아미노산 알라닌 및 프롤린은 4개의 트리플렛(triplet)에 의해 코딩되고, 세린 및 아르기닌은 6개의 트리플렛에 의해 코딩되는 한편, 트립토판 및 메티오닌은 단지 하나의 트리플렛에 의해 코딩된다. 이러한 축퇴성은 DNA에 의해 코딩되는 단백질의 아미노산 서열을 변경시키지 않고 넓은 범위에 걸쳐 DNA 염기 조성을 달라지게 할 수 있다.

[0183] [표 2a]

표준 유전자 암호

	T	C	A	G
T	TTT Phe (F) TTC " TTA Leu (L) TTG "	TCT Ser (S) TCC " TCA " TCG "	TAT Tyr (Y) TAC " TAA 종결 TAG 종결	TGT Cys (C) TGC TGA 종결 TGG Trp (W)
C	CTT Leu (L) CTC " CTA " CTG "	CCT Pro (P) CCC " CCA " CCG "	CAT His (H) CAC " CAA Gln (Q) CAG "	CGT Arg (R) CGC " CGA " CGG "
A	ATT Ile (I) ATC " ATA " ATG Met (M)	ACT Thr (T) ACC " ACA " ACG "	AAT Asn (N) AAC " AAA Lys (K) AAG "	AGT Ser (S) AGC " AGA Arg (R) AGG "
G	GTT Val (V) GTC " GTA " GTG "	GCT Ala (A) GCC " GCA " GCG "	GAT Asp (D) GAC " GAA Glu (E) GAG "	GGT Gly (G) GGC " GGA " GGG "

[0184]

[0185] 많은 유기체는 성장 중인 펩티드 쇠 내의 특정 아미노산의 삽입을 코딩하는 특정 코돈의 사용에 대한 바이어스를 나타낸다. 유기체 간의, 코돈 선호 또는 코돈 바이어스, 코돈 사용의 차이는 유전자 암호의 축퇴성에 의해 제공되며, 많은 유기체 사이에서 널리 입증되어 있다. 코돈 바이어스는 흔히 메신저 RNA (mRNA)의 번역의 효율과 상관되며, 이는 다시 특히, 번역되는 코돈의 성질 및 특정 트랜스퍼(transfer) RNA (tRNA) 분자의 이용가능성에 의존적인 것으로 여겨진다. 세포에서 선택된 tRNA의 우세함은 일반적으로 펩티드 합성에서 가장 자주 사용되는 코돈의 반영이다. 따라서, 유전자는 코돈 최적화를 기준으로 주어진 유기체에서의 최적의 유전자 발현을 위해 맞춤화될 수 있다.

[0186] 매우 다양한 동물, 식물 및 미생물 종에 이용가능한 다수의 유전자 서열이 주어질 경우, 코돈 사용의 상대적 빈도를 계산할 수 있다. 코돈 사용 표는 예를 들어 www.kazusa.or.jp/codon/ (2008년 3월 20일 방문)에서 이용가능한 "코돈 사용 데이터베이스(Codon Usage Database)"에서 용이하게 이용가능하며, 이들 표는 다양한 방식으로 수정될 수 있다. 문헌[Nakamura, Y., *et al. Nucl. Acids Res.* 28:292 (2000)]을 참조하라. 젠뱅크 릴리즈(GenBank Release) 128.0 [2002년 2월 15일]으로부터 계산된 효모에 대한 코돈 사용 표는 하기의 표 2로서 재현되어 있다. 이러한 표는 mRNA 명명법을 사용하며, 따라서, DNA에서 관찰되는 티민 (T) 대신에 상기 표는 RNA에서 관찰되는 우라실 (U)을 사용한다. 상기 표 2b는 빈도가 모든 64개 코돈에 대해서라기보다는 오히려 각 아미노산에 대해서 계산되도록 수정되었다.

[0187] [표 2b]

사카로마이세스 세레비자이에 있어서의 코돈 사용 표

아미노산	코돈	번호	1000 개당 빈도
Phe	UUU	170666	26.1
Phe	UUC	120510	18.4
Leu	UUA	170884	26.2
Leu	UUG	177573	27.2
Leu	CUU	80076	12.3
Leu	CUC	35545	5.4
Leu	CUA	87619	13.4
Leu	CUG	68494	10.5
Ile	AUU	196893	30.1
Ile	AUC	112176	17.2
Ile	AUA	116254	17.8
Met	AUG	136805	20.9
Val	GUU	144243	22.1
Val	GUC	76947	11.8
Val	GUA	76927	11.8
Val	GUG	70337	10.8
Ser	UCU	153557	23.5
Ser	UCC	92923	14.2
Ser	UCA	122028	18.7
Ser	UCG	55951	8.6
Ser	AGU	92466	14.2
Ser	AGC	63726	9.8
Pro	CCU	88263	13.5
Pro	CCC	44309	6.8
Pro	CCA	119641	18.3
Pro	CCG	34597	5.3
Thr	ACU	132522	20.3
Thr	ACC	83207	12.7
Thr	ACA	116084	17.8
Thr	ACG	52045	8.0
Ala	GCU	138358	21.2
Ala	GCC	82357	12.6
Ala	GCA	105910	16.2
Ala	GCG	40358	6.2
Tyr	UAU	122728	18.8
Tyr	UAC	96596	14.8
His	CAU	89007	13.6
His	CAC	50785	7.8

[0188]

아미노산	코돈	번호	1000 개당 빈도
Gln	CAA	178251	27.3
Gln	CAG	79121	12.1
Asn	AAU	233124	35.7
Asn	AAC	162199	24.8
Lys	AAA	273618	41.9
Lys	AAG	201361	30.8
Asp	GAU	245641	37.6
Asp	GAC	132048	20.2
Glu	GAA	297944	45.6
Glu	GAG	125717	19.2
Cys	UGU	52903	8.1
Cys	UGC	31095	4.8
Trp	UGG	67789	10.4
Arg	CGU	41791	6.4
Arg	CGC	16993	2.6
Arg	CGA	19562	3.0
Arg	CGG	11351	1.7
Arg	AGA	139081	21.3
Arg	AGG	60289	9.2
Gly	GGU	156109	23.9
Gly	GGC	63903	9.8
Gly	GGA	71216	10.9
Gly	GGG	39359	6.0
종결	UAA	6913	1.1
종결	UAG	3312	0.5
종결	UGA	4447	0.7

[0189]

[0190]

이러한 표 또는 유사한 표를 사용함으로써, 당업자는 임의의 주어진 폴리펩티드 서열에 빈도를 적용할 수 있으며, 폴리펩티드를 코딩하나 주어진 종에 대하여 최적의 코돈을 사용하는 코돈-최적화된 코딩 영역의 핵산 단편을 생성할 수 있다.

[0191]

주어진 폴리펩티드 서열을 코딩하도록 최적화된 빈도로 코돈을 랜덤하게 할당하는 것은 각 아미노산에 대한 코돈 빈도를 계산한 다음, 코돈을 랜덤하게 폴리펩티드 서열에 할당함으로써, 수동으로 행할 수 있다. 추가로, 다양한 알고리즘 및 컴퓨터 소프트웨어 프로그램을 당업자가 용이하게 이용할 수 있다. 예를 들어, 미국 위스콘신주 매디슨 소재의 디엔에이스타 인크.(DNASTar, Inc.)로부터 입수가 가능한 레이저진 패키지(Lasergene Package) 내의 "EditSeq" 기능, 미국 메릴랜드주 베테스다 소재의 인포맥스 인크.(InforMax, Inc.)로부터 입수가 가능한 벡터 엔티아이 스위트(VectorNTI Suite) 내의 백트랜슬레이션(backtranslation) 기능, 및 미국 캘리포니아주 샌디에고 소재의 액셀리스 인크.(Accelrys, Inc.)로부터 입수가 가능한 GCG-위스콘신 패키지(GCG-Wisconsin Package) 내의 "백트랜슬레이트(backtranslate)" 기능이 있다. 게다가, 코딩 영역 서열을 코돈-최적화시키기 위한 다양한 자료, 예를 들어 www.entelechon.com/bioinformatics/backtranslation.php?lang=eng (2008년 4월 15일 방문)에서의 "백트랜슬레이션" 기능 및 <http://bioinfo.pbi.nrc.ca:8090/EMBOSS/index.html> (2002년 7월 9일 방문)에서 이용가능한 "backtranseq" 기능이 공개적으로 이용가능하다. 주어진 빈도를 기준으로 코돈을 할당하기 위한 기본적인 알고리즘을 구축하는 것도 또한 당업자에 의해 기본 수학 함수로 용이하게 달성될 수 있다.

[0192]

코돈-최적화된 코딩 영역은 소프트웨어 패키지, 예를 들어 "합성 유전자 디자이너(synthetic gene designer)" (userpages.umbc.edu/~wug1/codon/sgd/, 2012년 3월 19일 방문)를 비롯한 당업자에게 공지되어 있는 다양한 방법으로 설계될 수 있다.

[0193]

온도 및 용액 이온 강도의 적절한 조건 하에 핵산 단편의 단일 가닥 형태가 다른 핵산 단편에 어닐링(annealing)될 수 있을 때, 폴리뉴클레오티드 또는 핵산 단편은 다른 핵산 단편, 예를 들어 cDNA, 게놈 DNA 또는 RNA 분자에 "혼성화가능"하다. 혼성화 및 세척 조건은 잘 알려져 있으며, 문헌[Sambrook, J., Fritsch, E. F. and Maniatis, T. Molecular Cloning: A Laboratory Manual, 2nd ed., Cold Spring Harbor Laboratory: Cold Spring Harbor, NY (1989)], 특히 이에 수록된 제11장 및 표 11.1에 예시되어 있다(전체적으로 본 명세서에 참고로 포함됨). 온도 및 이온 강도 조건은 혼성화의 "엄격성(stringency)"을 결정한다. 엄격성 조건을 조정하여, 중간 정도로 유사한 단편(예를 들어 관계가 먼 유기체로부터의 상동 서열) 내지 고도로 유사한 단편(예

를 들어 밀접하게 관련된 유기체로부터의 기능성 효소를 중복시키는 유전자)을 스크리닝할 수 있다. 혼성화 후 세척은 엄격한 조건을 결정한다. 한 세트의 조건은 실온에서 15분 동안 6X SSC, 0.5% SDS로 시작한 다음, 45℃에서 30분 동안 2X SSC, 0.5% SDS로 반복한 다음, 50℃에서 30분 동안 0.2X SSC, 0.5% SDS로 2회 반복한 일련의 세척을 사용한다. 다른 세트의 엄격한 조건은 더 높은 온도를 사용하며, 여기서, 세척은 0.2X SSC, 0.5% SDS에서의 최종 2회의 30분 세척의 온도가 60℃로 증가된 것을 제외하고는 상기의 것과 동일하다. 다른 세트의 고도로 엄격한 조건에서는 65℃에서 0.1X SSC, 0.1% SDS에서의 2회 최종 세척을 사용한다. 추가의 세트에서 엄격한 조건은, 예를 들어 0.1X SSC, 0.1% SDS, 65℃에서의 혼성화 및 2X SSC, 0.1% SDS에 의한 세척 후에 0.1X SSC, 0.1% SDS에 의한 세척을 포함한다.

[0194] 혼성화의 엄격성에 따라 염기들 사이의 미스매치가 가능하지만, 혼성화는 두 핵산이 상보성 서열을 포함하는 것을 필요로 한다. 핵산 혼성화에 적절한 엄격성은, 본 기술 분야에 잘 알려진 변수인 핵산의 길이 및 상보성의 정도에 의존한다. 두 뉴클레오타이드 서열 사이의 유사성 또는 상동성의 정도가 클수록, 이들 서열을 가진 핵산의 혼성체의 T_m 값이 커진다. 핵산 혼성화의 상대적 안정성(더욱 높은 T_m 에 상응함)은 하기의 순서로 감소한다: RNA:RNA, DNA:RNA, DNA:DNA. 길이가 100개 초과인 뉴클레오타이드의 혼성체에 있어서, T_m 을 계산하기 위한 식이 도출되어 있다 (문헌[Sambrook et al., 상기 문헌, 9.50-9.51] 참조). 더 짧은 핵산, 즉, 올리고뉴클레오타이드의 혼성화에 있어서는 미스매치의 위치가 더욱 중요해지며 올리고뉴클레오타이드의 길이가 그의 특이성을 결정한다(문헌[Sambrook et al., 상기 문헌, 11.7-11.8] 참조). 일 실시 형태에서, 혼성화가능한 핵산의 길이는 약 10개 이상의 뉴클레오타이드이다. 일 실시 형태에서, 혼성화가능한 핵산의 최소 길이는 약 15개 이상의 뉴클레오타이드; 약 20개 이상의 뉴클레오타이드이거나; 또는 상기 길이는 약 30개 이상의 뉴클레오타이드이다. 더욱이, 프로브의 길이와 같은 인자에 따라 필요한 경우에 온도 및 세척액 염 농도를 조정할 수 있음을 당업자는 인식할 것이다.

[0195] 본 명세서에 사용되는 바와 같이, 용어 "폴리펩티드"는 단수의 "폴리펩티드"뿐 아니라 복수의 "폴리펩티드"를 포함하고자 하며, 아마이드 결합 (펩티드 결합으로도 공지됨)에 의해 선상으로 연결된 단량체 (아미노산)로 구성된 분자를 말한다. 용어 "폴리펩티드"는 둘 이상의 아미노산의 임의의 쇠(들)를 말하며, 특정 길이의 생성물을 말하는 것은 아니다. 따라서, 둘 이상의 아미노산의 쇠(들)를 지칭하기 위해 사용되는 펩티드, 다이펩티드, 트라이펩티드, 올리고펩티드, "단백질", "아미노산 쇠" 또는 임의의 다른 용어는 "폴리펩티드"의 정의 내에 포함되며, 용어 "폴리펩티드"는 이들 용어의 임의의 것 대신에 또는 그와 상호교환가능하게 사용될 수 있다. 폴리펩티드는 천연의 생물학적 공급원으로부터 유래될 수 있거나, 또는 재조합 기술에 의해 생성될 수 있지만, 반드시 지정된 핵산 서열로부터 번역되는 것은 아니다. 이는 화학적 합성에 의한 것을 포함하는 임의의 방식으로 생성될 수 있다.

[0196] "단리된" 폴리펩티드, 또는 이의 단편, 변이체 또는 유도체는 그의 자연적 환경에서는 존재하지 않는 폴리펩티드를 의도한다. 특정 정제 수준은 필요하지 않다. 예를 들어, 단리된 폴리펩티드는 그의 천연 환경 또는 자연적인 환경으로부터 제거될 수 있다. 숙주 세포에서 발현되는 재조합으로 생성된 폴리펩티드 및 단백질은 임의의 적합한 기술에 의하여 분리되거나, 분획화되거나 또는 부분적으로 또는 사실상 정제된 천연 또는 재조합 폴리펩티드가 그러한 바와 같이, 본 발명의 목적을 위해 단리된 것으로 여겨진다.

[0197] 본 명세서에 사용되는 바와 같이, 용어 "변이체" 및 "돌연변이체"는 동의어이며, 예를 들어 재조합 DNA 기술, 예를 들어 돌연변이 유발을 이용하여 생성된 하나 이상의 아미노산 삽입, 결실, 돌연변이 및 치환에 의해 특정하게 나열된 폴리펩티드와 상이해진 폴리펩티드를 말한다. 어떤 아미노산 잔기가 관심대상의 활성을 무효화하지 않고서 대체되거나, 부가되거나 또는 결실될 수 있는지를 결정함에 있어서의 지침은 특정 폴리펩티드의 서열을 상동성 폴리펩티드의 것, 예를 들어 효모 또는 박테리아의 서열과 비교하고, 높은 상동성의 영역 (보존된 영역)에서 이루어지는 아미노산 서열 변화의 수를 최소화함으로써 또는 아미노산을 콘센서스 서열로 대체함으로써 찾아질 수 있다.

[0198] 본 명세서에 사용되는 바와 같이, "엔지니어링된 폴리펩티드"는 합성인, 즉, 자연에서 발견되는 폴리펩티드와 일부 방식으로 상이해진 폴리펩티드를 말한다.

[0199] 대안적으로, 이들 동일하거나 또는 유사한 폴리펩티드를 코딩하는 재조합 폴리뉴클레오타이드 변이체는 유전자 암호에서의 "중복성(redundancy)"의 사용을 이용함으로써 합성되거나 또는 선택될 수 있다. 다양한 제한 부위를 생성하는 사일런트(silent) 변화와 같은 다양한 코돈 치환을 도입하여, 발현을 위한 플라스미드 또는 바이러스 벡터 내로의 클로닝을 최적화시킬 수 있다. 폴리뉴클레오타이드 서열 내의 돌연변이는 폴리펩티드의 임의의 부분의 특성을 변형시키기 위하여 폴리펩티드에 부가되는 다른 펩티드의 도메인 또는 폴리펩티드에 반영될 수 있다.

예를 들어, 돌연변이는 표적 단백질의 발현을 감소시키거나 또는 제거하는데 사용될 수 있으며, 전체 유전자 또는 유전자의 일부의 결실, DNA 단편을 유전자 내로 (프로모터 또는 코딩 영역 중 어느 하나에서) 삽입하여 단백질이 발현되지 않게 하거나 더 낮은 수준으로 발현되도록 하는 것, 프레임 이동 또는 종결 코돈을 부가하는 코딩 영역 내로의 돌연변이를 도입하여 기능성 단백질이 발현되지 않게 하는 것, 및 하나 이상의 돌연변이를 코딩 영역 내로 도입하여 아미노산을 변경시켜서, 비기능성 또는 효소적으로 덜 활성인 단백질이 발현되게 하는 것을 포함하지만, 이에 한정되지 않는다.

[0200] 아미노산 "치환"은 하나의 아미노산을 유사한 구조 및/또는 화학적 특성을 가진 다른 아미노산으로 대체한 것, 즉 보존적 아미노산 대체의 결과이거나, 또는 상기 치환은 하나의 아미노산을 상이한 구조 및/또는 화학적 특성을 갖는 아미노산으로 대체한 것, 즉 비보존적 아미노산 대체의 결과일 수 있다. "보존적" 아미노산 치환은 관련 잔기의 극성, 전하, 용해성, 소수성, 친수성, 또는 양친매성 성질의 유사성을 바탕으로 이루어질 수 있다. 예를 들어, 비극성 (소수성) 아미노산에는 알라닌, 류신, 아이소류신, 발린, 프롤린, 페닐알라닌, 트립토판 및 메티오닌이 포함되고; 극성 중성 아미노산에는 글리신, 세린, 트레오닌, 시스테인, 티로신, 아스파라긴 및 글루타민이 포함되고; 양으로 하전된 (염기성) 아미노산에는 아르기닌, 라이신 및 히스티딘이 포함되고; 음으로 하전된 (산성) 아미노산에는 아스파르트산 및 글루탐산이 포함된다. 대안적으로, "비보존적" 아미노산 치환은 임의의 이들 아미노산의 극성, 전하, 용해도, 소수성, 친수성 또는 양친매성 성질의 차이를 선택함으로써 이루어질 수 있다. "삽입" 또는 "결실"은 재조합 단백질에 의해 구조적으로 또는 기능적으로 허용되는 변이의 범위 내에 있을 수 있다. 허용되는 변이는 재조합 DNA 기술을 사용하여 폴리펩티드 분자 내 아미노산의 삽입, 결실, 또는 치환을 체계적으로 생성하고 생성된 재조합 변이체를 활성에 대해 분석함으로써 실험적으로 결정할 수 있다.

[0201] 아미노산 또는 뉴클레오타이드 서열의 "실질적인 부분"은, 당업자에 의한 서열의 수동 평가에 의해, 또는 BLAST (문헌[Altschul, S. F., *et al.*, *J. Mol. Biol.*, 215:403-410(1993)])와 같은 알고리즘을 사용하는 컴퓨터-자동화 서열 비교 및 확인에 의해, 폴리펩티드 또는 유전자를 추정적으로 확인하기에 충분한 폴리펩티드의 아미노산 서열 또는 유전자의 뉴클레오타이드 서열을 포함하는 부분이다. 일반적으로, 폴리펩티드 또는 핵산 서열을 공지의 단백질 또는 유전자와 상동적인 것으로 추정적으로 확인하기 위해서는 10개 이상의 연결 아미노산 또는 30개 이상의 뉴클레오타이드의 서열이 필요하다. 더욱이, 뉴클레오타이드 서열과 관련하여, 서열-의존적인 유전자의 확인 방법(예를 들어, 서던 혼성화(Southern hybridization)) 및 단리 방법 (예를 들어, 박테리아 콜로니 또는 박테리오파지 플라크의 원위치(*in situ*) 혼성화)에서 20-30개의 연결 뉴클레오타이드를 포함하는 유전자 특이적 올리고뉴클레오타이드 프로브가 사용될 수 있다. 게다가, 12-15개 염기의 짧은 올리고뉴클레오타이드를 PCR에서 증폭 프라이머로 사용하여, 프라이머를 포함하는 특정 핵산 단편을 얻을 수 있다. 따라서, 뉴클레오타이드 서열의 "실질적인 부분"은, 서열을 포함하는 핵산 단편을 특이적으로 확인 및/또는 단리하기에 충분한 서열을 포함한다. 본 명세서에서는, 특정 단백질을 코딩하는 전 아미노산 및 뉴클레오타이드 서열을 교시한다. 본 명세서에 보고된 서열의 유의성을 가지고, 당업자는 지금에 와서야 개시된 서열의 전부 또는 실질적인 부분을 당업자에게 공지된 목적으로 사용할 수 있게 되었다. 따라서 본 발명은, 첨부하는 서열 목록에 보고된 전 서열과, 상기 정의된 서열들의 실질적인 부분을 포함한다.

[0202] 용어 "상보적"은, 서로 혼성화될 수 있는 뉴클레오타이드 염기들 사이의 관계를 설명하기 위해 사용된다. 예를 들어, DNA에 관하여, 아데닌은 티민에 대해 상보성이며, 시토신은 구아닌에 대하여 상보성이며, RNA에 관하여, 아데닌은 우라실에 대해 상보성이며, 시토신은 구아닌에 대해 상보성이다.

[0203] 본 기술 분야에 공지된 바와 같이, 용어 "동일성 (%)"은 둘 이상의 폴리펩티드 서열 또는 둘 이상의 폴리뉴클레오타이드 서열 사이의 관계로서, 이는 서열들을 비교함으로써 결정되는 바와 같다. 본 기술 분야에서, "동일성"은 또한 경우에 따라서, 폴리펩티드 또는 폴리뉴클레오타이드 서열 사이의 서열 관련성의 정도를 의미하며, 이는 이러한 서열들의 스트링(string)들 사이의 매치에 의해 결정되는 바와 같다. "동일성" 및 "유사성"은 하기의 문헌에 기재된 것을 포함하지만 이에 한정되지 않는 공지된 방법으로 용이하게 계산될 수 있다: 1.) 문헌 [Computational Molecular Biology (Lesk, A. M., ed.) Oxford University: NY (1988)]; 2.) 문헌 [Biocomputing: Informatics and Genome Projects (Smith, D. W., Ed.) Academic: NY (1993)]; 3.) 문헌 [Computer Analysis of Sequence Data, Part I (Griffin, A. M., and Griffin, H. G., eds.) Humana: NJ (1994)]; 4.) 문헌 [Sequence Analysis in Molecular Biology (von Heinje, G., Ed.) Academic (1987)]; 및 5.) 문헌 [Sequence Analysis Primer (Gribskov, M. and Devereux, J., Eds.) Stockton: NY (1991)].

[0204] 동일성을 결정하는 방법은 시험 서열들 사이에서 최상의 매치를 주도록 설계된다. 동일성 및 유사성을 결정하는 방법은 공개적으로 이용가능한 컴퓨터 프로그램에 성문화되어 있다. 서열 정렬과 동일성 (%)의 계산은 레이

저진 바이오인포매틱스 컴퓨팅 스위트 (LASERGENE bioinformatics computing suite)의 맥얼라인(Megalign)TM 프로그램 (미국 위스콘신주 매디슨 소재의 디엔에이스타 인크.(DNASTAR Inc.))을 이용하여 수행될 수 있다. 서열의 다중 정렬은 레이저진 바이오인포매틱스 컴퓨팅 스위트의 맥얼라인TM 프로그램 (디엔에이스타 인크.)에서 발견되고, 클러스탈 V (Clustal V) (문헌[Higgins and Sharp, CABIOS. 5:151-153 (1989)]; 문헌[Higgins, D.G. *et al.*, *Comput. Appl. Biosci.*, 8:189-191 (1992)]에 기재)로 표시되는 정렬 방법에 상응하는 "클러스탈 V 정렬 방법"을 비롯한 몇 가지 종류의 알고리즘을 포함하는 "클러스탈 정렬 방법"을 사용하여 수행된다. 다중 정렬의 경우, 디폴트 값은 갭 페널티(GAP PENALTY)=10 및 갭 길이 페널티(GAP LENGTH PENALTY)=10에 상응한다. 클러스탈 방법을 사용하는 단백질 서열의 동일성 (%)의 계산 및 쌍정렬을 위한 디폴트 파라미터는 KTUPLE=1, 갭 페널티=3, 윈도우(WINDOW)=5 및 DIAGONALS SAVED=5이다. 핵산의 경우, 이들 파라미터는 KTUPLE=2, 갭 페널티=5, 윈도우=4 및 DIAGONALS SAVED=4이다. 클러스탈 V 프로그램을 사용하여 서열들을 정렬한 후에, 동일한 프로그램 내의 "서열 거리" 표를 봄으로써 "동일성 (%)"을 얻을 수 있다. 부가적으로 "클러스탈 W 정렬 방법"이 이용 가능하며, 이는 클러스탈 W (문헌[Higgins and Sharp, CABIOS. 5:151-153 (1989)]; 문헌[Higgins, D.G. *et al.*, *Comput. Appl. Biosci.* 8:189-191 (1992)]에 기재)로 표시되고 레이저진 바이오인포매틱스 컴퓨팅 스위트(디엔에이스타 인크.)의 맥얼라인TM v6.1 프로그램에서 발견되는 정렬 방법에 상응한다. 다중 정렬을 위한 디폴트 파라미터(갭 페널티=10, 갭 길이 페널티=0.2, 발산 서열 지연율(Delay Divergen Seqs) (%)=30, DNA 전이 가중치=0.5, 단백질 가중 매트릭스(Protein Weight Matrix)=고넷 시리즈(Gonnet Series), DNA 가중 매트릭스(DNA Weight Matrix)=IUB). 클러스탈 W 프로그램을 사용하여 서열들을 정렬한 후에, 동일한 프로그램 내의 "서열 거리" 표를 봄으로써 "동일성 (%)"을 얻을 수 있다.

[0205] 당업자라면 많은 서열 동일성 수준들은 다른 종들로부터의 것과 같은 폴리펩티드들을 확인함에 있어서 또는 상응하는 폴리뉴클레오티드를 설명함에 있어서 유용함을 잘 알 것이며, 여기서, 이러한 폴리펩티드들은 동일하거나 또는 유사한 기능 또는 활성을 갖는다. 동일성 (%)의 유용한 예에는 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 또는 95%가 포함되지만, 이에 한정되지 않거나, 또는 55% 내지 100%의 임의의 정수 백분율, 예를 들어 55%, 56%, 57%, 58%, 59%, 60%, 61%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% 또는 99%가 본 발명을 설명하는데 유용할 수 있다. 적합한 폴리뉴클레오티드 단편은 상기 상동성을 가질 뿐 아니라, 전형적으로 적어도 50개 뉴클레오티드, 적어도 100개 뉴클레오티드, 적어도 150개 뉴클레오티드, 적어도 200개 뉴클레오티드 또는 적어도 250개 뉴클레오티드를 포함한다. 추가로, 상기 상동성을 갖는 적합한 폴리뉴클레오티드 단편은 적어도 50개 아미노산, 적어도 100개 아미노산, 적어도 150개 아미노산, 적어도 200개 아미노산, 또는 적어도 250개 아미노산을 갖는 폴리펩티드를 코딩한다.

[0206] 어구 "서열 분석 소프트웨어"는 뉴클레오티드 또는 아미노산 서열을 분석하는 데 유용한 임의의 컴퓨터 알고리즘 또는 소프트웨어 프로그램을 말한다. "서열 분석 소프트웨어"는 구매가능하거나 또는 독립적으로 개발될 수 있다. 전형적인 서열 분석 소프트웨어에는 하기의 것이 포함될 것이나, 이에 한정되지 않는다: 1.) GCG 프로그램 스위트(위스콘신 패키지 버전(Wisconsin Package Version) 9.0, 미국 위스콘신주 매디슨 소재의 제네틱스 컴퓨터 그룹(Genetics Computer Group; GCG)); 2.) BLASTP, BLASTN, BLASTX (문헌[Altschul *et al.*, *J. Mol. Biol.*, 215:403-410 (1990)]); 3.) DNASTAR (미국 위스콘신주 매디슨 소재의 디엔에이스타, 인크.); 4.) 시퀀처 (Sequencher) (미국 미시간주 앤아버 소재의 진 코즈 코퍼레이션 (Gene Codes Corporation)); 및 5.) 및 스미스-워터만(Smith-Waterman) 알고리즘 (문헌[W. R. Pearson, *Comput. Methods Genome Res.*, [Proc. Int. Symp.] (1994), Meeting Date 1992, 111-20. Editor (s): Suhai, Sandor. Plenum: New York, NY)을 포함하는 FASTA 프로그램. 본 출원의 맥락 내에서, 서열 분석 소프트웨어를 분석에 이용하는 경우에, 달리 특정하지 않는 한, 분석 결과는 언급된 프로그램의 "디폴트 값"에 기초할 것임이 이해될 것이다. 본 명세서에 사용되는 바와 같이, "디폴트 값"은 최초 초기화할 때에 소프트웨어에 원래 로딩된 임의의 값 또는 파라미터 세트를 의미할 것이다.

[0207] 표준 재조합 DNA 및 분자 클로닝 기술은 본 기술 분야에 잘 알려져 있으며, 문헌[Sambrook, J., Fritsch, E. F. and Maniatis, T., *Molecular Cloning: A Laboratory Manual*, Second Edition, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY (1989)] (이하, "마니아티스(Maniatis)"); 및 문헌[Silhavy, T. J., Bennis, M. L. and Enquist, L. W., *Experiments with Gene Fusions*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY (1984)]; 및 문헌[Ausubel, F. M. *et al.*, *Current Protocols in Molecular Biology*, published by Greene Publishing Assoc. and Wiley-Interscience (1987)]에 기재되어 있다. 본 발명에서 사용되는 추가의 방법은 문헌[*Methods in Enzymology*, Volume 194, *Guide to Yeast Genetics and Molecular and Cell Biology* (Part A, 2004, Christine Guthrie and Gerald R. Fink (Eds.), Elsevier

Academic Press, San Diego, CA)]에 있다. 다른 분자적 툴(tool) 및 기술은 본 기술 분야에 공지되어 있으며, 중첩 연장 폴리머라아제 연쇄 반응 (PCR)에 의한 스플라이싱 (문헌[Yu, et al. (2004) Fungal Genet. Biol. 41:973-981]), 사카로마이세스 세레비지에의 *URA3* 유전자좌에서의 돌연변이에 대한 양성 선발 (문헌[Boeke, J.D. et al. (1984) Mol. Gen. Genet. 197, 345-346]; 문헌[M A Romanos, et al. Nucleic Acids Res. 1991 January 11; 19(1): 187]), cre-lox 부위 특이적 재조합 시스템과, 돌연변이 lox 부위 및 FLP 기질 돌연변이 (문헌[Sauer, B. (1987) Mol Cell Biol 7: 2087-2096]; 문헌[Senecoff, et al. (1988) Journal of Molecular Biology, Volume 201, Issue 2, Pages 405-421]; 문헌[Albert, et al. (1995) The Plant Journal. Volume 7, Issue 4, pages 649-659]), "시임리스(seamless)" 유전자 결실 (문헌[Akada, et al. (2006) Yeast;23(5):399-405]), 및 겹 수복 방법 (문헌[Ma et al., *Genetics* 58:201-216; 1981])을 포함한다.

[0208] 본 명세서에 개시된 재조합 숙주 세포의 유전자 조작은 표준 유전자 기술 및 스크리닝을 이용하여 수행될 수 있으며, 유전자 조작에 적합한 임의의 숙주 세포에서 이루어질 수 있다 (문헌[*Methods in Yeast Genetics*, 2005, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, pp. 201-202]).

[0209] 실시 형태들에서, 본 명세서에 개시된 재조합 숙주 세포는 표 7에서와 같이 본 명세서의 다른 곳에서 언급된 효모를 비롯하여 유전자 변형 및 재조합 유전자 발현에 유용한 임의의 효모 또는 진균류 숙주일 수 있다. 다른 실시 형태들에서, 재조합 숙주 세포는 이사첸키아 속, 자이고사카로마이세스(*Zygosaccharomyces*) 속, 쉬조사카로마이세스 속, 데케라(*Dekkera*) 속, 토룰롭시스(*Torulopsis*) 속, 브레타노마이세스(*Brettanomyces*) 속, 토룰라스포라(*Torulaspora*) 속, 한세니아스포라(*Hanseniaspora*) 속, 클루베로마이세스(*Kluveromyces*) 속, 야로위아 속, 및 일부 칸디다 종의 구성원일 수 있다.

[0210] KARI 활성을 갖는 폴리펩티드

[0211] 일부 실시 형태에서, 본 발명에서 제공된 재조합 숙주 세포 및 방법은 KARI 효소가 필수적인 역할을 수행하는 아이소부탄올의 미생물에 의한 생성에서 발생하는 필요성을 처리한다. 도 1에 예시된 아이소부탄올 생합성 경로에서, 아세토라테이트의 다이하이드록시아이소발레레이트 (DHIV)로의, 기질에서 생성물로의 전환은 KARI 효소에 의해 촉매된다. 미국 특허 공개 제2011/0244536호에는 KARI의 SLSL 클레이드(Clade)의 구성원인, 케톨-산 리덕토아이소머라아제 활성을 갖는 폴리펩티드가 개시되어 있으며, 이는 참고로 포함된다. 상기 미국 특허 공개에서 개시된, KARI 활성을 갖는 폴리펩티드는 아이소부탄올 생성에 효과적인 것으로 밝혀졌다. KARI의 SLSL 클레이드는 표 3에 열거된 KARI 효소를 포함한다.

[0212] [표 3]

효과적인 KARI

설명	핵산의 서열 번호	아미노산의 서열 번호
스타필로코커스 카피티스 SK14	66	67
스타필로코커스 에피테르미디스 M23864-W1	68	69
스타필로코커스 호미니스 SK119	134	135
스타필로코커스 아우레우스 아종 아우레우스 TCH130	70	71
스타필로코커스 와르네리 L37603	72	73
스타필로코커스 에피테르미디스 W23144	74	75
스타필로코커스 사프로피티쿠스 아종 사프로피티쿠스 ATCC15305	76	77
스타필로코커스 카르노수스 아종 카르노수스 TM300	78	79
리스테리아 모노사이토게네스 EGD-e	80	81
리스테리아 그라이이 DSM 20601	82	83
엔테로코커스 카셀리플라부스 EC30	84	85
엔테로코커스 갈리나룸 EG2	86	87
마크로코커스 카세오라이트쿠스 JCSC5402	88	89
스트렙토코커스 베스티볼라리스	90	91
스트렙토코커스 뮤탄스 UA159	92	93
스트렙토코커스 코르도니이 str. 크가쿠스 sybstr. CH1	94	95
스트렙토코커스 수이스(스트렙토코커스 수이스) 89/1591	96	97
스트렙토코커스 인판타리우스 아종 인판타리우스 ATCC BAA-102	98	99
락토코커스 락티스 아종 크레모리스 MG1363	100	101
락토코커스 락티스	102	103
류코노스톡 메센테로이데스 아종 메센테로이데스 ATCC8293	104	105

[0213]

락토바실러스 부크네리 ATCC 11577	106	107
스타필로코커스 헤모라이트쿠스 JCSC1435	108	109
스타필로코커스 에피테르미디스 ATCC12228	110	111
스트렙토코커스 뉴모니아 CGSP14	112	113
스트렙토코커스 뉴모니아 TIGR4	114	115
스트렙토코커스 상귀니스 SK36	116	117
스트렙토코커스 살리마리우스 SK126	118	119
스트렙토코커스 서모필루스 LMD-9	120	121
스트렙토코커스 뉴모니아 CCR1 1974M2	122	123
락토코커스 락티스 아종 락티스 I11403	124	125
류코노스톡 메센테로이데스 아종 크레모리스 ATCC19254	126	127
류코노스톡 메센테로이데스 아종 크레모리스	128	129
락토바실러스 브레비스 아종 그라베센시스 ATCC27305	130	131
락토코커스 락티스 아종 락티스 NCD02118	132	133

[0214]

[0215] 본 명세서에서 기재되고 입증된 바와 같이, 본 출원인은 추가의 KARI 효소 및 상기 추가의 KARI의 변이체를 발

견하였는데, 이는 락토코커스 락티스 유래의 KARI에서 관찰되는 것에 비견되는 및/또는 이를 증가하는 아이소부탄을 생성으로 이어진다. 이러한 KARI 효소 및 변이체는 동일한 조건에서 락토코커스 락티스 KARI에서 관찰되는 것과 비교할 때 비견되는 또는 더 높은 아이소부탄을 역가 및/또는 더 높은 유효 아이소부탄을 생산성으로 이어진다. 따라서, 실시 형태들에서, 아이소부탄을 생성 경로에서 기능하는 KARI 활성을 갖는 폴리펩티드는 락토코커스 락티스 KARI (서열 번호 380)에서의 것에 비견되는 또는 이보다 더 우수한 아이소부탄을 역가 및/또는 유효 아이소부탄을 생산성을 갖는다.

[0216] 따라서 KARI 활성을 갖는 이러한 폴리펩티드는 아이소부탄을 생성에 적합할 수 있다. 본 기술 분야에서 입수가 가능한 구조 정보와 서열 정보의 조합을 이용하여, KARI 활성을 포함하고 예시된 서열에 대하여 100% 미만의 동일성을 포함하는 폴리펩티드를 아이소부탄을 생합성 경로에서의 사용용으로 제작할 수 있음을 인지할 것이다. 예를 들어, 2.6 Å의 해상도에서의 이. 콜라이 KARI 효소의 결정 구조가 해상되었으며 (문헌[Tyagi, et al., Protein Sci., 14: 3089-3100, 2005]), 이는 피. 아에루기노사 KARI (문헌[Ahn, et al., J. Mol. Biol., 328: 505-515, 2003]) 및 시금치 유래의 KARI 효소 (문헌[Biou V., et al. The EMBO Journal, 16: 3405-3415, 1997])의 구조가 해상된 바와 같았다. 더욱이, 표 1에 요약된 바와 같이, 실험적으로 검증된 기능을 갖는 25가지의 KARI 단백질의 아미노산 서열을 이용하여 작성된 프로파일 HMM (본 명세서에 제공됨; 표 2)이 본 명세서에 기재되어 있다. KARI는 슈도모나스 플루오레센스 Pf-5, 술폴로부스 술파타리쿠스 P2, 파이로바쿨룸 아에로필룸 str. IM2, 나트로노모나스 파라오니스 DSM 2160, 바실러스 서브틸리스 아종 서브틸리스 str. 168, 코리네박테리움 글루타미쿰 ATCC 13032, 파에오스프리틸룸 몰리쉬아눔, 랄스토니아 솔라나세아룸 GMI1000, 자이모모나스 모빌리스 아종 모빌리스 ZM4, 알칼리리니콜라 에르리케이 MLHE-, 캄필로박터 라리 RM2100, 마리노박터 아쿠아에올레이 VT8, 사이크로박터 아르키티쿠스 273-4, 하헬라 체주엔시스 KCTC 2396, 티오바실러스 데니트리피칸스 ATCC 25259, 아조토박터 비넬란디아 AvOP, 슈도모나스 시린가에 pv. 시린가에 B728a, 슈도모나스 시린가에 pv. 토마토 str. DC3000, 슈도모나스 퓨티다 KT2440, 슈도모나스 엔토모필라 L48, 슈도모나스 멘도시나 ymp, 슈도모나스 아에루기노사 PA01, 바실러스 세레우스 ATCC 10987, 바실러스 세레우스 ATCC 10987, 및 스피나시아 올레라세아 유래의 것이었다. HMMER 패키지 내의 hmmsearch 프로그램을 이용하여 프로파일 HMM을 $<10^{-3}$ 의 E 값과 매칭시키는 임의의 단백질이 기능성 KARI인 것으로 예상된다.

[0217] 아이소부탄의 생성은 숙주 미생물에 존재하는 당분해 경로를 이용하는 것으로 여겨진다. 당분해 동안의 글루코스로부터의 2개의 피루베이트 분자의 생성 중에, 글리세르알데히드-3-포스페이트 데하이드로게나아제 반응에 의한 NAD⁺로부터의 2개의 NADH 분자의 순생성이 있다. 2개의 피루베이트 분자로부터의 하나의 아이소부탄을 분자의 추가의 생성 동안, KARI 반응에 의한 하나의 NAD(P)H 분자의 순소비 및 아이소부탄을 데하이드로게나아제 반응에 의한 하나의 NAD(P)H 분자의 순소비가 있다. NADH와 NADPH의 상호 전환은 일반적으로 효모에 있어서 느리며 비효율적이고; 따라서, 소비되는 NADPH는 이 과정에서 기질을 소비하는 대사에 의해 (예를 들어, 펜토스 포스페이트 경로에 의해) 생성된다. 한편, 세포는 NAD⁺/NADH 비에 있어서 항상성을 유지하려고 노력하며, 이는 아이소부탄을 생성에서 생성되는 과도한 NADH가, 예를 들어 글리세롤의 생성에 의해 다른 대사 중간체의 낭비적인 환원에서 소비되게 한다. (문헌[Bakker, et al., 2001. Stoichiometry and compartmentation of NADH metabolism in Saccharomyces cerevisiae. FEMS Microbiol. Rev. 25:15-37.]). 따라서, 아이소부탄을 경로에 의해 소비되는 NADPH와 생성되는 NADH 사이의 불균형은 하기의 2가지 방식으로 글루코스로부터 생성되는 아이소부탄의 물 수율을 감소시킬 수 있다: 1) NADPH를 생성하기 위한 대사의 불필요한 작동, 및 2) NAD⁺/NADH 항상성을 유지하기 위한 대사 중간체의 낭비적인 반응. 아이소부탄을 경로에서 잘 기능하는 그리고 NADH에 대한 K_M 이 낮은, KARI 활성을 갖는 폴리펩티드를 사용하여 아이소부탄의 생성을 향상시킬 수 있다.

[0218] NADH를 보조 인자로 이용하는 다양한 능력을 갖는 변이체를 생성하기 위한, 표 3 및 표 10에 제공된 KARI 효소 서열에 대한 치환이 또한 본 명세서에 개시된다. 이러한 변이체는 특정 생성 조건에 있어서 아이소부탄을 생합성 경로와 같이 KARI를 이용하는 생합성 경로의 효율을 최적화하는 데 이용될 수 있는 대안을 제공한다. NADH를 이용하는 능력이 상이한 엔에어로스티페스 카카에로부터 유래된 K9 KARI 효소의 변이체에 있어서 호기성 조건으로부터 혐기성 조건으로 스위칭되는 조건 하에서의 아이소부탄을 생성이 실시예에서 입증된다. 따라서, 이 개시 내용이 갖추어지면, 당업자는 SLSL 클레이드 KARI 효소, 또는 엔테로코커스 갈리나룸, 스트렙토코커스 서모필루스 락토코커스 락티스 아종 크레모리스 MG1363, 비피도박테리움 안굴라툼, 비피도박테리움 덴티움, 또는 엔에어로스티페스 카카에, 락토코커스 락티스 KARI 효소 또는 이의 변이체 또는 활성 단편 - 일련의 생성 조건에 적합함 - 을 포함하는 재조합 숙주 세포를 생성할 수 있을 것이다.

[0219] 일부 실시 형태에서, KARI 활성을 가지며, 표 3 또는 표 10, 또는 실시예 16, 실시예 17, 실시예 21의 KARI 효소에 대하여 동일성이 약 80% 이상, 약 85% 이상, 약 90% 이상, 약 95% 이상, 약 98% 이상, 또는 약 99% 이상인

고, NADH에 대한 K_M 이 약 300 μM 미만, 100 μM 미만, 50 μM 미만, 20 μM 미만, 10 μM 미만, 또는 5 μM 미만인 폴리펩티드가 본 발명에서 제공된다. 일부 실시 형태에서, KARI 활성을 가지며, 표 3, 표 10 또는 실시예 16, 실시예 17, 실시예 21의 KARI 효소에 대하여 동일성이 약 80% 이상, 약 85% 이상, 약 90% 이상, 약 95% 이상, 약 98% 이상, 또는 약 99% 이상인 엔지니어링된 폴리펩티드가 본 발명에서 제공된다. 일부 실시 형태에서, 이러한 폴리펩티드는 NADH에 대한 K_M 이 상응하는 천연 효소의 것보다 더 작다. 일부 실시 형태에서, NADPH에 대한 K_M 에 대한 NADH에 대한 K_M 의 비는 0.1 미만이며, 일부 실시 형태에서는 1 미만이며, 일부 실시 형태에서는 2 미만이며, 일부 실시 형태에서는 4 미만이다.

[0220] 아이소부탄을 생성에 특히 적합한 KARI 효소 및 그의 변이체는 엔에어로스티페스 카카에 DSM 14662 유래의 케톨-산 리덕토아이스머라아제 (서열 번호 643)의 변이체들, 즉 "K9G9" (서열 번호 644) 및 "K9D3" (서열 번호 645)을 포함하지만, 이에 한정되지 않는데, 이들은 NADH에 대한 K_M 이 천연 효소 (서열 번호 643)의 것보다 더 낮다.

[0221] 본 발명에서 제공된 숙주 세포는 케톨-산 리덕토아이스머라아제 활성을 갖는 폴리펩티드를 포함할 수 있다. 실시 형태들에서, 이러한 폴리펩티드는 서열 번호 643의 서열 또는 그의 활성 변이체; 또는 엔에어로스티페스 카카에 DSM 14662로부터 유래된 KARI 또는 그의 활성 변이체에 대하여 동일성이 약 80% 이상, 약 85% 이상, 약 90% 이상, 약 95% 이상, 약 98% 이상, 또는 약 99% 이상이다. 실시 형태들에서, 상기 폴리펩티드는 서열 번호 645 또는 644의 서열에 대하여 동일성이 약 80% 이상, 약 85% 이상, 약 90% 이상, 약 95% 이상, 약 98% 이상, 또는 약 99% 이상이다. 실시 형태들에서, 폴리펩티드는 서열 번호 645 또는 644의 서열을 포함한다.

[0222] 일부 실시 형태에서, KARI 활성을 갖는 폴리펩티드는 서열 번호 419 [JB4P], 427 [SB2]의 아미노산 서열, 및 표 25 및 표 26에 열거된 모든 변이체에 대하여 약 80% 이상, 약 85% 이상, 약 90% 이상, 약 95% 이상, 약 98% 이상, 또는 약 99% 이상의 동일성을 포함한다. 이러한 변이체는 특정 생성 조건에 있어서 아이소부탄을 생합성 경로의 효율을 최적화하기 위한 대안을 제공한다. 조건들 하에서의 아이소부탄을 생성이 실시예에서 입증된다.

[0223] KARI 활성을 갖는 추가 폴리펩티드의 확인

[0224] 아이소부탄을 생성에 적합한 KARI를 나타낸 다양한 박테리아 및 진균류 중 유래의 KARI-코딩 유전자의 생물 다양성 스크린이 실시예 1에 설명되어 있다. 이 개시 내용이 갖추어지면, 당업자는 KARI 활성을 갖는 추가의 적합한 폴리펩티드를 용이하게 확인할 수 있을 것이다.

[0225] 다른 폴리뉴클레오티드, 유전자 및/또는 폴리펩티드의 서열은 본 발명에 개시된 그리고 본 기술 분야에서 입수 가능한 서열을 이용하여 당업자에게 잘 알려진 생물 정보 데이터베이스에서 그리고 문헌에서 확인할 수 있다. 예를 들어, 이러한 서열들은 본 발명에서 제공된 폴리뉴클레오티드 또는 폴리펩티드 서열을 이용하여 공개적으로 이용가능한 데이터베이스의 BLAST 검색을 통하여 확인할 수 있다. 이러한 방법에서, 동일성은 갭 페널티=10, 갭 길이 페널티=0.1, 및 고넷 250 시리즈의 단백질 가중 매트릭스의 디폴트 파라미터를 사용한 클러스탈 W 정렬 방법을 바탕으로 할 수 있다.

[0226] 부가적으로, 본 명세서에 개시된 폴리뉴클레오티드 또는 폴리펩티드 서열은 자연에서 다른 KARI 상동체를 확인하는 데 사용될 수 있다. 예를 들어, 본 명세서에 개시된 KARI 코딩 핵산 단편들 각각은 상동 단백질들을 코딩하는 유전자들을 단리하는 데 사용될 수 있다. 서열-의존성 프로토콜을 사용하는 상동 유전자의 단리는 본 기술 분야에 잘 알려져 있다. 서열-의존성 프로토콜의 예에는 (1) 핵산 혼성화 방법; (2) 핵산 증폭 기술 (예를 들어, 폴리머라아제 연쇄 반응 (PCR) (멀리스(Mullis) 등의 미국 특허 제4,683,202호); 라이가아제 연쇄 반응 (LCR) (문헌[Tabor et al., Proc. Acad. Sci. USA 82:1074 (1985)]); 또는 가닥 치환 증폭(strand displacement amplification; SDA) (문헌[Walker et al., Proc. Natl. Acad. Sci. U.S.A., 89:392 (1992)]))의 다양한 사용으로 예시되는 바와 같이, DNA 및 RNA 증폭 방법; 및 (3) 상보성에 의한 스크리닝 및 라이브러리 제작 방법이 포함되지만, 이에 한정되지 않는다.

[0227] 당업자라면, 이 개시 내용이 갖추어지면, 예를 들어 N-말단에서 서열 정렬을 바탕으로 하여 본 발명에서 제공된 폴리펩티드를 절단하고 KARI 활성을 확인함으로써 본 발명에서 제공된 폴리펩티드의 활성 단편들을 생성할 수 있음을 인지할 것이다. 실시 형태들에서, 본 발명에서 제공되는 엔에어로스티페스 카카에 KARI 및 그의 변이체는 N-말단에서 절단된다. 일 실시 형태에서, 처음 5개 이하의 아미노산이 본 발명에서 제공되는 폴리펩티드로부터 절단된다. 실시 형태들에서, 폴리펩티드는 서열 번호 27의 서열 또는 그의 변이체이다. 일 실시 형태에서, KARI 활성을 갖는 폴리펩티드는 서열 번호 635, 637의 서열 (각각 폴리뉴클레오티드 서열, 즉 서열 번호

636 및 638의 서열에 의해 코딩됨), K9_Annabel_SH (서열 번호 862, 단백질 서열 번호 863) 및 K9_Zeke_SH (서열 번호 860, 단백질 서열 번호 861), 또는 표 40에 열거된 임의의 변이체를 포함한다.

[0228] NADH에 대한 K_M 의 저하

[0229] 도 2 및 실시예에 나타난 바와 같이, 언에어로스티페스 카카에 유래의 KARI 효소에 있어서 슈도모나스 플루오레센스 KARI의 50, 52 및 53과, 선택적으로 47에 상응하는 위치에서의 돌연변이는 야생형과 비교하여 NADH에 대한 K_M 이 저하된 KARI로 이어지며, 이는 이들 위치에서의 돌연변이가 고도로 효과적인 KARI의 NADH 수용 변이체를 생성함을 검증해 주는 것이다. 언에어로스티페스 카카에 KARI의 추가의 돌연변이에 의하면, NADH에 대한 K_M 을 추가로 저하시키는 위치들이 나타났다.

[0230] 본 발명에서 입증된 바와 같이 (실시예 참조), 포스페이트 결합 영역, 특히 PF5 KARI (서열 번호 5)의 위치 47, 50, 52 및 53에 상응하는 2개 이상의 위치에서의 아미노산의 치환은 NADH에 대한 K_M 의 저하로 이어진다. 따라서, KARI 활성을 가지며 천연 아미노산 서열과 비교하여 PF5 KARI의 위치 47, 50, 52 및 53에 상응하는 4개의 위치 중 2개 이상의 위치에 치환을 포함하는, 본 명세서에서, 예를 들어 표 3 및 표 10에서 열거된 유기체로부터 유래된 폴리펩티드가 본 발명에서 제공된다. KARI 활성을 가지며 포스페이트 결합 영역 내에 치환을 포함하는 폴리펩티드가 본 발명에서 제공된다. KARI 활성을 가지며 K9 KARI (서열 번호 27)의 S56 및 S58에 상응하는 위치에서 치환을 포함하는 폴리펩티드가 본 발명에서 제공된다. 일부 실시 형태에서, S56에 상응하는 위치에서의 치환은 A이다. 일부 실시 형태에서, S58에 상응하는 위치에서의 치환은 D 또는 E이다. 일부 실시 형태에서, S53에 상응하는 위치에서의 치환은 Q, E, P, 또는 A이다. 일부 실시 형태에서, S56에 상응하는 위치에서의 치환은 V 또는 D이다. 일부 실시 형태에서, S58에 상응하는 위치에서의 치환은 D 또는 Q이다. 실시 형태들에서, 폴리펩티드는 K9 KARI (서열 번호 27)의 I86, N87, N107, T131, 또는 T191에 상응하는 하나 이상의 위치에서의 치환을 추가로 포함한다. 일부 실시 형태에서, 폴리펩티드는 지시된 위치들 중 2개 이상, 3개 이상, 4개 이상의 위치에서 또는 지시된 위치들 전부에서 치환을 포함한다. 일부 실시 형태에서, I86에 상응하는 위치에서의 치환은 T 또는 V이다. 일부 실시 형태에서, N87에 상응하는 위치에서의 치환은 P이다. 일부 실시 형태에서, N107에 상응하는 위치에서의 치환은 S이다. 일부 실시 형태에서, T131에 상응하는 위치에서의 치환은 C, L, A, M 또는 V이다. 일부 실시 형태에서, T191에 상응하는 위치에서의 치환은 A, S, D, C, 또는 G이다.

[0231] 실시 형태들에서, 폴리펩티드는 야생형 서열과 관련하여 10개 미만, 15개 미만, 또는 20개 미만의 치환을 포함한다. 실시 형태들에서, 폴리펩티드는 실험적으로 검증된 KARI를 바탕으로 한, 그리고 표 Z에 주어진 프로파일 HMM을 $< 10^{-3}$ 미만의 E 값과 매칭시킨다. 서열들은 hmmsearch (미국 버지니아주 애시번 소재의 자넬리아 팜 리서치 캠퍼스(Janelia Farm Research Campus)로부터 입수가능한 HMMER 소프트웨어 패키지)를 이용하여 표 Z에 주어진 프로파일 HMM과 비교될 수 있다.

[0232] KARI 활성을 가지며 NADH에 대한 K_M 이 저하된 추가의 폴리펩티드는 본 명세서에서 기재되고 입증된 방법을 이용하여 수득될 수 있다. 예를 들어, KARI 활성을 갖는 폴리펩티드는 본 명세서에 기재된 바와 같이 부위 포화 유전자 라이브러리의 제작에 이용될 수 있다. 이러한 유전자 라이브러리의 제작을 위한 키트는 (예를 들어, 미국 78480 오하이오주 클레블랜드 소재의 유에스비 코퍼레이션(USB Corporation)으로부터) 구매가능하다. 부위 지정 돌연변이 유발은 또한 구매가능한 키트 (예를 들어, 퀵체인지(QuickChange) II XL 부위 지정 돌연변이 유발 키트, 카탈로그 번호 200524, 미국 캘리포니아주 라 줄라 소재의 스트라타진(Stratagene))를 이용하여 실시될 수 있다. 돌연변이 유발을 위한 표적 부위용의 프라이머 설계는 본 기술 분야에 잘 알려져 있으며, 표적 부위를 확인하기 위한 다중 서열 정렬도 마찬가지로 잘 알려져 있다.

[0233] 일단 변이체가 생성되었으면, NADH 또는 NADPH에 의한 KARI 활성은 본 기술 분야에 공지되고/되거나 본 명세서에 개시된 방법을 이용하여 용이하게 평가될 수 있다. 예를 들어, KARI 활성은 HPLC/MS를 이용하여 2,3-다이하이드록시아이소발레레이트의 형성을 측정하는 것을 통하여 미하엘리스 상수를 결정함으로써 또는 340 nm에서 반응물로부터의 NADPH 또는 NADH의 사라짐을 측정함으로써 결정될 수 있다. 이와 마찬가지로, 변이주를 포함하는 주로부터의 아이소부탄올 생성이 확인될 수 있다.

[0234] 보조 인자 특이성

[0235] 보조 인자 특이성을 결정하기 위하여, V_{max}/K_M 비를 포화 아세토라테이트에서 각각의 보조 인자에 대하여 계산할

수 있으며; NADH에 대하여 더 높은 비를 갖는 변이체는 등몰 농도의 두 보조 인자 및 포화 아세토락테이트의 조건 하에서 NADPH보다 NADH에서 더 높은 비율로 반응할 것이다. NADH 및 NADPH에 대한 V_{max} 및 K_M 값은 본 기술 분야에 공지되고/되거나 본 명세서에 제공된 방법을 이용하여 결정될 수 있다 (실시에 16 참조). 예를 들어, NADH 및 NADPH에 대한 V_{max} 및 K_M 값을 결정하기 위하여, 부분적으로 정제된 단백질이 NADH 및 NADPH의 다양한 농도에서 분석될 수 있다.

[0236] 본 발명에서 입증된 바와 같이 (실시에 16 및 실시에 18과 도 8 참조), K9G9에서의 추가의 아미노산의 치환은 NADH에 대한 특이성이 증가된 변이체로 이어진다. 따라서, K9 KARI (서열 번호 27)의 K57, Y53, 및 E74에 상응하는 위치들 중 하나 이상에서 또는 상기 위치들 전부에서 치환을 포함하는 폴리펩티드가 본 발명에서 제공된다. Y53, K57, E74, N87 및 K90에 상응하는 위치들 중 하나 이상에서 또는 상기 위치들 전부에서 치환을 포함하는 폴리펩티드가 본 발명에서 또한 제공된다. 실시 형태들에서, Y53에 상응하는 위치에서의 치환은 F이다. 실시 형태들에서, K57에 상응하는 위치에서의 치환은 E이다. 실시 형태들에서, E74에 상응하는 위치에서의 치환은 G이다. 실시 형태들에서, N87에 상응하는 위치에서의 치환은 P이다. 실시 형태들에서, K90에 상응하는 위치에서의 치환은 M 또는 L이다. 실시 형태들에서, 변이체는 본 명세서에서 확인된, 서열 번호 27의 S56 또는 S58에 상응하는 적어도 하나의 위치의 치환을 포함하며, 서열 번호 27의 위치들에 상응하는 하나 이상, 2개 이상, 3개 이상, 또는 3개 초과와 추가의 치환(들)을 추가로 포함한다.

[0237] 실시 형태들에서, 폴리펩티드는 야생형 서열과 관련하여 2개 미만, 3개 미만, 4개 미만, 5개 미만, 10개 미만, 15개 미만, 또는 20개 미만의 치환을 포함한다. 실시 형태들에서, 폴리펩티드는 실험적으로 검증된 KARI를 바탕으로 한, 그리고 표 Z에 주어진 프로파일 HMM을 $< 10^{-3}$ 미만의 E 값에 매칭시킨다.

[0238] 실시예에서 입증된 바와 같이, K9SB2 (서열 번호 427)의 변이체를 생성하였으며, NADPH 친화성이 감소된 변이체에 대하여 스크리닝하였고, 이는 추가의 치환 위치를 나타냈다. 따라서, 실시 형태들에서, 폴리펩티드는 서열 번호 427의 F53, G55, A56, W59, F67, I84, L85, Q91, M94, 및 P135에 상응하는 하나 이상의 위치에서의 치환을 추가로 포함한다. 실시 형태들에서, 위치 G55에서의 치환은 D 또는 C이며, 위치 Q91에서의 치환은 L이며, 위치 A56에서의 치환은 T 또는 V이며, P135에서의 치환은 S이며, 위치 F53에서의 치환은 L이며, 위치 M94에서의 치환은 I이며, 위치 F67에서의 치환은 L 또는 I이며, 위치 W59에서의 치환은 C이며, 위치 I84에서의 치환은 L이며, 위치 L85에서의 치환은 M이다.

[0239] KARI 구조

[0240] KARI 활성을 갖는 폴리펩티드의 확인 및 변형에 유용한 구조 정보가 본 기술 분야, 예를 들어 본 명세서에 기재된 참고 문헌에 제공되어 있으며, 이외에도 표 Z에 본 출원과 함께 제공된 프로파일 HMM에 그리고 본 명세서에 참고로 포함된 미국 특허 공개 제20100197519호 및 미국 특허 공개 제20090163376호에 제공되어 있다.

[0241] NADPH의 포스페이트 p2' 산소 원자는 시금치 KARI의 Arg162, Ser165 및 Ser167의 측쇄와 수소 결합을 형성하는 것으로 보고되었다 (문헌[Biou V., et al. The EMBO Journal, 16: 3405-3415, 1997]). 안(Ahn) 등 (문헌[J. Mol. Biol., 328: 505-515, 2003])에 의한 연구에 의하면, 슈도모나스 아에루기노사 (PAO-KARI)에 대하여, 그의 구조를 시금치 KARI의 것과 비교한 이후, 3개의 NADPH 포스페이트 결합 부위 (Arg47, Ser50 및 Thr52)가 확인되었다. NADPH, 아세토락테이트 및 마그네슘 이온이 결합된 PF5-KARI의 구조는, PF5 KARI에 대하여 아미노산 서열 상동성이 92%인 피. 아에루기노사 PAO1-KARI (PDB ID 1NP3, 문헌[Ahn H. J. et al., J. Mol. Biol., 328: 505-515, 2003])의 결정 구조를 바탕으로 하여 만들었다. PAO1-KARI 구조는 호모도데카머(homo-dodecamer)이며, 각각의 도데카머는 광범위한 다이머 계면을 갖는 6개의 호모다이머(homo-dimer)로 이루어진다. KARI의 활성 부위는 이 다이머 계면에 위치한다. 생물학적 조립체는 사면체의 모서리 상에 위치하는 6개의 호모다이머에 의해 형성되며, 이는 23 포인트 그룹 대칭의 고도로 대칭적인 도데카머를 생성한다.

[0242] PF5-KARI 다이머의 모델은 딥뷰(DeepView)/스위스(Swiss) PDB 뷰어를 사용하여 PF5-KARI의 서열과 PAO1-KARI의 단량체 A 및 단량체 B의 좌표를 바탕으로 하여 만들었다 (문헌[Guex, N. and Peitsch, M.C., Electrophoresis, 18: 2714-2723, 1997]). 그 후, 이 모델은 추가의 변형을 위하여 실리콘 그래픽스(Silicon Graphics) 시스템 상의 프로그램 0로 임포트(import)하였다 (문헌[Jones, T.A. et al, Acta Crystallogr. A 47: 110-119, 1991]).

[0243] PAO1-KARI의 구조는 NADPH, 기질 또는 저해제 또는 마그네슘을 활성 부위 내에 갖지 않는다. 따라서, 아세토락테이트 결합 부위 내에 마그네슘 이온, NADPH 및 저해제 (N-하이드록시-N-아이소프로필옥사메이트)를 갖는 상추 KARI 구조 (PDB ID 1yve, 문헌[Biou V. et al., The EMBO Journal, 16: 3405-3415, 1997.])를 사용하여 활성

부위에서 이들 분자를 모델링하였다. 식물 KARI는 PF5- 또는 PA01 KARI (<20%의 아미노산 동일성) 중 어느 하나에 대하여 매우 적은 서열 상동성을 갖지만, 이들 두 KARI 효소의 활성 부위 영역에서의 구조는 매우 유사하다. 이들 두 KARI 구조의 활성 부위를 오버레이(overlay)하기 위하여, 프로그램 0에서의 명령어 LSQ_ext, LSQ_improve, LSQ_mol을 사용하여 상추 KARI의 단량체 A의 활성 부위를 PF5 KARI 모델의 단량체 A에 대하여 정렬(line up)하였다. 상추 KARI의 활성 부위 내에 결합된 NADPH, 2개의 마그네슘 이온 및 저해제의 좌표를 추출하고, PF5 KARI의 분자 A에 대하여 포함시켰다. 이들 분자의 좌표 세트는 상기 프로그램에 의해 계산된, 단량체 A로부터 단량체 B로의 변환 오퍼레이터(operator)를 적용함으로써 PF5 KARI의 단량체 B에 대하여 생성하였다.

[0244] PA01 KARI 결정 구조의 활성 부위 내에는 NADPH가 없기 때문에, NADPH 결합 부위 (PA01 KARI에서 잔기 44-45, 상추 KARI에서 157-170)에서의 포스페이트 결합 루프 영역의 구조는 상기 둘 사이에서 매우 상이하다. NADPH 결합된 형태를 모델링하기 위하여, PF5-KARI 포스페이트 결합 루프 (44-55)의 모델을 lyve (157-170)의 것으로 대체하였다. 이들 둘 사이의 측쇄들의 임의의 불일치는 프로그램 0에서 mutate_replace 명령어를 사용하여 PF5-KARI 서열에서의 측쇄로 변환시키고, 대체된 측쇄들의 배좌를 수동으로 조정하였다. NADPH/Mg/저해제 결합된 전체 다이머형 PF5-KARI 모델은 프로그램 CNX (미국 캘리포니아주 샌디에고 소재의 액셀리스(ACCELRY), 문헌[Burnger, A.T. and Warren, G.L., Acta Crystallogr., D 54: 905-921, 1998])를 이용하여 1 라운드의 에너지 최소화화를 조사하고, 그 후 상기 모델에서 저해제를 기질, 아세토락테이트 (AL)로 대체하였다.

[0245] 아이소부탄올 생성

[0246] 본 발명에서 제공된 숙주 세포는 케틀-산 리덕토아이스모라아제 활성을 갖는 폴리펩티드를 포함할 수 있다. 본 명세서에서 기재되고 입증된 바와 같이, 본 출원인은 추가의 KARI 효소 및 상기 추가의 KARI의 변이체를 발견하였는데, 이는 락토코커스 락티스 유래의 KARI에서 관찰되는 것에 비견되는 및/또는 이를 능가하는 아이소부탄올 생성으로 이어진다 (실시에 참조). 따라서, 실시 형태들에서, 아이소부탄올 생성 경로에서 기능하는 KARI 활성을 갖는 폴리펩티드는 락토코커스 락티스 KARI (서열 번호 380)에서의 것에 비견되는 또는 이보다 더 우수한 역가로 아이소부탄올을 생성하고/생성하거나 유효 아이소부탄올 생산성을 갖는다. 따라서, 이러한 폴리펩티드는 특히 본 명세서에 기재된 아이소부탄올 생성 경로를 포함하는 세포에서, 아이소부탄올 생성에 유용한 것으로 간주된다. 실시 형태들에서, 본 발명에서 제공된 폴리펩티드는 동일한 조건 하에서 락토코커스 락티스 KARI (서열 번호 380)에서 관찰되는 것보다 더 큰 또는 대략적으로 동일한 역가로 아이소부탄올을 생성하고/하거나 유효 아이소부탄올 생산성을 갖는다. 실시 형태들에서, 본 발명에서 제공된 폴리펩티드는 약 48시간 후 세포 1 g당 약 3 g 초과, 약 4 g 초과 또는 약 5 g 초과, 또는 세포 1 g 당 약 6 g 초과와 유효 아이소부탄올 생산성을 가지며, 여기서, 상기 48시간 중 적어도 마지막의 약 24시간은 혐기성 조건 하에 있다.

[0247] 더욱이, 본 출원인은 NADH에 대한 K_M 이 비치환 폴리펩티드의 것보다 더 낮은 것을 비롯하여, 상기에 기재된 KARI 활성을 갖는 폴리펩티드의 변이체가 혐기성 조건 하에서 아이소부탄을 생성에 있어서 이점을 제공함을 발견하였다. 이론에 의해 구해되고자 함이 없이, 이러한 변이체는 환원 등가물로서의 NADH의 더욱 효과적인 사용으로 인하여 개선된 아이소부탄을 생성을 제공하는 것으로 여겨진다. 실시 형태들에서, 이러한 변이체를 이용한 아이소부탄을 생성은 감소된 글리세롤 축적을 제공한다. 실시 형태들에서, 글리세롤에 대한 아이소부탄올의 몰비는 NADH에 대한 K_M 이 비치환 폴리펩티드의 것보다 더 낮은, 상기에 기재된 KARI 활성을 갖는 폴리펩티드의 변이체에 있어서 증가된다. 실시 형태들에서, 글리세롤에 대한 아이소부탄올의 몰비는 1 초과이다. 실시 형태들에서, 글리세롤에 대한 아이소부탄올의 몰비는 2 초과이다. 실시 형태들에서, 상기 몰비는 3 초과이다. 실시 형태들에서, 상기 몰비는 4 초과, 5 초과, 6 초과, 7 초과, 8 초과, 9 초과, 10 초과, 12 초과, 또는 14 초과이다. 실시 형태들에서, 상기 몰비는 약 1 내지 5, 약 1 내지 10, 약 2 내지 8, 약 5 내지 10, 약 5 내지 15, 약 10 내지 15 또는 약 12 내지 15의 범위이다.

[0248] 본 명세서에서 실시예에서 입증된 바와 같이, $(\text{NADH } V_{\text{max}}/K_M)/(\text{NADPH } V_{\text{max}}/K_M)$ 로 정의되는, NADH 보조 인자에 대한 생화학적 특이성이 증가함에 따라 아이소부탄올/글리세롤 비의 증가가 관찰되며, 이는 변경된 보조 인자 특이성이 감소된 NADPH 이용 및 부산물 형성에 이르게 되었음을 시사하는 것이다.

[0249] 알데히드 탈하이드로게나제의 변형

[0250] 본 발명의 실시 형태들에서, 재조합 숙주 세포는 감소되거나 또는 제거된 알데히드 데하이드로게나아제 활성 및 아이소부탄올 생합성 경로를 포함할 수 있으며, 여기서, 숙주 세포는 부탄올을 생성한다. 다른 실시 형태들에서, 재조합 숙주 세포는 아이소부탄올 또는 1-부탄올 생합성 경로를 포함할 수 있으며, 이는 본 명세서에 추가

로 기재된 바와 같다. 다른 실시 형태들에서, 아이소부탄올 생합성 경로는 하기로 이루어진 군으로부터 선택되는 기질의 생성물로의 전환을 촉매하는 폴리펩티드를 코딩하는 폴리뉴클레오티드를 포함할 수 있다: (a) 피루베이트의 아세토락테이트로의 전환; (b) 아세토락테이트의 2,3-다이하이드록시아이소발레레이트로의 전환; (c) 2,3-다이하이드록시아이소발레레이트의 2-케토아이소발레레이트로의 전환; (d) 2-케토아이소발레레이트의 아이소부티르알데히드로의 전환; 및 (e) 아이소부티르알데히드의 아이소부탄올로의 전환. 다른 실시 형태들에서, 아이소부탄올 생합성 경로는 아세토락테이트 신타아제, 케토산 리덕토아이소머라아제, 다이하이드록시산 데하이드라타아제, 케토아이소발레이트 데카르복실라아제, 및 알코올 데하이드로게나아제 활성을 갖는 폴리펩티드를 코딩하는 폴리뉴클레오티드를 포함할 수 있다. 다른 실시 형태들에서, 재조합 세포는 1-부탄올 생합성 경로를 포함한다. 다른 실시 형태들에서, 1-부탄올 생합성 경로는 하기로 이루어진 군으로부터 선택되는 기질의 생성물로의 전환을 촉매하는 폴리펩티드를 코딩하는 폴리뉴클레오티드를 포함한다: (a) 아세틸-CoA의 아세토아세틸-CoA로의 전환; (b) 아세토아세틸-CoA의 3-하이드록시부티릴-CoA로의 전환; (c) 3-하이드록시부티릴-CoA의 크로토닐-CoA로의 전환; (d) 크로토닐-CoA의 부티릴-CoA로의 전환; (e) 부티릴-CoA의 부티르알데히드로의 전환; (f) 부티르알데히드의 1-부탄올로의 전환. 다른 실시 형태들에서, 1-부탄올 생합성 경로는 활성을 갖는 폴리펩티드를 코딩하는 폴리뉴클레오티드를 포함할 수 있다.

[0251] 본 발명의 실시 형태들에서, 재조합 숙주 세포는 알데히드 데하이드로게나아제 활성을 갖는 폴리펩티드를 코딩하는 폴리뉴클레오티드 또는 유전자의 변형 또는 파괴, 또는 알데히드 데하이드로게나아제 활성을 갖는 폴리펩티드의 변형 또는 파괴를 포함할 수 있다. 발현의 감소 또는 제거를 위한 표적 유전자의 많은 유전자 변형 및 파괴 방법이 당업자에게 공지되어 있으며, 상기 많은 방법을 이용하여 본 명세서에 개시된 재조합 숙주 세포를 생성할 수 있다. 다른 실시 형태들에서, 재조합 숙주 세포는 알데히드 데하이드로게나아제 활성을 갖는 폴리펩티드를 코딩하는 내인성 폴리뉴클레오티드 또는 유전자 내에 또는 알데히드 데하이드로게나아제 활성을 갖는 내인성 폴리펩티드 내에 결실, 돌연변이 및/또는 치환을 포함할 수 있다. 이러한 변형, 파괴, 결실, 돌연변이 및/또는 치환은 감소되거나 또는 제거된 알데히드 데하이드로게나아제 활성으로 이어질 수 있다. 사용될 수 있는 변형은 알데히드 데하이드로게나아제 단백질을 코딩하는 전체 유전자 또는 상기 유전자의 일부분의 결실, DNA 단편을 코딩 유전자 내로 (프로모터 또는 코딩 영역 중 어느 하나에서) 삽입하여 단백질이 발현되지 않게 하거나 더 낮은 수준으로 발현되도록 하는 것, 프레임 이동 또는 종결 코돈을 추가하는 코딩 영역 내로 돌연변이를 도입하여 기능성 단백질이 발현되지 않게 하는 것, 및 하나 이상의 돌연변이를 코딩 영역 내로 도입하여 아미노산을 변경시켜서, 비기능성 단백질 또는 덜 활성인 단백질이 발현되게 하는 것을 포함하지만, 이에 한정되지 않는다. 다른 실시 형태들에서, 표적 유전자의 발현은 안티센스 RNA 또는 간섭 RNA의 발현에 의해 차단될 수 있으며, 동시 억제(cosuppression)로 이어지는 제작물이 도입될 수 있다. 다른 실시 형태들에서, 전사체의 합성 또는 안정성이 돌연변이에 의해 감퇴될 수 있다. 실시 형태들에서, mRNA로부터 단백질이 번역되는 효율을 돌연변이에 의해 조절할 수 있다. 이들 방법 전부는 표적 단백질을 코딩하는 공지되거나 또는 확인된 서열을 이용하여 당업자에 의해 용이하게 실시될 수 있다.

[0252] 다른 실시 형태들에서, 표적 알데히드 데하이드로게나아제 코딩 서열을 둘러싸고 있는 DNA 서열들이 또한 일부 변형 절차에서 유용하며, 이는 예를 들어 GOPID #13838로 확인되는, NCBI(미국 국립 생물공학 정보 센터)에 의해 조직화된 게놈 프로젝트(Genome Projects)의 게놈 프로젝트 ID9518에 의해 조직화된 전 게놈 서열에 있어서 사카로마이세스 세레비지애와 같은 효모를 위하여 이용가능하다. 효모 게놈 서열의 추가의 비제한적 예로는 칸디다 알비칸스(*Candida albicans*)의 것이 있으며, 이는 GPID #10771, #10701 및 #16373에 포함된다. 다른 효모 게놈 서열은 공개적으로 이용가능한 데이터베이스에서 당업자에 의해 용이하게 찾아질 수 있다.

[0253] 다른 실시 형태들에서, 표적 알데히드 데하이드로게나아제 코딩 서열을 둘러싸고 있는 DNA 서열은 상동 재조합을 이용한 변형 방법에 유용할 수 있다. 이 방법의 비제한적 예에서, 알데히드 데하이드로게나아제 유전자 측면 서열은 선발가능한 마커 유전자에 접하여 두어져서, 상기 마커 유전자가 알데히드 데하이드로게나아제 유전자를 대체하는 상동 재조합을 매개할 수 있다. 다른 비제한적 예에서, 부분적인 알데히드 데하이드로게나아제 유전자 서열 및 선발가능한 마커 유전자에 접해 있는 알데히드 데하이드로게나아제 유전자 측면 서열은 마커 유전자가 표적 알데히드 데하이드로게나아제 유전자의 일부분을 대체하는 상동 재조합을 매개하는 데 사용될 수 있다. 실시 형태들에서, 선발가능한 마커는 부위 특이적 재조합 부위와 접해 있을 수 있어서, 상응하는 부위 특이적 리콤비나아제의 발현 이후, 내성 유전자가 알데히드 데하이드로게나아제 유전자로부터, 후자의 재활성화 없이 절단되게 된다. 실시 형태들에서, 부위 특이적 재조합은 알데히드 데하이드로게나아제 단백질의 발현을 파괴하는 재조합 부위를 뒤에 남긴다. 다른 실시 형태들에서, 상동 재조합 벡터가 또한 선발가능한 마커의 절단 이후 알데히드 데하이드로게나아제 유전자에 있어서 결실을 남기도록 제작될 수 있으며, 이는 당업자에게 잘

알려진 바와 같다.

- [0254] 다른 실시 형태들에서, 결실은 유사 분열 재조합을 이용하여 알데히드 데하이드로게나아제 표적 유전자에 대하여 이루어질 수 있으며, 이는 문헌[Wach et al., *Yeast*, 10:1793-1808; 1994]에 기재된 바와 같다. 이러한 방법은 20 bp만큼 짧을 수 있고 표적 DNA 서열에 접해 있는 게놈 영역들 사이에 선발가능한 마커를 함유하는 DNA 단편을 제조하는 것을 포함할 수 있다. 다른 실시 형태들에서, 이 DNA 단편은 마커 유전자의 말단에 혼성화되고 효모 게놈과 재조합될 수 있는 게놈 영역을 포함하는 올리고뉴클레오타이드를 프라이머로 사용하여, 선발가능한 마커 유전자의 PCR 증폭에 의해 제조될 수 있다. 실시 형태들에서, 선행 DNA 단편은 효율적으로 효모를 형질전환시키고 게놈 내로 재조합되어서 표적 DNA 서열의 결실을 함께 포함하는 유전자 대체로 이어질 수 있다 (이는 예를 들어 문헌[Methods in Enzymology, Volume 194, Guide to Yeast Genetics and Molecular and Cell Biology (Part A, 2004, Christine Guthrie and Gerald R. Fink (Eds.), Elsevier Academic Press, San Diego, CA)]에 개시된 바와 같음).
- [0255] 게다가, 프로모터 대체 방법을 이용하여 내인성 전사 제어 요소를 교체하여 문헌[Mnaimneh et al., (2004) *Cell* 118(1):31-44]에 기재된 것과 같이 발현을 조정하는 다른 수단을 허용할 수 있다.
- [0256] 다른 실시 형태들에서, 알데히드 데하이드로게나아제 표적 유전자 코딩된 활성은 랜덤 돌연변이 유발을 이용하여 파괴될 수 있으며, 그 후, 이것에 이어서 스크리닝을 하여 감소되거나 또는 사실상 제거된 활성을 갖는 주를 확인할 수 있다. 이러한 유형의 방법에서, 표적 유전자 코딩 영역의 DNA 서열, 또는 성장에 있어서 탄소 기질 의존성에 영향을 주는 게놈의 임의의 다른 영역은 공지되어 있을 필요가 없다. 실시 형태들에서, 알데히드 데하이드로게나아제 활성이 감소된 세포, 또는 알데히드 데하이드로게나아제 활성이 감소된 다른 돌연변이체에 대한 스크리닝은 본 발명의 재조합 숙주 세포에 유용할 수 있다.
- [0257] 유전자 돌연변이를 생성하는 방법은 보편적이며, 본 기술 분야에 잘 알려져 있고, 돌연변이체를 생성하는 활동에 적용될 수 있다. 보편적으로 사용되는 랜덤 유전자 변형 방법(문헌[Methods in Yeast Genetics, 2005, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY]에 개관됨)은 자연적 돌연변이 유발, 돌연변이 유발 유전자에 의해 야기되는 돌연변이, 화학적 돌연변이, UV 또는 X선을 사용한 조사, 또는 트랜스포존 돌연변이를 포함한다.
- [0258] 숙주 세포의 화학적 돌연변이는 하기 DNA 돌연변이원들 중 하나로 처리하는 것을 포함할 수 있지만, 이에 한정되지 않는다: 에틸 메탄설포네이트 (EMS), 아질산, 다이메틸 설페이트 또는 N-메틸-N'-니트로-N-니트로소-구아니딘 (MNNG). 이러한 돌연변이 유발 방법은 문헌[Spencer et al., *Mutagenesis in Yeast*, 1996, *Yeast Protocols: Methods in Cell and Molecular Biology*. Humana Press, Totowa, NJ]에 개관되었다. 실시 형태들에서, EMS에 의한 화학적 돌연변이 유발은 문헌[Methods in Yeast Genetics, 2005, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY]에 개시된 바와 같이 수행될 수 있다. 자외광(UV light) 또는 X선을 사용한 조사가 또한 효모 세포에서 랜덤 돌연변이 유발을 생성하는데 사용될 수 있다. UV 조사에 의한 돌연변이 유발의 일차 효과는 DNA 복제의 충실도(fidelity)를 파괴하는 피리미딘 다이머의 형성이다. 효모의 UV-돌연변이 유발에 대한 프로토콜은 문헌[Spencer et al., *Mutagenesis in Yeast*, 1996, *Yeast Protocols: Methods in Cell and Molecular Biology*. Humana Press, Totowa, NJ]에서 찾아볼 수 있다. 실시 형태들에서, 돌연변이 유발 유전자 표현형의 도입이 또한 숙주 세포에서의 랜덤 염색체 돌연변이의 생성에 사용될 수 있다. 실시 형태들에서, 보편적인 돌연변이 유발 유전자 표현형은 하기 유전자들 중 하나 이상의 파괴를 통하여 수득될 수 있다: PMS1, MAG1, RAD18 또는 RAD51. 다른 실시 형태들에서, 비-돌연변이 유발 유전자 표현형의 복구는 야생형 대립 유전자의 삽입에 의해 수득될 수 있다. 다른 실시 형태들에서, 이들 또는 다른 공지된 랜덤 돌연변이 유발 공정 중 임의의 것으로부터 생성된 변형된 세포의 컬렉션은 감소되거나 또는 제거된 알데히드 데하이드로게나아제 활성에 대하여 스크리닝될 수 있다.
- [0259] 게놈들을 완전히 서열결정하고 주석을 달았으며, 이는 하기 효모 주에 대하여 공개적으로 입수가가능하다: 아쉬비아 고시피이(*Ashbya gossypii*) ATCC 10895, 칸디다 글라브라타(*Candida glabrata*) CBS 138, 클루이베로마이세스 락티스 NRRL Y-1140, 피키아 스티피티스 CBS 6054, 사카로마이세스 세레비지에 S288c, 쉬조사카로마이세스 폼베(*Schizosaccharomyces pombe*) 972h-, 및 야로위아 리포라이티카(*Yarrowia lipolytica*) CLIB122. 전형적으로, 본 명세서에 제공된 것과 같은, 공지된 알데히드 데하이드로게나아제 폴리뉴클레오타이드 또는 폴리펩티드 서열을 이용한 공개적으로 입수가가능한 데이터베이스의 BLAST (상기에 기재됨) 검색은 효모 세포와 같은 다른 숙주 세포의 알데히드 데하이드로게나아제-코딩 서열을 확인하는 데 사용된다.
- [0260] 다른 실시 형태들에서, 알데히드 데하이드로게나아제 활성을 갖는 폴리펩티드는 아이소부티르알데히드의 아이소

부티르산으로의 전환을 촉매할 수 있다. 다른 실시 형태들에서, 재조합 숙주 세포에서의 아이소부티르알데히드의 아이소부티르산으로의 전환은 감소되거나 또는 제거된다. 또 다른 실시 형태에서, 알데히드 데하이드로게나아제 활성을 갖는 폴리펩티드, 유전자 또는 폴리뉴클레오티드는 효소 위원회 번호 EC 1.2.1.3, EC 1.2.1.4, 및/또는 EC 1.2.1.5에 상응할 수 있다.

[0261] 실시 형태들에서, 본 발명의 재조합 숙주 세포는 에스. 세레비지애이며, 알데히드 데하이드로게나아제 활성을 갖는 폴리펩티드는 ALD2, ALD3, ALD4, ALD5, ALD6, 또는 이의 조합일 수 있다. 다른 실시 형태들에서, 재조합 숙주 세포는 클루이베로마이세스 락티스일 수 있으며, 알데히드 데하이드로게나아제 활성을 갖는 폴리펩티드는 KLLA0F00440, KLLA0E23057, KLLA0D10021, KLLA0D09999G, 또는 이의 조합일 수 있다. 다른 실시 형태들에서, 재조합 숙주 세포는 피키아 스티피티스일 수 있으며, 알데히드 데하이드로게나아제 활성을 갖는 폴리펩티드는 ALD2, ALD3, ALD4, ALD5, ALD7, 또는 이의 조합일 수 있다. 다른 실시 형태들에서, 재조합 숙주 세포는 락토바실러스 플란타룸일 수 있으며, 알데히드 데하이드로게나아제 활성을 갖는 폴리펩티드는 AldH일 수 있다. 다른 실시 형태들에서, 재조합 숙주 세포는 이. 콜라이일 수 있으며, 알데히드 데하이드로게나아제 활성을 갖는 폴리펩티드는 aldA, aldB, aldH 또는 이의 조합일 수 있다.

[0262] 본 발명의 실시 형태들에서, 재조합 숙주 세포는 에스. 세레비지애일 수 있으며, 알데히드 데하이드로게나아제 활성을 갖는 폴리펩티드를 코딩하는 내인성 폴리뉴클레오티드 또는 유전자는 ALD2, ALD3, ALD4, ALD5, ALD6 또는 이의 조합일 수 있다. 본 발명의 실시 형태들에서, 재조합 숙주 세포는 에스. 세레비지애일 수 있으며, 알데히드 데하이드로게나아제 활성을 갖는 폴리펩티드를 코딩하는 내인성 폴리뉴클레오티드 또는 유전자는 ALD6일 수 있다. 다른 실시 형태들에서, 재조합 숙주 세포는 클루이베로마이세스 락티스일 수 있으며, 알데히드 데하이드로게나아제 활성을 갖는 폴리펩티드를 코딩하는 내인성 폴리뉴클레오티드 또는 유전자는 KLLA0F00440, KLLA0E23057, KLLA0D10021, KLLA0D09999G 또는 이의 조합일 수 있다. 다른 실시 형태들에서, 재조합 숙주 세포는 피키아 스티피티스일 수 있으며, 알데히드 데하이드로게나아제 활성을 갖는 폴리펩티드를 코딩하는 내인성 폴리뉴클레오티드 또는 유전자는 ALD2, ALD3, ALD4, ALD5, ALD7 또는 이의 조합일 수 있다. 실시 형태들에서, 알데히드 데하이드로게나아제 활성을 갖는 폴리펩티드는 사카로마이세스 세레비지애 유래의 ALD6의 상동체이다. kanMX 카세트에 의한 알데히드 데하이드로게나아제 유전자 결실을 함유하는 에스. 세레비지애 결실 주는 아메리칸 타입 컬처 컬렉션(American Type Culture Collection) [카탈로그 번호: 4000753]으로부터 구매가능하다.

[0263] 다른 실시 형태들에서, 재조합 숙주 세포는 락토바실러스 플란타룸일 수 있으며, 알데히드 데하이드로게나아제 활성을 갖는 폴리펩티드를 코딩하는 내인성 폴리뉴클레오티드는 AldH일 수 있다. 다른 실시 형태들에서, 재조합 숙주 세포는 이. 콜라이일 수 있으며, 알데히드 데하이드로게나아제 활성을 갖는 폴리펩티드를 코딩하는 내인성 폴리뉴클레오티드는 aldA, aldB, aldH 또는 이의 조합일 수 있다.

[0264] 본 명세서에 개시된 재조합 숙주 세포에서 변형 또는 불활성화용으로 표적화될 수 있는 알데히드 데하이드로게나아제 폴리뉴클레오티드, 유전자 및 폴리펩티드의 예는 하기 표 4의 것을 포함하지만, 이에 한정되지 않는다.

[0265] [표 4]

알데히드 데하이드로게나아제 표적 유전자 코딩 영역 및 단백질

	핵산 서열 번호	아미노산 서열 번호
에스. 세레비자애 유래의 ALD2	732	733
에스. 세레비자애 유래의 ALD3	734	735
에스. 세레비자애 유래의 ALD4	736	737
에스. 세레비자애 유래의 ALD5	738	739
에스. 세레비자애 유래의 ALD6	740	741
클루이베로마이세스 락티스 유래의 KLLA0F00440	742	743
클루이베로마이세스 락티스 유래의 KLLA0E23057	744	745
클루이베로마이세스 락티스 유래의 KLLA0D10021	746	747
클루이베로마이세스 락티스 유래의 KLLA0D09999	748	749
피키아 스티피티스 유래의 ALD2	750	751
피키아 스티피티스 유래의 ALD3	752	753
피키아 스티피티스 유래의 ALD4	754	755
피키아 스티피티스 유래의 ALD5	756	757
피키아 스티피티스 유래의 ALD7	758	759
이. 콜라이 유래의 aldA	760	761
이. 콜라이 유래의 aldB	762	763
이. 콜라이 유래의 aldH (puuC)	764	765

[0266]

[0267]

본 명세서에 개시된 재조합 숙주 세포에서 변형 또는 불활성화용으로 표적화될 수 있는 알데히드 데하이드로게나아제 폴리뉴클레오티드, 유전자 및 폴리펩티드의 다른 예에는 표 4의 서열들 중 임의의 하나에 대하여 서열 동일성이 적어도 약 70% 내지 약 75%, 약 75% 내지 약 80%, 약 80% 내지 약 85%, 약 85% 내지 약 90%, 약 90% 내지 약 95%, 약 96%, 약 97%, 약 98%, 또는 약 99%인 알데히드 데하이드로게나아제 폴리뉴클레오티드, 유전자 및/또는 폴리펩티드가 포함되지만, 이에 한정되지 않으며, 여기서, 이러한 폴리뉴클레오티드 또는 유전자는 알데히드 데하이드로게나아제 활성을 코딩하거나 또는 이러한 폴리펩티드는 알데히드 데하이드로게나아제 활성을 갖는다. 본 명세서에 개시된 재조합 숙주 세포에서 변형 또는 불활성화용으로 표적화될 수 있는 알데히드 데하이드로게나아제 폴리뉴클레오티드, 유전자 및 폴리펩티드의 또 다른 예에는 표 4의 서열들 중 임의의 하나의 활성 변이체, 단편 또는 유도체가 포함되지만, 이에 한정되지 않으며, 여기서, 이러한 폴리뉴클레오티드 또는 유전자는 알데히드 데하이드로게나아제 활성을 코딩하거나 또는 이러한 폴리펩티드는 알데히드 데하이드로게나아제 활성을 갖는다.

[0268]

실시 형태들에서, 다른 알데히드 데하이드로게나아제 폴리뉴클레오티드, 유전자 및/또는 폴리펩티드의 서열은 본 발명에 개시된 그리고 본 기술 분야에서 입수가능한 서열을 이용하여 당업자에게 잘 알려진 생물 정보 데이

터베이스에서 그리고 문헌에서 확인될 수 있다. 예를 들어, 이러한 서열들은 공지된 알데히드 데하이드로게나아제 코딩 폴리뉴클레오티드 또는 폴리펩티드 서열을 이용하여 공개적으로 이용가능한 데이터베이스의 BLAST 검색을 통하여 확인될 수 있다. 이러한 방법에서, 동일성은 갭 페널티=10, 갭 길이 페널티=0.1, 및 고넷 250 시리즈의 단백질 가중 매트릭스의 디폴트 파라미터를 사용한 클러스탈 W 정렬 방법을 바탕으로 할 수 있다.

[0269] 부가적으로, 본 명세서에 개시되거나 또는 본 기술 분야에 공지된 알데히드 데하이드로게나아제 폴리뉴클레오티드 또는 폴리펩티드 서열은 자연에서 다른 알데히드 데하이드로게나아제 상동체를 확인하는 데 사용될 수 있다. 예를 들어, 본 명세서에 개시된 알데히드 데하이드로게나아제 코딩 핵산 단편 각각은 상동 단백질을 코딩하는 유전자를 단리하는 데 사용될 수 있다. 서열-의존성 프로토콜을 사용하는 상동 유전자의 단리는 본 기술 분야에 잘 알려져 있다. 서열-의존성 프로토콜의 예에는 (1) 핵산 혼성화 방법; (2) 핵산 증폭 기술 (예를 들어, 폴리머라아제 연쇄 반응 (PCR) (멀리스(Mullis) 등의 미국특허 제4,683,202호); 라이가아제 연쇄 반응 (LCR) (문헌[Tabor et al., Proc. Acad. Sci. USA 82:1074 (1985)]; 또는 가닥 대체 증폭(SDA) (문헌[Walker et al., Proc. Natl. Acad. Sci. U.S.A., 89:392 (1992)]으로 예시되는 바와 같이, DNA 및 RNA 증폭 방법; 및 (3) 상보성에 의한 스크리닝 및 라이브러리 제작 방법이 포함되지만, 이에 한정되지 않는다.

[0270] 따라서, 본 명세서에 개시된 알데히드 데하이드로게나아제 폴리뉴클레오티드 또는 폴리펩티드 (예를 들어, 표 4의 서열 번호 732-765) 중 임의의 것에 대하여 서열 동일성이 적어도 약 70% 내지 약 75%, 약 75% 내지 약 80%, 약 80% 내지 약 85%, 약 85% 내지 약 90%, 약 90% 내지 약 95%, 약 96%, 약 97%, 약 98%, 또는 약 99%인 알데히드 데하이드로게나아제 폴리뉴클레오티드, 유전자 및 폴리펩티드를 제공하는 것은 본 발명의 범주 이내이다. 동일성은 갭 페널티=10, 갭 길이 페널티=0.1, 및 단백질 가중 매트릭스의 고넷 250 시리즈의 디폴트 파라미터를 사용한 클러스탈 W 정렬 방법을 바탕으로 한다.

[0271] 알데히드 데하이드로게나아제 활성이 감소되거나 또는 제거되도록 본 명세서에 개시된 재조합 숙주 세포에서 알데히드 데하이드로게나아제를 변형시키는 것은 본 기술 분야에 공지된 방법을 이용하여 확인될 수 있다. 예를 들어, 특정 알데히드 데하이드로게나아제의 파괴는 알데히드 데하이드로게나아제 유전자에 대하여 내재적인 그리고 외재적인 프라이머를 이용한 PCR 스크리닝에 의해 또는 알데히드 데하이드로게나아제 유전자 서열에 대하여 설계된 프로브를 이용한 서던 블롯에 의해 확인될 수 있다. 대안적으로, 아이소부티르산의 형성 감소를 위하여 아이소부티르알데히드에 노출된 주를 스크리닝하기 위하여 가스 크로마토그래피-질량 분광법 또는 액체 크로마토그래피를 이용할 수 있다. 따라서, 아이소부티르산 형성이 감소된 주에 대한 스크리닝 방법이 본 발명에서 제공되며, 이는 a) 알데히드 데하이드로게나아제 활성을 갖는 폴리펩티드를 코딩하는 폴리뉴클레오티드에서의 변형 및/또는 알데히드 옥시다아제 활성을 갖는 폴리펩티드를 코딩하는 폴리뉴클레오티드에서의 변형을 포함하는 주를 제공하는 단계; b) 상기 세포를 아이소부티르알데히드와 접촉시키는 단계; 및 c) 아이소부티르산 형성량을 측정하는 단계를 포함하고, 여기서, 아이소부티르산 형성은 변형되지 않은 대조 주와 비교하여 감소된다. 일부 실시 형태에서, 변형은 결실, 돌연변이 및/또는 치환이다. 일부 실시 형태에서, 상기 측정된 가스 크로마토그래피-질량 분광법을 이용하여 실시된다. 일부 실시 형태에서, 아이소부티르산은 약 10% 이상, 20% 이상, 30% 이상, 40% 이상, 50% 이상, 60% 이상, 70% 이상, 80% 이상, 또는 90% 이상만큼 감소된다. 일부 실시 형태에서, 아이소부티르산 형성은 사실상 제거된다.

[0272] 알데히드 옥시다아제의 변형

[0273] 본 발명의 실시 형태들에서, 본 명세서에 개시된 재조합 숙주 세포는 알데히드 옥시다아제를 코딩하는 폴리뉴클레오티드, 유전자, 또는 폴리펩티드의 변형 또는 파괴를 가질 수 있다. 실시 형태들에서, 재조합 숙주 세포는 알데히드 옥시다아제 활성을 갖는 폴리펩티드를 코딩하는 내인성 폴리뉴클레오티드 또는 유전자 내에 또는 알데히드 옥시다아제 활성을 갖는 내인성 폴리펩티드 내에 결실, 돌연변이 및/또는 치환을 포함한다. 이러한 변형, 파괴, 결실, 돌연변이 및/또는 치환은 감소되거나 또는 제거된 알데히드 옥시다아제 활성으로 이어질 수 있다.

[0274] 본 발명의 실시 형태들에서, 알데히드 옥시다아제 활성을 갖는 폴리펩티드는 아이소부티르알데히드의 아이소부티르산으로의 전환을 촉매할 수 있다. 다른 실시 형태들에서, 재조합 숙주 세포에서의 아이소부티르알데히드의 아이소부티르산으로의 전환은 감소되거나 또는 제거된다. 다른 실시 형태들에서, 알데히드 옥시다아제 활성을 갖는 폴리펩티드, 유전자 또는 폴리뉴클레오티드는 효소 위원회 번호 EC 1.2.3.1에 상응할 수 있다.

[0275] 실시 형태들에서, 본 발명의 재조합 숙주 세포는 피키아 스티피티스일 수 있으며, 알데히드 옥시다아제 활성을 갖는 폴리펩티드, 유전자 또는 폴리뉴클레오티드는 AOX1 및/또는 AOX2일 수 있다.

[0276] 본 명세서에 개시된 재조합 숙주 세포에서 변형 또는 불활성화용으로 표적화될 수 있는 알데히드 옥시다아제 폴

리뉴클레오티드, 유전자 및 폴리펩티드의 예는 하기 표 5의 것을 포함하지만, 이에 한정되지 않는다.

[표 5]

알데히드 옥시다아제 표적 유전자 코딩 영역 및 단백질

	핵산 서열 번호	아미노산 서열 번호
피키아 스티피티스 유래의 AOX1	864	866
피키아 스티피티스 유래의 AOX2	867	868

본 명세서에 개시된 재조합 숙주 세포에서 변형 또는 불활성화용으로 표적화될 수 있는 알데히드 옥시다아제 폴리뉴클레오티드, 유전자 및 폴리펩티드의 다른 예에는 표 5의 서열들 중 임의의 하나에 대하여 서열 동일성이 적어도 약 70% 내지 약 75%, 약 75% 내지 약 80%, 약 80% 내지 약 85%, 약 85% 내지 약 90%, 약 90% 내지 약 95%, 약 96%, 약 97%, 약 98%, 또는 약 99%인 알데히드 옥시다아제 폴리뉴클레오티드, 유전자 및/또는 폴리펩티드가 포함되지만, 이에 한정되지 않으며, 여기서, 이러한 폴리뉴클레오티드 또는 유전자는 알데히드 옥시다아제 활성을 갖는 폴리펩티드를 코딩하거나 또는 이러한 폴리펩티드는 알데히드 옥시다아제 활성을 갖는다. 본 명세서에 개시된 재조합 숙주 세포에서 변형 또는 불활성화용으로 표적화될 수 있는 알데히드 옥시다아제 폴리뉴클레오티드, 유전자 및 폴리펩티드의 또 다른 예에는 표 5의 서열들 중 임의의 하나의 활성 변이체, 단편 또는 유도체가 포함되지만, 이에 한정되지 않으며, 여기서, 이러한 폴리뉴클레오티드 또는 유전자는 알데히드 옥시다아제 활성을 코딩하거나 또는 이러한 폴리펩티드는 알데히드 옥시다아제 활성을 갖는다.

실시 형태들에서, 본 명세서에 개시되거나 또는 본 기술 분야에 공지된 알데히드 옥시다아제 서열을 코딩하는 폴리뉴클레오티드, 유전자 및/또는 폴리펩티드는 아세트락테이트 리덕타아제 또는 알데히드 데하이드로게나아제에 대하여 상기에 개시된 바와 같이, 변형될 수 있다. 다른 실시 형태들에서, 알데히드 옥시다아제를 코딩하는 폴리뉴클레오티드, 유전자 및/또는 폴리펩티드는 다른 알데히드 옥시다아제 폴리뉴클레오티드, 유전자 및/또는 폴리펩티드 서열을 확인하는 데 사용될 수 있고/있거나 다른 세포에서 알데히드 옥시다아제 상동체를 확인하는 데 사용될 수 있으며, 이는 알데히드 데하이드로게나아제에 대하여 상기에 개시된 바와 같다. 이러한 알데히드 옥시다아제 코딩 서열은 예를 들어 당업자에게 잘 알려진 생물 정보 데이터베이스에서 및/또는 문헌에서 확인될 수 있다. 예를 들어, 생물 정보학을 이용하여 다른 세포 유형에서 알데히드 옥시다아제 코딩 서열을 확인하는 것은 본 명세서에 제공된 것들 중 임의의 것과 같이, 공지된 핵소스 키나아제 코딩 DNA 및 폴리펩티드 서열에 의해 공개적으로 이용가능한 데이터베이스의 BLAST (상기에 개시된 바와 같음) 검색을 통하여 달성될 수 있다. 동일성은 갭 페널티=10, 갭 길이 페널티=0.1, 및 단백질 가중 매트릭스의 고넷 250 시리즈의 디폴트 파라미터를 사용한 클러스탈 W 정렬 방법을 바탕으로 한다.

알데히드 옥시다아제 활성이 감소되거나 또는 제거되도록 본 명세서에 개시된 재조합 숙주 세포에서 알데히드 옥시다아제를 변형시키는 것은 본 기술 분야에 공지된 방법을 이용하여 확인될 수 있다. 예를 들어, 특정 알데히드 옥시다아제의 파괴는 알데히드 옥시다아제 유전자에 대하여 내재적인 그리고 외재적인 프라이머를 이용한 PCR 스크리닝에 의해 또는 알데히드 옥시다아제 유전자 서열에 대하여 설계된 프로브를 이용한 서던 블롯에 의해 확인될 수 있다. 대안적으로, (실시예에서 기재되고 입증된 바와 같이) 아이소부티르산의 형성 감소를 위하여 아이소부티르알데히드에 노출된 주를 스크리닝하기 위하여 가스 크로마토그래피 또는 다른 분석 방법을 이용할 수 있다. 일부 실시 형태에서, 아이소부티르산은 약 10% 이상, 20% 이상, 30% 이상, 40% 이상, 50% 이상, 60% 이상, 70% 이상, 80% 이상, 또는 90% 이상만큼 감소된다. 일부 실시 형태에서, 아이소부티르산 형성은 사실상 제거된다.

본 출원인은 감소되거나 또는 제거된 알데히드 데하이드로게나아제 및/또는 알데히드 옥시다아제 활성을 포함하는 재조합 숙주 세포를 제공하였다. 실시 형태들에서, 본 명세서에 개시된 재조합 숙주 세포는 피루베이트 데카르복실라아제 활성을 갖는 폴리펩티드를 코딩하는 폴리뉴클레오티드에서의 변형 및/또는 핵소키나아제 2 활성을 갖는 폴리펩티드를 코딩하는 폴리뉴클레오티드에서의 변형을 추가로 포함할 수 있다. 실시 형태들에서, 본 발명의 재조합 숙주 세포는 생합성 경로의 생성물 (예를 들어 아이소부탄올)을 생성할 수 있으며, 하기로 이루어진 군으로부터 선택되는 기질의 생성물로의 전환을 촉매하는 폴리펩티드를 코딩하는 폴리뉴클레오티드를 포함할 수 있다: (a) 피루베이트의 아세트락테이트로의 전환; (b) 아세트락테이트의 2,3-다이하이드록시아이소발레레이트로의 전환; (c) 2,3-다이하이드록시아이소발레레이트의 2-케토아이소발레레이트로의 전환; (d) 2-케토아이소발레레이트의 아이소부티르알데히드로의 전환; 및 (e) 아이소부티르알데히드의 아이소부탄올로의 전환. 다

른 실시 형태들에서, 이러한 재조합 숙주 세포는 생합성 경로의 생성물 (예를 들어, 아이소부탄올)을, 감소되거나 또는 제거된 알데히드 데하이드로게나아제 활성 및/또는 알데히드 옥시다아제 활성을 포함하지 않는 재조합 숙주 세포에 의해 생성되는 동일 생성물의 수율 또는 양보다 더 큰 수율 또는 양으로 생성할 수 있다. 다른 실시 형태들에서, 본 발명의 재조합 숙주 세포는 아이소부티르알데히드의 아이소부티르산으로의 전환을 감소시키거나 또는 제거할 수 있으며, 알데히드 데하이드로게나아제 및/또는 알데히드 옥시다아제 활성을 갖는 후보 폴리펩티드의 스크리닝에 사용될 수 있다. 이와 같이, 본 출원인은 생합성 경로의 생성물 (예를 들어, 아이소부탄올)의 수율 또는 역가를 증가시키는 방법, 아이소부티르알데히드의 아이소부티르산으로의 전환을 감소시키거나 또는 제거하는 방법, 및 알데히드 데하이드로게나아제 및/또는 알데히드 옥시다아제 활성을 갖는 후보 폴리펩티드의 스크리닝 방법을 또한 제공하였다.

[0283] 본 발명의 실시 형태들에서, 재조합 숙주 세포를 제조하는 방법으로서, (a) 본 명세서에 개시된 재조합 숙주 세포를 제공하는 단계; 및 (b) 생합성 경로 (예를 들어, 아이소부탄올 생합성 경로)의 폴리펩티드를 코딩하는 폴리뉴클레오티드로 상기 숙주 세포를 형질전환시키는 단계를 포함하는 방법이 제공된다. 다른 실시 형태들에서, 재조합 숙주 세포를 제조하는 방법으로서, (a) 알데히드 데하이드로게나아제 활성을 갖는 폴리펩티드를 코딩하는 폴리뉴클레오티드 내에 또는 알데히드 데하이드로게나아제 활성을 갖는 폴리펩티드 내에 변형을 포함하는 재조합 숙주 세포를 제공하는 단계; 및 (b) 생합성 경로 (예를 들어, 아이소부탄올 생합성 경로)의 폴리펩티드를 코딩하는 폴리뉴클레오티드로 상기 숙주 세포를 형질전환시키는 단계를 포함하는 방법이 제공된다. 다른 실시 형태들에서, 재조합 숙주 세포를 제조하는 방법으로서, (a) 알데히드 옥시다아제 활성을 갖는 폴리펩티드를 코딩하는 폴리뉴클레오티드 내에 또는 알데히드 옥시다아제 활성을 갖는 폴리펩티드 내에 변형을 포함하는 재조합 숙주 세포를 제공하는 단계; 및 (b) 아이소부탄올 생합성 경로의 폴리펩티드를 코딩하는 폴리뉴클레오티드로 상기 숙주 세포를 형질전환시키는 단계를 포함하는 방법이 제공된다.

[0284] 실시 형태들에서, 아이소부티르알데히드의 아이소부티르산으로의 전환을 감소시키거나 또는 제거하는 방법이 제공되며, 이는 (a) 본 명세서에 개시된 재조합 숙주 세포를 제공하는 단계; 및 (b) 아이소부티르알데히드의 아이소부티르산으로의 전환이 감소되거나 또는 제거된 알데히드 데하이드로게나아제 및/또는 알데히드 옥시다아제 활성을 포함하지 않는 재조합 숙주 세포와 비교하여 감소되거나 또는 제거되는 조건 하에서 상기 숙주 세포를 성장시키는 단계를 포함한다. 본 명세서에 개시된 재조합 숙주 세포의 아이소부티르알데히드의 아이소부티르산으로의 전환은 본 기술 분야에 공지된 방법 (예를 들어, 문헌[Methods in Yeast Genetics, 2005, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, pp. 201-202]) 및/또는 본 명세서에 기재된 방법에 의해 측정될 수 있다.

[0285] DHMB의 감소

[0286] 아이소부탄올 생합성 경로를 포함하는 숙주 세포에서의 DHMB의 생성은 이 경로의 기질들 전부가 원하는 생성물로 전환되고 있는 것은 아님을 나타낸다. 따라서, 수율은 저하된다. 게다가, DHMB는 생성물 생성에 대하여 저해 효과를 가질 수 있다. 예를 들어, DHMB는 생합성 경로에서 효소의 활성을 감소시키거나 또는 발효 동안 효모 성장 및/또는 생산성에 대하여 다른 저해 효과를 가질 수 있다. 따라서, 본 명세서에 기재된 방법들은 발효 동안 DHMB를 감소시키는 방법을 제공한다. 본 방법들은 DHMB의 생성을 감소시키는 방법 및 발효 중인 조성물로부터 DHMB를 제거하는 방법 둘 모두를 포함한다.

[0287] DHMB 생성의 감소

[0288] 본 명세서에 기재된 일부 실시 형태에서, 재조합 숙주 세포는 아세트락테이트를 DHMB로 전환시키는 능력이 감소되거나 또는 제거된 것을 포함할 수 있다. 숙주 세포가 아세트락테이트를 DHMB로 전환시키는 능력은 예를 들어 아세트락테이트 리덕타아제 활성을 갖는 폴리펩티드를 코딩하는 폴리뉴클레오티드 또는 유전자의 변형 또는 파괴에 의해 또는 아세트락테이트 리덕타아제 활성을 갖는 폴리펩티드의 변형 또는 파괴에 의해 감소되거나 또는 제거될 수 있다. 다른 실시 형태들에서, 재조합 숙주 세포는 아세트락테이트 리덕타아제 활성을 갖는 폴리펩티드를 코딩하는 내인성 폴리뉴클레오티드 또는 유전자 내에 또는 아세트락테이트 리덕타아제를 갖는 내인성 폴리펩티드 내에 결실, 돌연변이 및/또는 치환을 포함할 수 있다. 이러한 변형, 파괴, 결실, 돌연변이 및/또는 치환은 감소되거나, 사실상 제거되거나, 또는 제거된 아세트락테이트 리덕타아제 활성으로 이어질 수 있다. 본 발명의 일부 실시 형태에서, 생합성 경로의 생성물은 아세트락테이트를 DHMB로 전환시키는 능력이 감소되거나 또는 제거된 것을 포함하지 않는 재조합 숙주 세포에서의 동일 생성물의 생성과 비교하여 더 큰 수율 또는 양으로 생성된다.

[0289] 따라서, 상기 생성물은 DHMB가 사실상 없거나, 또는 없는 부탄올을 포함하는 조성물일 수 있다. 일부 실시 형

태에서, 부탄올을 포함하는 조성물은 약 5 mM, 약 4 mM, 약 3 mM, 약 2 mM, 약 1 mM, 약 0.5 mM, 약 0.4 mM, 약 0.3 mM 이하의 DHMB, 또는 약 0.2 mM 이하의 DHMB를 함유한다.

[0290] 또한 본 생성물은 DHMB가 사실상 없거나 또는 없는 2,3-부탄다이올 (BDO)을 포함하는 조성물일 수 있다. 일부 실시 형태에서, BDO를 포함하는 조성물은 약 5 mM, 약 4 mM, 약 3 mM, 약 2 mM, 약 1 mM, 약 0.5 mM, 약 0.4 mM, 약 0.3 mM 이하의 DHMB, 또는 약 0.2 mM 이하의 DHMB를 함유한다.

[0291] 아세트락테이트를 DHMB 이외의 기질로 전환시키는 것을 포함하는 생합성 경로의 임의의 생성물은 아세트락테이트 리덕타아제 활성의 기재된 변형을 갖는 본 명세서에 개시된 재조합 숙주 세포에서 더욱 큰 유효성으로 생성될 수 있다. 이러한 생성물은 부탄올, 예를 들어, 아이소부탄올, 2-부탄올, 및 BDO와, 분지쇄 아미노산을 포함하지만, 이에 한정되지 않는다.

[0292] 일부 실시 형태에서, 숙주 세포는 아세트락테이트 리덕타아제 활성을 갖는 폴리펩티드를 코딩하는 적어도 하나의 내인성 폴리뉴클레오티드 내에 적어도 하나의 결실, 돌연변이 및/또는 치환을 포함한다. 일부 실시 형태에서, 숙주 세포는 아세트락테이트 리덕타아제 활성을 갖는 폴리펩티드를 코딩하는 둘 이상의 각각의 내인성 폴리뉴클레오티드 내에 적어도 하나의 결실, 돌연변이 및/또는 치환을 포함한다.

[0293] 일부 실시 형태에서, 아세트락테이트 리덕타아제 활성을 갖는 폴리펩티드는 아세트락테이트의 DHMB로의 전환을 촉매할 수 있다. 일부 실시 형태에서, 아세트락테이트 리덕타아제 활성을 갖는 폴리펩티드는 아세트락테이트의 2S,3S-DHMB (팩스트 DHMB) 및/또는 2S,3R-DHMB (슬로우 DHMB)로의 환원을 촉매할 수 있다.

[0294] [표 6]

유전자	서열 번호 (핵산, 아미노산)
YMR226C	676, 677
YIL074C (Chr 9)	678, 679
YIR036C (Chr 9)	680, 681
YPL061W (ALD6)(Chr 16)	682, 683
YPL088W(Chr 16)	684, 685
YCR105W (ADH7)(Chr 3)	686, 687
YDR541C(Chr 4)	688, 689
YER081 (SER3)(Chr 5)	690, 691
YPL275W (FDH2)(Chr 16)	692, 693
YBR006W (UGA5)(Chr2)	694, 695
YOL059W (Chr 15)	696, 697
YER081W (Chr 5)	869, 870
YOR375C (Chr 15)	871, 872

[0295]

[0296] 일부 실시 형태에서, 재조합 숙주 세포에 있어서 아세트락테이트의 DHMB로의 전환은 감소되거나, 사실상 제거되거나 또는 제거된다. 일부 실시 형태에서, 아세트락테이트 리덕타아제 활성을 갖는 폴리펩티드는 YMR226C, YER081W, YIL074C, YBR006W, YPL275W, YOL059W, YIR036C, YPL061W, YPL088W, YCR105W, YOR375C 및 YDR541C로 이루어진 군으로부터 선택된다. 일부 실시 형태에서, 아세트락테이트 리덕타아제 활성을 갖는 폴리펩티드는 표 6에 열거된 서열, 또는 표 6에 열거된 폴리펩티드 서열에 대하여 동일성이 약 70% 이상, 약 75% 이상, 약 80% 이상, 약 85% 이상, 약 90% 이상, 약 95% 이상, 또는 약 99% 이상인 서열을 포함하는 폴리펩티드이다. 일부 실시 형태에서, 아세트락테이트 리덕타아제 활성을 갖는 폴리펩티드는 표 6에 열거된 폴리뉴클레오티드 서열 또는 표 6에 열거된 폴리뉴클레오티드 서열에 대하여 동일성이 약 70% 이상, 약 75% 이상, 약 80% 이상, 약 85% 이상, 약 90% 이상, 약 95% 이상, 또는 약 99% 이상인 서열에 의해 코딩되는 폴리펩티드이다.

[0297] [표 7]

예시적인 YMR226C 효소 상동체

종	등록 번호	서열 번호 (핵산, 아미노산)
사카로마이세스 파라독수스(<i>Saccharomyces paradoxus</i>)	AABY01000127	698, 699
사카로마이세스 바야누스(<i>Saccharomyces bayanus</i>)	AACA01000631	700, 701 MSQGRKAAERLANKTVLITGASAGIGKAT ALEYLEASNGNMKILIAARRLEKLEELKK TIDEEFPNAKVHVGLDITQAEKIKPFIE NLPEAFKDIDILINNAGKALGSERVEIA TQDIQDVFDNTVTALINVTQAVLPIFQAK NSGDI VNLGLGGRQRRIPHRLHLLCFQVC RRCVH*QFEKGTQHQEDQSYLDRAGAG*D RVL.TGQIQR**GTS*KRL.QGHYAVDGRRR G*LNRIFHFQKAEHRGCRHPDLPHPQPSLA LPHLSRL* (서열 번호 701) 이 서열은 7배 커버리지(coverage)에 의한 "드래프트(draft)" 게놈 서열을 사용한 비교 게놈학 연구의 결과임 (문헌[Kellis et al, Nature 423:241-254 (2003)]).
사카로마이세스 카스텔리이(<i>Saccharomyces castellii</i>)	AACF01000116	702, 703
사카로마이세스 미카타(<i>Saccharomyces mikatae</i>)	AACH01000019	704, 705
야취미아 교시페이	AE016819	706, 707
칸디다 클라브라타	CR380959	708, 709
데바료마이세스 한세나이(<i>Debaryomyces hansenii</i>)	CR382139	710, 711
셰페르소마이세스 스티피티스(<i>Scheffersomyces stipitis</i>) (이전에는 피키아 스티피티스)	XM_001387479	712, 713
메이예로자이마 구일리에르몽디이(<i>Meyerozyma guilliermondii</i>) (이전에는 피키아 구일리에르몽디이(<i>Pichia guilliermondii</i>))	XM_001482184	714, 715

[0298]

종	등록 번호	서열 번호 (핵산, 아미노산)
반테르발토자이마 폴리스포라(<i>Vanderwaltozyma polyspora</i>) (이전에는 클루이베로마이세스 폴리스포루스(<i>Kluyveromyces polysporus</i>))	XM_001645671	716, 717
칸디다 두블리니엔시스(<i>Candida dubliniensis</i>)	XM_002419771	718, 719
자이코사카로마이세스 로옥시이(<i>Zygosaccharomyces rouxii</i>)	XM_002494574	720, 721
라칸세아 서모톨레란스(<i>Lachancea thermotolerans</i>) (이전에는 클루이베로마이세스 서모톨레란스(<i>Kluyveromyces thermotolerans</i>))	XM_002553230	722, 723
클루이베로마이세스 락티스	XM_451902	724, 725
사카로마이세스 클루이베리(<i>Saccharomyces kluyveri</i>)	SAKL0H04730	726, 727
야로위야 리포라이티카	XM_501554	728, 729
취조사카로마이세스 쉼베	NM_001018495	730, 731

[0299]

- [0300] 일부 실시 형태에서, 아세토락테이트 리덕타아제 활성을 갖는 폴리펩티드는 YMR226C 또는 YMR226C의 상동체이다. 따라서, 일부 실시 형태에서, 아세토락테이트 리덕타아제 활성을 갖는 폴리펩티드는 표 7에 열거된 서열, 또는 표 7에 열거된 폴리펩티드 서열에 대하여 동일성이 약 70% 이상, 약 75% 이상, 약 80% 이상, 약 85% 이상, 약 90% 이상, 약 95% 이상, 또는 약 99% 이상인 서열을 포함하는 폴리펩티드이다. 일부 실시 형태에서, 아세토락테이트 리덕타아제 활성을 갖는 폴리펩티드는 표 7에 열거된 폴리뉴클레오티드 서열 또는 표 7에 열거된 폴리뉴클레오티드 서열에 대하여 동일성이 약 70% 이상, 약 75% 이상, 약 80% 이상, 약 85% 이상, 약 90% 이상, 약 95% 이상, 또는 약 99% 이상인 서열에 의해 코딩되는 폴리펩티드이다.
- [0301] 아세토락테이트를 DHMB로 전환시킬 수 있는 아세토락테이트 리덕타아제는 예를 들어 아세토락테이트 소비의 변화, DHMB 생성의 변화, DHIV 생성의 변화, 또는 다른 하류 생성물 (예를 들어, 부탄올) 생성의 변화에 대하여 유전자 변경 효모를 스크리닝함으로써 확인될 수 있다.
- [0302] DHMB 생성에 연루된 유전자를 확인하는 한 가지 방법은 효모 닉아웃 라이브러리에서 개별 효모 주에 의해 생성되는 DHMB의 양을 측정하는 것을 포함한다. 닉아웃 라이브러리는 예를 들어 오픈 바이오시스템즈(Open Biosystems)(등록상표) (미국 매사추세츠주 월섬 소재의 서모 피셔 사이언티픽(Thermo Fisher Scientific)의 부)로부터 입수가 가능하다. 이 방법에서, DHMB 생성의 감소는 닉아웃된 유전자가 DHMB 생성을 증가시키는 기능을 한다는 것을 나타내며, DHMB 생성의 증가는 닉아웃된 유전자가 DHMB 생성을 감소시키는 기능을 한다는 것을 나타낸다.
- [0303] 닉아웃 ("KO") 라이브러리를 사용하여 DHMB 합성에서의 연루에 대하여 후보 유전자들을 확인할 수 있는 2가지 방법은 하기를 포함한다: (1) 내인성 기질 (아세토락테이트 및 NADPH 또는 NADH)로부터의, 성장 동안 배양물에 축적된 DHMB 및 DHIV가 배양물로부터의 샘플에서 분석될 수 있다. 이들 샘플은 고온 (80-100°C) 수조 내에 10-20분 동안 두어질 수 있거나, 또는 세포를 사멸시키고 세포를 투과성화할 2% 포름산과 같은 용액 내에 희석될 수 있다. 어느 하나의 처리 후, 원심분리(5 min, 1100 x g) 후 작은 분자들이 상청액에서 발견될 것이다. 대조주 (예를 들어, BY4743)의 DHMB/DHIV 비는 상이한 KO 유도체의 것과 비교될 수 있으며, 유난히 낮은 DHMB/DHIV 비를 갖는 임의의 주(들)에서 빠진 유전자(들)는 아세토락테이트 리덕타아제 (ALR)를 코딩할 수 있다. (2) 외인성 기질 (아세토락테이트 및 NADH 및/또는 NADPH)로부터의 시험관 내에서의 DHMB 및/또는 DHIV의 형성 비는 투과성화된 세포의 현탁물로부터 취해진 그리고 상기에 기재된 방법들 중 어느 하나로 불활성화된 시기 적절한 (timed) 샘플에서 측정될 수 있다. DHMB 및 DHIV 합성에 있어서의 기질은 동일하기 때문에, 이는 샘플에서의 ALR 및 KARI 활성의 상대적인 수준의 측정을 허용한다.
- [0304] DHMB 생성에 연루된 유전자를 확인하는 다른 방법은 효모 과다발현 라이브러리에서 개별 효모 주에 의해 생성되는 DHMB의 양을 측정하는 것을 포함한다. 과다발현 라이브러리는 예를 들어 오픈 바이오시스템즈(등록상표) (미국 매사추세츠주 월섬 소재의 서모 피셔 사이언티픽의 부)로부터 입수가 가능하다. 이 방법에서, DHMB 생성의 감소는 과다발현된 유전자가 DHMB 생성을 감소시키는 기능을 한다는 것을 나타내며, DHMB 생성의 증가는 과다발현된 유전자가 DHMB 생성을 증가시키는 기능을 한다는 것을 나타낸다.
- [0305] DHMB 생성에 연루된 유전자를 확인하는 다른 방법으로는 DHMB-생성 효모 주를 생화학적으로 분석하는 것이 있다. 예를 들어, DHMB-생성 세포는 파괴될 수 있다. 이 파괴는 낮은 pH 및 냉온에서 수행될 수 있다. 세포 용해물은 예를 들어 황산암모늄의 첨가에 의해 또는 당업자에게 공지된 다른 기술에 의해 분획으로 분리될 수 있으며, 생성된 분획들은 효소 활성에 대하여 분석될 수 있다. 예를 들어, 분획들은 아세토락테이트를 DHMB로 전환시키는 능력에 대하여 분석될 수 있다. 효소 활성을 갖는 분획을 본 기술 분야에 공지된 방법에 의해 처리하여 효소를 정제하고 농축시킬 수 있다 (예를 들어, 투석 및 크로마토그래피에 의한 분리). 충분한 순도 및 농축이 달성될 때, 효소를 서열 결정할 수 있고, 아세토락테이트를 DHMB로 전환시킬 수 있는 아세토락테이트 리덕타아제를 코딩하는 상응하는 유전자를 확인할 수 있다.
- [0306] 더욱이, 아세토락테이트의 DHMB로의 환원이 효모에서 일어나지만, 이는 이. 콜라이에서는 동일한 정도로 일어나지 않기 때문에, 효모에서는 발현되지만 이. 콜라이에서는 발현되지 않는 아세토락테이트 리덕타아제가 스크리닝용으로 선택될 수 있다. 선택된 효소는 효모 또는 다른 단백질 발현 시스템에서 발현될 수 있으며, 아세토락테이트를 DHMB로 전환시키는 능력에 대하여 스크리닝될 수 있다.
- [0307] 아세토락테이트의 DHMB로의 전환을 촉매할 수 있는 효소는, 아세토락테이트 수준에 대하여 분석함으로써, DHMB 수준에 대하여 분석함으로써, DHIV 수준에 대하여 분석함으로써, 또는 DHIV를 아이소부탄올을 비롯한 부탄올로 전환시키는 데 있어서의 하류 생성물들 중 임의의 것에 대하여 분석함으로써 스크리닝될 수 있다.

- [0308] DHMB는 당업자에게 공지된 임의의 기술을 이용하여 측정될 수 있다. 예를 들어, DHMB는 당업자에게 공지된 방법 및 본 명세서에 제공된 실시예에서 설명된 기술에 의해 분리되고 정량화될 수 있다. 예를 들어, DHMB는 액체 크로마토그래피-질량 분광법, 액체 크로마토그래피-핵자기 공명 (NMR), 박층 크로마토그래피, 및/또는 UV/Vis 검출 장치를 갖춘 HPLC를 사용하여 분리되고 정량화될 수 있다.
- [0309] 실시 형태들에서, 본 명세서에 개시된 선택된 아세토락테이트 리덕타아제 폴리뉴클레오티드, 유전자 및/또는 폴리펩티드는 변형되거나 또는 파괴될 수 있다. 많은 적합한 방법이 당업자에게 공지되어 있으며, 이는 알데히드 데하이드로게나아제 (상기)에 대하여 기재된 것을 포함한다.
- [0310] 아세토락테이트 리덕타아제 활성이 감소되거나 또는 제거되도록 본 명세서에 개시된 재조합 숙주 세포에서 아세토락테이트 리덕타아제를 변형시키는 것은 본 기술 분야에 공지된 방법을 이용하여 확인될 수 있다. 예를 들어, 아세토락테이트 리덕타아제-코딩 폴리뉴클레오티드 서열의 존재 또는 부재는 PCR 스크리닝을 이용하여 결정될 수 있다. 아세토락테이트 리덕타아제 활성의 감소는 또한 아세토락테이트의 DHMB로의 전환의 감소를 바탕으로 하여 결정될 수 있다. 아세토락테이트 리덕타아제 활성의 감소는 또한 DHMB 생성의 감소를 바탕으로 하여 결정될 수 있다. 아세토락테이트 리덕타아제 활성의 감소는 또한 부탄올 생성의 증가를 바탕으로 하여 결정될 수 있다.
- [0311] 따라서, 일부 실시 형태에서, 부탄올을 생성할 수 있는 효모는 약 5 mM, 약 4 mM, 약 3 mM, 약 2 mM, 약 1 mM, 약 0.9 mM, 약 0.8 mM, 약 0.7 mM, 약 0.6 mM, 약 0.5 mM, 약 0.4 mM 또는 약 0.3 mM 이하의 DHMB를 생성한다. 일부 실시 형태에서, 부탄올을 생성하는 효모는 대략적으로 약 5 mM, 약 4 mM, 약 3 mM, 약 2 mM, 약 1 mM, 약 0.9 mM, 약 0.8 mM, 약 0.7 mM, 약 0.6 mM, 약 0.5 mM, 약 0.4 mM 또는 약 0.3 mM 이하의 DHMB를 생성한다. 일부 실시 형태에서, 부탄올을 생성하는 효모는 약 0.2 mM 이하 또는 0.2 mM 이하의 DHMB를 생성한다.
- [0312] 일부 실시 형태에서, 부탄올을 생성할 수 있는 효모는 약 50시간 이상 동안 발효 조건 하에 배양될 때 대략적으로 약 10 mM 이하의 DHMB를 생성한다. 일부 실시 형태에서, 부탄올을 생성할 수 있는 효모는 약 20시간 이상, 약 25시간 이상, 약 30시간 이상, 약 35시간 이상, 약 40시간 이상, 약 45시간 이상, 또는 약 50시간 이상 동안 발효 조건 하에 배양될 때 대략적으로 약 5 mM 이하의 DHMB를 생성한다. 일부 실시 형태에서, 부탄올을 생성할 수 있는 효모는 약 5시간 이상, 약 10시간 이상, 약 15시간 이상, 약 20시간 이상, 약 25시간 이상, 약 30시간 이상, 약 35시간 이상, 약 40시간 이상, 약 45시간 이상, 또는 약 50시간 이상 동안 발효 조건 하에 배양될 때 약 3 mM 이하의 DHMB를 생성하였다. 일부 실시 형태에서, 부탄올을 생성할 수 있는 효모는 약 1시간 이상, 약 5시간 이상, 약 10시간 이상, 약 15시간 이상, 약 20시간 이상, 약 25시간 이상, 약 30시간 이상, 약 35시간 이상, 약 40시간 이상, 약 45시간 이상, 또는 약 50시간 이상 동안 발효 조건 하에 배양될 때 약 1 mM 이하의 DHMB를 생성하였다. 일부 실시 형태에서, 부탄올을 생성할 수 있는 효모는 약 1시간 이상, 약 5시간 이상, 약 10시간 이상, 약 15시간 이상, 약 20시간 이상, 약 25시간 이상, 약 30시간 이상, 약 35시간 이상, 약 40시간 이상, 약 45시간 이상, 또는 약 50시간 이상 동안 발효 조건 하에 배양될 때 약 0.5 mM 이하의 DHMB를 생성하였다.
- [0313] 일부 실시 형태에서, 아세토락테이트 리덕타아제를 코딩하는 내인성 폴리뉴클레오티드 내에 적어도 하나의 결실, 돌연변이 및/또는 치환을 포함하는 효모는 상기 양의 시간 동안 발효 조건 하에 배양될 때 아세토락테이트 리덕타아제를 코딩하는 내인성 폴리뉴클레오티드를 함유하는 효모에 의해 생성되는 DHMB의 양의 약 0.5배, 약 0.4배, 약 0.3배, 약 0.2배, 약 0.1배, 약 0.05배 이하를 생성한다.
- [0314] 일부 실시 형태에서, 부탄올을 생성할 수 있는 효모는 약 5 mM 이상, 약 6 mM 이상, 약 7 mM 이상, 약 8 mM 이상, 약 9 mM 이상, 또는 약 10 mM 이상인 양의 DHIV를 생성한다.
- [0315] 일부 실시 형태에서, 부탄올을 생성할 수 있는 효모는 적어도 대략적으로 DHMB의 생성량인 양의 DHIV를 생성한다. 일부 실시 형태에서, 부탄올을 생성할 수 있는 효모는 DHMB의 생성량의 약 2배, 약 3배, 약 5배, 약 10배, 약 15배, 약 20배, 약 25배, 약 30배, 약 35배, 약 40배, 약 45배, 또는 약 50배 이상인 양의 DHIV를 생성한다.
- [0316] 일부 실시 형태에서, 부탄올을 생성할 수 있는 효모는 적어도 대략적으로 DHMB의 생성 속도와 동일한 속도로 DHIV를 생성한다. 일부 실시 형태에서, 부탄올을 생성할 수 있는 효모는 DHMB의 생성 속도의 약 2배, 약 3배, 약 5배, 약 10배, 약 15배, 약 20배, 약 25배, 약 30배, 약 35배, 약 40배, 약 45배, 또는 약 50배 이상인 속도로 DHIV를 생성한다.
- [0317] 일부 실시 형태에서, 부탄올을 생성할 수 있는 효모는 소비되는 글루코스 1몰당 0.010몰 미만의 DHMB를 생성한

다. 일부 실시 형태에서, 효모는 소비되는 글루코스 1몰당 약 0.009몰 미만, 약 0.008몰 미만, 약 0.007몰 미만, 약 0.006몰 미만, 또는 약 0.005몰 미만의 DHMB를 생성한다. 일부 실시 형태에서, 효모는 소비되는 글루코스 1몰당 약 0.004몰 미만, 약 0.003몰 미만, 약 0.002몰 미만, 또는 약 0.001몰 미만의 DHMB를 생성한다.

[0318] 일부 실시 형태에서, 아세토락테이트 리덕타아제 활성은 화학적 수단에 의해 저해된다. 예를 들어, 아세토락테이트 리덕타아제는 다른 공지된 기질, 예를 들어 후지사와(Fujisawa) 등의 문헌에서 열거된 것을 이용하여 저해될 수 있으며, 상기 기질은 L-세린, D-세린, 2-메틸-DL-세린, D-트레오닌, L-알로-트레오닌, L-3-하이드록시아이소부티레이트, D-3-하이드록시아이소부티레이트, 3-하이드록시프로피오네이트, L-3-하이드록시부티레이트, 및 D-3-하이드록시부티레이트를 포함한다. 전체적으로 본 명세서에 참고로 포함된 문헌[Biochimica et Biophysica Acta 1645:89-94 (2003)].

[0319] DHMB 제거

[0320] 본 명세서에 기재된 다른 실시 형태들에서, DHMB의 감소는 발효물로부터 DHMB를 제거함으로써 성취될 수 있다. 따라서, DHMB 농도가 감소된 발효물이 또한 본 발명에서 개시된다. DHMB의 제거는 예를 들어 더욱 큰 순도의 생성물, 또는 원하는 순도를 성취하기 위하여 더욱 적은 프로세싱을 필요로 하는 생성물로 이어질 수 있다. 따라서, 순도가 증가된 에탄올 또는 부탄올과 같은 생합성 경로의 생성물을 포함하는 조성물이 또한 제공된다.

[0321] DHMB는 발효 공정 동안 또는 발효 공정 후 제거될 수 있으며, 본 기술 분야에 공지된 임의의 수단에 의해 제거될 수 있다. DHMB는 예를 들어 유기 상 내로의 추출 또는 반응 추출에 의해 제거될 수 있다.

[0322] 일부 실시 형태에서, 발효 브로쓰는 약 0.5 mM 미만의 DHMB를 포함한다. 일부 실시 형태에서, 발효 브로쓰는 약 5시간, 약 10시간, 약 15시간, 약 20시간, 약 25시간, 약 30시간, 약 35시간, 약 40시간, 약 45시간, 또는 약 50시간의 발효 후 약 1.0 mM 미만의 DHMB를 포함한다. 일부 실시 형태에서, 발효 브로쓰는 약 20시간, 약 25시간, 약 30시간, 약 35시간, 약 40시간, 약 45시간, 또는 약 50시간의 발효 후 약 5.0 mM 미만의 DHMB를 포함한다.

[0323] 부탄올 생합성 경로

[0324] 소정의 적합한 아이소부탄올 생합성 경로가 미국 특허 제7,851,188호 및 미국 특허 제7,993,889호에 개시되어 있으며, 상기 미국 특허 각각은 본 명세서에 참고로 포함된다. 개시된 아이소부탄올 생합성 경로의 다이어그램이 도 1에 제공되어 있다. 미국 특허 제7,851,188호에 기재된 바와 같이, 예시적인 아이소부탄올 생합성 경로에서의 단계는 하기의 전환을 포함한다:

[0325] - 예를 들어 아세토락테이트 신타아제 (ALS)에 의해 촉매되는 피루베이트의 아세토락테이트로의 전환 (상기 미국 특허의 도 1, 경로 단계 a 참조),

[0326] - 예를 들어, 아세토하이드록시산 아이소머리덕타아제 (KARI)에 의해 촉매되는 아세토락테이트의 2,3-다이하이드록시아이소발레레이트로의 전환 (상기 미국 특허의 도 1, 경로 단계 b 참조);

[0327] - 예를 들어, 다이하이드록시산 데하이드라타아제 (DHAD)로도 지칭되는 아세토하이드록시산 데하이드라타아제에 의해 촉매되는 2,3-다이하이드록시아이소발레레이트의 2-케토아이소발레레이트로의 전환 (상기 미국 특허의 도 1, 경로 단계 c 참조);

[0328] - 예를 들어, 분지쇄 2-케토산 데카르복실라아제에 의해 촉매되는 2-케토아이소발레레이트의 아이소부티르알데히드로의 전환 (상기 미국 특허의 도 1, 경로 단계 d 참조);

[0329] - 예를 들어, 분지쇄 알코올 데하이드로게나아제에 의해 촉매되는 아이소부티르알데히드의 아이소부탄올로의 전환 (상기 미국 특허의 도 1, 경로 단계 e 참조).

[0330] 다른 실시 형태에서, 아이소부탄올 생합성 경로는 하기의 기질의 생성물로의 전환을 포함한다:

[0331] - 예를 들어 아세토락테이트 신타아제에 의해 촉매될 수 있는 피루베이트의 아세토락테이트로의 전환;

[0332] - 예를 들어 케톨-산 리덕토아이소머라아제에 의해 촉매될 수 있는 아세토락테이트의 2,3-다이하이드록시아이소발레레이트로의 전환;

[0333] - 예를 들어 다이하이드록시산 데하이드라타아제에 의해 촉매될 수 있는 2,3-다이하이드록시아이소발레레이트의 α-케토아이소발레레이트로의 전환;

[0334] - 예를 들어 트랜스아미나아제 또는 발린 데하이드로게나아제에 의해 촉매될 수 있는 α-케토아이소발레레이트

의 발린으로의 전환;

- [0335] - 예를 들어 발린 데카르복실라아제에 의해 촉매될 수 있는 발린의 아이소부틸아민으로의 전환;
- [0336] - 예를 들어 오메가 트랜스아미나아제에 의해 촉매될 수 있는 아이소부틸아민의 아이소부티르알데히드로의 전환; 및
- [0337] - 예를 들어 분지쇄 알코올 데하이드로게나아제에 의해 촉매될 수 있는 아이소부티르알데히드의 아이소부탄올로의 전환.
- [0338] 다른 실시 형태에서, 아이소부탄을 생합성 경로는 하기의 기질의 생성물로의 전환을 포함한다:
- [0339] - 예를 들어 아세토락테이트 신타아제에 의해 촉매될 수 있는 피루베이트의 아세토락테이트로의 전환;
- [0340] - 예를 들어 아세토하이드록시산 리덕토아이소머라아제에 의해 촉매될 수 있는 아세토락테이트의 2,3-다이하이드록시아이소발레레이트로의 전환;
- [0341] - 예를 들어, 아세토하이드록시산 데하이드라타아제에 의해 촉매될 수 있는 2,3-다이하이드록시아이소발레레이트의 α -케토아이소발레이트로의 전환;
- [0342] - 예를 들어 분지쇄 케토산 데하이드로게나아제에 의해 촉매될 수 있는 α -케토아이소발레이트의 아이소부티릴-CoA로의 전환;
- [0343] - 예를 들어 아세틸화 알데히드 데하이드로게나아제에 의해 촉매될 수 있는 아이소부티릴-CoA의 아이소부티르알데히드로의 전환; 및
- [0344] - 예를 들어 분지쇄 알코올 데하이드로게나아제에 의해 촉매될 수 있는 아이소부티르알데히드의 아이소부탄올로의 전환.
- [0345] 다른 실시 형태에서, 아이소부탄을 생합성 경로는 도 1에서 단계 k, g 및 e로 나타난 기질의 생성물로의 전환을 포함한다.
- [0346] 상기에 기재된 기질의 생성물로의 전환과, 추가의 아이소부탄을 경로를 위한 것에 사용될 수 있는 유전자 및 폴리펩티드가 미국 특허 공개 제2007/0092957호 및 국제특허 공개 WO 2011/019894호에 기재되어 있으며, 이들 둘 모두는 본 명세서에 참고로 포함된다. 본 명세서에 참고로 포함된 미국 특허 공개 제2011/019894호, 미국 특허 공개 제2007/0092957호, 미국 특허 공개 제2010/0081154호에는 락토코커스 락티스 및 스트렙토코커스 뮈탄스 유래의 것을 포함하는 다이하이드록시산 데하이드라타아제가 기재되어 있다. 케토아이소발레이트 데카르복실라아제는 락토코커스 락티스, 마크로코커스 카세오라이트쿠스 (서열 번호 542) 및 리스테리아 그라이이 (서열 번호 543)로부터 유래된 것을 포함한다. 참고로 포함된 미국 특허 공개 제2009/0269823호 및 미국 특허 공개 제2011/0269199호에는 NADH를 보조 인자로 이용하는 것을 비롯한 알코올 데하이드로게나아제가 기재되어 있다. 알코올 데하이드로게나아제는 아크로모박터 자일로속시단스(*Achromobacter xylosoxidans*) 유래의 SadB를 포함한다. 추가의 알코올 데하이드로게나아제는 말 간 ADH 및 베이예린키아 인디카(*Beijerinckia indica*) ADH를 포함한다. 알코올 데하이드로게나아제는 NADH를 보조 인자로 이용하는 것을 포함한다. 일 실시 형태에서, 부탄을 생합성 경로는 a) NADH에 대한 K_m 이 약 300 μ M 미만, 약 100 μ M 미만, 약 50 μ M 미만, 약 20 μ M 미만 또는 약 10 μ M 미만인 케톨-산 리덕토아이소머라아제; b) NADH를 보조 인자로 이용하는 알코올 데하이드로게나아제; 또는 c) a)와 b) 둘 모두를 포함한다.
- [0347] 전체적으로 본 명세서에 참고로 포함된 국제특허 공개 WO 2011/019894호 및 미국 특허 공개 제2011/019894호, 미국 특허 공개 제2007/0092957호, 미국 특허 공개 제2010/0081154호에는 적합한 다이하이드록시산 데하이드라타아제가 기재되어 있다. DHAD 활성을 증가시키는 방법이 예를 들어 2011년 2월 17일자로 출원된 미국 특허 출원 제13/029,558호 및 미국 특허 공개 제2010/0081173호에 기재되어 있으며, 이들은 전체적으로 본 명세서에 참고로 포함된다.
- [0348] 적합한 케톨-산 리덕토아이소머라아제 (KARI) 효소가 미국 특허 공개 제2008/0261230 A1호, 미국 특허 공개 제2009/0163376호, 미국 특허 공개 제2010/0197519호, 미국 특허 공개 제2010/0143997호 및 미국 특허 공개 제2011/0244536호에 기재되어 있으며, 이들은 전체적으로 본 명세서에 참고로 포함된다. 본 명세서에 개시된 KARI의 예로는 *비브리오 콜레라*, *슈도모나스 아에루기노사* PA01, 및 슈도모나스 플루오레센스 PF5 유래의 것이 있다. 일부 실시 형태에서, KARI 효소의 비활성은 약 0.1 마이크로몰/min/mg 이상, 약 0.2 마이크로몰/min/mg 이상, 약 0.3 마이크로몰/min/mg 이상, 약 0.4 마이크로몰/min/mg 이상, 약 0.5 마이크로몰/min/mg 이상, 약

0.6 마이크로몰/min/mg 이상, 약 0.7 마이크로몰/min/mg 이상, 약 0.8 마이크로몰/min/mg 이상, 약 0.9 마이크로몰/min/mg 이상, 약 1.0 마이크로몰/min/mg 이상, 또는 약 1.1 마이크로몰/min/mg 이상이다. 기질의 생성물로의 전환인 아세토락테이트의 2,3-다이하이드록시아이소발레레이트로의 전환을 촉매하는 데 적합한 폴리펩티드는 NADH에 대한 K_M 이 약 300 μM 미만, 약 100 μM 미만, 약 50 μM 미만, 약 25 μM 미만 또는 약 10 μM 미만인 것을 포함한다.

- [0349] 일부 실시 형태에서, KARI는 NADPH를 이용한다. NADPH 소비량을 측정하는 방법은 본 기술 분야에 공지되어 있다. 예를 들어, 전체적으로 본 명세서에 참고로 포함된 미국 특허 공개 제2008/0261230호는 NADPH의 소비량을 측정하는 방법을 제공한다. 일부 실시 형태에서, NADPH 소비량 분석은 340 nm에서 아세토락테이트의 α - β -다이하이드록시-아이소발레레이트로의 효소적 전환 동안의 보조 인자, NADPH의 사라짐을 측정하는 방법이다. 활성은 NADPH에 대하여 $6220 M^{-1}cm^{-1}$ 의 몰 흡광 계수를 이용하여 계산되며, 세포 추출물 중 총 단백질 1 mg당 분당 소비되는 NADPH의 마이크로몰로서 보고된다 (문헌[Aulabaugh and Schloss, Biochemistry 29: 2824-2830, 1990] 참조).
- [0350] 일부 실시 형태에서, KARI는 NADH를 이용할 수 있다. 일부 실시 형태에서, KARI는 혐기성 조건 하에서 NADH를 이용할 수 있다. NADH를 사용하는 KARI 효소는 예를 들어 미국 특허 공개 제2009/0163376호에 기재되어 있으며, 이는 전체적으로 본 명세서에 참고로 포함된다.
- [0351] 사용될 수 있는 추가의 유전자는 본 기술 분야에 잘 알려져 있는 방법을 이용하여 또는 생물정보학을 통하여 당업자에 의해 확인될 수 있다.
- [0352] 개시된 생합성 경로를 이용한 아이소부탄올 생성을 위한 박테리아 및 효모의 유전자 엔지니어링 및 키메라 유전자의 제작이 미국 특허 공개 제2007/0092957 A1호에 부가적으로 기재되어 있으며, 이는 본 명세서에 참고로 포함된다.
- [0353] 사용될 수 있는 1-부탄올의 생성을 위한 생합성 경로는 본 명세서에 참고로 포함된 미국 특허 공개 제2008/0182308호에 기재된 것을 포함한다. 일 실시 형태에서, 1-부탄올 생합성 경로는 하기의 기질의 생성물로의 전환을 포함한다:
- [0354] - a) 예를 들어 아세틸-CoA 아세틸 트랜스퍼라아제에 의해 촉매될 수 있는 아세틸-CoA의 아세토아세틸-CoA로의 전환;
- [0355] - b) 예를 들어 3-하이드록시부티릴-CoA 데하이드로게나아제에 의해 촉매될 수 있는 아세토아세틸-CoA의 3-하이드록시부티릴-CoA로의 전환;
- [0356] - c) 예를 들어 크로토나아제에 의해 촉매될 수 있는 3-하이드록시부티릴-CoA의 크로토닐-CoA로의 전환;
- [0357] - d) 예를 들어 부티릴-CoA 데하이드로게나아제에 의해 촉매될 수 있는 크로토닐-CoA의 부티릴-CoA로의 전환;
- [0358] - e) 예를 들어 부티르알데히드 데하이드로게나아제에 의해 촉매될 수 있는 부티릴-CoA의 부티르알데히드로의 전환; 및
- [0359] - f) 예를 들어 부탄올 데하이드로게나아제에 의해 촉매될 수 있는 부티르알데히드의 1-부탄올로의 전환.
- [0360] 사용될 수 있는 2-부탄올의 생성을 위한 생합성 경로는 미국 특허 공개 제2007/0259410호 및 미국 특허 공개 제2009/0155870호에 기재된 것을 포함하며, 상기 미국 특허 공개들은 본 명세서에 참고로 포함된다. 일 실시 형태에서, 2-부탄올 생합성 경로는 하기의 기질의 생성물로의 전환을 포함한다:
- [0361] - a) 예를 들어 아세토락테이트 신타아제에 의해 촉매될 수 있는 피루베이트의 알파-아세토락테이트로의 전환;
- [0362] - b) 예를 들어 아세토락테이트 데카르복실라아제에 의해 촉매될 수 있는 알파-아세토락테이트의 아세토인으로의 전환;
- [0363] - c) 예를 들어, 아세토닌 아미나아제에 의해 촉매될 수 있는 아세토인의 3-아미노-2-부탄올로의 전환;
- [0364] - d) 예를 들어 아미노부탄올 키나아제에 의해 촉매될 수 있는 3-아미노-2-부탄올의 3-아미노-2-부탄올 포스페이트로의 전환;
- [0365] - e) 예를 들어 아미노부탄올 포스페이트 포스포릴라아제에 의해 촉매될 수 있는 3-아미노-2-부탄올 포스페이트의 2-부탄올로의 전환; 및

- [0366] - f) 예를 들어 부탄올 데하이드로게나아제에 의해 촉매될 수 있는 2-부타논의 2-부탄올로의 전환.
- [0367] 다른 실시 형태에서, 2-부탄올 생합성 경로는 하기의 기질의 생성물로의 전환을 포함한다:
- [0368] - a) 예를 들어 아세트락테이트 신타아제에 의해 촉매될 수 있는 피루베이트의 알파-아세트락테이트로의 전환;
- [0369] - b) 예를 들어 아세트락테이트 데카르복실라아제에 의해 촉매될 수 있는 알파-아세트락테이트의 아세트인으로의 전환;
- [0370] - c) 예를 들어 부탄다이올 데하이드로게나아제에 의해 촉매될 수 있는 아세트인의 2,3-부탄다이올로의 전환;
- [0371] - d) 예를 들어 다이알 데하이드라타아제에 의해 촉매될 수 있는 2,3-부탄다이올의 2-부타논으로의 전환; 및
- [0372] - e) 예를 들어 부탄올 데하이드로게나아제에 의해 촉매될 수 있는 2-부타논의 2-부탄올로의 전환.
- [0373] 본 발명의 일부 실시 형태에서, 재조합 숙주 세포는 생합성 경로를 포함한다. 생합성 경로는 감소되거나 또는 제거된 알데히드 데하이드로게나아제 활성 및 아이소부탄올 또는 1-부탄올 생합성 경로를 포함할 수 있으며, 여기서, 상기 경로는 기질의 생성물로의 전환인 피루베이트의 아세트락테이트로의 전환을 포함한다. 일부 실시 형태에서, 피루베이트를 아세트락테이트로 전환시킬 수 있는 생합성 경로를 포함하는 숙주 세포는 아세트락테이트 신타아제 활성을 갖는 폴리펩티드를 코딩하는 폴리뉴클레오티드를 포함한다. 예를 들어, 상기 생합성 경로는 부탄올 생성 경로 또는 부탄다이올 생성 경로일 수 있다. 또한 상기 생합성 경로는 분지쇄 아미노산 (예를 들어, 류신, 아이소류신, 발린) 생성 경로일 수 있다.
- [0374] 다른 실시 형태들에서, 재조합 숙주 세포는 본 명세서에 기재된 바와 같이 아이소부탄올, 1-부탄올, 또는 2-부탄올 생합성 경로를 포함할 수 있다. 일부 실시 형태에서, 부탄올 생합성 경로는 아이소부탄올 생합성 경로이다. 본 명세서에 개시된 재조합 숙주 세포에서의 아이소부탄올 또는 2-부탄올의 생성은 아세트락테이트 리덕타아제 활성의 감소, 상당한 제거 또는 제거로부터 이익을 얻을 수 있다.
- [0375] 변형
- [0376] 피루베이트 데카르복실라아제 유전자의 기능적 결실을 이용하여 생합성 생성 경로에서 이용하기 위한 피루베이트의 이용가능성을 증가시켰다. 예를 들어, 전체적으로 본 명세서에 참고로 포함된 미국 특허 공개 제 2007/0031950 A1호에는 D-락탄의 생성에 사용되는 D-락테이트 데하이드로게나아제 유전자의 발현 및 하나 이상의 피루베이트 데카르복실라아제 유전자의 파괴를 갖는 효모 주가 개시되어 있다. 전체적으로 본 명세서에 참고로 포함된 미국 특허 공개 제 2005/0059136 A1호에는 외인성 락테이트 데하이드로게나아제 유전자를 가질 수 있는, 피루베이트 데카르복실라아제 활성이 없는 글루코스 내성 C2 공급원 독립성(glucose tolerant two carbon source independent; GCSI) 효모 주가 개시되어 있다. 문헌[Nevoigt and Stahl, *Yeast* 12:1331-1337 (1996)]에는 글리세롤 수율에 대한, 사카로마이세스 세레비지에에서의 감소된 피루베이트 데카르복실라아제 및 증가된 NAD-의존성 글리세롤-3-포스페이트 데하이드로게나아제의 영향이 기재되어 있다. 전체적으로 본 명세서에 참고로 포함된 미국 특허 공개 제 2009/0305363호에는 세포질-국소화된 아세트락테이트 신타아제의 발현을 위한 효모의 엔지니어링에 의한, 피루베이트의 아세트락테이트로의 증가된 전환 및 피루베이트 데카르복실라아제 활성의 상당한 감소가 개시되어 있다.
- [0377] 본 발명의 실시 형태들에서, 본 명세서에 개시된 재조합 숙주 세포는 피루베이트 데카르복실라아제 (PDC) 활성을 갖는 폴리펩티드를 코딩하는 내인성 폴리뉴클레오티드에서의 변형 또는 PDC 활성을 갖는 내인성 폴리펩티드에서의 변형을 포함할 수 있다. 실시 형태들에서, 본 명세서에 개시된 재조합 숙주 세포는 PDC를 코딩하는 폴리뉴클레오티드, 유전자 및/또는 폴리펩티드의 변형 또는 파괴를 가질 수 있다. 실시 형태들에서, 재조합 숙주 세포는 PDC 활성을 갖는 폴리펩티드를 코딩하는 내인성 폴리뉴클레오티드 또는 유전자 내에 또는 PDC 활성을 갖는 내인성 폴리펩티드 내에 결실, 돌연변이 및/또는 치환을 포함한다. 이러한 변형, 파괴, 결실, 돌연변이 및/또는 치환은 감소되거나 제거된 PDC 활성으로 이어져서, 예를 들어 PDC 킥아웃 (PDC-KO) 표현형으로 이어질 수 있다.
- [0378] 본 발명의 실시 형태들에서, 본 명세서에 개시된 재조합 숙주 세포의 내인성 피루베이트 데카르복실라아제 활성은 피루베이트를 아세트알데히드로 전환시키며, 이는 그 후 아세테이트를 통하여 아세틸-CoA로 또는 에탄올로 전환될 수 있다. 다른 실시 형태들에서, 재조합 숙주 세포는 피루베이트 데카르복실라아제를 코딩하는 하나의 유전자를 함유하는 클루이베로마이세스 락티스, 피루베이트 데카르복실라아제를 코딩하는 하나의 유전자를 함유하는 칸디다 글라브라타, 또는 피루베이트 데카르복실라아제를 코딩하는 하나의 유전자를 함유하는 쉬조사카로

마이세스 폼베이다.

[0379] 다른 실시 형태들에서, 재조합 숙주 세포는 피루베이트 데카르복실라아제 조절 유전자, PDC2 뿐만 아니라 PDC1, PDC5, 및 PDC6에 의해서도 코딩되는 피루베이트 데카르복실라아제의 3가지의 아이소자임을 함유하는 사카로마이세스 세레비지에이다. 에스. 세레비지에에서의 비제한적 예에서, PDC1 및 PDC5 유전자, 또는 PDC1, PDC5, 및 PDC6 유전자는 파괴된다. 에스. 세레비지에에서의 다른 비제한적 예에서, 피루베이트 데카르복실라아제 활성은 PDC2 조절 유전자의 파괴에 의해 감소될 수 있다. 에스. 세레비지에에서의 다른 비제한적 예에서, PDC1, PDC2, PDC5 및/또는 PDC6에 대하여 서열 동일성이 약 70% 내지 약 75%, 약 75% 내지 약 80%, 약 80% 내지 약 85%, 약 85% 내지 약 90%, 약 90% 내지 약 95%, 약 96%, 약 97%, 약 98%, 또는 약 99%인 것과 같은 피루베이트 데카르복실라아제 단백질을 코딩하는 폴리뉴클레오티드 또는 유전자가 파괴될 수 있다.

[0380] 실시 형태들에서, PDC 활성을 갖는 폴리펩티드 또는 PDC 활성을 갖는 폴리펩티드를 코딩하는 폴리뉴클레오티드 또는 유전자는 효소 위원회 번호 EC 4.1.1.1에 상응한다. 다른 실시 형태들에서, 본 명세서에 개시된 재조합 숙주 세포의 PDC 유전자는 사용되는 발효 조건 하에서 활성을 갖지 않으며, 따라서 이러한 유전자는 변형될 필요가 없거나 또는 불활성화될 필요가 없을 것이다.

[0381] 피루베이트 데카르복실라아제 코딩 유전자의 파괴로 인하여 피루베이트 데카르복실라아제 활성이 감소된 재조합 숙주 세포의 예가 보고되었으며, 예를 들어 사카로마이세스의 경우 문헌[Flikweert *et al.*, *Yeast* (1996) 12:247-257]에, 클루이베로마이세스의 경우 문헌[Bianchi *et al.*, *Mol Microbiol.* (1996) 19(1):27-36]에, 그리고 조절 유전자의 파괴는 문헌[Hohmann, *Mol. Gen. Genet.* (1993) 241:657-666]에 보고되었다. 피루베이트 데카르복실라아제 활성을 갖지 않는 사카로마이세스 주는 ATCC로부터 등록 번호 200027 및 200028로 입수가 가능하다.

[0382] 본 명세서에 개시된 재조합 숙주 세포에서 변형 또는 불활성화용으로 표적화될 수 있는 PDC 폴리뉴클레오티드, 유전자 및/또는 폴리펩티드의 예는 하기 표 8의 것을 포함하지만, 이에 한정되지 않는다.

[0383] [표 8]

피루베이트 데카르복실라아제 표적 유전자 코딩 영역 및 단백질

설명	서열 번호 핵산	서열 번호 아미노산
사카로마이세스 세레비지에 유래의 <i>PDC1</i> 피루베이트 데카르복실라아제	648	649
사카로마이세스 세레비지에 유래의 <i>PDC5</i> 피루베이트 데카르복실라아제	650	651
사카로마이세스 세레비지에 유래의 <i>PDC6</i> 피루베이트 데카르복실라아제	652	653
칸디다 글라브라타 유래의 피루베이트 데카르복실라아제	654	655
피키아 스티피티스 유래의 <i>PDC1</i> 피루베이트 데카르복실라아제	656	657
피키아 스티피티스 유래의 <i>PDC2</i> 피루베이트 데카르복실라아제	658	659
클루이베로마이세스 락티스 유래의 피루베이트 데카르복실라아제	660	661
야로위아 리포라이트카 유래의 피루베이트 데카르복실라아제	662	663
쉬조사카로마이세스 폼베 유래의 피루베이트 데카르복실라아제	664	665
자이코사카로마이세스 로옥시이 유래의 피루베이트 데카르복실라아제	666	667

[0384]

[0385] 본 명세서에 개시된 재조합 숙주 세포에서 변형 또는 불활성화용으로 표적화될 수 있는 PDC 폴리뉴클레오티드, 유전자 및 폴리펩티드의 다른 예에는 표 8의 서열들 중 임의의 하나에 대하여 서열 동일성이 적어도 약 70% 내지 약 75%, 약 75% 내지 약 80%, 약 80% 내지 약 85%, 약 85% 내지 약 90%, 약 90% 내지 약 95%, 약 96%, 약 97%, 약 98%, 또는 약 99%인 PDC 폴리뉴클레오티드, 유전자 및/또는 폴리펩티드가 포함되지만, 이에 한정되지 않으며, 여기서, 이러한 폴리뉴클레오티드 또는 유전자는 PDC 활성을 코딩하거나 또는 이러한 폴리펩티드는 PDC 활성을 갖는다. 본 명세서에 개시된 재조합 숙주 세포에서 변형 또는 불활성화용으로 표적화될 수 있는 PDC 폴리뉴클레오티드, 유전자 및 폴리펩티드의 또 다른 예에는 표 8의 서열들 중 임의의 하나의 활성 변이체, 단편 또는 유도체가 포함되지만, 이에 한정되지 않으며, 여기서, 이러한 폴리뉴클레오티드 또는 유전자는 PDC 활성을 코딩하거나 또는 이러한 폴리펩티드는 PDC 활성을 갖는다.

[0386] 실시 형태들에서, 본 명세서에 개시되거나 또는 본 기술 분야에 공지된 PDC 서열을 코딩하는 폴리뉴클레오티드, 유전자 및/또는 폴리펩티드는 알데히드 데하이드로게나아제에 대하여 상기에 개시된 바와 같이, 변형될 수 있다. 다른 실시 형태들에서, PDC를 코딩하는 폴리뉴클레오티드, 유전자 및/또는 폴리펩티드는 다른 PDC 폴리

뉴클레오타이드, 유전자 및/또는 폴리펩티드 서열을 확인하는 데 사용될 수 있거나 또는 다른 세포에서 PDC 상동체를 확인하는 데 사용될 수 있으며, 이는 아세트락테이트 데하이드로게나아제에 대하여 상기에 개시된 바와 같다. 이러한 PDC 코딩 서열은 예를 들어 당업자에게 잘 알려진 생물정보 데이터베이스에서 및/또는 문헌에서 확인될 수 있다. 예를 들어, 생물 정보학을 이용하여 다른 세포 유형에서 PDC 코딩 서열을 확인하는 것은 본 명세서에 제공된 것과 같이, 공지된 PDC 코딩 DNA 및 폴리펩티드 서열에 의해 공개적으로 이용가능한 데이터베이스의 BLAST (상기에 개시된 바와 같음) 검색을 통하여 달성될 수 있다. 동일성은 갭 페널티=10, 갭 길이 페널티=0.1, 및 단백질 가중 매트릭스의 고넷 250 시리즈의 디폴트 파라미터를 사용한 클러스탈 W 정렬 방법을 바탕으로 한다.

[0387] PDC 활성이 감소되거나 또는 제거되도록 본 명세서에 개시된 재조합 숙주 세포에서 PDC를 변형시키는 것은 본 기술 분야에 공지된 방법을 이용하여 확인될 수 있다. 예를 들어, 특정 피루베이트 데카르복실라아제의 파괴는 그 유전자 서열에 대하여 외재적인 프라이머를 이용한 PCR 스크리닝에 의해, 또는 피루베이트 데카르복실라아제 유전자 서열에 대하여 설계된 프로브를 이용한 서던 블롯에 의해 확인될 수 있다. 대안적으로, 분석 방법, 예를 들어 가스 크로마토그래피 또는 HPLC를 이용하여 주를 아세트알데히드 및/또는 에탄올의 감소되거나 또는 제거된 생성에 대하여 스크리닝할 수 있다.

[0388] 헥소키나아제 2 유전자의 기능적 결실을 이용하여 생합성 경로에서 이용하기 위한 피루베이트의 이용가능성을 증가시키고 글루코스 억제를 감소시켰다. 예를 들어, 전체적으로 본 명세서에 참고로 포함된 국제특허 공개 WO 2000/061722 A1호에는 하나 이상의 기능적으로 결실된 헥소키나아제 2 유전자 또는 유사체를 갖는 효모를 호기적으로 성장시킴에 의한 효모 바이오매스의 생성이 개시되어 있다. 게다가, 문헌[Rossell et al., Yeast Research 8:155-164 (2008)]에서는 헥소키나아제 2 유전자가 결실된 사카로마이세스 세레비지에가, 과량의 당 및 혐기성 조건 하에서 이산화탄소의 비생성률로서 정의되는 발효능의 75% 감소를 나타낸다는 것이 밝혀졌다. 고갈(starvation) 후, 발효능은 헥소키나아제 2 유전자 결실이 없는 주의 것과 유사하였다. 문헌[Diderich et al., Applied and Environmental Microbiology 67:1587-1593 (2001)]에서는 헥소키나아제 2 유전자가 결실된 에스. 세레비지에가 더욱 낮은 피루베이트 데카르복실라아제 활성을 가짐이 밝혀졌다.

[0389] 실시 형태들에서, 본 명세서에 개시된 재조합 숙주 세포는 헥소키나아제 2 활성을 갖는 폴리펩티드를 코딩하는 내인성 폴리뉴클레오타이드에서의 변형 및/또는 헥소키나아제 2 활성을 갖는 폴리펩티드에서의 변형을 포함할 수 있다. 실시 형태들에서, 본 명세서에 개시된 재조합 숙주 세포는 헥소키나아제 2를 코딩하는 폴리뉴클레오타이드, 유전자 또는 폴리펩티드의 변형 또는 파괴를 가질 수 있다. 실시 형태들에서, 재조합 숙주 세포는 헥소키나아제 2 활성을 갖는 폴리펩티드를 코딩하는 내인성 폴리뉴클레오타이드 또는 유전자 내에 또는 헥소키나아제 2 활성을 갖는 내인성 폴리펩티드 내에 결실, 돌연변이 및/또는 치환을 포함한다. 이러한 변형, 파괴, 결실, 돌연변이 및/또는 치환은 감소되거나 또는 제거된 헥소키나아제 2 활성으로 이어져서, 예를 들어 헥소키나아제 2 넥아웃 (HXK2-KO) 표현형으로 이어질 수 있다. 실시 형태들에서, 숙주 세포는 전체적으로 본 명세서에 참고로 포함된 미국 특허 공개 제2011/0124060 A1호 또는 미국 특허 공개 제2012/0015416 A1호에 기재된 바와 같이 변형을 포함한다.

[0390] 실시 형태들에서, 헥소키나아제 2 활성을 갖는 폴리펩티드는 헥소스의 헥소스-6-포스페이트로의 전환을 촉매할 수 있고/있거나 D-글루코스의 D-글루코스 6-포스페이트로의 전환, D-프룩토스의 D-프룩토스 6-포스페이트로의 전환, 및/또는 D-만노스의 D-만노스 6-포스페이트로의 전환을 촉매할 수 있다. 다른 실시 형태들에서, 헥소키나아제 2 활성을 갖는 폴리펩티드, 유전자 또는 폴리뉴클레오타이드는 효소 위원회 번호 EC 2.7.1.1에 상응할 수 있다.

[0391] 본 발명의 실시 형태들에서, 재조합 숙주 세포는 에스. 세레비지에일 수 있으며, 헥소키나아제 2 활성을 갖는 폴리펩티드, 유전자 또는 폴리뉴클레오타이드는 HXK2일 수 있다. 다른 실시 형태들에서, 재조합 숙주 세포는 케이. 락티스일 수 있으며, 헥소키나아제 2 활성을 갖는 폴리펩티드, 유전자 또는 폴리뉴클레오타이드는 RAG5일 수 있다. 다른 실시 형태들에서, 재조합 숙주 세포는 에이치. 폴리모르파(H. polymorpha)일 수 있으며, 헥소키나아제 2 활성을 갖는 폴리펩티드, 유전자 또는 폴리뉴클레오타이드는 HPGLK1일 수 있다. 다른 실시 형태들에서, 재조합 숙주 세포는 에스. 폼베일 수 있으며, 헥소키나아제 2 활성을 갖는 폴리펩티드, 유전자 또는 폴리뉴클레오타이드는 HXK2일 수 있다.

[0392] 본 명세서에 개시된 재조합 숙주 세포에서 변형 또는 불활성화용으로 표적화될 수 있는 헥소키나아제 2 폴리뉴클레오타이드, 유전자 및 폴리펩티드의 예는 하기 표 9의 것을 포함하지만, 이에 한정되지 않는다.

[0393] [표 9]

헥소키나아제 2 표적 유전자 코딩 영역 및 단백질

에스. 세레비지에 유래의 <i>HXX2</i>	핵산 (서열 번호 668)
	아미노산 (서열 번호 669)
케이. 락티스 유래의 <i>RAG5</i>	핵산 (서열 번호 670):
	아미노산 (서열 번호 671):
에이치. 폴리모르파 유래의 <i>HPGLK1</i>	핵산 (서열 번호 672)
	아미노산 (서열 번호 673)
에스. 폼베 유래의 <i>HXX2</i>	핵산 (서열 번호 674):
	아미노산 (서열 번호 675):

[0394]

[0395] 본 명세서에 개시된 재조합 숙주 세포에서 변형 또는 불활성화용으로 표적화될 수 있는 헥소키나아제 2 폴리뉴클레오티드, 유전자 및 폴리펩티드의 다른 예에는 표 9의 서열들 중 임의의 하나에 대하여 서열 동일성이 적어도 약 70% 내지 약 75%, 약 75% 내지 약 80%, 약 80% 내지 약 85%, 약 85% 내지 약 90%, 약 90% 내지 약 95%, 약 96%, 약 97%, 약 98%, 또는 약 99%인 헥소키나아제 2 폴리뉴클레오티드, 유전자 및/또는 폴리펩티드가 포함되지만, 이에 한정되지 않으며, 여기서, 이러한 폴리뉴클레오티드 또는 유전자는 헥소키나아제 2 활성을 코딩하거나 또는 이러한 폴리펩티드는 헥소키나아제 2 활성을 갖는다. 본 명세서에 개시된 재조합 숙주 세포에서 변형 또는 불활성화용으로 표적화될 수 있는 헥소키나아제 2 폴리뉴클레오티드, 유전자 및 폴리펩티드의 또 다른 예에는 표 9의 서열들 중 임의의 하나의 활성 변이체, 단편 또는 유도체가 포함되지만, 이에 한정되지 않으며, 여기서, 이러한 폴리뉴클레오티드 또는 유전자는 헥소키나아제 2 활성을 코딩하거나 또는 이러한 폴리펩티드는 헥소키나아제 2 활성을 갖는다.

[0396] 실시 형태들에서, 본 명세서에 개시되거나 또는 본 기술 분야에 공지된 헥소키나아제 2 서열을 코딩하는 폴리뉴클레오티드, 유전자 및/또는 폴리펩티드는 알데히드 데하이드로게나아제에 대하여 상기에 개시된 바와 같이, 변형되거나 또는 파괴될 수 있다. 다른 실시 형태들에서, 헥소키나아제 2를 코딩하는 폴리뉴클레오티드, 유전자 및/또는 폴리펩티드는 다른 헥소키나아제 2 폴리뉴클레오티드, 유전자 및/또는 폴리펩티드 서열을 확인하는 데 사용될 수 있거나 또는 다른 세포에서 헥소키나아제 2 상동체를 확인하는 데 사용될 수 있으며, 이는 알데히드 데하이드로게나아제에 대하여 상기에 개시된 바와 같다. 이러한 헥소키나아제 2 코딩 서열은 예를 들어 당업자에게 잘 알려진 생물정보 데이터베이스에서 및/또는 문헌에서 확인될 수 있다. 예를 들어, 생물 정보학을 이용하여 다른 세포 유형에서 헥소키나아제 2 코딩 서열을 확인하는 것은 본 명세서에 제공된 것과 같이, 공지된 헥소키나아제 2 코딩 DNA 및 폴리펩티드 서열에 의해 공개적으로 이용가능한 데이터베이스의 BLAST (상기에 개시된 바와 같음) 검색을 통하여 달성될 수 있다. 동일성은 갭 페널티=10, 갭 길이 페널티=0.1, 및 단백질 가중 매트릭스의 고넷 250 시리즈의 디폴트 파라미터를 사용한 클러스탈 W 정렬 방법을 바탕으로 한다.

[0397] 헥소키나아제 2 활성이 감소되거나 또는 제거되도록 본 명세서에 개시된 재조합 숙주 세포에서 헥소키나아제 2를 변형시키는 것은 본 기술 분야에 공지된 방법을 이용하여 확인될 수 있다. 예를 들어, 헥소키나아제 2의 파괴는 헥소키나아제 2 유전자 서열에 대하여 외재적인 프라이머를 이용한 PCR 스크리닝에 의해, 또는 헥소키나아제 2 유전자 서열에 대하여 설계된 프로브를 이용한 서던 블롯에 의해 확인될 수 있다. 대안적으로, 글루코스 함유 배지 상에서의 바이오매스 수율 증가에 대하여 추정 헥소키나아제 2 녹아웃 주를 조사할 수 있다.

[0398] 본 명세서에 제공된 세포에 유용할 수 있는 추가의 변형의 예에는 글리세롤-3-포스페이트 데하이드로게나아제 활성을 감소시키기 위한 변형 및/또는 피루베이트 데카르복실라아제 활성을 갖는 폴리펩티드를 코딩하는 적어도 하나의 유전자에서의 파괴 또는 미국 특허 공개 제2009/0305363호(본 명세서에 참고로 포함됨)에 기재된 바와 같이 피루베이트 데카르복실라아제 유전자 발현을 제어하는 조절 요소를 코딩하는 적어도 하나의 유전자에서의 파괴, 미국 특허 공개 제2010/0120105호 (본 명세서에 참고로 포함됨)에 기재된 바와 같이 엔트너-두도로프 (Entner-Doudoroff) 경로를 통하여 탄소 유동 증가를 제공하거나 또는 등가물 균형을 감소시키는, 숙주 세포에 대한 변형이 포함된다. 다른 변형은 전체적으로 본 명세서에 참고로 포함된 국제특허 공개 제WO 2012/033832호

에 기재된 피루베이트 이용 생합성 경로 내의 단계를 촉매하는 폴리펩티드를 코딩하는 적어도 하나의 폴리뉴클레오티드의 통합을 포함한다. 효모 생성 숙주 세포가 *pdC*-인, 글루코스 억제를 감소시키는 효과를 갖는 유전자 변형이 전체적으로 본 명세서에 참고로 포함된 미국 특허 공개 제2011/0124060호에 기재되어 있다.

[0399] 본 명세서에 참고로 포함된 미국 특허 공개 제20120064561A1호에는 (a) 다이하이드록시산 데하이드라타아제 활성을 갖는 폴리펩티드를 코딩하는 적어도 하나의 이중성 폴리뉴클레오티드; 및 (b) (i) Fe-S 클러스터 생합성에 영향을 미치는 폴리펩티드를 코딩하는 내인성 유전자 내의 적어도 하나의 결실, 돌연변이 및/또는 치환, 및/또는 (ii) Fe-S 클러스터 생합성에 영향을 미치는 폴리펩티드를 코딩하는 적어도 하나의 이중성 폴리뉴클레오티드를 포함하는 재조합 숙주 세포가 개시되어 있다. 실시 형태들에서, Fe-S 클러스터 생합성에 영향을 미치는 폴리펩티드는 *AFT1*, *AFT2*, *FRA2*, *GRX3*, 또는 *CCC1*에 의해 코딩된다. 실시 형태들에서, Fe-S 클러스터 생합성에 영향을 미치는 폴리펩티드는 구성성 돌연변이 *AFT1* L99A, *AFT1* L102A, *AFT1* C291F, 또는 *AFT1* C293F이다.

[0400] 부가적으로, 숙주 세포는 포스포케톨라아제 활성을 갖는 폴리펩티드를 코딩하는 이중성 폴리뉴클레오티드 및/또는 포스포트랜스아세틸라아제 활성을 갖는 폴리펩티드를 코딩하는 이중성 폴리뉴클레오티드, 예를 들어 서열 번호 962 및 963의 서열에 의해 코딩되는 그리고 전체적으로 본 명세서에 참고로 포함된 국제특허 공개 WO 2011/159853호에 기재된 것을 포함할 수 있다.

[0401] 아이소부탄올 및 기타 생성물

[0402] 본 발명의 실시 형태들에서, 생합성 경로의 생성물의 생성 방법이 제공되며, 이는 (a) 본 명세서에 개시된 재조합 숙주 세포를 제공하는 단계; 및 (b) 생합성 경로의 생성물을 생성하는 조건 하에서 숙주 세포를 성장시키는 단계를 포함한다. 다른 실시 형태들에서, 생성물은 에탄올과 함께 동시 생성물(co-product)로서 생성된다. 또 다른 실시 형태들에서, 생합성 경로의 생성물은 아이소부탄올이다.

[0403] 본 발명의 다른 실시 형태들에서, 생합성 경로의 생성물은 감소되거나 또는 제거된 알데히드 데하이드로게나아제 및/또는 알데히드 옥시다아제 활성 및/또는 아세토라테이트 리덕타아제 활성을 포함하지 않는 재조합 숙주 세포에서의 동일 생성물의 생성과 비교하여 더 큰 수율 또는 양으로 생성된다. 실시 형태들에서, 수율은 약 2% 이상만큼, 약 5% 이상만큼 또는 약 10% 이상만큼 증가된다. 실시 형태들에서, 이러한 더 큰 수율은 이론치의 약 10% 초과, 이론치의 약 20% 초과, 이론치의 약 25% 초과, 이론치의 약 30% 초과, 이론치의 약 40% 초과, 이론치의 약 50% 초과, 이론치의 약 60% 초과, 이론치의 약 70% 초과, 이론치의 약 75% 초과, 이론치의 약 80% 초과, 이론치의 약 85% 초과, 이론치의 약 90% 초과, 이론치의 약 95% 초과, 이론치의 약 96% 초과, 이론치의 약 97% 초과, 이론치의 약 98% 초과, 이론치의 약 99% 초과, 또는 이론치의 약 100%의 수율의 생성을 포함한다. 다른 실시 형태들에서, 생성물은 에탄올과 함께 동시 생성물로서 생성된다. 또 다른 실시 형태들에서, 생합성 경로의 생성물은 아이소부탄올이다.

[0404] 경로 부산물로서, 아이소부티랄데히드의 아이소부티르산으로의 전환을 갖는 생합성 경로의 임의의 생성물이, 알데히드 데하이드로게나아제 및/또는 알데히드 옥시다아제 활성의 기재된 변형을 갖는 본 명세서에 개시된 재조합 숙주 세포에서 더욱 큰 유효성으로 생성될 수 있다. 이러한 생성물의 목록은 아이소부탄올을 포함하지만, 이에 한정되지 않는다.

[0405] 아이소부탄올 생성용 미생물 숙주

[0406] 아이소부탄올 생성용 미생물 숙주는 박테리아, 시아노박테리아, 사상균 및 효모로부터 선택될 수 있다. 부탄올 생성에 사용되는 미생물 숙주는 수율이 부탄올 독성에 의해 제한되지 않도록 아이소부탄올에 대하여 내성이어야 한다. 부탄올-내성 돌연변이체가 용매생성(solventogenic) 클로스트리디아(*Clostridia*)로부터 단리되었지만, 다른 잠재적으로 유용한 박테리아 주의 부탄올 내성에 관하여 이용가능한 정보는 거의 없다. 박테리아에서의 알코올 내성의 비교에 관한 연구의 대부분은 부탄올이 에탄올보다 더욱 독성임을 시사하며(문헌[de Cavalho et al., Microsc., Res. Tech., 64: 215-22, 2004]), 문헌[Kabelitz, et al., FEMS Microbiol. Lett., 220: 223-227, 2003], 문헌[Tomas, et al., J. Bacteriol., 186: 2006-2018, 2004]은 클로스트리듐 아세토부틸리쿰(*Clostridium acetobutylicum*)에서 발효 동안에 1-부탄올의 수율이 1-부탄올 독성에 의해 제한될 수 있음을 보고한다. 클로스트리듐 아세토부틸리쿰에 미치는 1-부탄올의 1차 효과는 막 기능의 파괴이다(문헌[Hermann et al., Appl. Environ. Microbiol., 50: 1238-1243, 1985]).

[0407] 아이소부탄올 생성용으로 선택된 미생물 숙주는 아이소부탄올에 대하여 내성이어야 하며, 탄수화물을 아이소부탄올로 전환시킬 수 있어야 한다. 적합한 미생물 숙주의 선택 기준은 하기를 포함한다: 아이소부탄올에 대한

고유 내성, 높은 탄수화물 이용률, 유전자 조작을 위한 유전학적 도구의 이용가능성 및 안정한 염색체 변화를 생성하는 능력.

[0408] 아이소부탄올에 대한 내성이 있는 적합한 숙주 주는 이 주의 고유 내성을 바탕으로 한 스크리닝에 의해 확인될 수 있다. 아이소부탄올에 대한 미생물의 고유 내성은 최소 배지에서 성장될 때 성장 속도의 50% 저해를 초래하는 아이소부탄올의 농도 (IC_{50})를 결정함으로써 측정될 수 있다. IC_{50} 값은 본 기술 분야에 공지된 방법을 사용하여 결정될 수 있다. 예를 들어, 관심대상의 미생물은 다양한 양의 아이소부탄올의 존재 하에 성장될 수 있고, 성장 속도는 600 nm에서 광학 밀도를 측정함으로써 모니터링될 수 있다. 배가 시간(doubling time)은 성장 곡선의 로그 부분으로부터 계산되고 성장 속도의 척도로서 사용될 수 있다. 성장의 50% 저해를 생성하는 아이소부탄올의 농도는 아이소부탄올 농도에 대한 성장 저해율 (%)의 그래프로부터 결정될 수 있다. 일 실시 형태에서, 숙주 주는 아이소부탄올에 대하여 IC_{50} 이 약 0.5% 초과이다.

[0409] 아이소부탄올 생성용 미생물 숙주는 또한 높은 비율의 글루코스를 이용하여야 한다. 대부분의 미생물은 탄수화물을 대사할 수 있다. 그러나, 소정의 환경 미생물은 탄수화물을 고효율로 대사할 수 없으며, 따라서 적합한 숙주가 되지 못할 것이다.

[0410] 숙주를 유전적으로 변형시키는 능력은 임의의 재조합 미생물의 제조에 필수적이다. 유전자 전달 기술의 양식은 전기 천공, 접합(conjugation), 형질 도입 또는 자연적 형질 전환에 의한 것일 수 있다. 광범위한 숙주 접합성 플라스미드 및 약물 내성 마커가 이용가능하다. 클로닝 벡터는 숙주 미생물에서 기능할 수 있는 항생제 내성 마커의 성질에 기초하여 숙주 유기체에 맞추어진다.

[0411] 또한, 미생물 숙주는 다양한 유전자를 결실시킴으로써 탄소 흐름에 대한 경쟁 경로를 불활성화시키기 위하여 조작되어야 한다. 이는, 직접적인 불활성화를 위한 트랜스포존 또는 염색체 통합 벡터의 이용가능성을 필요로 한다. 부가적으로, 생산 숙주는 고유 아이소부탄올 내성을 향상시키는 돌연변이가 수득될 수 있도록 화학적 돌연변이에 따라야 한다.

[0412] 상기에 기재된 기준을 바탕으로 하면, 아이소부탄올 제조에 적합한 미생물 숙주는 클로스트리듐, 자이모모나스, 에스케리키아, 살모넬라, 로도코커스, 슈도모나스, 바실러스, 비브리오(Vibrio), 락토바실러스, 엔테로코커스, 알칼리게네스, 클렙시엘라, 파에니바실러스, 아트로박터, 코리네박테리움, 브레비박테리움, 피키아, 칸디다, 이사첸키아, 한세놀라, 클루이베로마이세스, 및 사카로마이세스 속의 구성원을 포함하지만, 이에 한정되지 않는다. 적합한 숙주는 에스케리키아 콜라이, 알칼리게네스 에우트로푸스(*Alcaligenes eutrophus*), 바실러스 리케니포르미스(*Bacillus licheniformis*), 파에니바실러스 마세란스(*Paenibacillus macerans*), 로도코커스 에리트로폴리스(*Rhodococcus erythropolis*), 슈도모나스 퓨티다, 락토바실러스 플란타룸, 엔테로코커스 파에시움(*Enterococcus faecium*), 엔테로코커스 갈리나룸(*Enterococcus gallinarum*), 엔테로코커스 파에칼리스(*Enterococcus faecalis*), 바실러스 서브틸리스 및 사카로마이세스 세레비지애를 포함한다. 일부 실시 형태에서, 숙주 세포는 사카로마이세스 세레비지애이다. 에스. 세레비지애 효모는 본 기술 분야에 공지되어 있으며, 아메리칸 타입 컬처 컬렉션 (미국 메릴랜드주 록빌 소재), 센트라알부류 부어 쉴멜컬처즈(Centraalbureau voor Schimmelcultures; CBS) 진균류 생물다양성 센터(Fungal Biodiversity Centre), 르사프르(LeSaffre), 게르트 스트랜드(Gert Strand) AB, 펴 솔루션즈(Ferm Solutions), 노스 아메리칸 바이오프로덕츠(North American Bioproducts), 마트렉스(Martrex), 및 알레만드(Lallemand)를 포함하지만 이에 한정되지 않는 다양한 공급처로부터 입수가 가능하다. 에스. 세레비지애는 BY4741, CEN.PK 113-7D, 에탄올 레드(Ethanol Red)(등록상표) 효모, 펴 프로(Ferm Pro)TM 효모, 바이오-펙(Bio-Ferm)(등록상표) XR 효모, 게르트 스트랜드 프레스티지 배치 터보(Prestige Batch Turbo) 알코올 효모, 게르트 스트랜드 포트 디스틸러스(Pot Distillers) 효모, 게르트 스트랜드 디스틸러스 터보(Distillers Turbo) 효모, 퍼맥스(FerMax)TM 그린(Green) 효모, 퍼맥스TM 골드(Gold) 효모, 서모삭(Thermosacc)(등록상표) 효모, BG-1, PE-2, CAT-1, CBS7959, CBS7960, 및 CBS7961을 포함하지만, 이에 한정되지 않는다.

[0413] 생산 숙주의 제작

[0414] 발효성 탄소 기질의 부탄올로의 전환을 위한 효소 경로를 코딩할 필요한 유전자를 함유하는 재조합 미생물이 본 기술 분야에 잘 알려진 기술을 이용하여 제작될 수 있다. 본 발명에서, 본 발명의 아이소부탄올 생합성 경로들 중 하나의 효소, 예를 들어 아세트락테이트 신타아제, 아세트하이드록시산 아이소머로리덕타아제, 아세트하이드록시산 데하이드라타아제, 분지쇄 α -케토산 데카르복실라아제, 및 분지쇄 알코올 데하이드로게나아제를 코딩하는 유전자가 다양한 공급원으로부터 단리될 수 있으며, 이는 상기에 기재된 바와 같다.

- [0415] 게놈으로부터 원하는 유전자를 수득하는 방법은 일반적이며, 분자 생물학 분야에 잘 알려져 있다. 예를 들어, 유전자의 서열이 공지되어 있을 경우, 적합한 게놈 라이브러리는 제한 엔도뉴클레아제 절단에 의해 생성될 수 있으며, 원하는 유전자 서열에 대하여 상보성인 프로브를 이용하여 스크리닝될 수 있다. 일단 그 서열이 단리되면, 적절한 벡터를 이용한 형질전환에 적합한 양의 DNA를 수득하기 위하여 폴리머라아제 연쇄 반응 (미국 특허 제4,683,202호)과 같은 표준 프라이머-유도된 증폭 방법을 이용하여 DNA를 증폭시킬 수 있다. 이중성 숙주에서의 발현에 있어서의 코돈 최적화를 위한 도구는 용이하게 이용가능하다. 코돈 최적화를 위한 일부 도구는 숙주 미생물의 GC 함량을 바탕으로 하여 이용가능하다.
- [0416] 일단 관련 경로 유전자가 확인되고 단리되면, 본 기술 분야에 잘 알려진 수단에 의해 상기 유전자로 적합한 발현 숙주를 형질전환시킬 수 있다. 다양한 숙주 세포의 형질전환에 유용한 벡터 또는 카세트는 일반적이며, 에피센터(EPICENTRE)(등록상표) (미국 위스콘신주 매디슨 소재), 인비트로젠 코포레이션(Invitrogen Corp.)(미국 캘리포니아주 칼스배드 소재), 스트라타진(Stratagene) (미국 캘리포니아주 라 줄라 소재) 및 뉴 잉글랜드 바이오랩스, 인크.(New England Biolabs, Inc.) (미국 매사추세츠주 비벌리 소재)와 같은 회사로부터 구매가능하다. 전형적으로, 벡터 또는 카세트는 관련 유전자의 전사 및 번역을 유도하는 서열, 선발가능한 마커, 및 자가 복제 또는 염색체 통합을 허용하는 서열을 포함한다. 적절한 벡터는 전사 개시 제어부를 품고 있는 유전자의 5'의 영역 및 전사 종결을 제어하는 DNA 단편의 3'의 영역을 포함한다. 둘 모두의 제어 영역은 형질전환된 숙주 세포에 대하여 상동성인 유전자로부터 유래될 수 있지만, 그러한 제어 영역은 생산 숙주로서 선택된 특정 종에 천연적이지 않은 유전자로부터 또한 유래될 수 있음이 이해되어야 한다.
- [0417] 원하는 숙주 세포에서 상기 관련 경로 코딩 영역의 발현을 구동시키기에 유용한 개시 제어 영역 또는 프로모터가 많으며, 이는 당업자에게 친숙하다. 실시예에서 사용되는 것을 비롯하여 이들 유전자적 요소를 구동시킬 수 있는 실질적으로 임의의 프로모터가 본 발명에 적합하며, 이는 *CYC1*, *HIS3*, *GAL1*, *GAL10*, *ADH1*, *PGK*, *PHO5*, *GAPDH*, *ADC1*, *TRP1*, *URA3*, *LEU2*, *ENO*, *TPI* (사카로마이세스에서의 발현에 유용함); *AOX1* (피키아에서의 발현에 유용함); 및 *lac*, *ara*, *tet*, *trp*, *lpl*, *lpr*, *T7*, *tac*, 및 *trc* (에스케리키아 콜라이, 알칼리게네스, 및 슈도모나스에서의 발현에 유용함)와, *amy*, *apr*, *npr* 프로모터 및 다양한 파지 프로모터 - 바실러스 서브틸리스, 바실러스 리케니포르미스, 및 파에니바실러스 마세란스에서의 발현에 유용함 - 를 포함하지만, 이에 한정되지 않는다. 효모 재조합 숙주 세포에 있어서, 다수의 프로모터가 유전자의 발현 카세트의 제작에서 사용될 수 있으며, 이는 효모에서 사용하기에 적합한 하기의 구성성 프로모터들, 즉 *FBA1*, *TDH3* (*GPD*), *ADH1*, *ILV5*, 및 *GPM1*과; 효모에서 사용하기에 적합한 하기 유도성 프로모터, 즉 *GAL1*, *GAL10*, *OLE1*, 및 *CUP1*을 포함하지만, 이에 한정되지 않는다. 다른 효모 프로모터는 하이브리드(hybrid) 프로모터 *UAS(PGK1)-FBA1p* (서열 번호 406), *UAS(PGK1)-ENO2p* (서열 번호 538), *UAS(FBA1)-PDC1p* (서열 번호 539), *UAS(PGK1)-PDC1p* (서열 번호 540), 및 *UAS(PGK1)-OLE1p* (서열 번호 541)를 포함한다.
- [0418] 프로모터, 전사 종결 서열 및 코딩 영역은 효모 2 마이크로 플라스미드 내로 클로닝되고 효모 세포를 형질전환시킬 수 있다 (문헌[Ludwig, et al. Gene, 132: 33-40, 1993]; 미국 특허 공개 제20080261861A1호).
- [0419] 주어진 숙주에서의 유전자 발현량의 조절은 천연 또는 인공 프로모터의 선택을 통한 것과 같이 전사 수준을 다양하게 함으로써 성취될 수 있다. 게다가, 원하는 수준의 유전자 전사를 성취하기 위한 프로모터 라이브러리의 사용과 같은 기술은 본 기술 분야에 잘 알려져 있다. 이러한 라이브러리는 예를 들어 유전자 카세트 앞에 랜덤 cDNA 단편을 클로닝함으로써 (문헌[Goh et al. (2002) AEM 99, 17025]), 프로모터 내에 존재하는 조절 서열을 조정함으로써 (문헌[Ligr et al. (2006) Genetics 172, 2113]), 또는 공지된 프로모터 서열의 돌연변이 유발에 의해 (문헌[Alper et al. (2005) PNAS, 12678]; 문헌[Nevoigt et al. (2006) AEM 72, 5266]) 본 기술 분야에 공지된 기술을 이용하여 생성될 수 있다.
- [0420] 종결 제어 영역은 또한 숙주의 다양한 천연의 유전자로부터 유래될 수 있다. 선택적으로, 종결 부위는 불필요할 수 있거나 또는 포함될 수 있다.
- [0421] 특정 벡터는 광범위한 숙주 박테리아에서 복제될 수 있으며, 접합에 의해 전달될 수 있다. *pRK404* 및 3가지의 관련 벡터들-*pRK437*, *pRK442*, 및 *pRK442(H)*의 주석을 단 전 서열은 입수가 가능하다. 이들 유도체는 그람(Gram) 음성 박테리아에서의 유전자 조작에 유익한 도구인 것으로 입증되었다 (문헌[Scott et al., Plasmid 50:] 74-79, 2003]). 일련의 그람 음성 박테리아에서 기능할 수 있는 프로모터를 갖는 광범위한 숙주 범위의 Inc P4 플라스미드 RSF1010의 몇몇 플라스미드 유도체가 또한 이용가능하다. 플라스미드 *pAYC36* 및 *pAYC37*은 그람 음성 박테리아에서 이중성 유전자 발현을 허용하기 위한 다중 클로닝 부위와 함께 활성 프로모터를 갖는다.
- [0422] 염색체 유전자 치환 도구도 또한 광범위하게 이용할 수 있다. 예를 들어, 광범한 숙주 범위의 레플리콘

(replicon)의 감열성 변이체 pWV101을 변형시켜 일련의 그람 양성 박테리아에서 유전자 치환을 초래하기 위하여 사용될 수 있는 플라스미드 pVE6002를 제작하였다 (문헌[Maguin et al., J. Bacteriol., 174: 5633-5638, 1992]). 부가적으로, 다양한 계통 내에서 랜덤 돌연변이를 생성하기 위하여, 시험관 내 트랜스포좀(transposome)이 에피센터(EPICENTRE)(등록상표)와 같은 상업적 공급처로부터 입수가 가능하다.

[0423] 다양한 미생물 숙주에서의 부탄올 생합성 경로의 발현은 하기에 더욱 상세하게 기재되어 있다.

[0424] 이. 콜라이에서의 부탄올 생합성 경로의 발현

[0425] 이. 콜라이의 형질전환에 유용한 벡터 또는 카세트는 일반적이며, 상기에 열거된 회사로부터 구매가능하다. 예를 들어, 아이소부탄올 생합성 경로의 유전자는 다양한 공급원으로부터 단리되고, 변형된 pUC19 벡터 내로 클로닝되고, 이. 콜라이 NM522를 형질전환시킬 수 있다.

[0426] 로도코쿠스 에리트르폴리스에서의 부탄올 생합성 경로의 발현

[0427] 일련의 이. 콜라이-로도코쿠스 서틀 벡터가 알. 에리트르폴리스에서의 발현에 이용가능하며, 이는 pRhBR17 및 pDA71 (문헌[Kostichka et al., Appl. Microbiol. Biotechnol., 62: 61-68, 2003])을 포함하지만, 이에 한정되지 않는다. 부가적으로, 일련의 프로모터가 알. 에리트르폴리스에서의 이중성 유전자 발현에 이용가능하다 (문헌[Nakashima et al., Appl. Environ. Microbiol., 70: 5557-5568, 2004] 및 문헌[Tao et al., Appl. Microbiol. Biotechnol., 68: 346-354, 2005]). 알. 에리트르폴리스에서의 염색체 유전자의 표적화된 유전자 파괴는 타오(Tao) 등의 상기 문헌 및 문헌[Brans et al., Appl. Environ. Microbiol., 66: 2029-2036, 2000]에 기재된 방법을 이용하여 생성될 수 있다.

[0428] 상기에 기재된 바와 같이, 아이소부탄올의 생성에 필요한 이중성 유전자는 처음에 pDA71 또는 pRhBR71 내에 클로닝되고, 이. 콜라이를 형질전환시킬 수 있다. 그 후 상기 벡터는 전기천공에 의해 알. 에리트르폴리스를 형질전환시킬 수 있으며, 이는 코스티크카(Kostichka) 등의 상기 문헌에 기재된 바와 같다. 상기 재조합체는 글루코스를 함유하는 합성 배지에서 성장시킬 수 있으며, 아이소부탄올 생성은 본 기술 분야에 공지된 방법을 이용하여 뒤따를 수 있다.

[0429] 비. 서브틸리스에서의 부탄올 생합성 경로의 발현

[0430] 비. 서브틸리스에서의 유전자 발현 및 돌연변이 생성 방법이 또한 본 기술 분야에 잘 알려져 있다. 예를 들어, 아이소부탄올 생합성 경로의 유전자는 다양한 공급원으로부터 단리되고, 변형된 pUC19 벡터 내로 클로닝되고, 바실러스 서브틸리스 BE1010을 형질전환시킬 수 있다. 부가적으로, 아이소부탄올 생합성 경로의 5개의 유전자는 발현을 위하여 두 오페론으로 분할될 수 있다. 상기 경로의 3개의 유전자 (*bubB*, *ilvD*, 및 *kivD*)는 바실러스 서브틸리스 BE1010의 염색체 내로 통합될 수 있다 (문헌[Payne, et al., J. Bacteriol., 173, 2278-2282, 1991]). 나머지 2개의 유전자 (*ilvC* 및 *bdbB*)는 발현 벡터 내로 클로닝되고, 통합된 아이소부탄올 유전자를 지닌 바실러스 주를 형질전환시킬 수 있다.

[0431] 비. 리케니포르미스에서의 부탄올 생합성 경로의 발현

[0432] 비. 서브틸리스에서 복제되는 플라스미드 및 서틀 벡터 대부분은 원형질체 형질전환 또는 전기 천공 중 어느 하나에 의해 비. 리케니포르미스를 형질전환시키는 데 사용될 수 있다. 아이소부탄올의 생성에 필요한 유전자는 플라스미드 pBE20 또는 pBE60 유도체 내에 클로닝될 수 있다 (문헌[Nagarajan et al., Gene, 114: 121-126, 1992]). 비. 리케니포르미스의 형질전환 방법은 본 기술 분야에 공지되어 있다 (문헌[Fleming et al. Appl. Environ. Microbiol., 61: 3775-3780, 1995]). 비. 서브틸리스에서의 발현용으로 제작된 플라스미드로 비. 리케니포르미스를 형질전환시켜 아이소부탄올을 생성하는 재조합 미생물 숙주를 생성할 수 있다.

[0433] 파에니바실러스 마세란스에서의 부탄올 생합성 경로의 발현

[0434] 플라스미드는 비. 서브틸리스에서의 발현용으로 상기에 기재된 바와 같이 제작되고, 이를 사용하여 원형질체 형질전환에 의해 파에니바실러스 마세란스를 형질전환시켜 아이소부탄올을 생성하는 재조합 미생물 숙주를 생성할 수 있다.

[0435] 알칼리게네스 (랄스토니아(Ralstonia)) 에우트로푸스에서의 부탄올 생합성 경로의 발현

[0436] 알칼리게네스 에우트로푸스에서의 유전자 발현 및 돌연변이 생성 방법은 본 기술 분야에 공지되어 있다 (문헌[Taghavi et al., Appl. Environ. Microbiol., 60: 3585-3591, 1994]). 아이소부탄올 생합성 경로의 유전자는 상기에 기재된 임의의 광범위한 숙주 범위의 벡터 내에 클로닝되고, 전기 천공되어 아이소부탄올을 생성하는 재

조합체를 생성할 수 있다. 알칼리케네스에서의 폴리(하이드록시부티레이트) 경로가 상세하게 기재되었으며, 알칼리케네스 에우트로푸스 계놈을 변형시키기 위한 다양한 유전자적 기술이 공지되어 있고, 상기 도구는 아이소부탄을 생합성 경로의 엔지니어링에 적용될 수 있다.

[0437] 슈도모나스 퓨티다에서의 부탄올 생합성 경로의 발현

[0438] 슈도모나스 퓨티다에서의 유전자 발현 방법은 본 기술 분야에 공지되어 있다 (예를 들어, 본 명세서에 참고로 포함된 벤-바사트(Ben-Bassat) 등의 미국 특허 제6,586,229호 참조). 부탄올 경로 유전자는 pPCU18 내로 삽입될 수 있으며, 이 라이게이션된 DNA는 전기적격성(electrocompetent) 슈도모나스 퓨티다 DOT-T1 C5aAR1 세포 내로 전기천공되어 아이소부탄올을 생성하는 재조합체를 생성할 수 있다.

[0439] 사카로마이세스 세레비지에에서의 부탄올 생합성 경로의 발현

[0440] 사카로마이세스 세레비지에에서의 유전자 발현 방법이 본 기술 분야에 공지되어 있다 (예를 들어, 문헌[Methods in Enzymology, Volume 194, *Guide to Yeast Genetics and Molecular and Cell Biology*, Part A, 2004, Christine Guthrie and Gerald R. Fink, eds., Elsevier Academic Press, San Diego, CA]). 효모에서의 유전자의 발현은 전형적으로 프로모터, 이어서 관심대상의 유전자 및 전사 종결 서열을 필요로 한다. 본 발명에서 실시예에서 사용되는 것을 비롯한 다수의 효모 프로모터가 아이소부탄올 생합성 경로를 코딩하는 유전자의 발현 카세트의 제작에서 사용될 수 있으며, 이는 구성성 프로모터 FBA, GPD, ADH1, 및 GPM과, 유도성 프로모터 GAL1, GAL10, 및 CUP1을 포함하지만, 이에 한정되지 않는다. 적합한 전사 종결 서열은 FBAt, GPDt, GPMt, ERG10t, GAL1t, CYC1, 및 ADH1을 포함하지만, 이에 한정되지 않는다. 예를 들어, 아이소부탄올 생합성 경로의 적합한 프로모터, 전사 종결 서열 및 유전자는 이. 콜라이-효모 서틀 벡터 내로 클로닝되고, 효모 세포를 형질전환시킬 수 있는데, 이는 미국 특허 공개 제20100129886호에 기재된 바와 같다. 이들 벡터는 에스케리키아 콜라이 및 효모 주 둘 모두에서 주 증식을 가능하게 한다. 전형적으로, 벡터는 선발가능한 마커 및 원하는 숙주에서 자가 복제 또는 염색체 통합을 허용하는 서열을 함유한다. 효모에서 전형적으로 사용되는 플라스미드는 서틀 벡터 pRS423, pRS424, pRS425, 및 pRS426 (미국 메릴랜드주 록빌 소재의 아메리칸 타입 컬처 컬렉션)이며, 이는 이. 콜라이 복제 기점 (예를 들어, pMB1), 효모 2 μ 복제 기점, 및 영양 선발을 위한 마커를 함유한다. 이들 4가지의 벡터에 있어서의 선발 마커는 His3 (벡터 pRS423), Trp1 (벡터 pRS424), Leu2 (벡터 pRS425) 및 Ura3 (벡터 pRS426)이다. 관심대상의 폴리펩티드를 코딩하는 유전자를 포함하는 발현 벡터의 제작은 이. 콜라이에서의 표준 분자 클로닝 기술에 의해 또는 효모에서의 겹 수복 재조합 방법에 의해 수행될 수 있다.

[0441] 겹 수복 클로닝 접근법은 효모에서의 고효율 상동 재조합을 이용한다. 통상적으로, 효모 벡터 DNA를 절단하여 (예를 들어, 그의 다중 클로닝 부위에서) 그의 서열 내에 "갭"을 만든다. 다수의 관심대상의 삽입 DNA들이 생성되며, 이는 5' 말단 및 3' 말단 둘 모두에 ≥ 21 bp 서열을 함유하는데, 이는 순차적으로 서로 간에 그리고 벡터 DNA의 5' 및 3' 말단과 중첩된다. 예를 들어, "유전자 X"를 위한 효모 발현 벡터를 제작하기 위해서는, 발현 카세트를 위한 효모 프로모터 및 효모 종결 서열을 선택한다. 프로모터 및 종결 서열은 효모 계놈 DNA로부터 증폭되고, 유전자 X는 그의 소스 유기체로부터 PCR 증폭되거나 또는 유전자 X 서열을 포함하는 클로닝 벡터로부터 수득된다. 선형화된 벡터의 5' 말단과 프로모터 서열의 사이, 프로모터와 유전자 X의 사이, 유전자 X와 종결 서열의 사이, 및 종결 서열과 선형화된 벡터의 3' 말단 사이에는 적어도 21 bp의 중첩 서열이 있다. 그 다음에 "갭이 있는" 벡터 및 삽입 DNA를 효모 균주에 동시-형질전환시키고, 플라스미드 상의 영양 선발 마커의 상보성을 허용하는 적절한 화합물 혼합물을 함유하는 배지상에 플레이팅한다. 선발된 세포로부터 제조된 플라스미드 DNA를 사용하여 PCR 지도화에 의해 정확한 삽입물 조합의 존재를 확인할 수 있다. 그 후, 효모로부터 단리된 플라스미드 DNA (통상 농도가 낮음)로 에스케리키아 콜라이 주, 예를 들어 *TOP10*을 형질전환시킨 다음, 플라스미드 제작물을 추가로 검증하기 위하여 미니 프랩(mini prep) 및 제한효소 지도화를 할 수 있다. 마지막으로, 서열 분석에 의해 제작물을 확인할 수 있다.

[0442] 겹 수복 기술과 같이, 효모 계놈 내로의 통합은 또한 효모 내에서 상동 재조합 시스템을 이용한다. 전형적으로, 코딩 영역 플러스 제어 요소(프로모터 및 종결 서열) 및 영양요구성 마커를 함유하는 카세트는 카세트에 혼성화하는 프라이머를 사용하여, 고충실도 DNA 폴리머라아제를 이용하여 PCR-증폭되며, 삽입이 요구되는 계놈 영역의 5' 및 3' 영역에 대하여 40-70개 염기쌍의 서열 상동성을 함유한다. 그 후, PCR 생성물로 효모를 형질전환시키고, 통합된 영양요구성 마커에 대한 선발을 허용하는 적절한 화합물 혼합물을 함유하는 배지상에 플레이팅한다. 예를 들어, "유전자 X"를 염색체 위치 "Y" 내로 통합시키기 위하여, *SOE PCR* 또는 공통의 제한효소 절단 및 클로닝에 의하여, 프로모터-코딩 영역X-종결 서열 제작물을 플라스미드 DNA 제작물로부터 PCR 증폭시키고, 독립영양성 마커(예를 들어 *URA3*)에 연결시킨다. 프로모터-코딩 영역X-종결 서열-*URA3* 영역을 함

유하는 전체 카세트는 효모 염색체 상의 위치 "Y"의 5' 및 3' 영역에 대하여 40-70 bp의 상동성을 함유하는 프라이머 서열을 사용해 PCR 증폭시킨다. 상기 PCR 생성물로 효모를 형질전환되고, 우라실이 결여된 성장 배지 상에서 선발한다. 형질전환체는 콜로니 PCR에 의해서, 또는 염색체 DNA의 직접적 서열 결정에 의해 검증될 수 있다.

[0443] 락토바실러스 플란타룸에서의 부탄올 생합성 경로의 발현

[0444] 락토바실러스 속은 락토바실랄레스(*Lactobacillales*) 과에 속하며, 바실러스 서브틸리스 및 스트렙토코커스의 형질전환에 사용되는 많은 플라스미드 및 벡터가 락토바실러스에 사용될 수 있다. 적합한 벡터의 비제한적 예에는 pAM β 1 및 이의 유도체 (문헌[Renault et al., Gene 183:175-182, 1996]; 및 문헌[O'Sullivan et al., Gene, 137: 227-231, 1993]); pMBB1 및 pMBB1의 유도체, pHW800 (문헌[Wyckoff et al., Appl. Environ. Microbiol., 62: 1481-1486, 1996]); 접합 플라스미드, pMG1 (문헌[Tanimoto et al., J. Bacteriol., 184: 5800-5804, 2002]); pNZ9520 (문헌[Kleerebezem et al., Appl. Environ. Microbiol., 63: 4581-4584, 1997]); pAM401 (문헌[Fujimoto et al., Appl. Environ. Microbiol., 67: 1262-1267, 2001]); 및 pAT392 (문헌[Arthur et al., Antimicrob. Agents Chemother., 38: 1899-1903, 1994])이 포함된다. 락토바실러스 플란타룸 유래의 몇몇 플라스미드가 또한 보고되었다 (문헌[van Kranenburg R, et al. Appl. Environ. Microbiol., 71: 1223-1230, 2005]).

[0445] 다양한 엔테로코커스 중 (이. 파에시움, 이. 갈리나룸, 및 이. 파에칼리스)에서의 부탄올 생합성 경로의 발현

[0446] 엔테로코커스 속은 락토바실랄레스 과에 속하며, 락토바실러스(*Lactobacilli*), 바실러스(*Bacilli*) 및 스트렙토코커스(*Streptococci*) 종의 형질전환에서 사용되는 많은 플라스미드 및 벡터가 엔테로코커스 종에 사용될 수 있다. 적합한 벡터의 비제한적 예에는 pAM β 1 및 이의 유도체 (문헌[Renault et al., Gene, 183: 175-182, 1996]; 및 문헌[O'Sullivan et al., Gene, 137: 227-231, 1993]); pMBB1 및 pMBB1의 유도체, pHW800 (문헌[Wyckoff et al. Appl. Environ. Microbiol., 62: 1481-1486, 1996]); 접합 플라스미드, pMG1 (문헌[Tanimoto et al., J. Bacteriol., 184: 5800-5804, 2002]); pNZ9520 (문헌[Kleerebezem et al., Appl. Environ. Microbiol., 63: 4581-4584, 1997]); pAM401 (문헌[Fujimoto et al., Appl. Environ. Microbiol., 67: 1262-1267, 2001]); 및 pAT392 (문헌[Arthur et al., Antimicrob. Agents Chemother., 38: 1899-1903, 1994])가 포함된다. 락토코커스 유래의 *nisa* 유전자를 사용하는 이. 파에칼리스의 발현 벡터가 또한 사용될 수 있다 (문헌[Eichenbaum et al., Appl. Environ. Microbiol., 64: 2763-2769, 1998]). 부가적으로, 이. 파에시움 염색체에서의 유전자 대체용 벡터가 사용될 수 있다 (문헌[Nallaapareddy et al., Appl. Environ. Microbiol., 72: 334-345, 2006]).

[0447] 발효 배지

[0448] 본 발명의 발효 배지는 적합한 탄소 기질을 함유해야 한다. 적합한 탄소 기질은 단당류, 예를 들어 글루코스 및 프룩토스, 올리고당류, 예를 들어 락토오스, 말토오스, 갈락토스, 수크로스, 다당류, 예를 들어 전분 또는 셀룰로스 또는 이들의 혼합물과, 재생가능한 공급재료로부터의 정제되지 않은 혼합물, 예를 들어 치즈 유장 투과물, 옥수수 침지액, 사탕무 당밀 및 보리 누룩을 포함할 수 있지만, 이에 한정되지 않는다. 부가적으로, 탄소 기질은 또한 일원자-탄소 기질, 예를 들어 이산화탄소, 또는 메탄올일 수 있으며, 이것에 있어서 주요한 생화학적 중간체로의 대사적 전환이 입증되었다. 메틸로트로픽 미생물(methylotrophic microorganism)은 또한, 일원자 탄소 및 이원자 탄소 기질 외에 다수의 다른 탄소 함유 화합물, 예를 들어 메틸아민, 글루코사민 및 다양한 아미노산을 대사 활성화에 이용하는 것으로 공지되어 있다. 예를 들어, 메틸로트로픽 효모는 메틸아민으로부터의 탄소를 이용하여 트레할로스 또는 글리세롤을 형성하는 것으로 공지되어 있다 (문헌[Bellion et al., Microb. Growth C1 Compd., [Int. Symp.], 7th (1993), 415-32. (eds): Murrell, J. Collin; Kelly, Don P. Publisher: Intercept, Andover, UK]). 이와 유사하게, 다양한 칸디다 종은 알라닌 또는 올레산을 대사할 것이다 (문헌[Sulter et al., Arch. Microbiol., 153: 485-489, 1990]). 그러므로, 본 발명에 이용되는 탄소 공급원은 매우 다양한 탄소 함유 기질을 포함할 수 있으며, 미생물의 선택에 의해서만 한정될 것임이 고려된다.

[0449] 상기에 언급된 탄소 기질 및 이의 혼합물 전부가 본 발명에 적합한 것으로 고려되지만, 몇몇 실시 형태에서, 탄소 기질은 글루코스, 프룩토스 및 수크로스, 또는 C5 당을 사용하도록 변형된 효모 세포를 위한 자일로스 및/또는 아라비노스와 같은 C5 당과의 그의 혼합물이다. 수크로스는 재생가능한 당 공급원, 예를 들어 사탕수수, 사탕무, 카사바(cassava), 단 수수(sweet sorghum) 및 이들의 혼합물로부터 유래될 수 있다. 글루코스 및 텍스트로스는 옥수수, 밀, 호밀, 보리, 귀리 및 이들의 혼합물과 같은 곡물을 비롯한 전분 기재의 공급원료의 당화를 통하여 재생가능한 곡물 공급원으로부터 유래될 수 있다. 게다가, 발효성 당은 예를 들어, 전체적으로 본 명세

서에 참고로 포함된 미국 특허 공개 제2007/0031918A1호에 기재된 바와 같이, 전처리 및 당화 과정을 통하여 재생 가능한 셀룰로오스계 또는 리그노셀룰로오스계 바이오매스로부터 유래될 수 있다. 바이오매스는 임의의 셀룰로오스계 또는 리그노셀룰로오스계 물질을 말하며, 셀룰로오스를 포함하며, 선택적으로 헤미셀룰로스, 리그닌, 전분, 올리고당류 및/또는 단당류를 추가로 포함하는 물질을 포함한다. 바이오매스는 또한 단백질 및/또는 지질과 같은 추가의 성분을 포함할 수 있다. 바이오매스는 단일 공급원으로부터 유래할 수 있거나, 또는 바이오매스는 1 가지 초과 공급원으로부터 유래된 혼합물을 포함할 수 있으며; 예를 들어, 바이오매스는 옥수수 속대 및 옥수수 대의 혼합물, 또는 풀과 잎의 혼합물을 포함할 수 있다. 바이오매스는 바이오에너지 작물, 농업적 잔사, 도시 고형 폐기물, 산업 고형 폐기물, 종이 제조로부터의 슬러지, 정원 폐기물, 나무 및 삼림 폐기물을 포함하지만, 이에 한정되지 않는다. 바이오매스의 예는, 옥수수 낱알, 옥수수 속대, 옥수수 잔사, 예를 들어 옥수수 껍질, 옥수수 대, 풀, 밀, 밀짚, 보리, 보릿짚, 건초, 벼짚, 스위치그래스, 폐지, 사탕수수 버캐스, 수수, 대두, 낱알의 제분으로부터 수득된 성분, 나무, 가지, 뿌리, 잎, 나뭇조각, 톱밥, 관목 및 떨기나무, 야채, 과일, 꽃, 동물 퇴비 및 이들의 혼합물을 포함하지만, 이에 한정되지 않는다.

[0450] 적절한 탄소 공급원에 더하여, 발효 배지는 적합한 무기물, 염, 보조 인자, 완충제, 및 본 명세서에 기재된 부탄올 생성에 필요한 효소 경로의 촉진 및 배양물의 성장에 적합한, 당업자에게 공지된 다른 성분들을 함유해야 한다.

[0451] 배양 조건

[0452] 전형적으로, 세포는 적절한 배지에서 약 20°C 내지 약 40°C 범위의 온도에서 성장시킨다. 본 발명에서 적합한 성장 배지는 일반적인, 상업적으로 제조된 배지, 예를 들어 루리아 베르타니(Luria Bertani; LB) 브로쓰, 사부로트 텍스트로스(Sabouraud Dextrose; SD) 브로쓰 또는 이스트 미디움(Yeast Medium; YM) 브로쓰 또는 효모 질소 염기, 황산암모늄 및 텍스트로스 (탄소/에너지 공급원으로서)를 포함하는 브로쓰 또는 대부분의 사카로마이세스 세레비지에 주 성장을 위한 최적 비율의 펩톤, 효모 추출물 및 텍스트로스의 블렌드인 YPD 배지이다. 다른 규정 또는 합성 성장 배지가 또한 사용될 수 있으며, 특정 미생물의 성장에 적절한 배지는 미생물학 또는 발효 과학 분야의 숙련자에 의해 공지될 것이다. 예를 들어 사이클릭 아데노신 2',3'-모노포스페이트(cAMP)와 같이 직접적으로 또는 간접적으로 이화대사물 억제제를 조정하는 것으로 공지된 제제의 사용이 발효 배지 내에 또한 혼입될 수 있다.

[0453] 발효에 적합한 pH 범위는 pH 5.0 내지 pH 9.0이며, 여기서 pH 6.0 내지 pH 8.0이 초기 조건에 바람직하다. 효모의 발효에 적합한 pH 범위는 전형적으로 약 pH 3.0 내지 약 pH 9.0이다. 일 실시 형태에서, 약 pH 5.0 내지 약 pH 8.0이 초기 조건을 위해 사용된다. 다른 미생물의 발효에 적합한 pH 범위는 약 pH 3.0 내지 약 pH 7.5이다. 일 실시 형태에서, 약 pH 4.5 내지 약 pH 6.5가 초기 조건에 사용된다.

[0454] 호기성 또는 혐기성 조건 하에서 발효가 수행될 수 있다. 일 실시 형태에서, 혐기성 또는 미세호기성(microaerobic) 조건이 발효에 사용된다.

[0455] 산업적 배치(Batch) 및 연속 발효

[0456] 본 발명의 방법은 배치식 발효법을 이용할 수 있다. 고전적인 배치 발효는 배지의 조성이 발효 시작시에 정해지며, 발효 중에 인공적으로 변경되지 않는 폐쇄 시스템이다. 따라서, 발효의 시작시에 배지에 원하는 미생물(들)이 접종되며, 상기 시스템에 어떠한 것도 첨가하지 않고서 발효가 일어나게 한다. 그러나, 전형적으로 "배치" 발효는 탄소 공급원의 첨가와 관련하여 배치식이며, pH 및 산소 농도와 같은 인자를 조절하려는 시도가 흔히 이루어진다. 배치 시스템에서, 이 시스템의 대사산물 및 바이오매스 조성물은 발효가 중단되는 시점까지 끊임없이 변화한다. 배치 배양 내에서 세포는 정적 유도기(static lag phase)를 통하여 고 성장 대수기(log phase)로, 그리고 마지막으로 성장 속도가 감소되거나 정지되는 정체기(stationary phase)로 완화된다. 미처리될 경우, 정체기의 세포는 결국 사멸할 것이다. 대수기의 세포는 일반적으로 최종 생산물 또는 중간체의 대량 생산에 책임이 있다.

[0457] 표준 배치식 시스템의 한 변형은 유가식(Fed-Batch) 시스템이다. 또한, 유가식 발효 공정이 본 발명에서 적합하며, 이는 발효가 진행됨에 따라 기질이 증분식으로 첨가된다는 것을 제외하고는 전형적인 배치 시스템을 포함한다. 유가식 시스템은 이화 대사산물 억제가 세포의 대사 작용을 저해하는 경향이 있을 때 그리고 배지 중에 한정된 양의 기질이 있는 것이 바람직한 경우 유용하다. 유가식 시스템 내의 실제 기질 농도의 측정은 어려우므로, pH, 용존 산소 및 CO₂와 같은 폐 가스(waste gas)의 분압과 같이 측정가능한 인자의 변화를 기초로 추산된다. 배치식 및 유가식 발효는 통상적이며, 본 기술 분야에 잘 알려져 있고, 예는 본 명세서에 참고로

포함된, 문헌[Thomas D. Brock, *Biotechnology: A Textbook of Industrial Microbiology*, Second Edition (1989) Sinauer Associates, Inc., Sunderland, MA], 또는 문헌[Deshpande, Mukund, *Appl. Biochem. Biotechnol.*, 36: 227, 1992]에서 찾아볼 수 있다.

- [0458] 본 발명이 배치 방식으로 수행되더라도, 방법은 연속 발효 방법에 적응가능할 것임이 고려된다. 연속식 발효는 규정 발효 배지가 연속적으로 생물 반응기에 첨가되고, 동일한 양의 조절 배지가 프로세싱을 위하여 동시에 제거되는 개방 시스템이다. 연속 발효는 일반적으로, 세포가 주로 대수기 성장에 있는 경우 일정한 고밀도로 배양물을 유지시킨다.
- [0459] 연속식 발효는 세포 성장 또는 최종 산물 농도에 영향을 주는 하나의 인자 또는 임의의 개수의 인자의 조절을 허용한다. 예를 들어, 한 가지 방법은 제한 영양소, 예를 들어 탄소 공급원 또는 질소의 수준을 고정된 비율로 유지하고 모든 다른 파라미터는 조정되게 할 것이다. 다른 시스템에서, 성장에 영향을 미치는 많은 인자는 계속하여 변경될 수 있는 반면, 배치 탁도에 의해 측정되는 세포 농도는 일정하게 유지된다. 연속식 시스템은 정상 상태 성장 조건을 유지하도록 노력하며, 따라서 배출되는 배지로 인한 세포 손실은 발효에서 세포 성장 속도에 대하여 균형이 맞추어져야 한다. 연속식 발효 방법에 있어서 영양소 및 성장 인자를 조정하는 방법 뿐만 아니라, 생성물 형성 속도를 최대화하는 기술도 산업 미생물학 분야에서 주지되어 있으며, 다양한 방법이 브록(Brock)의 상기 문헌에 상술되어 있다.
- [0460] 본 발명은 회분식, 유가식 또는 연속식 방법 중 어느 하나를 사용하여 실시할 수 있으며 임의의 알려진 발효 양식이 적합할 것임이 고려된다. 부가적으로, 세포는 전세포 촉매로서 기질 상에 고정되고, 아이소부탄올 생성을 위한 발효 조건이 가해질 수 있음이 고려된다.
- [0461] 발효 배지로부터 부탄올을 분리하는 방법
- [0462] 생물학적으로 생성된 부탄올은 ABE 발효에 대하여 본 기술 분야에 공지되어 있는 방법을 사용하여 발효 배지로부터 분리될 수 있다 (예를 들어, 문헌[Durre, *Appl. Microbiol. Biotechnol.* 49:639-648 (1998)], 문헌[Groot *et al.*, *Process. Biochem.* 27:61-75 (1992)] 및 상기 문헌의 참고 문헌 참조). 예를 들어, 원심분리, 여과, 경사분리(decantation) 등에 의해 발효 배지로부터 고체를 제거할 수 있다. 그 후, 부탄올은 증류, 공비 증류, 액체-액체 추출, 흡착, 가스 스트리핑(gas stripping), 막 증발, 또는 투과증발(pervaporation)과 같은 방법을 사용하여 발효 배지로부터 분리될 수 있다.
- [0463] 부탄올이 물과 저비점의 공비성 혼합물을 형성하기 때문에, 증류를 사용하여 이의 공비 조성까지 혼합물을 분리할 수 있다. 증류를 다른 분리 방법과 조합하여 사용하여 공비 혼합물에 맞추어 분리를 얻을 수 있다. 부탄올의 분리 및 정제를 위하여 증류와 조합하여 사용할 수 있는 방법은 경사분리, 액체-액체 추출, 흡착 및 막-기재의 기술을 포함하지만, 이에 한정되지 않는다. 부가적으로, 부탄올을 엔트레이너(entrainer)를 사용하는 공비 증류를 사용하여 분리할 수 있다 (예를 들어, 문헌[Doherty and Malone, *Conceptual Design of Distillation Systems*, McGraw Hill, New York, 2001] 참조).
- [0464] 증류를 경사분리와 조합하여 사용하여, 부탄올을 분리 및 정제할 수 있도록 부탄올-물 혼합물은 이질적인 공비 혼합물을 형성한다. 이 방법에서는 부탄올을 함유하는 발효 브로쓰를 공비 조성 부근까지 증류한다. 그 다음에, 공비혼합물을 응축시키고 경사분리에 의해 발효 배지로부터 부탄올을 분리한다. 경사분리된 수성 상은 환류물로서 제1 증류 컬럼으로 되돌아갈 수 있다. 부탄올이 풍부한 경사분리된 유기 상은 제2 증류 컬럼에서 증류에 의하여 추가로 정제될 수 있다.
- [0465] 또한 증류와 조합한 액체-액체 추출을 사용하여 발효 배지로부터 부탄올을 분리할 수 있다. 이 방법에서는 적합한 용매로 액체-액체 추출하여 발효 브로쓰로부터 부탄올을 추출한다. 그 후에 부탄올-함유 유기 상을 증류하여 용매로부터 부탄올을 분리한다.
- [0466] 또한, 흡착과 조합된 증류를 사용하여, 발효 배지로부터 부탄올을 분리할 수 있다. 이 방법에서는, 부탄올을 함유하는 발효 브로쓰를 공비 조성 부근까지 증류한 후에 분자체와 같은 흡착제를 사용하여 잔류하는 물을 제거한다 (문헌[Aden *et al.*, *Lignocellulosic Biomass to Ethanol Process Design and Economics Utilizing Co-Current Dilute Acid Prehydrolysis and Enzymatic Hydrolysis for Corn Stover*, Report NREL/TP-510-32438, National Renewable Energy Laboratory, June 2002]).
- [0467] 부가적으로, 투과증발과 조합한 증류를 사용하여 발효 배지로부터 부탄올을 분리하고 정제할 수 있다. 이 방법에서, 부탄올을 함유하는 발효 브로쓰는 공비 조성 부근까지 증류되고, 그 후 잔류하는 물은 친수성 막을 통한

투과증발에 의해 제거된다 (문헌[Guo *et al.*, *J. Membr. Sci.* 245, 199-210 (2004)]).

[0468] 원위치 생성물 제거(In situ product removal; ISPR) (추출 발효로도 칭해짐)는, 부탄올 (또는 다른 발효 알코올)을 이것이 생성될 때 발효 용기로부터 제거하고 이럼으로써 미생물이 부탄올을 고수율로 생성하게 하는 데 사용될 수 있다. 본 기술 분야에서 설명된 발효 알코올을 제거하기 위한 하나의 ISPR 방법으로는 액체-액체 추출이 있다. 일반적으로, 부탄올 발효와 관련하여, 예를 들어, 미생물을 포함하는 발효 배지는 부탄올 농도가 독성 수준에 도달하기 전의 소정의 시점에서 유기 추출제와 접촉된다. 유기 추출제 및 발효 배지는 2상 혼합물을 형성한다. 부탄올은 유기 추출제 상 내로 분배되어, 미생물을 함유하는 수성 상 중의 농도가 감소되고, 이럼으로써 저해성 부탄올에 대한 미생물의 노출이 제한된다.

[0469] 액체-액체 추출은 예를 들어 미국 특허 공개 제2009/0305370호에 기재된 방법에 따라 수행될 수 있으며, 상기 미국 특허 공개의 내용은 전체적으로 본 명세서에 포함된다. 미국 특허 공개 제2009/0305370호에는 부탄올을 생성하고 액체-액체 추출을 이용하여 발효 브로쓰로부터 회수하는 방법이 기재되어 있으며, 상기 방법은 발효 브로쓰를 수불혼화성 추출제와 접촉시켜 수성 상 및 유기 상을 포함하는 2상 혼합물을 형성하는 단계를 포함한다. 전형적으로, 추출제는 포화, 단일불포화, 다중불포화 (및 이의 조합의) C_{12} 내지 C_{22} 지방 알코올, C_{12} 내지 C_{22} 지방산, C_{12} 내지 C_{22} 지방산의 에스테르, C_{12} 내지 C_{22} 지방 알데히드 및 이들의 혼합물로 이루어진 군으로부터 선택되는 유기 추출제일 수 있다. ISPR을 위한 추출제(들)는 비-알코올 추출제일 수 있다. ISPR 추출제는 외인성 유기 추출제, 예를 들어 올레일 알코올, 베헤닐 알코올, 세틸 알코올, 라우릴 알코올, 미리스틸 알코올, 스테아릴 알코올, 1-운데카놀, 올레산, 라우르산, 미리스트산, 스테아르산, 메틸 미리스테이트, 메틸 올레에이트, 운데카날, 라우릭 알데히드, 20-메틸운데카날 및 이들의 혼합물일 수 있다.

[0470] 일부 실시 형태에서, 상기 에스테르는 발효 배지 중의 알코올을 카르복실산 (예를 들어, 지방산), 및 카르복실산에 의한 알코올의 에스테르화가 가능한 촉매와 접촉시킴으로써 형성될 수 있으며, 이는 전체적으로 본 명세서에 참고로 포함된 국제특허 공개 WO/2011/159998호에 기재된 바와 같다. 이러한 실시 형태들에서, 카르복실산은 ISPR 추출제로서의 역할을 할 수 있는데, 상기 추출제 내로 알코올 에스테르가 분배된다. 카르복실산은 발효 용기로 공급되고/되거나 바이오매스로부터 유래되어, 발효 용기로 공급되는 발효성 탄소를 공급할 수 있다. 공급원료에 존재하는 지질은 촉매에 의해 카르복실산으로 가수분해될 수 있으며, 동일 촉매 (예를 들어, 효소)가 알코올에 의해 카르복실산을 에스테르화할 수 있다. 촉매는 발효 이전에 공급원료에 공급될 수 있거나, 또는 공급원료의 공급 전 또는 공급원료의 공급과 동시에 발효 용기에 공급될 수 있다. 촉매가 발효 용기에 공급될 때, 지질의 카르복실산으로의 가수분해 및 발효 용기에 존재하는 부탄올에 의한 상기 카르복실산의 사실상 동시적인 에스테르화에 의해 알코올 에스테르가 수득될 수 있다. 공급원료로부터 유래되지 않은 천연 오일 및/또는 카르복실산이 또한 발효 용기에 공급될 수 있으며, 이때 천연 오일은 카르복실산으로 가수분해된다. 알코올에 의해 에스테르화되지 않은 임의의 카르복실산이 ISPR 추출제의 일부로서의 역할을 할 수 있다. 알코올 에스테르를 함유하는 추출제는 발효 배지로부터 분리될 수 있으며, 알코올은 추출제로부터 회수될 수 있다. 추출제는 발효 용기로 재순환될 수 있다. 따라서, 예를 들어, 부탄올 생성의 경우에 부탄올의 에스테르로의 전환은 발효 배지 중 자유 부탄올 농도를 감소시킬 수 있으며, 이는 부탄올 농도 증가의 독성 효과로부터 미생물을 차폐한다. 게다가, 분획화되지 않은 곡물을, 그 안의 지질의 분리 없이 공급 원료로 사용할 수 있으며, 그 이유는 지질이 촉매에 의해 카르복실산으로 가수분해되고 이럼으로써 ISPR 추출제 중 지질의 증가 속도가 감소될 수 있기 때문이다.

[0471] 원위치 생성물 제거가 배치식으로 또는 연속식으로 실시될 수 있다. 연속식의 원위치 생성물 제거에서, 생성물은 반응기로부터 연속적으로 제거된다. 배치식의 원위치 생성물 제거에서, 소정 부피의 유기 추출제를 발효 용기에 첨가하고, 상기 추출제를 이 과정 동안 제거하지 않는다. 원위치 생성물 제거에 있어서, 유기 추출제는 발효 시작시에 발효 배지와 접촉하여 2상 발효 배지를 형성할 수 있다. 대안적으로, 유기 추출제는 미생물이 배양물의 광학 밀도를 측정함으로써 결정될 수 있는 원하는 양의 성장을 이룬 후에 발효 배지와 접촉할 수 있다. 또한, 유기 추출제는 발효 배지 중 알코올 생성물 수준이 소정의 수준에 도달하는 시점에 발효 배지와 접촉할 수 있다. 본 발명의 일부 실시 형태에 따른 부탄올 생성의 경우에, 카르복실산에 의해 부탄올을 에스테르화하여 부탄올 에스테르를 생성하고 그 결과 발효 용기 중 부탄올의 농도가 감소되도록, 카르복실산 추출제를 부탄올 농도가 독성 수준에 도달하기 전의 시점에 발효 배지와 접촉시킬 수 있다. 그 후, 에스테르-함유 유기 상은 부탄올 에스테르의 원하는 유효 역가가 성취된 후 발효 용기로부터 제거될 수 있다 (그리고 수성 상을 구성하는 발효 브로쓰로부터 분리될 수 있다). 일부 실시 형태에서, 에스테르-함유 유기 상은 발효 용기 내의 이용가능한 발효성 당의 발효가 사실상 완료된 후 수성 상으로부터 분리된다.

- [0472] 실시예
- [0473] 본 발명을 하기 실시예로 추가로 정의한다. 이들 실시예는 본 발명의 바람직한 실시 형태를 나타내면서도 단지 예로서 주어지는 것임을 이해해야 한다. 당업자라면, 상기 논의 및 이러한 실시예로부터 본 발명의 본질적인 특징을 확인할 수 있으며, 본 발명의 사상 및 범주를 벗어나지 않고서도 본 발명을 다양하게 변경하고 수정하여 본 발명을 다양한 용도 및 조건에 적합하게 할 수 있다.
- [0474] 일반 방법:
- [0475] 본 실시예에 사용되는 표준 제조법 DNA 및 분자 클로닝 기술은 본 기술 분야에 잘 알려져 있으며, 문헌 [Sambrook, J., Fritsch, E. F. and Maniatis, T., Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y., 1989], 문헌[T. J. Silhavy, M. L. Bannan, and L. W. Enquist, Experiments with Gene Fusions, Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y., 1984], 및 문헌[Ausubel, F. M. et al., Current Protocols in Molecular Biology, Greene Publishing Assoc. and Wiley-Interscience, N.Y., 1987]에 기재되어 있다.
- [0476] 또한, 박테리아 배양물의 유지 및 성장에 적합한 재료 및 방법이 본 기술 분야에 잘 알려져 있다. 하기 실시예에서의 이용에 적합한 기술은 문헌[Manual of Methods for General Bacteriology, Philipp Gerhardt, R. G. E. Murray, Ralph N. Costilow, Eugene W. Nester, Willis A. Wood, Noel R. Krieg and G. Briggs Phillips, eds., American Society for Microbiology, Washington, DC., 1994], 또는 문헌[Thomas D. Brock, Biotechnology: A Textbook of Industrial Microbiology, Second Edition, Sinauer Associates, Inc., Sunderland, MA, 1989]에서 찾을 수 있다. 다르게 특정하지 않는 한, 박테리아 세포의 성장 및 유지를 위해 사용된 모든 시약, 제한효소 및 재료는 알드리치 케미칼즈(Aldrich Chemicals) (미국 위스콘신주 밀워키 소재), 비디 다이아그노스틱 시스템즈(BD Diagnostic Systems) (미국 메릴랜드주 스파크 소재), 라이프 테크놀로지스(Life Technologies) (미국 메릴랜드주 락빌 소재), 또는 시그마 케미칼 컴퍼니(Sigma Chemical Company) (미국 미주리주 세인트 루이스 소재)로부터 획득하였다.
- [0477] 사용되는 약어의 의미는 하기와 같다: "Å"는 앙스트롬을 의미하며, "min"은 분을 의미하며, "h"는 시간을 의미하며, "μl"는 마이크로리터를 의미하며, "ng/μl"는 마이크로리터당 나노그램을 의미하며, "pmol/μl"는 마이크로리터당 피코몰을 의미하며, "ml"는 밀리리터를 의미하며, "L"는 리터를 의미하며, "g/L"는 리터당 그램을 의미하며, "ng"은 나노그램을 의미하며, "sec"는 초를 의미하며, "ml/min"은 분당 밀리리터를 의미하며, "w/v"는 부피당 중량을 의미하며, "v/v"는 부피당 부피를 의미하며, "nm"는 나노미터를 의미하며, "mm"은 밀리미터를 의미하며, "cm"은 센티미터를 의미하며, "mM"은 밀리몰을 의미하며, "M"은 몰을 의미하며, "g"은 그램을 의미하며, "μg"은 마이크로그램을 의미하며, "mg"은 밀리그램을 의미하며, "g"는 중력 상수를 의미하며, "rpm"은 분당 회전수를 의미하며, "HPLC"는 고성능 액체 크로마토그래피를 의미하며, "MS"는 질량 분광법을 의미하며, "HPLC/MS"는 고성능 액체 크로마토그래피/질량 분광법을 의미하며, "EDTA"는 에틸렌다이아민-테트라아세트산을 의미하며, "dNTP"는 데옥시뉴클레오타이드 트라이포스페이트를 의미하며, "°C"는 섭씨 온도를 의미하며, "V"는 전압을 의미한다.
- [0478] 유전자 라이브러리의 고 처리량 스크리닝 분석
- [0479] 돌연변이 KARI 효소의 유전자 라이브러리의 고 처리량 스크리닝을 (실시예 16 및 실시예 21을 제외하고서) 본 명세서에 설명된 바와 같이 수행하였다: 554.4 g/L의 글리세롤, 68 mM의 (NH₄)₂SO₄, 4 mM의 MgSO₄, 17 mM의 시트르산나트륨, 132 mM의 KH₂PO₄, 36 mM의 K₂HPO₄를 함유하는 10x 냉동 배지를 분자적으로 순수한 물로 제조하고, 여과 살균하였다. 냉동 배지는 10x 냉동 배지를 LB 배지로 희석시킴으로써 준비하였다. 분취량 (200 μl)의 냉동 배지를 96웰 아카이브(archive) 플레이트 (카탈로그 번호: 3370, 미국 뉴욕주 코닝 소재의 코닝 인크.(Corning Inc.))의 각각의 웰에 사용하였다.
- [0480] LB 한천 배지로부터의 클론을 선별하고, 냉동 배지를 함유하는 96웰 아카이브 플레이트 내에 접종하고, 진탕 없이 37°C에서 하룻밤 성장시켰다. 그 후, 상기 아카이브 플레이트를 -80°C에서 보관하였다. pBAD-HisB (인비트로젠(Invitrogen))로 형질전환시킨 이. 콜라이 주 Bw25113을 항상 음성 대조군으로 사용하였다. 실시예 3, 실시예 4 및 실시예 5에서 라이브러리에 대한 양성 대조군은 각각 야생형 K9-KARI, AB1D3, AB1D3이다.
- [0481] 아카이브 플레이트로부터의 클론을 깊은 96웰 플레이트 내에 접종하였다. 각각의 웰은 해동된 아카이브 플레이트로부터의 세포 3.0 μl, 유도제로서 0.02%(w/v) 아라비노스 및 암피실린 100 μg/ml을 함유하는 LB 배지 200

μl를 함유하였다. 세포를 진탕시키면서 (900 rpm) 80% 습도를 이용하여 37℃에서 하룻밤 성장시키고, 원심분리 (4000 rpm, 25℃에서 5 min) (에펜도르프(Eppendorf) 원심분리기, 미국 뉴욕주 웨스트베리 소재의 브린크만 인스트루먼트, 인크.(Brinkmann Instruments, Inc.))에 의해 수확하고, 세포 펠렛을 이후의 분석을 위하여 -20℃에서 보관하였다.

[0482] 분석 기질, (R,S)-아세토락테이트를 올라바우프(Aulabaugh) 및 술로스(Schloss)에 의해 설명된 바와 같이 합성하였다 (문헌[Aulabaugh and Schloss, Biochemistry, 29: 2824-2830, 1990]). 이 분석법에서 사용한 모든 다른 화학물질은 시그마로부터 구매하였다.

[0483] KARI에 의한 아세토락테이트의 α, β-다이하이드록시-아이소발레레이트로의 효소적 전환에 이어, 보조 인자, NADPH 또는 NADH가 상기 반응물로부터 사라지는 것을 플레이트 판독기 (미국 캘리포니아주 서니베일 소재의 몰레큘러 디바이스(Molecular Device))를 이용하여 340 nm에서 측정하였다. NADPH 또는 NADH 중 어느 하나에 대하여 6220 M⁻¹cm⁻¹의 몰 흡광 계수를 이용하여 활성을 계산하였다. 사용한 스톡 용액은 K₂HPO₄ (0.2 M); KH₂PO₄ (0.2 M); EDTA (0.5 M); MgCl₂ (1.0 M); NADPH (2.0 mM); NADH (2.0 mM) 및 아세토락테이트 (45 mM)였다. 2.0 ml의 K₂HPO₄, 3.0 ml의 KH₂PO₄, 4.0 ml의 MgCl₂, 0.1 ml의 EDTA 및 90.9 ml의 물을 함유하는 100 ml의 반응 완충제 (pH 6.8)를 제조하였다.

[0484] 깊은 웰의 플레이트 중의 냉동 세포 펠렛 및 버그버스터(BugBuster)를 실온에서 30분 동안 동시에 가온하였다. 96웰 분석 플레이트의 각각의 웰에 120 μl의 상기 반응 완충제 및 20 μl의 NADH (2.0 mM)를 충전시켰다. 75 μl의 50% 버그버스터 (v/v, 물 중)를 30분간의 가온 후 각각의 웰에 첨가하고, 세포를 플레이트 진탕기를 이용하여 현탁시켰다. 플레이트를 실온에서 20분 동안 인큐베이션하였다. 분취량 (예상되는 활성에 따라 15 내지 25 μl)의 세포 용해물을 96웰 분석 플레이트의 각각의 웰 내로 옮겼다. 340 nm에서의 흡광도를 배경으로서 기록하고, 16 μl의 아세토락테이트 (4.5 mM, 반응 완충제로 희석시킴)를 각각의 웰에 첨가하고, 플레이트 판독기에 의해 진탕시키면서 혼합하였다. 340 nm에서의 흡광도를 기질 첨가 후 예상되는 활성에 따라 0분, 및 10분 내지 30분에 기록하였다. 흡광도에서의 차이 (기질 첨가 전 및 후)를 이용하여 돌연변이체들의 활성을 결정하였다. 양성 대조군과 비교하여 더 높은 KARI 활성을 갖는 돌연변이체를 제스크리닝용으로 선발하였다.

[0485] 실시예 1, 실시예 2 및 실시예 3에서 라이브러리에 대하여 스크리닝하는 클론의 수는 각각 약 12,000개, 12,000개 및 92개이다. 각각의 라이브러리로부터의 최고 퍼포머(performer)를 이차 분석으로서 하기에 설명된 바와 같이 제스크리닝하였다.

[0486] 활성 돌연변이체의 이차 분석

[0487] 고 처리량 스크리닝 (상기)에 의해 확인한 선발된 돌연변이체를 함유하는 세포를, 250 rpm에서 진탕시키면서, 유도제로서 0.025%(w/v) 아라비노스 및 100 μg/ml의 암피실린을 함유하는 LB 배지 3.0 ml에서 37℃에서 하룻밤 성장시켰다. 그 후, 세포를 깊은 96웰 플레이트 내에 분취하고 (200 μl/웰), 실온에서 5분 동안 4,000 xg에서 원심분리에 의해 수확하였다. 75 μl의 50% 버그버스터 (v/v, 물 중)를 각각의 웰에 첨가하고, 세포를 플레이트 진탕기를 이용하여 현탁시켰다. 플레이트를 실온에서 20분 동안 인큐베이션하였다. 분취량 (예상되는 활성에 따라 15 내지 25 μl)의 세포 용해물을 96웰 분석 플레이트의 각각의 웰 내로 옮겼는데, 상기 플레이트는 웰당 120 μl의 반응 완충제 및 20 μl의 NADH (2.0 mM)를 함유한다. 340 nm에서의 흡광도를 배경으로서 기록하고, 16 μl의 아세토락테이트 (4.5 mM, 반응 완충제로 희석시킴)를 각각의 웰에 첨가하고, 플레이트 판독기에 의해 진탕시키면서 혼합하였다. 340 nm에서의 흡광도를 기질 첨가 후 예상되는 활성에 따라 0분, 및 5분 내지 10분에 기록하였다. 흡광도에서의 차이 (기질 첨가 전 및 후)를 이용하여 돌연변이체들의 활성을 결정하였다. 양성 대조군과 비교하여 더 높은 KARI 활성을 갖는 돌연변이체를 추가의 특성화용으로 선발하였다.

[0488] NADH 및 NADPH의 미하엘리스 상수의 측정

[0489] KARI 효소 활성은 상기에 설명된 바와 같이 NADH 또는 NADPH 산화에 의해 일상적으로 측정할 수 있지만, 이들 피리딘 뉴클레오티드의 미하엘리스 상수(K_M)를 측정하기 위하여 2,3-다이하이드록시아이소발레레이트 생성물의 형성량을 HPLC/MS를 이용하여 직접적으로 측정하였다.

[0490] 버그버스터 용해 세포 (상기에 설명한 바와 같음)로부터의 조 세포 추출물의 단백질 농도를 바이오라드(BioRad) 단백질 분석 시약 (미국 94547 캘리포니아주 허큘리스 소재의 바이오라드 래보러토리즈, 인크.(BioRad Laboratories, Inc.))을 이용하여 측정하였다. 0.2 내지 1.0 마이크로그램의 조 추출 단백질을, 100 mM의

MOPS KOH, pH 6.8, 10 mM의 $MgCl_2$, 1 mM의 EDTA, 1 mM의 글루코스-6-포스페이트 (시그마-알드리치), 0.2 유닛의 류코노스톡 메센테로이데스 글루코스-6-포스페이트 데하이드로게나아제 (시그마-알드리치) 및 다양한 농도의 NADH 또는 NADPH로 이루어진 반응 완충제에 90 μ l의 부피로 첨가하였다. 반응은 10 μ l의 /S/- 아세토락테이트를 2.5 mM의 최종 농도로 그리고 100 μ l의 최종 부피로 첨가함으로써 개시하였다. 30°C에서의 10분간의 인큐베이션 후, 50 μ l의 반응 혼합물을 인출하고 이것에 150 μ l의 0.1% 포름산을 첨가함으로써 반응을 쉐킹하였다. NADH 및 NADPH의 K_M 을 측정하기 위하여, 사용한 농도는 0.0003, 0.001, 0.003, 0.01, 0.03, 0.1, 0.3 및 1 mM이었다.

[0491] 2,3-다이하이드록시아이소발레레이트에 대하여 분석하기 위하여, 2 μ l의 포름산으로 쉐킹한 반응 혼합물을 워터스(Waters) SQD 질량 분석계 (미국 매사추세츠주 밀포드 소재의 워터스 코퍼레이션(Waters Corporation))를 갖춘 워터스 어퀴티(Acquity) HPLC 내에 주입하였다. 크로마토그래피 조건은 다음과 같았다: 워터스 어퀴티 HSS T3 컬럼 (2.1 mm의 직경, 100 mm의 길이)에서의 유량 (0.5 ml/min). 완충제 A는 물 중 0.1% (v/v)로 이루어지며, 완충제 B는 아세트니트릴 중 0.1% 포름산이었다. 1분 동안 1% 완충제 B (완충제 A 중), 이어서 1분에서 1% 완충제 B로부터 1.5분에서 75% 완충제 B까지의 선형 구배를 이용하여 샘플을 분석하였다. 반응 생성물, 2,3-다이하이드록시아이소발레레이트를 전기분무 이온화 -30 V 콘 전압(cone voltage)을 이용하여 $m/z=133$ 에서 이온화에 의해 검출하였다. 생성물인 2,3-다이하이드록시아이소발레레이트의 양을 인증된 표준물과의 비교에 의해 계산하였다.

[0492] NADH 및 NADPH에 대한 K_M 을 계산하기 위하여, S-아세토락테이트의 고정된 농도 (2.5 mM)에서 분석법에서 측정된 DHIV 형성 속도 데이터들, 포화 아세토락테이트 농도를 가정하여 마이크로소프트 엑셀(Microsoft Excel)에서의 최소 자승 회귀를 이용하여 단일 기질 미하엘리스-멘텐 등식(Michaelis-Menten equation)에 피팅시켰다.

[0493] 플라스미드 pYZ058, pLH550, pLH556, 및 pLH702의 제작

[0494] pYZ058 (pHR81-P_{CUP1}-AIsS-P_{ILV5}-효모 KARI; 서열 번호 176)을 pYZ090 (pHR81-P_{CUP1}-AIsS-P_{ILV5}-락티스 KARI; 서열 번호 195)으로부터 유도하였다. pYZ090을 PmeI 및 SfiI 효소로 절단하고, 효모 KARI의 PCR 생성물과 라이게이션시켰다. 상기 PCR 생성물을 사카로마이세스 세레비지에 BY4741 (리서치 제네틱스 인크.(Research Genetics Inc.)) 주의 게놈 DNA로부터, 어퍼 프라이머(upper primer) 5'-catcatcacagtttaaacagtatgttgaagcaaatcaacttcggtgg-3' (서열 번호 272) 및 로워 프라이머(lower primer) 5'-ggacggggccctgcaggccttattggttttctggtctcaactttctgac-3' (서열 번호 273)를 사용하여 증폭시켰으며, PmeI 및 SfiI 효소로 절단하였다. pYZ058을 서열결정에 의해 확인하였다.

[0495] pLH550 (pHR81-PCUP1-AIsS-P_{ILV5}-Pf5.KARI, 서열 번호 175)을 pYZ058 (서열 번호 176)로부터 유도하였다. 야생형 Pf5.KARI 유전자를 OT1349 (5'-catcatcacagtttaaacagtatgaaagttttctacgataaagactgcgacc-3'; 서열 번호 177) 및 OT1318 (5'-gcacttgatagccctgcaggcccttagttcttggtttgtcgacattttg-3'; 서열 번호 178)을 이용하여 PCR 증폭시키고, PmeI 및 SfiI 효소로 절단하고, PmeI 및 SfiI로 절단한 pYZ058 벡터와 라이게이션시켰다. 생성된 벡터, pLH550을 서열결정에 의해 확인하였다.

[0496] pLH556 (서열 번호 138; 도 4)은 상기 벡터를 SpeI 및 NotI 효소로 절단하고, SpeI 및 NotI 부위의 오버행(overhang) 서열을 함유하는, OT1383 (5'-ctagtcaccggtggc-3', 서열 번호 179) 및 OT1384 (5'-ggccgccaccggtga-3', 서열 번호 180)에 의해 어닐링된 링커와 라이게이션시킴으로써 pLH550으로부터 유도하였다. 이 클로닝 단계는 AIsS 유전자와, 큰 PCUP1 프로모터 단편을, 기능성이 아닌 160 bp의 나머지 상류 서열과 함께 제거한다. pLH556을 서열결정에 의해 확인하였다.

[0497] pHR81::ILV5p-K9D3 (pLH702, 서열 번호 181)을 pLH556으로부터 유도하였다. K9D3 돌연변이 KARI 유전자를 PmeI 및 SfiI 효소를 이용하여 벡터 pBAD-K9D3으로부터 잘라내고, PmeI 및 SfiI 부위에서 pLH556과 라이게이션시켜 Pf5.KARI 유전자를 K9D3 유전자로 대체하였다. 제작된 벡터를 서열결정에 의해 확인하였다.

[0498] 실시예 1

[0499] 다양한 KARI 유전자를 포함하는 효모 아이소부탄올 경로 주의 제작

[0500] KARI 활성을 갖는 폴리펩티드 및 효모 아이소부탄올 생성에서의 성능을 확인하기 위하여, 다양한 박테리아 및 진균류 중 유래의 KARI-코딩 유전자의 생물다양성 스크리닝을 실시하였다. 표 10에 나타낸, 사카로마이세스 세레비지에 유전자의 코돈 선호성을 바탕으로 하여 KARI 유전자를 코돈 최적화하였다. 각각의 KARI 유전자에 있어서, PmeI 제한효소 부위 및 추가의 3 bp (AGT)를 ATG 시작 코돈 앞에 서열 5'-GTTTAAACAGT-3' (서열 번호

136)을 갖는 5' 말단에 부가하고, SfiI 제한효소 부위를 서열 5'-GGCCCTGCAGGCC-3' (서열 번호 137)을 갖는 3' 말단에 부가하였다. KARI 유전자들 전부는 젠스크립트 유에스에이 인크.(GenScript USA Inc.; 미국 뉴저지주 피스카타웨이 소재)에 의해 합성되었다. 각각의 KARI 유전자를 pHR81-P_{CUP1}-AlsS-P_{I1v5}-Pf5.I1v5 벡터 (서열 번호 175) 내에, PmeI 및 SfiI 부위를 통하여 서브클로닝하였다 (I1v5는 효모 케톨-산 리덕토아이소머라아제를 코딩한다). 이 벡터는 2개의 발현 카세트, 즉 효모 CUP1 프로모터 하의 바실러스 서브틸리스 아세트락테이트 신타아제 (AlsS) 유전자, 및 I1v5 프로모터에 의해 제어되는 효모 I1v5 유전자를 함유한다. 서열 분석을 수행하여 KARI 유전자 서열을 확인하였다.

[0501] KARI 유전자를 지닌 pHR81-PCUP1-AlsS-PI_{1v5}-KARI 벡터들은 pLH468 (pRS423-P_{FBA1}-DHAD- P_{TDH3}-kivD-P_{GPM1}-hADH1; 서열 번호 139)과 함께 숙주 주 BP1135 (PNY1505; 실시예 8) (CEN.pk 113-7D 델타 ura3::loxP 델타 his3 델타 pdc6 델타 pdc1::ilvD.Sm 델타 pdc5::sadB 델타 gpd2::loxP 델타 fra2)를 동시 형질전환시켰다. 효모 형질전환체를 30℃에서의 5-7일 후 최소 드롭아웃(drop-out) 배지 플레이트 (SE-Ura-His, 2% 에탄올) 상에서 선발하고, SE-Ura-His 상에 재도말하여, 추가로 3일간 인큐베이션한 후 세포 패치를 수득하였다. 상기 세포 패치를 진탕 플라스크 접종에 사용하였다.

[0502] 실시예 2

[0503] 아이소부탄을 생성에 대한 KARI 다양성 컬렉션의 스크리닝

[0504] 다양한 KARI 유전자를 효모에서 그의 "유효 생산성"을 바탕으로 하여 평가하였다. 유효 생산성은 점진적 산소 제한 조건 하에서 소정 기간의 성장 (예를 들어, 48시간) 후 결정하였다. 효모 바이오매스는 효모 세포의 1 OD₆₀₀이 0.3 g/L와 동등하다고 가정하여 계산하였다.

[0505] 다양한 KARI 유전자를 지닌 효모 아이소부탄을 경로 주를 0.2% 글루코스 및 0.2% 에탄올을 포함하는 10 mL의 SEG-Ura,His 배지 내에 접종하고, 약 2 OD가 되도록 30℃에서 하룻밤 호기적으로 성장시켰다. 상기 배양물을 원심분리하고, 세포의 일부분을 125 mL 진탕 플라스크에서 25 mL의 총 부피로 0.4의 초기 OD₆₀₀가 되도록 SEG-Ura,His (2% 글루코스, 1% 에탄올)에 재현탁시켰다. 상기 진탕 플라스크를 나사식 중실 플라스틱 캡(screw-on solid plastid cap)으로 닫고, 배양물을 외부 환경과의 최소 공기 및 산소 교환 하에서 플라스크에서 점진적 산소 제한 조건 하에서 성장시켰다. 30℃, 250 RPM에서의 48시간의 인큐베이션 후, 배양물을 OD₆₀₀ 측정 및 HPLC 분석용으로 꺼내어 아이소부탄을 생성량을 측정하였다.

[0506] 하기에 나타난 바와 같이, 스크리닝한 KARI 유전자로부터, 다수가 락토코커스 락티스 KARI에 비견되는 또는 그보다 더 우수한 아이소부탄을 역가를 가졌다. 특히, K9 (엔에어로스티페스 카카에 DSM 14662) KARI 클론은 점진적 산소 제한 조건 하에서의 48시간의 성장 후 측정할 때 높은 아이소부탄을 역가 및 유효 아이소부탄을 생산성을 나타냈다 (표 10).

[0507] [표 10]

30℃에서 진탕 플라스크에서 점진적으로 산소가 제한되는 조건 하에서 48 시간의 성장 후 측정함, 다양한 KARI 유전자를 지니는 효모 아이소부탄올 생산주로부터의 아이소부탄올 역가 및 유효 생산성

KARI 클론	서열 번호 (핵산, 아미노산) *LTS 및 S2를 제외한 모든 핵산 서열은 코돈 최적화됨	아이소부탄올 역가 (g/L)	유효 아이소부탄올 생산성 (g/g)	소스 유기체
B3K01 ("K1")	140, 141	2.6	4.1	비피도박테리움 안굴라툼 DSM 20098
B3K02 ("K2")	142, 143	3.5	3.7	비피도박테리움 덴티움 ATCC 27678
B3K09 ("K9")	26, 27	4.3	5.2	엔에어로스티페스 카카에 DSM 14662
B3K25 ("K25")	375, 376	3.6	4.4	엔테로코커스 갈리나툼 EG2
B3K26 ("K26")	381, 382	4.4	3.2	스트렙토코커스 서모필루스 LMD-9
B3K29 ("K29")	377, 378	4.1	3.3	락토코커스 락티스 아종 크레모리스 MG1363
LTS	379, 380	2.7	3.1	락토코커스 락티스
B3K07 ("K7")	274, 275	3.7	2.8	칼로스트리둠 베이체렌키이 NCIMB 8052
S2	276, 277	3.6	1.5	자이모모나스 모빌리스

[0508]

[0509] 실시예 3

[0510] 효모 아이소부탄올 경로 주의 KARI 효소 분석

[0511] IpOHA (N-아이소프로필 옥살릴하이드록삼산)는 KARI 효소에 의해 촉매되는 반응을 위한 반응 중간체의 모방체이다. 이것은 KARI 효소의 활성 부위에 결합하는 타이트한(tight) 결합 저해제이다. IpOHA의 합성 및 이. 콜라이 유래의 KARI에의 그의 타이트한 결합은 문헌[A. Aulabaugh and J. V. Schloss, Biochemistry, 1990, 29, 2824-2830]에 기재되어 있다. 활성 부위 적정에 있어서의 그의 사용은 이전에 보고되지 않았다. IpOHA를 문헌에 따라 [^{14} C]-옥살레이트로부터 합성하였다.

[0512] 실시예 2로부터의 효모 배양물을 수확하고, KARI 효소 활성에 대하여 분석하였다. 25 mL의 상기 배양물을 펠렛화하고, 10 mL의 pH 7.5의 50 mM 트리스(Tris)-HCl에 재현탁시켰다. 세포를 다시 원심분리하여 상기 완충제를 제거하였으며, 세포 펠렛을 -70℃에 보관한다. 세포 펠렛을 1 mL의 pH 7.5의 50 mM 트리스-HCl에 재현탁시키고, 초음파 처리하였다. 용해성 조 세포 추출물을 사용하여 효소 분석을 수행하였다. 효소의 일부분을 물 과량의 [^{14} C]-IpOHA, 및 포화 농도의 NAD(P)H 및 Mg^{2+} 와 함께 인큐베이션하였다. 처음에 가역적, 희석-민감성 복합체가 형성되기 때문에, 추출물 농도를 높게 유지하여 복합체 형성을 유리하게 하고, 그에 따라 타이트한 복합체 형성에 걸리는 시간을 감소시켰다. 각각의 KARI가 타이트한 복합체를 형성하는 데 얼마나 긴 시간이 걸리는지가 선택적으로 공지되어 있지 않기 때문에, 2개의 시점을 각각의 샘플에 대하여 취하여 결과들이 일치되는지를 검증하였다. 인큐베이션 시간의 마지막에, 마이크로콘(Microcon)(등록상표) (미국 매사추세츠주 빌레리카 소재의 밀리포어 인크.(Millipore Inc.))을 이용한 한외여과에 의해 단백질 분자로부터 작은 분자를 분리하고, 고분자량 분획을 카운팅하였다. μ M 또는 mg/ml 중 어느 하나의 단위의, 샘플 중 KARI의 농도를 ^{14} C dpm, 부피 및 KARI 서브유닛 분자량으로부터 역계산하였다. 고정 시간 효소 분석을 동시에 실행하였으며, 이 데이터를 사용

하여 U/ml을 계산하였다. 주어진 샘플에 있어서 U/ml을 mg/ml로 나눔으로써 비활성을 계산하였다. 전체 활성과, IpOHA에 결합하는 능력은 절대적으로 상관된다고 가정하였다. 이렇게 측정한 KARI 효소의 비활성이 표 11에 열거되어 있다. "유닛/mg" 단위의 KARI 활성은 IpOHA 분석법을 이용하여 정량화한 KARI 효소 1 mg당 활성을 나타낸다. 총 단백질 농도를 브래드포드법(Bradford method)으로 결정하였으며, KARI의 발현 수준은 KARI 효소의 양을 전체 용해성 세포 단백질의 양으로 나눔으로써 계산한다.

[표 11]

IPOHA 분석법에 의해 측정한 KARI 효소 활성

KARI 클론	KARI 활성 (U/mg)	KARI 총 단백질 (%)	유기체
B3K01	0.56	21	비피도박테리움 안굴라툼 DSM 20098
B3K02	0.44	28	비피도박테리움 덴티움 ATCC 27678
B3K09	2.4	15	엔에어로스티페스 카카에 DSM 14662
B3K25	1.3	21	엔테로코커스 갈리나룸 EG2
B3K26	1.5	17	스트렙토코커스 서모필루스 LMD-9
B3K29	1.6	14	락토코커스 락티스 아종 크레모리스 MG1363
LTS	0.8	23.0	락토코커스 락티스

실시예 4

야생형보다 더 낮은 K_m 을 갖는 NADH를 이용하여 변이체를 확인하기 위한 부위 포화 유전자 라이브러리의 제작

K9 KARI용의 pBAD-기반 박테리아 발현 벡터를 제작하기 위하여, K9 KARI 유전자 (미국 뉴저지주 피스카타웨이 소재의 젠스크립트에 의해 합성됨)를 PmeI 및 SfiI 부위를 통하여 pBAD-ps-JEA1 벡터 (서열 번호 905) 내로 서브클로닝하였다. 엔에어로스티페스 카카에 유래의 케톨-산 리덕토아이스머라아제 (KARI) (K9-KARI로 칭해짐)를 라이브러리 제작에 사용하였다. 하나의 유전자 라이브러리를 구매가능한 키트, T4 폴리뉴클레오티드 키나아제 (PNK) (미국 오하이오주 클레블랜드 소재의 유에스비 코퍼레이션(USB Corporation), #70031Z) 및 Chang_IT 다중 돌연변이 부위 유도 돌연변이 유발(Multiple Mutation Site Directed Mutagenesis) 키트 (미국 오하이오주 클레블랜드 소재의 유에스비 코퍼레이션, #78480)를 사용하여 제작하였다.

올리고뉴클레오티드 (K9_56_58_060210f: GAAGGANNKAAANNKTGGAAGAGAGC, 서열 번호 144; 및 K9_56_58_060210r: GCTCTCTCCAMNNTTTMNTCTTC, 서열 번호 145)는 인테그레이티드 디엔에이 테크놀로지스, 인크.(Integrated DNA Technologies, Inc; 미국 아이오와주 코랄빌 소재)에 의해 합성되었다. 상기 올리고뉴클레오티드들을 먼저 T4 PNK로 포스포릴화하였다. 간략하게는, 30 μ l의 반응 혼합물은 키트와 함께 공급된 3.0 μ l의 10x T4 PNK 완충제, 4.0 μ l의 프라이머 (약 35 μ M), 0.8 μ l의 100 mM ATP 믹스, 0.6 μ l의 T4 PNK 및 22 μ l의 물을 함유하였다. 반응 혼합물을 37°C에서 1.0시간 동안 인큐베이션하고, 그 후 T4 PNK를 65°C에서 20분 동안 불활성화시켰다.

그 후, 포스포릴화된 프라이머들을 후속 PCR 반응에 직접적으로 사용하여 두 부위에서의 돌연변이를 상기 키트를 사용하여 K9 KARI 야생형 내로 도입하였다. 간략하게는, 30 μ l의 반응 혼합물은 키트와 함께 공급된 3.0 μ l의 10x 반응 완충제, 3.0 μ l의 포스포릴화 정방향 프라이머 및 역방향 프라이머, 2.0 μ l의 K9 KARI 야생형 (50 ng/ μ l), 1.2 μ l의 Chang_IT 효소 및 17.8 μ l의 물을 함유하였다. 이 반응 혼합물을 얇은 벽의 200 μ l-용량의 PCR 튜브 내에 넣고, 하기 PCR 반응 프로그램을 PCR에 사용하였다: 출발 온도는 95°C, 2분이었으며, 이어서 30회의 가열/냉각 사이클이었다. 각각의 사이클은 95°C, 30초; 55°C, 30초; 및 68°C, 20분으로 이루어졌다. 상기 온도 사이클링의 완료시에, 샘플들을 68°C에서 25분 더 유지하고, 그 후 이후의 프로세싱을 위하여 4°C에서 유지하였다. PCR 반응물을 자이모(Zymo) DNA 클린업(clean-up) 키트 (미국 캘리포니아주 오렌지 소재의 자이모 리서치 코퍼레이션(Zymo Research Corporation), #D4004)를 사용하여 클린업하였다. DNA를 84 μ l의 물을 사용하여 막으로부터 용출시켰다. DNA 주형을 Dpn I (미국 위스콘신주 매디슨 소재의 프로메가(Promega), #R6231)을 이용하여 37°C에서 3시간 동안 제거하였다 (반응 혼합물: 10 μ l의 10x 반응 완충제, 1.0 μ l의 BSA, 6.0 μ l의 Dpn I 및 83 μ l의 세정된 PCR DNA). Dpn I 절단된 DNA를 자이모 DNA 클린업 키트

로 다시 클린업하고, Dpn I로 다시 절단하여 DNA 주형을 완전히 제거하였다 (반응 혼합물: 1.5 μ l의 10x 반응 완충제, 0.15 μ l의 BSA, 0.85 μ l의 Dpn I 및 83 μ l의 세정된 PCR DNA). 반응 혼합물을 전기-적격성 이. 콜라이 Bw25113($\Delta i/vC$) 주 (전체적으로 본 명세서에 참고로 포함된 미국 특허 제8,129,162호에 기재됨)의 형질전환에 직접적으로 사용하였으며, 이는 바이오라드 진 펄서(Gene Pulser) II (미국 캘리포니아주 허큘리스 소재의 바이오-라드 래보라토리즈 인크.)를 이용하였다. 형질전환된 클론을 LB 배지 및 100 μ g/ml의 암피실린 (카탈로그 번호: L1004, 미국 캘리포니아주 홀리스터 소재의 테크노바 인크.(Teknova Inc.))을 함유하는 한천 플레이트 상에 도말하고, 37°C에서 하룻밤 인큐베이션하였다. 클론을 NADH를 이용하여 활성에 대하여 스크리닝하였다. 변이체들에 있어서의 K_M 을 측정하였다 (표 12).

[표 12]

DHIV 형질량 분석을 통하여 결정한, 이. 콜라이 추출물에서의 K9 KARI 변이체의 동력학적 값

돌연변이체	서열 번호 (핵산, 아미노산)	돌연변이	K_M (μ M) (NADH)	K_M (μ M) (NADPH)
K9 WT	26, 27		326	0.2
AB1D1	28, 29	S56A	164	1
495B5	30, 31	S56A/S58H	44	4
AB1D3 ("K9D3" 으로도 칭해짐)	32, 33	S56A/S58D	38	9
AB1G9 ("K9G9" 로도 칭해짐)	34, 35	S56AS58E	47	23

실시예 5

NADH에 대한 K_M 을 저하시키기 위한 부위 포화 유전자 라이브러리의 제작

슈도모나스 플루오레센스 KARI (PF5-KARI)를 이용한 연구를 바탕으로 하여, 위치 24, 33, 61, 80, 156 및 170을 K9 KARI에 있어서의 돌연변이 유발 표적으로서 표적화하였다. PF5-KARI와 K9 KARI 사이의 다중 서열 정렬 (MSA) (도 2)에 의하면, 상응하는 위치는 30, 39, 67, 86, 162, 및 176이다.

더 많은 돌연변이 유발 표적을 확인하기 위하여, 부탄올 생성 주(butanologen strain) (다른 실시예 참조)에서 아이소부탄올을 생성하는 것으로 결정된 기존의 KARI 효소 (K1, K2, K7, K9, K25, K26, 엘. 락티스 및 S2)의 MSA를 이용하여 더욱 많은 돌연변이 유발 표적을 확인하였다. 위치 41, 87, 131, 191, 227, 및 246을 돌연변이 유발 표적으로서 선택하였다.

위치 30, 39, 41, 67, 86, 87, 131, 162, 176, 191, 227, 및 246을 표적화하는 올리고뉴클레오티드들은 인테그레이티드 디엔에이 테크놀로지스, 인크. (미국 아이오와주 코탈빌 소재)에 의해 상업적으로 합성되었다 (표 13). 위치 30, 67, 131, 162, 176, 191, 227, 및 246을 표적화하는 8쌍의 올리고뉴클레오티드를 사용하여 메가프라이머(Megaprimer)들을 생성하였으며, 이는 인비트로젠으로부터의 슈퍼믹스(Supremix)를 사용하였다 (카탈로그 번호: 10572-014, 미국 캘리포니아주 칼스배드 소재의 인비트로젠). 각각의 PCR 반응에 있어서, 프라이머 쌍, 상기 8쌍의 올리고뉴클레오티드로부터의 상이한 위치를 코딩하는 하나의 정방향 프라이머와 하나의 역방향 프라이머의 임의의 조합 (예를 들어 K9_30_101110f 및 K9_67_101110r)을 이용하였다. 총 P_{8^2} 또는 56가지의 조합이 있다. 25 μ l의 반응 혼합물은 22.5 μ l의 슈퍼믹스 용액, 1.0 μ l의 정방향 프라이머 및 1.0 μ l의 역방향 프라이머, 0.5 μ l의 AB1D3 DNA 주형 (50 ng/ μ l)을 함유하였다. 혼합물을 마스터사이클러 (Mastercycler) 구매 장비(미국 뉴욕주 웨스트베리 소재의 브링크만 인스트루먼트, 인크.)에서의 PCR 반응을 위하여 얇은 벽의 200 μ l 튜브 내에 넣었다. 하기 조건을 PCR 반응에 이용하였다: 출발 온도는 95°C, 1.0분이었으며, 이어서 35회의 가열/냉각 사이클이었다. 각각의 사이클은 95°C, 20초; 55°C, 20초; 및 72°C, 1.0분으로 이루어졌다. 상기 온도 사이클링의 완료시에, 샘플들을 72°C에서 2.0분 더 유지하고, 그 후 4°C에서 샘플 회수를 기다리면서 유지하였다. PCR 생성물을 제조업자가 권고하는 바와 같이 DNA 세정 키트 (카탈로그 번호: D4003, 미국 캘리포니아주 오렌지 소재의 자이모 리서치)를 이용하여 클린업하였다.

[0527] 그 후, 퀵체인지(QuickChange) II XL 부위 지정 돌연변이 유발 키트 (카탈로그 번호 200524, 미국 캘리포니아주 라 줄라 소재의 스트라타진)를 이용하여, 상기 메가프라이머를 이용하여 유전자 라이브러리를 생성하였다. 25 μ l의 반응 혼합물은 2.5 μ l의 10x 반응 완충제, 1.0 μ l의 50 ng/ μ l의 주형, 20.5 μ l의 메가프라이머, 0.5 μ l의 40 mM dNTP 믹스, 0.5 μ l의 pfu-울트라(ultra) DNA 폴리머라아제를 함유하였다. 메가프라이머 및 주형을 제외하고는, 여기서 사용한 모든 시약은 상기에 나타난 키트와 함께 제공되었다. 이 반응 혼합물을 얇은 벽의 200 μ l-용량의 PCR 튜브 내에 넣고, 하기 반응을 PCR에 이용하였다: 출발 온도는 95℃, 30초였으며, 이어서 25회의 가열/냉각 사이클이었다. 각각의 사이클은 95℃, 30초; 55℃, 1분; 및 68℃, 6분으로 이루어졌다. 상기 온도 사이클링의 완료시에, 샘플들을 68℃에서 8분 더 유지하고, 그 후 이후의 프로세싱을 위하여 4℃에서 유지하였다. PCR 반응 혼합물을 실시예 4에서 이용한 것과 동일하게 Dpn I 제한 효소로 프로세싱하였다.

[0528] 그 후, 퀵체인지 II XL 부위 지정 돌연변이 유발 키트를 이용하여, 올리고뉴클레오타이드 K9_37&39_101110f, K9_37&39_101110r 및 K9_86&87_101110f, K9_86&87_101110r을 직접적으로 사용하여 유전자 라이브러리를 생성하였다. 상기 두 올리고뉴클레오타이드 세트들을 위한 2가지의 25 μ l 반응 혼합물. 각각의 25 μ l 반응 혼합물은 2.5 μ l의 10x 반응 완충제, 1.0 μ l의 50 ng/ μ l 주형, 1.0 μ l의 정방향 프라이머, 1.0 μ l의 역방향 프라이머, 0.5 μ l의 40 mM dNTP 믹스, 0.5 μ l의 pfu-울트라 DNA 폴리머라아제 및 18.5 μ l의 물을 함유하였다. PCR 프로그램 및 후속 Dpn I 프로세싱은 동일하다.

[0529] Dpn I 프로세싱된 DNA 혼합물을 제조업자의 프로토콜에 따라 자이모 DNA 클린업 키트를 사용하여 클린업하였다. 클린업된 DNA를 사용하여, 바이오라드 진 펄서 II (미국 캘리포니아주 허큘리스 소재의 바이오-라드 레보라토리즈 인크.)를 이용하여 전기-적격성 이. 콜라이 Bw25113 ($\Delta ilvC$) 주를 형질전환시켰다. 형질전환된 클론을 LB 배지 및 100 μ g/ml의 암피실린 (카탈로그 번호: L1004, 미국 캘리포니아주 홀리스터 소재의 테크노바 인크.)을 함유하는 한천 플레이트 상에 도말하고, 37℃에서 하룻밤 인큐베이션하였다. 클론을 NADH를 이용하여 개선된 활성에 대하여 스크리닝하였다. 개선된 돌연변이체에 있어서의 K_M 를 측정하였다 (표 14).

[0530] [표 13]

프라이머

K9-KARI의 표적화된 위치(들)	프라이머
30	K9_30_101110f: gactatcgccgttatcggtNNKggttctcaaggtcac 서열 번호 146 K9_30_101110r: GTGACCTTGAGAACCMMNACCGATAACGGCGATAGTC 서열 번호 147
67	K9_67_101110f: gagctgaagaacaaggtNNKgaagtctacacgcgtgc 서열 번호 148 K9_67_101110r: GCAGCGGTGTAGACTTCMNNACCTTGTCTTCAGCTC 서열 번호 149
131	K9_131_101110f: caaaggacgttgatgtcNNKatgatcgctccaaag 서열 번호 150 K9_131_101110r: CTTTGGAGCGATCATMNNACATCAACGTCCTTT 서열 번호 151
162	K9_162_101110f: gctgtcgaacaagacNNKactggcaagcgtttg 서열 번호 152 K9_162_101110r: CAAAGCCTTGCCAGTMNNGTCTTGTTCGACAGC 서열 번호 153
176	K9_176_101110f: gctttggcctacgctttaNNKatcggtggtagtagc 서열 번호 154 K9_176_101110r: GCTCTAGCACCCGATMNNNTAAAGCGTAGGCCAAAGC 서열 번호 155
191	K9_191_101110f: gaaactaccttcagannKgaactgaaaccgac 서열 번호 156 K9_191_101110r: GTCGGTTTCAGTTTCMNNCTCTGAAGGTAGTTTC 서열 번호 157
227	K9_227_101110f: gccggttacgacccaNNKaacgcttacttcgaatg 서열 번호 158 K9_227_101110r: CATTCGAAGTAAGCGTTMNNNTGGGTCGTAACCGGC 서열 번호 159
246	K9_246_101110f: gttgacttgatctacNNKtctggtttctccggtatgc 서열 번호 160 K9_246_101110r: GCATACCGGAGAAACCAGAMNNGTAGATCAAGTCAAC 서열 번호 161
39, 41	K9_37&39_101110f: gttctcaaggtcacgctNNKgccNNKaatgctaaggatcc 서열 번호 162 K9_37&39_101110r: GGATTCCTTAGCATMNNNGCMNAGCGTGACCTTGAGAAC 서열 번호 163
86, 87	K9_86&87_101110f: gacatcattatgatctctgNNKNNKgatgaaaagcaggc 서열 번호 164 K9_86&87_101110r: GCCTGCTTTTCATCMNNMNNCAAGATCATAATGATGTC 서열 번호 165

[0531]

[0532] [표 14]

측정된 K_M 값을 갖는 일부 돌연변이체의 목록

돌연변이체	서열 번호 (핵산, 아미노산)	돌연변이	K_M (μM) (NADH)	K_M (μM) (NADPH)
A07A9	38, 39	S56A/S58D/I86T/N87P	15	7
A07B5	36, 37	S56A/S58D/I86V/N87P	8	4
A07H8	40, 41	S56A/S58D/N87P	8	6
A07D8	42, 43	S56A/S58D/T131C/T191S	26	6
A07F7	44, 45	S56A/S58D/T131V/T191A	28	7
A07H7	46, 47	S56A/S58D/T191S	29	8

[0533]

[0534] 실시예 6

[0535] NADH에 대한 K_M 을 저하시키기 위한 조합 라이브러리의 제작

[0536] 돌연변이 유발 결과 (실시예 4)를 바탕으로 하면, T131L, T131A, T131V, T131M, T131C, T191D, T191C, T191S, 및 T191G는 NADH에 대한 K_M 을 개선시키는 유익한 돌연변이인 것으로 여겨진다. A07B5 내로 이들 유익한 돌연변이를 도입하기 위한 조합 라이브러리를 만들었다.

[0537] 모든 올리고뉴클레오타이드는 인테그레이티드 디엔에이 테크놀로지스, 인크.(미국 아이오와주 코랄빌 소재)에 의해 합성되었다. 상기 올리고뉴클레오타이드들을 먼저 T4 PNK로 포스포릴화하였다. 간략하게는, 20 μl 의 반응 혼합물은 키트와 함께 공급된 2.0 μl 의 10x T4 PNK 완충제, 2.85 μl 의 프라이머 (약 35 μM), 0.6 μl 의 100 mM ATP 믹스, 0.4 μl 의 T4 PNK 및 14.15 μl 의 물을 함유하였다. 반응 혼합물을 37°C에서 1.0시간 동안 인큐베이션하고, 그 후 T4 PNK를 65°C에서 20분 동안 불활성화시켰다.

[0538] 그 후, 포스포릴화된 프라이머들을 후속 PCR 반응에 직접적으로 사용하여 두 부위에서의 돌연변이를 상기 키트를 사용하여 A07B5 내로 도입하였다. 간략하게는, 50 μl 의 반응 혼합물은 키트와 함께 공급된 5.0 μl 의 10x 반응 완충제, 2.5 μl 의 포스포릴화 정방향 프라이머 (표 15에 나타난 각각의 정방향 프라이머 0.5 μl), 2.5 μl 의 역방향 프라이머 (표 15에 나타난 각각의 정방향 프라이머 0.625 μl), 2.5 μl 의 A07B5 (50 ng/ μl), 2.5 μl 의 Chang_IT 효소 및 35 μl 의 물을 함유하였다. 이 반응 혼합물을 얇은 벽의 200 μl -용량의 PCR 튜브 내에 넣고, 하기 PCR 반응 프로그램을 PCR에 사용하였다: 출발 온도는 95°C, 2분이었으며, 이어서 30회의 가열/냉각 사이클이었다. 각각의 사이클은 95°C, 30초; 55°C, 30초; 및 68°C, 20분으로 이루어졌다. 상기 온도 사이클링의 완료시에, 샘플들을 68°C에서 25분 더 유지하고, 그 후 이후의 프로세싱을 위하여 4°C에서 유지하였다. PCR 반응물을 자이모 DNA 클린업 키트 (미국 캘리포니아주 오렌지 소재의 자이모 리서치 코포레이션, #D4004)를 사용하여 클린업하였다. DNA를 84 μl 의 물을 사용하여 막으로부터 용출시켰다. DNA 주형을 Dpn I (미국 위스콘신주 매디슨 소재의 프로메가, #R6231)을 이용하여 37°C에서 3시간 동안 제거하였다 (반응 혼합물: 10 μl 의 10x 반응 완충제, 1.0 μl 의 BSA, 6.0 μl 의 Dpn I 및 83 μl 의 세정된 PCR DNA). Dpn I 절단된 DNA를 자이모 DNA 클린업 키트로 다시 클린업하고, Dpn I로 다시 절단하여 DNA 주형을 완전히 제거하였다 (반응 혼합물: 1.5 μl 의 10x 반응 완충제, 0.15 μl 의 BSA, 0.85 μl 의 Dpn I 및 83 μl 의 세정된 PCR DNA).

[0539] Dpn I 프로세싱된 DNA 혼합물을 제조업자의 프로토콜에 따라 자이모 DNA 클린업 키트를 사용하여 클린업하였다. 클린업된 DNA를 사용하여, 바이오라드 진 펄서 II (미국 캘리포니아주 허큘리스 소재의 바이오-라드 래보러토리즈 인크.)를 이용하여 전기-적격성 이. 콜라이 Bw25113 ($\Delta i/vC$) 주를 형질전환시켰다. 형질전환된 클론을 LB 배지 및 100 $\mu g/ml$ 의 암피실린 (카탈로그 번호: L1004, 미국 캘리포니아주 홀리스터 소재의 테크노바 인크.)을 함유하는 한천 플레이트 상에 도말하고, 37°C에서 하룻밤 인큐베이션하였다. 클론을 NADH를 이용하여 개선된 활성에 대하여 스크리닝하였다. 개선된 돌연변이체에 있어서의 K_M 을 측정하였다 (표 16).

[0540] [표 15]

실시예 6에 있어서의 프라이머

K9-KARI의 표적화된 위치(들)	프라이머
131	K9_131L_112210f: ggacgttgatgtcTTGatgatcgctcc 서열 번호 166 K9_131A_112210f: ggacgttgatgtcGCAatgatcgctcc 서열 번호 167 K9_131V_112210f: ggacgttgatgtcGTTatgatcgctcc 서열 번호 168 K9_131M_112210f: ggacgttgatgtcATGatgatcgctcc 서열 번호 169 K9_131C_112210f: ggacgttgatgtcTGAatgatcgctcc 서열 번호 170
191	K9_191D_112210r: GGTTCAGTTTCGTCTCTGAAGGTAGTTTC 서열 번호 171 K9_191C_112210r: GGTTCAGTTTCGCATCTGAAGGTAGTTTC 서열 번호 172 K9_191S_112210r: GGTTCAGTTTCCGATCTGAAGGTAGTTTC 서열 번호 173 K9_191G_112210r: GGTTCAGTTTCGCCTCTGAAGGTAGTTTC 서열 번호 174

[0541]

[0542] [표 16]

측정된 K_M 값을 갖는 일부 돌연변이체의 목록

돌연변이체	서열 번호 (핵산, 아미노산)	돌연변이	K _M (μM) (NADH)	K _M (μM) (NADPH)
AWB9	52, 53	S56A/S58D/I86V/N87P/T131A	10	4
AWC1	54, 55	S56A/S58D/I86V/N87P/T131V	9	3
AWD6	62, 63	S56A/S58D/I86V/N87P/T131V/T191S	5	2
AWD10	64, 65	S56A/S58D/I86V/N87P/T131A/T191C	8	3
AWF4	56, 57	S56A/S58D/I86V/N87P/N107S/T131V	7	3
AWF6	58, 59	S56A/S58D/I86V/N87P/T131V/T191D	7	2
AWG4	50, 51	S56A/S58D/I86V/N87P/T131M	7	3
AWH3	60, 61	S56A/S58D/I86V/N87P/T131V/T191G	6	2
AS6F1	48, 49	S56A/S58D/I86V/N87P/T131M/T191G	4	1

[0543]

[0544] 실시예 7

[0545] K9 KARI 변이체로부터의 아이소부탄올의 생성

[0546] K9 KARI의 하기 변이체들을 상기에 설명한 바와 같이 생성하였다.

[0547] [표 17]

KARI 변이체 및 상응하는 효모 발현 벡터

클론명	효모 벡터명	뉴클레오티드 점 돌연변이 위치					아미노산 돌연변이 위치
		166	168	172	173	174	
WT K9 KARI	pHR81- PI1v5- KARI-K9	T	T	T	C	C	S56, S58
AB1G9	pHR81- PI1v5- KARI-K9.G9	G	G	G	A	G	S56A, S58E
495B5	pHR81- PI1v5- KARI-K9.B5	G	t	C	A	T	S56A, S58H
AB1D3	pHR81- PI1v5- KARI-K9.D3	G	t	G	A	t	S56A, S58D
AB1D1	pHR81- PI1v5- KARI-K9.D1	G	t	t	C	C	S56A

[0548]

[0549]

효모 발현 플라스미드는 이. 콜라이 벡터 (pBAD.KARI) 유래의 변이형 KARI 유전자를 PmeI 및 SfiI 부위에서 pHR81-PI1v5-Pf5.KARI 벡터 pLH556 (도 4, 서열 번호 138) 내로 서브클로닝함으로써 만들었다. 효모 경로 주는 PNY2204 숙주 (MATa ura3Δ::loxP his3Δ pdc6Δ pdc1Δ::P[PDC1]-DHAD|ilvD_Sm-PDC1t-pUC19-loxP-kanMX-loxP-P[FBA1]-ALS|alsS_Bs-CYC1t pdc5Δ::P[PDC5]-ADH|sadB_Ax-PDC5t gpd2Δ::loxP fra2Δ adh1Δ::UAS(PGK1)P[FBA1]-kivD_L1(y)-ADH1t; 실시예 13)에서, 경로 플라스미드 #1로서의 상기 KARI 벡터, 및 경로 플라스미드 #2로서의 pBP915 (pRS423-PFBA1-DHAD-PGPM1-hADH1; 서열 번호 182)의 동시 형질전환에 의해 만들었다. 형질전환체들을 탄소 공급원으로서 0.1% 에탄올 및 2% 글루코스를 함유하는 동일 배지에 패치하였다 (patched). 3개의 패치를 세럼 바이알에서 미세호기성 조건 하에서 아이소부탄을 생성에 대하여 검사하였다. K9D3을 발현하는, pBP915 및 pLH702 플라스미드로 형질전환된 클론을 PNY1910으로 표기하였다.

[0550]

SE-Ura-His 플레이트 상의, 형질전환으로부터의 효모 콜로니는 5-7일 후에 나타났다. 상기 콜로니들을 신선한 SE-Ura-His 플레이트 상에 패치하고, 30℃에서 3일 동안 인큐베이션하였다. 패치된 세포를 0.2% 글루코스 및 0.2% 에탄올을 포함하는 25 mL의 SEG-Ura,HIS 배지 내에 접종하고, 2-3의 OD가 되도록 30℃에서 1-2일 동안 호기적으로 성장시켰다. 세포를 원심분리하고, 1 mL의 SEG-Ura,HIS 배지 (2% 글루코스, 0.1% 에탄올, 10 mg/L의 에르고스테롤, 50 mM MES, pH 5.5, 티아민 30 mg/L, 니코틴산 30 mg/L)에 재현탁시켰다. 60 mL의 세럼 바이알에서, 출발 OD=0.2에 있어서 45 mL의 총 부피의 동일 배지로 계산된 양의 세포를 옮겼으며, 이때 뚜껑을 크림퍼 (crimper)로 타이트하게 닫았다. 이 단계를 공기 중에서 보통의 바이오후드(bio-hood)에서 행하였다. 상기 세럼 바이알을 30C, 200 rpm에서 2일 동안 인큐베이션하였다. 48시간에, 샘플들을 OD 및 글루코스, 아이소부탄을 및 경로 중간체의 HPLC 분석을 위하여 꺼냈다. 24시간 샘플들을 혐기성 챔버에서 취하여 세럼 바이알에서의 혐기성 조건을 유지하였다. 상기 48 h 인큐베이션의 처음 시기에, 상부 공간 (대략 15 mL)에 존재하는 공기 및 액체 배지는 성장 중인 효모 세포에 의해 소비된다. 상기 상부 공간 중의 산소가 소비된 후, 배양물은 혐기성으로 된다. 따라서, 이 실험은 호기성 조건으로부터 산소 제한 및 혐기성 조건으로의 스위칭을 포함한다.

[0551]

4가지의 K9 변이체 중, AB1G9 및 AB1D3은 상대적으로 높은 아이소부탄을 역가를 생성한 반면, 495B5 및 AB1D1은 더 낮은 역가를 갖는다. 야생형 K9 KARI 주는 최저 역가를 생성하였다. 이론에 의해 구해되고자 함이 없이, 상기 더욱 낮은 역가는 세포가 호기성 조건으로부터 혐기성 조건으로 스위칭될 때 NADH와 NADPH의 균형 이동으로 인한 것이라고 믿어진다. 이 근거에 의하면, 혐기성 조건 하에서 NADH 농도 및 이용가능성은 유의하게 증가하여 NADH를 이용하는 변이형 KARI 효소에 유리하게 되었다. 동력학적 분석을 바탕으로 하면, AB1G9 ("K9G9") 및 AB1D3 ("K9D3") 돌연변이체는 NADPH에 대한 K_m 이 상대적으로 높으며 (23 및 9.2 μ M), 이외에도 이들은 NADH

에 대한 K_M 이 상대적으로 낮다 (47 및 38 μM). 비교로서, 495B5 및 AB1D1의 K_M 은 NADPH에 대하여 각각 2.5 및 1.1 μM 이며, 야생형 K9의 K_M 은 0.10 μM 이다. AB1G9 및 AB1D3의 낮은 NADH K_M 은, AB1G9 및 AB1D3의 높은 NADPH K_M 과 함께, 혐기성 조건 하에서 NADPH 이용성의 감소 및 상대적으로 높은 NADH 이용성에 이르게 되었을 수 있다. 증거로서, AB1G9 및 AB1D3은 더 낮은 글리세롤 축적성 (아이소부탄올:글리세롤 = 3.3)을 가지며, 이는 495B5 및 AB1D1 (2-3)과 비교하여 그러하다. 아이소부탄올:글리세롤의 비는 야생형 K9의 경우 호기성 조건에서 혐기성 조건으로 스위칭되는 동일 조건 하에서 1:1이다.

[0552] [표 18]

야생형 및 변이형 K9 KARI 효소의 동력학적 특성과, 세럼 바이알에서 호기성에서 혐기성으로의 스위치 실험으로부터 측정된 아이소부탄올 역가 및 생산성

클론 명	K_M (NADPH)	NADPH V_{max} (U/mg)	K_M (NADH)	NADH V_{max} (U/mg)	아이소부탄올 (g/L)	유효 아이소부탄올 생산성 (세포 1 g당 g)
WT K9 KARI	0.19	2.0	326	1.5	0.9	3.1
AB1G9	23	2.4	47	2.0	3.4	10.5
495B5	3.5	2.5	44	1.9	2.2	9.2
AB1D3	9.2	2.1	38	1.9	3.3	10.8
AB1D1	1.1	2.8	164	2.1	2.3	9.5

[0553]

[0554] 실시예 8

[0555] 사카로마이세스 세레비지에 주 BP1135 (PNY1505) 및 PNY1507과 아이소부탄올-생성 유도체의 제작

[0556] 이 실시예는 사카로마이세스 세레비지에 주 BP1135 및 PNY1507의 제작을 설명한다. 이들 주는 PNY1503 (BP1064)으로부터 유도하였다. PNY1503은 CEN.PK 113-7D (CBS 8340; CBS (Centraalbureau voor Schimmelcultures) 진균류 생물다양성 센터, 네덜란드)로부터 유도하였다. BP1135는 *FRA2* 유전자의 추가의 결실을 함유한다. PNY1507은 사카로마이세스 세레비지에에서의 발현을 위하여 코돈 최적화된, 락토코커스 락티스 유래의 *kivD* 유전자의 *ADH1* 유전자와 내로의 통합에 의해, *ADH1* 유전자를 추가로 결실시켜 BP1135로부터 유도하였다.

[0557] 결실/통합은 형질전환체의 선발을 위한 *URA3* 유전자 및 표적 유전자의 상류 및 하류의 상동성 영역을 함유하는 PCR 단편을 이용하여 상동 재조합에 의해 생성하였다. *URA3* 유전자를 상동 재조합에 의해 제거하여 스카리스 (scarless) 결실/통합을 생성하였다.

[0558] 스카리스 결실/통합 절차를 문헌[Akada *et al.*, *Yeast*, 23:399, 2006]으로부터 채용하였다. 일반적으로, 각각의 결실/통합을 위한 PCR 카세트는, 4개의 단편, A-B-U-C 및 통합될 유전자를 상기 개별 단편들의 플라스미드 내로의 클로닝에 의해 조합한 후 전체 카세트를 결실/통합 절차를 위하여 PCR로 증폭시킴으로써 만들었다. 통합될 유전자를 카세트 내에 단편 A와 단편 B 사이에 포함시켰다. 상기 PCR 카세트는 프로모터 (*URA3* 유전자의 250 bp 상류) 및 종결 서열 (*URA3* 유전자의 150 bp 하류) 영역과 함께, 천연 CEN.PK 113-7D *URA3* 유전자로 이루어진, 선발가능/역선발가능(counter-selectable) 마커, *URA3* (단편 U)을 함유하였다. 단편 A 및 C (각각은 대략적으로 100 내지 500 bp의 길이)는 표적 영역의 바로 상류의 서열 (단편 A) 및 표적 영역의 3' 서열 (단편 C)에 상응하였다. 단편 A 및 C를 상동 재조합에 의하여 염색체 내로의 카세트의 통합을 위하여 사용하였다. 단편 B (500 bp 길이)는 표적 유전자의 바로 하류의 500 bp에 상응하였으며, *URA3* 마커 및 단편 C를 상동 재조합에 의해 염색체로부터 잘라내는 데 사용하였고, 그 이유는 단편 B에 상응하는 서열의 직접 반복이 염색체 내로의 카세트의 통합시에 생성되었기 때문이다.

[0559] *FRA2* 결실

[0560] *FRA2* 결실은 코딩 서열의 3' 말단으로부터 250개의 뉴클레오티드를 결실시켜 *FRA2* 코딩 서열의 처음 113개 뉴

클레오티드를 온전하게 남기도록 설계하였다. 인-프레임(in-frame) 종결 코돈은 결실의 7개 뉴클레오티드 하류에 존재하였다. 스카리스 *FRA2* 결실을 위한 PCR 카세트에 있어서의 4개의 단편을 퓨전 하이 퍼델리티(Phusion High Fidelity) PCR 마스터 믹스(Master Mix) (미국 매사추세츠주 입스위치 소재의 뉴 잉글랜드 바이오랩스(New England BioLabs))를 사용하여 그리고 젠트라 퓨어진 이스트/백트 (Gentra Puregene Yeast/Bact) 키트 (미국 캘리포니아주 발렌시아 소재의 퀴아젠(Qiagen))로 제조된 주형으로서의 CEN.PK 113-7D 게놈 DNA를 사용하여 증폭시켰다. *FRA2* 단편 A를, 프라이머 oBP594 (서열 번호 183) 및 프라이머 oBP595 (서열 번호 184) - *FRA2* 단편 B의 5' 말단에 대하여 상동성을 갖는 5' 테일을 함유함 - 를 사용하여 증폭시켰다. *FRA2* 단편 B를, *FRA2* 단편 A의 3' 말단에 대하여 상동성을 갖는 5' 테일을 함유하는 프라이머 oBP596 (서열 번호 185), 및 *FRA2* 단편 U의 5' 말단에 대하여 상동성을 갖는 5' 테일을 함유하는 프라이머 oBP597 (서열 번호 186)을 사용하여 증폭시켰다. *FRA2* 단편 U를, *FRA2* 단편 B의 3' 말단에 대하여 상동성을 갖는 5' 테일을 함유하는 프라이머 oBP598 (서열 번호 187), 및 *FRA2* 단편 C의 5' 말단에 대하여 상동성을 갖는 5' 테일을 함유하는 프라이머 oBP599 (서열 번호 188)를 사용하여 증폭시켰다. *FRA2* 단편 C를, *FRA2* 단편 U의 3' 말단에 대하여 상동성을 갖는 5' 테일을 함유하는 프라이머 oBP600 (서열 번호 189), 및 프라이머 oBP601 (서열 번호 190)을 사용하여 증폭시켰다. PCR 생성물을 PCR 정제 키트 (퀴아젠)를 사용하여 정제하였다. *FRA2* 단편 AB는 *FRA2* 단편 A 및 *FRA2* 단편 B를 혼합하고 프라이머 oBP594 (서열 번호 183) 및 oBP597 (서열 번호 186)을 사용하여 증폭시킴으로써 중첩 PCR에 의해 생성하였다. *FRA2* 단편 UC는 *FRA2* 단편 U 및 *FRA2* 단편 C를 혼합하고 프라이머 oBP598 (서열 번호 187) 및 oBP601 (서열 번호 190)을 사용하여 증폭시킴으로써 중첩 PCR에 의해 생성하였다. 생성된 PCR 생성물을 아가로스 겔에 이어서 겔 추출 키트(Gel Extraction kit) (퀴아젠)에서 정제하였다. *FRA2* ABUC 카세트는 *FRA2* 단편 AB 및 *FRA2* 단편 UC를 혼합하고 프라이머 oBP594 (서열 번호 183) 및 oBP601 (서열 번호 190)을 사용하여 증폭시킴으로써 중첩 PCR에 의해 생성하였다. PCR 생성물을 PCR 정제 키트 (퀴아젠)로 정제하였다.

[0561] PNY1503의 적격성 세포를 만들고, 프로즌-이지 이스트 트랜스포메이션(Frozen-EZ Yeast Transformation) II 키트 (자이모 리서치; 미국 캘리포니아주 오렌지 소재)를 사용하여 *FRA2* ABUC PCR 카세트에 형질전환시켰다. 형질전환 혼합물을 30℃에서 1% 에탄올을 보충한 우라실이 결여된 합성 완전 배지 상에 플레이팅하였다. *fra2* Nick 아웃이 있는 형질전환체를 젠트라 퓨어진 이스트/백트 키트 (퀴아젠)로 제조된 게놈 DNA를 사용하여, 프라이머 oBP602 (서열 번호 191) 및 oBP603 (서열 번호 192)을 사용한 PCR에 의해 스크리닝하였다. 올바른 형질전환체를 YPE (효모 추출물, 펩톤, 1% 에탄올)에서 성장시키고, 5-플루오로-오로트산 (0.1%)을 함유하는 합성 완전 배지 상에서 30℃에서 플레이팅하여, URA3 마커가 상실된 단리체를 선발하였다. 결실 및 마커 제거는 젠트라 퓨어진 이스트/백트 키트 (퀴아젠)로 제조된 게놈 DNA를 사용하여, 프라이머 oBP602 (서열 번호 191) 및 oBP603 (서열 번호 192)을 사용한 PCR에 의해 확인하였다. 상기 단리체 유래의 *FRA2* 유전자의 부재는 *FRA2*의 결실된 코딩 서열에 특이적인 프라이머, oBP606 (서열 번호 194) 및 oBP605 (서열 번호 193)를 사용한 음성 PCR 결과에 의해 입증하였다. 올바른 단리체를 주 CEN.PK 113-7D MATa *ura3Δ::loxP his3Δ pdc6Δ pdc1Δ::P[PDC1]-DHAD|ilvD_Sm-PDC1t pdc5Δ::P[PDC5]-ADH|sadB_Ax-PDC5t gpd2Δ::loxP fra2Δ*로서 선발하고, PNY1505 (BP1135)로 표기하였다.

[0562] 이 주를 아이소부탄올 경로 플라스미드 (pYZ090, 서열 번호 195) 및 pLH468 (서열 번호 139)로 형질전환시키고, 하나의 클론을 BP1168 (PNY1506)로 표기하였다.

[0563] pYZ090 (서열 번호 195)을 ALS의 발현을 위한, 효모 CUP1 프로모터 (뉴클레오티드 2-449)로부터 발현되며 CYC1 종결 서열 (뉴클레오티드 2181-2430)이 뒤따르는 바실러스 서브틸리스 유래의 *alsS* 유전자의 코딩 영역 (뉴클레오티드 위치 457-2172)을 갖는 키메라 유전자, 및 KARI의 발현을 위한, 효모 ILV5 프로모터 (2433-3626)로부터 발현되며 ILV5 종결 서열 (뉴클레오티드 4682-5304)이 뒤따르는 락토코커스 락티스 유래의 *ilvC* 유전자의 코딩 영역 (뉴클레오티드 3634-4656)을 갖는 키메라 유전자를 함유하도록 제작하였다.

[0564] *ADH1* 결실 및 *kivD* L1(y) 통합

[0565] *ADH1* 유전자를 결실시키고, 사카로마이세스 세레비지에에서의 발현용으로 코돈 최적화된 락토코커스 락티스 유래의 *kivD* 코딩 영역으로 대체하였다. *ADH1* 결실-*kivD* L1(y) 통합을 위한 스카리스 카세트를 먼저 플라스미드 pUC19-URA3MCS 내로 클로닝하였으며, 이는 본 명세서에 참고로 포함된, 2010년 6월 18일자로 출원된 미국 특허 출원 제61/356379호에 설명된 바와 같았다. 상기 벡터는 pUC19를 기반으로 하며, 다중 클로닝 부위 (MCS) 내에 위치하는, 사카로마이세스 세레비지에 CEN.PK 113-7D 유래의 *URA3* 유전자의 서열을 함유한다. pUC19는 pMB1 레플리콘(replicon) 및 에스케리키아 콜라이에서의 복제 및 선발을 위한 베타-락타마아제를 코딩하는 유전자를 함유한다. *URA3*의 코딩 서열에 더하여, 이 유전자의 상류 (250 bp) 및 하류 (150 bp)의 서열이 효모에서의 *URA3* 유전자의 발현을 위하여 존재한다. 벡터는 클로닝 목적을 위해 사용될 수 있으며, 효모 통합 벡터로서 사용될

수 있다.

[0566] 사카로마이세스 세레비지에에서의 발현용으로 코돈 최적화된 락토코커스 락티스 유래의 *kivD* 코딩 영역은 pLH468 (서열 번호 139)을 주형으로 사용하여, *PmeI* 제한효소 부위를 함유하는 프라이머 oBP562 (서열 번호 197), 및 *ADH1* 단편 B의 5' 말단에 대하여 상동성을 갖는 5' 테일을 함유하는 프라이머 oBP563 (서열 번호 198)을 사용하여 증폭시켰다. *ADH1* 단편 B는 *kivD_L1(y)*의 3' 말단에 대하여 상동성을 갖는 5' 테일을 함유하는 프라이머 oBP564 (서열 번호 199), 및 *FseI* 제한효소 부위를 함유하는 프라이머 oBP565 (서열 번호 200)를 사용하여, 상기와 같이 제조한 게놈 DNA로부터 증폭시켰다. PCR 생성물을 PCR 정제 키트 (퀴아젠)를 사용하여 정제하였다. *kivD_L1(y)*-*ADH1* 단편 B는 *kivD_L1(y)* 및 *ADH1* 단편 B PCR 생성물을 혼합하고 프라이머 oBP562 (서열 번호 197) 및 oBP565 (서열 번호 200)를 사용하여 증폭시킴으로써 중첩 PCR에 의해 생성하였다. 생성된 PCR 생성물을 *PmeI* 및 *FseI*으로 절단하고, T4 DNA 라이가아제를 사용하여 적절한 효소를 이용한 절단 후에 pUC19-URA3MCS의 상응하는 부위 내로 라이게이션시켰다. *ADH1* 단편 A는 *SacI* 제한효소 부위를 함유하는 프라이머 oBP505 (서열 번호 201), 및 *AscI* 제한효소 부위를 함유하는 프라이머 oBP506 (서열 번호 202)을 사용하여 게놈 DNA로부터 증폭시켰다. *ADH1* 단편 A PCR 생성물을 *SacI* 및 *AscI*로 절단하고, T4 DNA 라이가아제를 사용하여 *kivD_L1(y)*-*ADH1* 단편 B를 함유하는 플라스미드의 상응하는 부위 내로 라이게이션시켰다. *ADH1* 단편 C는 *PacI* 제한효소 부위를 함유하는 프라이머 oBP507 (서열 번호 203), 및 *SalI* 제한효소 부위를 함유하는 프라이머 oBP508 (서열 번호 204)을 사용하여 게놈 DNA로부터 증폭시켰다. *ADH1* 단편 C PCR 생성물을 *PacI* 및 *SalI*로 절단하고, T4 DNA 라이가아제를 사용하여 *ADH1* 단편 A-*kivD_L1(y)*-*ADH1* 단편 B를 함유하는 플라스미드의 상응하는 부위 내로 라이게이션시켰다. 하이브리드 프로모터 UAS(PGK1)-P_{FBA1}은 *AscI* 제한효소 부위를 함유하는 프라이머 oBP674 (서열 번호 205), 및 *PmeI* 제한효소 부위를 함유하는 프라이머 oBP675 (서열 번호 206)를 사용하여 벡터 pRS316-UAS(PGK1)-P_{FBA1}-GUS (서열 번호 209)로부터 증폭시켰다. UAS(PGK1)-P_{FBA1} PCR 생성물을 *AscI* 및 *PmeI*로 절단하고, T4 DNA 라이가아제를 사용하여 *kivD_L1(y)*-*ADH1* 단편 ABC를 함유하는 플라스미드의 상응하는 부위 내로 라이게이션시켰다. 상기 전체 통합 카세트를 생성된 플라스미드로부터, 프라이머 oBP505 (서열 번호 201) 및 oBP508 (서열 번호 204)을 사용하여 증폭시키고, PCR 정제 키트 (퀴아젠)를 사용하여 정제하였다.

[0567] PNY1505의 적격성 세포를 만들고, 프로즌-이지 이스트 트랜스포메이션 II 키트 (자이모 리서치)를 사용하여, 상기에서 제작한 *ADH1*-*kivD_L1(y)* PCR 카세트에 형질전환시켰다. 형질전환 혼합물을 30℃에서 1% 에탄올을 보충한 우라실이 결여된 합성 완전 배지 상에 플레이팅하였다. 형질전환체를 YPE (1% 에탄올)에서 성장시키고, 5-플루오로-오로트산 (0.1%)을 함유하는 합성 완전 배지 상에서 30℃에서 플레이팅하여, URA3 마커가 상실된 단리체를 선별하였다. *ADH1*의 결실 및 *kivD_L1(y)*의 통합은 젠트라 퓨어인 이스트/백트 키트 (퀴아젠)를 이용하여 제조한 게놈 DNA를 사용하여, 외재적 프라이머 oBP495 (서열 번호 207) 및 oBP496 (서열 번호 208)을 사용하여 그리고 *kivD_L1(y)* 특이적 프라이머 oBP562 (서열 번호 197) 및 외재적 프라이머 oBP496 (서열 번호 208)을 사용하여 PCR에 의해 확인하였다. 올바른 단리체를 주 CEN.PK 113-7D MATa *ura3Δ::loxP* *his3Δ* *pdv6Δ* *pdv1Δ::P[PDC1]-DHAD|ilvD_Sm-PDC1tpdc5Δ::P[PDC5]-ADH|sadB_Ax-PDC5t* *gpd2Δ::loxP* *fra2Δ* *adh1Δ::UAS(PGK1)P[FBA1]-kivD_L1(y)-ADH1*로서 선별하고, PNY1507 (BP1201)로 표기하였다. PNY1507을 아이소부탄올 경로 플라스미드 pYZ090 (서열 번호 195) 및 pBP915 (서열 번호 182)로 형질전환시키고, 생성된 주를 PNY1513으로 명명하였다.

[0568] pRS316-UAS(PGK1)-FBA1p-GUS 벡터의 제작

[0569] 카세트 UAS(PGK1)-FBA1p (서열 번호 766)를 클로닝하기 위하여, 먼저 602 bp FBA1 프로모터 (FBA1p)를 프라이머 T-FBA1(*SalI*) (서열 번호 767) 및 B-FBA1(*SpeI*) (서열 번호 768)을 사용하여 CEN.PK의 게놈 DNA로부터 PCR 증폭시키고, PGK1p 프로모터를 플라스미드 pWS358-PGK1p-GUS (서열 번호 769)의 *SalI*/*SpeI* 절단에 의해 제거한 후 상기 플라스미드 상의 *SalI* 및 *SpeI* 부위 내로 클로닝하여 pWS358-FBA1p-GUS를 생성하였다. pWS358-PGK1p-GUS 플라스미드는 PGK1p 및 베타-글루쿠로니다아제 유전자 (GUS) DNA 단편을, pRS423 벡터 (문헌[Christianson et al., Gene, 110:119-122, 1992])로부터 유래된 pWS358의 다중 클로닝 부위 내로 삽입함으로써 생성하였다. 둘째, 생성된 pWS358-FBA1p-GUS 플라스미드를 *SalI* 및 *SacI*로 절단하고, FBA1p 프로모터, GUS 유전자 및 FBAt 종결 서열을 함유하는 DNA 단편을 겔 정제하고, pRS316 상의 *SalI*/*SacI* 부위 내로 클로닝하여 pRS316-FBA1p-GUS를 생성하였다. 셋째, 3-포스포글리세레이트 키나아제 (PGK1) 오픈 리딩 프레임(open reading frame)의 상류의 위치 -519와 위치 -402 사이에 위치하는 상류 활성화 서열(upstream activation sequence; UAS), 즉 UAS(PGK1)를 함유하는 118 bp DNA 단편을 프라이머 T-U/PKG1(*KpnI*) (서열 번호 770) 및 B-U/PKG1(*SalI*) (서열 번호 771)을 사용하여 CEN.PK의 게놈 DNA로부터 PCR 증폭시켰다. 상기 PCR 생성물을 *KpnI* 및 *SalI*로 절단하고,

pRS316-FBA1p-GUS 상의 KpnI/SalI 부위 내로 클로닝하여 pRS316-UAS(PGK1)-FBA1p-GUS를 생성하였다.

[0570] 실시예 9

[0571] ALD6의 제거를 포함하는 개선된 재조합 숙주 세포

[0572] 이 실시예의 목적은 개선된 아이소부탄을 생성을 위한 효모 숙주 주의 변형 방법을 설명하는 것이다. 이들 변형은 아이소부틸알데히드 리덕타아제 활성을 코딩하는 유전자의 통합과, 각각 NADP⁺-의존성 아세트알데히드 데하이드로게나아제 및 NADPH-의존성 데하이드로게나아제를 코딩하는 천연 유전자 ALD6 및 YMR226c의 제거를 포함한다.

[0573] 에스. 세레비지에 주 PNY2211의 제작

[0574] PNY2211은 하기 단락에서 설명하는 바와 같이 에스. 세레비지에 주 PNY1507 (실시예 8)로부터 몇몇 단계로 제작하였다. 먼저, PNY1507을 포스포케톨라아제 유전자를 함유하도록 변형시켰다. 다음, 포스포케톨라아제 유전자에 인접한 서열에 표적화되는 통합 벡터를 사용하여 아세트락테이트 신타아제 유전자 (*alsS*)를 상기 주에 추가하였다. 마지막으로, 상동 재조합을 이용하여 포스포케톨라아제 유전자 및 통합 벡터 서열을 제거하여, *pdc1* Δ::*ilvD* (실시예 12에 설명함)와 염색체 XII의 천연 TRX1 유전자 사이의 유전자간 영역(intergenic region) 내의 *alsS*의 스카리스 삽입을 생성한다. PNY2211의 생성된 유전자형은 MATa *ura3* Δ::loxP *his3* Δ *pdc6* Δ *pdc1* Δ::P[PDC1]-DHAD|*ilvD*_Sm-PDC1t-P[FBA1]-ALS|*alsS*_Bs-CYC1t *pdc5* Δ::P[PDC5]-ADH| *sadB*_Ax-PDC5t *gpd2* Δ::loxP *fra2* Δ *adh1* Δ::UAS(PGK1)P[FBA1]-*kivD*_L1(y)-ADH1t이다.

[0575] 포스포케톨라아제 유전자 카세트를 상동 재조합에 의해 PNY1507 내로 도입하였다. 상기 통합 제작물을 하기와 같이 생성하였다. 플라스미드 pRS423::CUP1-*alsS*+FBA-*budA* (전체적으로 본 명세서에 참고로 포함된 미국 특허 공개 제2009/0305363호에 이전에 공개됨)를 *NotI* 및 *XmaI*로 절단하여 1.8 kb FBA-*budA* 서열을 제거하고, 상기 벡터를 클레나우(Klenow) 단편으로 처리한 후 재라이게이션시켰다. 다음, CUP1 프로모터는, DNA2.0 (미국 캘리포니아주 멘로 파크 소재)으로부터의 DNA 합성 및 벡터 제작 서비스를 통한 TEF1 프로모터 변이체 (전체적으로 본 명세서에 참고로 포함된 문헌[Nevoigt *et al. Appl. Environ. Microbiol.* 72: 5266-5273 (2006)]에 이전에 기재된 M4 변이체)로 대체하였다. 생성된 플라스미드, pRS423::TEF(M4)-*alsS*를 *StuI* 및 *MluI*로 절단하고 (이는 CYC1 종결 서열 및 *alsS* 유전자의 일부를 함유하는 1.6 kb 부분을 제거함), 프라이머 N1176 (서열 번호 282) 및 N1177 (서열 번호 283)을 사용하여 pRS426::GPD-*xpk1*+ADH-*eutD* (서열 번호 383)로부터 생성한 4 kb PCR 생성물, 및 프라이머 N822 (서열 번호 285) 및 N1178 (서열 번호 286)을 사용하여 효모 게놈 DNA (ENO1 프로모터 영역)로부터 생성한 0.8 kb PCR 생성물 DNA (서열 번호 284)와 조합하고, 에스. 세레비지에 주 BY4741 (ATCC #201388)을 형질전환시켰으며; 겐 수복 클로닝 방법은 문헌[Ma *et al. Gene* 58:201-216 (1987)]을 참조한다. 히스티딘을 포함하지 않는 합성 완전 배지 상에 세포를 플레이트함으로써 형질전환체를 수득하였다. 예상되는 플라스미드 (pRS423::TEF(M4)-*xpk1*+ENO1-*eutD*, 서열 번호 293)의 적당한 조립을 PCR에 의해 (프라이머 N821 (서열 번호 287) 및 N1115 (서열 번호 288)) 그리고 제한효소 절단에 의해 (BglI) 확인하였다. 후속적으로 2개의 클론을 서열결정하였다. 3.1 kb TEF(M4)-*xpk1* 유전자를 *SacI* 및 *NotI*을 이용한 절단에 의해 단리하고, pUC19-URA3::*ilvD*-TRX1 벡터 내로 클로닝하였다 (클론 A, *Af*/II로 절단). 클로닝 단편들을 클레나우 단편으로 처리하여 라이게이션을 위한 블런트(blunt) 말단을 생성하였다. 라이게이션 반응물들로 이. 콜라이 Stb13 세포를 형질전환시켜 암피실린 내성에 대하여 선발하였다. TEF(M4)-*xpk1*의 삽입을 PCR에 의해 (프라이머 N1110 (서열 번호 367) 및 N1114 (서열 번호 290)) 확인하였다. 상기 벡터를 *Af*/II로 선형화하고, 클레나우 단편으로 처리하였다. 1.8 kb KpnI-HincII 제네티신 내성 카세트 (서열 번호 384)는 클레나우 단편 처리 후 라이게이션에 의해 클로닝하였다. 라이게이션 반응물들로 이. 콜라이 Stb13 세포를 형질전환시켜 암피실린 내성에 대하여 선발하였다. 상기 제네티신 카세트의 삽입을 PCR에 의해 (프라이머 N160SeqF5 (서열 번호 210) 및 BK468 (서열 번호 368)) 확인하였다. 상기 플라스미드의 서열을 서열 번호 291 (pUC19-URA3::*pdc1*::TEF(M4)-*xpk1*::kan)로 제공한다.

[0576] 생성된 통합 카세트 (*pdc1*::TEF(M4)-*xpk1*::KanMX::TRX1)를 단리하고 (*AscI* 및 *NaeI* 절단에 의해 5.3 kb 밴드를 생성하고, 이를 겔 정제함), 자이모 리서치 프로즌-이지 이스트 트랜스포메이션 키트 (카탈로그 번호 T2001)를 사용하여 PNY1507을 형질전환시켰다. 형질전환체를 YPE + 50 μg/ml의 G418 상에 플레이트함으로써 선발하였다. 예상된 유전자좌에서의 통합을 PCR에 의해 (프라이머 N886 (서열 번호 211) 및 N1214 (서열 번호 281)) 확인하였다. 다음, Cre 리코비나아제를 코딩하는 플라스미드 pRS423::GAL1p-Cre (서열 번호 271)를 사용하여 loxP가 측면에 위치하는 KanMX 카세트를 제거하였다. 상기 카세트의 적당한 제거를 PCR에 의해 (프라이머 oBP512 (서열 번호 337) 및 N160SeqF5 (서열 번호 210)) 확인하였다. 마지막으로, 실시예 13에 설명되어 있는

alsS 통합 플라스미드 (pUC19-kan::pdc1::FBA-*alsS*::TRX1, 클론 A)로 이 주를 형질전환시켰으며, 이는 포함된 제네티신 선발 마커를 이용한 것이었다. 두 통합체는 플라스미드 pYZ090Δ*alsS* (서열 번호 371) 및 pBP915 (서열 번호 182)를 이용한 형질전환 (문헌[Amberg, Burke and Strathern "Methods in Yeast Genetics" (2005)]의 프로토콜 #2를 이용하여 형질전환시킴), 및 글루코스 함유 배지에서 성장 및 아이소부탄올 생성의 평가에 의해 아세트락테이트 신타아제 활성에 대하여 시험하였다 (성장 및 아이소부탄올 측정 방법은 하기와 같다: 모든 주는 탄소 공급원으로서 0.3% 글루코스 및 0.3% 에탄올을 함유하는, 히스티딘 및 우라실을 제외한 합성 완전 배지 (125 mL의 배기 에rlenmeyer(Erlenmeyer) 플라스크 (브이더블유알(VWR) 카탈로그 번호 89095-260) 내의 10 mL의 배지)에서 성장시켰다. 하룻밤 인큐베이션 (이노바(Innova)(등록상표)40 뉴 브룬스위크 사이언티픽(New Brunswick Scientific) 진탕기에서 30°C, 250 rpm) 후, 배양물들을 2% 글루코스 및 0.05% 에탄올을 함유하는 합성 완전 배지 (125 mL의 타이트하게 뚜껑을 덮은 에rlenmeyer 플라스크 (브이더블유알 카탈로그 번호 89095-260) 내의 20 mL의 배지)에서 0.2의 OD (에펜도르프 바이오포토미터(Eppendorf BioPhotometer) 측정)로 다시 희석시켰다. 48시간의 인큐베이션 (이노바(등록상표)40 뉴 브룬스위크 사이언티픽 진탕기에서 30°C, 250 rpm) 후, 배양 상청액들 (스핀-엑스(Spin-X) 원심분리 튜브 필터 유닛 (코스타(Costar) 카탈로그 번호 8169)을 사용하여 수집함)을 전체적으로 본 명세서에 참고로 포함된 미국 특허 공개 제2007/0092957호에 기재된 방법에 따라 HPLC로 분석하였다. 두 클론 중 하나는 양성이었으며, 이를 PNY2218로 명명하였다.

[0577] PNY2218을 Cre 리코미니아제로 처리하고, 생성된 클론들을 PCR에 의해 (프라이머 N886 (서열 번호 211) 및 N160SeqR5 (서열 번호 388)) pUC19 통합 벡터 서열 및 *xpk1* 유전자의 상실에 대하여 스크리닝하였다. 이는, 통합 벡터의 삽입 (벡터, 마커 유전자 및 loxP 서열이 상실되기 때문에 "스카리스" 삽입) 동안 도입된 상동 DNA 및 *xpk1*의 상류의 DNA의 재조합 후 pdc1-TRX1 유전자간 영역 내에 통합된 *alsS* 유전자만을 남겼다. 이 재조합은 임의의 지점에서 일어날 수 있었지만, 벡터 통합은 심지어 제네티신 선발 없이도 안정한 것으로 나타났으며, 재조합 사건은 단지 Cre 리코미니아제의 도입 후에 관찰되었다. 하나의 클론을 PNY2211로 표기하였다.

[0578] 플라스미드 pYZ090Δ*alsS* 및 pBP915를 함유하는 PNY2218의 단리체를 PNY2209로 표기하였다.

[0579] PNY1528 (PNY2211 내의 hADH의 통합)

[0580] 결실/통합은 형질전환체의 선발을 위한 *URA3* 유전자 및 표적 영역의 상류 및 하류의 상동성 영역을 함유하는 PCR 생성물을 이용하여 상동 재조합에 의해 생성하였다. *URA3* 유전자를 상동 재조합에 의해 제거하여 스카리스 결실/통합을 생성하였다.

[0581] YPRCΔ15 결실 및 말 간 *adh* 통합

[0582] YPRCΔ15 유전자좌를 결실시키고, 사카로마이세스 세레비지에 유래의 PDC5 프로모터 영역 (538 bp) 및 사카로마이세스 세레비지에 유래의 ADH1 종결 영역 (316 bp)과 함께, 사카로마이세스 세레비지에에서의 발현용으로 코돈 최적화된 말 간 *adh* 유전자로 대체하였다. YPRCΔ15 결실- P[PDC5]-*adh*_{HL}(y)-ADH1t 통합을 위한 스카리스 카세트들 먼저 플라스미드 pUC19-URA3MCS (실시예 8에서 설명함) 내로 클로닝하였다.

[0583] 단편 A-B-U-C를 퓨전 하이 피델리티 PCR 마스터 믹스 (미국 매사추세츠주 입스위치 소재의 뉴 잉글랜드 바이오랩스)를 사용하고, 젠트라 퓨어진 이스트/백트 키트 (미국 캘리포니아주 발렌시아 소재의 퀴아젠)로 제조된 주형으로서의 CEN.PK 113-7D 게놈 DNA를 사용하여 증폭시켰다. YPRCΔ15 단편 A를, KpnI 제한효소 부위를 함유하는 프라이머 oBP622 (서열 번호 212), 및 YPRCΔ15 단편 B의 5' 말단에 대하여 상동성을 갖는 5' 테일을 함유하는 프라이머 oBP623 (서열 번호 213)을 사용하여 게놈 DNA로부터 증폭시켰다. YPRCΔ15 단편 B를, YPRCΔ15 단편 A의 3' 말단에 대하여 상동성을 갖는 5' 테일을 함유하는 프라이머 oBP624 (서열 번호 214), 및 FseI 제한효소 부위를 함유하는 프라이머 oBP625 (서열 번호 215)를 사용하여 게놈 DNA로부터 증폭시켰다. PCR 생성물을 PCR 정제 키트 (퀴아젠)를 사용하여 정제하였다. YPRCΔ15 단편 A - YPRCΔ15 단편 B는 YPRCΔ15 단편 A 및 YPRCΔ15 단편 B의 PCR 생성물을 혼합하고, 프라이머 oBP622 (서열 번호 212) 및 oBP625 (서열 번호 215)를 사용하여 증폭시킴으로써 중첩 PCR에 의해 생성하였다. 생성된 PCR 생성물을 KpnI 및 FseI로 절단하고, T4 DNA 라이가아제를 사용하여 적절한 효소를 이용한 절단 후에 pUC19-URA3MCS의 상응하는 부위 내로 라이게이션시켰다. YPRCΔ15 단편 C를, NotI 제한효소 부위를 함유하는 프라이머 oBP626 (서열 번호 216), 및 PacI 제한효소 부위를 함유하는 프라이머 oBP627 (서열 번호 217)을 사용하여 게놈 DNA로부터 증폭시켰다. YPRCΔ15 단편 C PCR 생성물을 NotI 및 PacI로 절단하고, T4 DNA 라이가아제를 사용하여 YPRCΔ15 단편 AB를 함유하는 플라스미드의 상응하는 부위 내로 라이게이션시켰다. PDC5 프로모터 영역을, AscI 제한효소 부위를 함유하는 프라이머 HY21 (서열 번호 218), 및 *adh*_{HL}(y)의 5' 말단에 대하여 상동성을 갖는 5' 테일을 함유하는 프라이머 HY24 (서열 번호 219)를 사용하여 CEN.PK 113-7D 게놈 DNA로부터 증폭시켰다. *adh*_{HL}(y)-ADH1t를

pBP915 (서열 번호 182)로부터 증폭시켰으며, 이는 P[PDC5]의 3' 말단에 대하여 상동성을 갖는 5' 테일을 함유하는 프라이머 HY25 (서열 번호 220), 및 PmeI 제한효소 부위를 함유하는 HY4 (서열 번호 221)를 이용하였다. PCR 생성물을 PCR 정제 키트 (퀴아젠)를 사용하여 정제하였다. P[PDC5]-adh_HL(y)-ADH1t는 P[PDC5] 및 adh_HL(y)-ADH1t의 PCR 생성물을 혼합하고, 프라이머 HY21 (서열 번호 218) 및 HY4 (서열 번호 221)를 사용하여 증폭시킴으로써 중첩 PCR에 의해 생성하였다. 생성된 PCR 생성물을 AscI 및 PmeI로 절단하고, T4 DNA 라이가아제를 사용하여 YPRCΔ15 단편 ABC를 함유하는 플라스미드의 상응하는 부위 내로 라이게이션시켰다. 전체 통합 카세트를 생성된 플라스미드로부터, 프라이머 oBP622 (서열 번호 212) 및 oBP627 (서열 번호 217)을 사용하여 증폭시켰다.

[0584] PNY2211의 적격성 세포를 만들고, 프로즌-이지 이스트 트랜스포메이션 II 키트 (미국 캘리포니아주 오렌지 소재의 자이모 리서치)를 사용하여 YPRCΔ15 결실- P[PDC5]-adh_HL(y)-ADH1t 통합 카세트 PCR 생성물로 형질전환시켰다. 형질전환 혼합물을 30℃에서 1% 에탄올을 보충한 우라실이 결여된 합성 완전 배지 상에 플레이팅하였다. 형질전환체를 프라이머 URA3-end F (서열 번호 222) 및 oBP637 (서열 번호 224)을 사용하여 PCR에 의해 스크리닝하였다. 올바른 형질전환체를 YPE (1% 에탄올)에서 성장시키고, 1% EtOH가 보충되고 5-플루오로-오로트산 (0.1%)을 함유하는 합성 완전 배지 상에서 30℃에서 플레이팅하여, URA3 마커가 상실된 단리주를 선발하였다. YPRCΔ15의 결실 및 P[PDC5]-adh_HL(y)-ADH1t의 통합을 외재적 프라이머 oBP636 (서열 번호 223) 및 oBP637 (서열 번호 224)을 사용하여, 이스타 게노믹(YeaStar Genomic) DNA 키트 (자이모 리서치)를 이용하여 제조한 게놈 DNA를 사용하여 PCR에 의해 확인하였다. 하기 유전자형의 올바른 단리체를 추가의 변형용으로 선발하였다: CEN.PK 113-7D MATa ura3Δ::loxP his3Δ pdc6Δ pdc1Δ::P[PDC1]-DHAD|ilvD_Sm-PDC1t-P[FBA1]-ALS|alsS_Bs-CYC1t pdc5Δ::P[PDC5]-ADH|sadB_Ax-PDC5t gpd2Δ::loxP fra2Δ adh1Δ::UAS(PGK1)P[FBA1]-kivD_L1(y)-ADH1t yprcΔ15Δ::P[PDC5]-ADH|adh_H1-ADH1t.

[0585] fra2Δ에서의 말 간 adh 통합

[0586] 사카로마이세스 세레비지에 유래의 PDC1 프로모터 영역 (870 bp) 및 사카로마이세스 세레비지에 유래의 ADH1 종결 영역 (316 bp)과 함께, 사카로마이세스 세레비지에에서의 발현용으로 코돈 최적화된 말 간 adh 유전자를 fra2 결실 부위 내로 통합시켰다. fra2Δ - P[PDC1]-adh_HL(y)-ADH1t 통합을 위한 스카리스 카세트를 먼저 플라스미드 pUC19-URA3MCS 내로 클로닝하였다.

[0587] 단편 A-B-U-C를 퓨전 하이 피델리티 PCR 마스터 믹스 (미국 매사추세츠주 입스위치 소재의 뉴 잉글랜드 바이오랩스)를 사용하여 그리고 젠트라 퓨어진 이스트/백트 키트 (미국 캘리포니아주 발렌시아 소재의 퀴아젠)로 제조된 주형으로서의 CEN.PK 113-7D 게놈 DNA를 사용하여 증폭시켰다. fra2Δ 단편 C를, NotI 제한효소 부위를 함유하는 프라이머 oBP695 (서열 번호 229), 및 PacI 제한효소 부위를 함유하는 프라이머 oBP696 (서열 번호 230)를 사용하여 게놈 DNA로부터 증폭시켰다. fra2Δ 단편 C PCR 생성물을 NotI 및 PacI로 절단하고, T4 DNA 라이가아제를 사용하여 pUC19-URA3MCS의 상응하는 부위 내로 라이게이션시켰다. fra2Δ 단편 B를, PmeI 제한효소 부위를 함유하는 프라이머 oBP693 (서열 번호 227), 및 FseI 제한효소 부위를 함유하는 프라이머 oBP694 (서열 번호 228)를 사용하여 게놈 DNA로부터 증폭시켰다. 생성된 PCR 생성물을 PmeI 및 FseI으로 절단하고, T4 DNA 라이가아제를 사용하여, 적절한 효소를 이용한 절단 후에 fra2Δ 단편 C를 함유하는 플라스미드의 상응하는 부위 내로 라이게이션시켰다. fra2Δ 단편 A를, BamHI 및 AsiSI 제한효소 부위를 함유하는 프라이머 oBP691 (서열 번호 225), 및 AscI 및 SwaI 제한효소 부위를 함유하는 프라이머 oBP692 (서열 번호 226)를 사용하여 게놈 DNA로부터 증폭시켰다. fra2Δ 단편 A PCR 생성물을 BamHI 및 AscI으로 절단하고, T4 DNA 라이가아제를 사용하여, 적절한 효소를 이용한 절단 후에 fra2Δ 단편 BC를 함유하는 플라스미드의 상응하는 부위 내로 라이게이션시켰다. PDC1 프로모터 영역을, AscI 제한효소 부위를 함유하는 프라이머 HY16 (서열 번호 231), 및 adh_H1(y)의 5' 말단에 대하여 상동성을 갖는 5' 테일을 함유하는 프라이머 HY19 (서열 번호 232)를 사용하여 CEN.PK 113-7D 게놈 DNA로부터 증폭시켰다. adh_H1(y)-ADH1t을, P[PDC1]의 3' 말단에 대하여 상동성을 갖는 5' 테일을 함유하는 프라이머 HY20 (서열 번호 233), 및 PmeI 제한효소 부위를 함유하는 HY4 (서열 번호 221)를 사용하여 pBP915로부터 증폭시켰다. PCR 생성물을 PCR 정제 키트 (퀴아젠)를 사용하여 정제하였다. P[PDC1]-adh_HL(y)-ADH1t는, P[PDC1] 및 adh_HL(y)-ADH1t의 PCR 생성물을 혼합하고, 프라이머 HY16 (서열 번호 231) 및 HY4 (서열 번호 221)를 사용하여 증폭시킴으로써 중첩 PCR에 의해 생성하였다. 생성된 PCR 생성물을 AscI 및 PmeI로 절단하고, T4 DNA 라이가아제를 사용하여 fra2Δ 단편 ABC를 함유하는 플라스미드의 상응하는 부위 내로 라이게이션시켰다. 전체 통합 카세트를 생성된 플라스미드로부터, 프라이머 oBP691 (서열 번호 225) 및 oBP696 (서열 번호 230)을 사용하여 증폭시켰다.

[0588] *YPRCΔ15*에서 통합된 *adh_H1(y)*를 포함하는 PNY2211 변이체의 적격 세포를 만들고, 프로즌-이지 이스트 트랜스포메이션 II 키트 (자이모 리서치)를 사용하여 *fra2Δ*-P[PDC1]-*adh_H1(y)*-ADH1t 통합 카세트 PCR 생성물로 형질전환시켰다. 형질전환 혼합물을 30℃에서 1% 에탄올을 보충한 우라실이 결여된 합성 완전 배지 상에 플레이팅하였다. 형질전환체를 프라이머 URA3-말단 F (서열 번호 222) 및 oBP731 (서열 번호 235)을 사용하여 PCR에 의해 스크리닝하였다. 올바른 형질전환체를 YPE (1% 에탄올)에서 성장시키고, 1% EtOH가 보충되고 5-플루오로-오로트산 (0.1%)을 함유하는 합성 완전 배지 상에서 30℃에서 플레이팅하여, *URA3* 마커가 상실된 단리주를 선발하였다. P[PDC1]-*adh_H1(y)*-ADH1t의 통합을 내재적 프라이머 HY31 (서열 번호 236) 및 외재적 프라이머 oBP731 (서열 번호 235)을 사용한 콜로니 PCR, 및 외재적 프라이머 oBP730 (서열 번호 234) 및 oBP731 (서열 번호 235)을 이용한 PCR - 이스타 게놈 DNA 키트 (자이모 리서치)를 사용하여 제조한 게놈 DNA를 사용함 -에 의해 확인하였다. 하기 유전자형의 올바른 단리체를 PNY1528로 표기하였다: CEN.PK 113-7D MATa *ura3Δ::loxP his3Δ pdc6Δ pdc1Δ::P[PDC1]-DHAD1ilvD_Sm-PDC1t-P[FBA1]-ALS|alsS_Bs-CYC1t pdc5Δ::P[PDC5]-ADH|sadB_Ax-PDC5t gpd2Δ::loxP fra2Δ::P[PDC1]-ADH|adh_H1-ADH1t adh1Δ::UAS(PGK1)P[FBA1]-kivD_L1(y)-ADH1t yprcΔ15Δ::P[PDC5]-ADH|adh_H1-ADH1t*.

[0589] PNY2237 (스카리스 YMR226c 결실)

[0590] PCR 증폭된 2.0 kb 선형 스카리스 결실 카세트를 이용하여 상동 재조합에 의해 유전자 YMR226c를 에스. 세레비지에 주 PNY1528로부터 결실시켰다. 상기 카세트는 선발가능한 마커로서 그의 천연 프로모터 및 종결 서열과 함께인 *URA3* 유전자, 결실 카세트의 통합 및 천연 개재 서열의 제거를 촉진하기 위한, YMR226c 유전자의 염색체 유전자좌의 측면에 위치하는 상류 및 하류 상동 서열 및 *URA3* 마커의 재조합 및 제거를 촉진하기 위한 반복 서열로 이루어진 스플라이싱된 PCR 증폭 단편으로부터 제작하였다. 정방향 및 역방향 PCR 프라이머 (N1251 및 N1252, 각각 서열 번호 247 및 248)에 의해 1,208 bp *URA3* 발현 카세트를 증폭시켰는데, 상기 발현 카세트는 pLA33 (pUC19::loxP-*URA3*-loxP (서열 번호 268))에서 기원하였다. 정방향 및 역방향 프라이머 (N1253 및 N1254, 각각 서열 번호 249 및 250)에 의해, 에스. 세레비지에 주 PNY2211 (상기)의 게놈 DNA 체제로부터 3' *URA3* 중첩 서열 태그를 포함하는 250 bp 하류 상동 서열을 증폭시켰다. 정방향 및 역방향 PCR 프라이머 (N1255 및 N1256, 각각 서열 번호 251 및 252)에 의해, 에스. 세레비지에 주 PNY2211의 게놈 DNA 체제로부터 5' *URA3* 중첩 서열 태그를 포함하는 250 bp 반복 서열을 증폭시켰다. 정방향 및 역방향 PCR 프라이머 (N1257 및 N1258, 각각 서열 번호 253 및 254)에 의해, 에스. 세레비지에 주 PNY2211의 게놈 DNA 체제로부터 5' 반복 중첩 서열 태그를 포함하는 250 bp 상류 상동성 서열을 증폭시켰다.

[0591] 자이모 리서치 프로즌 이스트 트랜스포메이션 키트를 사용하여 적격성으로 만들어진 주 PNY1528 (상기)을 대략 1.5 μg의 PCR 증폭된 카세트로 형질전환시키고, 상기 형질전환 믹스를 SE 1.0% -우라실 상에 플레이팅하고, 통합된 *ymr226cΔ::URA3* 카세트를 포함하는 세포의 선발을 위하여 30℃에서 인큐베이션하였다. 후속적으로, 72 내지 96시간 후에 나타나는 형질전환체를 동일 배지 상에 짧게 도말하고, 30℃에서 24 내지 48시간 동안 인큐베이션한다. 짧게 도말한 것(short-streak)을 PCR에 의해 *ymr226cΔ::URA3*에 대하여 스크리닝하며, 이는 밖으로 향하는 5' *URA3* 결실 카세트-특이적 내재적 프라이머 (N1249, 서열 번호 245)가 측면에 위치하는, 안으로 향하는 염색체 특이적 프라이머 (N1239, 서열 번호 243)와 쌍을 형성한 것, 및 밖으로 향하는 3' *URA3* 결실 카세트-특이적 프라이머 (N1250, 서열 번호 246)가 측면에 위치하는, 안으로 향하는 염색체-특이적 프라이머 (N1242, 서열 번호 244)와 쌍을 형성한 것을 이용한다. 양성 PNY1528 *ymr226cΔ::URA3* PCR 스크린은 각각 598 및 726 bp의 5' 및 3' PCR 생성물로 이어졌다.

[0592] 3개의 양성 PNY1528 *ymr226cΔ::URA3* 클론을 골라내고, YPE 1% 배지에서 하룻밤 배양하였으며, 이 중 100 μl를 YPE 1% + 5-FOA 상에 플레이팅하였는데, 이는 마커 제거를 위한 것이었다. 24 내지 48시간 후에 나타나는 콜로니를 5' 및 3' 염색체-특이적 프라이머 (N1239 및 N1242)를 사용하여 마커 상실에 대하여 PCR 스크리닝하였다. 양성 PNY1528 *ymr226cΔ* 마커리스(markerless) PCR 스크린은 801 bp의 PCR 생성물로 이어졌다. 다수의 클론을 수득하고, 하나를 PNY2237로 표기하였다.

[0593] PNY2238 및 PNY2243 (ALD6 결실 주)

[0594] *ALD6* 코딩 서열이 Cre-lox 재활용성 *URA3* 선발 마커로 대체되도록 벡터를 설계하였다. *ALD6*의 5' 및 3' 서열을 PCR에 의해 증폭시켰다 (각각 프라이머 쌍 N1179 및 N1180과, N1181 및 N1182; 각각 서열 번호 237, 238, 239, 및 240). 이들 단편을 TOPO 벡터 (인비트로젠 카탈로그 번호 K2875-J10) 내로 클로닝하고 서열결정한 후 (M13 정방향 프라이머 (서열 번호 269) 및 역방향 프라이머 (서열 번호 270)), 5' 및 3' 측면부를 각각 *EcoRI* 및 *SphI* 부위에서 pLA33 (pUC19::loxP::*URA3*::loxP) (서열 번호 268) 내로 클로닝하였다. 각각의 라이게이션 반

응물로 이. 콜라이 Stb13 세포를 형질전환시키고, 이를 LB Amp 플레이트 상에서 인큐베이션하여 형질전환체를 선발하였다. 서열들의 적당한 삽입을 PCR에 의해 (각각 M13 정방향 프라이머 (서열 번호 269) 및 N1180 (서열 번호 238) 및 M13 역방향 프라이머 (서열 번호 270) 및 N1181 (서열 번호 239)) 확인하였다.

[0595] 상기에 설명한 벡터 (pUC19::ald6Δ::loxP-URA3-loxP)를 *AhdI*로 선형화하고, (DNA를 포함하는 세포의 인큐베이션을 2.5시간까지 연장시킨 것을 제외하고서) 표준 아세트산리튬법을 이용하여 PNY1528 및 PNY2237을 형질전환시켰다. 탄소 공급원으로서 1% 에탄올이 제공된, 우라실이 제외된 합성 완전 배지 상에 플레이팅함으로써 형질전환체를 수득하였다. 패치된 형질전환체를 PCR에 의해 스크리닝하여 결실/통합을 확인하였는데, 이는 프라이머 N1212 (서열 번호 241) 및 N1180 (5' 말단) (서열 번호 238) 및 N1181 (서열 번호 239) 및 N1213 (서열 번호 242) (3' 말단)을 이용하였다. 히스티딘 마커 선발을 이용하여, Cre 리코비나아제를 지닌 플라스미드 (pRS423::GAL1p-Cre = 서열 번호 271)로 상기 주를 형질전환시켰다. 형질전환체를 0.5% 갈락토스로 보충된 YPE 상에서 계대하였다. 콜로니를 5-FOA에 대한 내성 (URA3 마커의 상실)에 대하여 그리고 히스티딘 영양요구성 (Cre 플라스미드의 상실)에 대하여 스크리닝하였다. 측면에 위치하는 loxP 부위를 통한 URA3 유전자의 적당한 제거는 PCR (프라이머 N1262 및 N1263, 각각 서열 번호 255 및 256)에 의해 확인하였다. 부가적으로, ALD6 유전자에 대하여 내재적인 프라이머 (N1230 및 N1231; 각각 서열 번호 261 및 262)를 사용하여 부분 이배체 (merodiploid)가 전혀 존재하지 않음을 보장하였다. 마지막으로, *ald6Δ::loxP* 클론을 PCR로 스크리닝하여 *ura3Δ::loxP* (N1228 및 N1229, 서열 번호 259 및 260)와 *gpd2Δ::loxP* (N1223 및 N1225, 서열 번호 257 및 258) 사이의 전좌가 일어나지 않았음을 확인하였다. 2개의 양성 클론을 PNY1528의 형질전환체의 스크리닝에 의해 확인하였다. 클론 B를 PNY2243으로 표기하였다. 3개의 양성 클론을 PNY2237의 형질전환체의 스크리닝에 의해 확인하였다. 클론 E 및 K 둘 모두를 소규모로 아이소부탄올 생성에 대하여 평가하였다 (하기). 대부분의 파라미터에서 통계적으로 동일하지는 않지만, 클론 E를 추가의 개발을 위하여 선발하였다 (PNY2238).

[0596] 실시예 10

[0597] 아이소부탄올 경로 플라스미드

[0598] 이 실시예의 목적은 숙주 주에서 아이소부탄올을 생성하기 위한 아이소부탄올 경로 플라스미드의 제작 또는 변형을 설명하기 위한 것이다.

[0599] pYZ067 (서열 번호 374)을 하기 키메라 유전자를 함유하도록 제작하였다: 1) 효모 FBA1 프로모터로부터 발현되는, C-말단 루미오(Lumio) 태그를 포함하는 에스. 뮤탄스 UA159 유래의 *ilvD* 유전자의 코딩 영역에 이어서 FBA1 종결 서열 - 다이하이드록시산 테하이드라타아제의 발현을 위한 것임 -, 2) 효모 GPM1 프로모터로부터 발현되는 말 간 ADH의 코딩 영역에 이어서 ADH1 종결 서열 - 알코올 테하이드로게나아제의 발현을 위한 것임 -, 및 3) 효모 TDH3 프로모터로부터 발현되는, 락토코커스 락티스 유래의 *KivD* 유전자의 코딩 영역에 이어서 TDH3 종결 서열 - 케토아이소발레이트 테카르복실라아제의 발현을 위한 것임 - .

[0600] pYZ067Δ*kivD*Δ*hADH* (서열 번호 385)를 pYZ067 (서열 번호 374)로부터 제작하였으며, 이는 *kivD* 및 *adh* 둘 모두를 위한 프로모터-유전자-종결 서열 카세트를 결실시킴에 의한 것이었다. pYZ067을 BamHI 및 SacI (미국 매사추세츠주 입스위치 소재의 뉴 잉글랜드 바이오랩스)으로 절단하고, 7934 bp 단편을 아가로스 겔, 이어서 겔 추출 키트 (미국 캘리포니아주 발렌시아 소재의 퀴아젠) 상에서 정제하였다. 단리된 DNA 단편을 DNA 폴리머라아제 I, 큰 (클레나우) 단편 (미국 매사추세츠주 입스위치 소재의 뉴 잉글랜드 바이오랩스)으로 처리하고, 그 후 T4 DNA 라이가아제를 이용하여 자기 라이게이션시키고, 이를 이용하여 적격성 TOP10 에스케리키아 콜라이 (미국 캘리포니아주 칼스배드 소재의 인비트로젠)를 형질전환시켰다. 형질전환체로부터의 플라스미드를 단리하고, 서열 분석에 의해 적당한 결실에 대하여 점검하였다. 올바른 플라스미드 단리체를 pYZ067Δ*kivD*Δ*hADH*로 표기하였다.

[0601] pYZ067Δ*kivD*Δ*ilvD* (서열 번호 772)는 ADH의 발현을 위한, 사카로마이세스 세레비지에에서의 발현용으로 코돈 최적화되고, 효모 GPM 프로모터 (뉴클레오티드 3916-3160)로부터 발현되고, ADH1 종결 서열 (뉴클레오티드 2012-1697)이 이어지는, 말 간 유래의 *adh* 유전자의 코딩 영역 (뉴클레오티드 위치 3148-2021)을 갖는 키메라 유전자를 함유하도록 제작하였다. pYZ067Δ*kivD*Δ*ilvD*는 *kivD* 및 *ilvD* 둘 모두를 위한 프로모터-유전자-종결 서열 카세트를 결실시킴으로써 pYZ067로부터 제작하였다. pYZ067을 AatII 및 SacI (미국 매사추세츠주 입스위치 소재의 뉴 잉글랜드 바이오랩스)으로 절단하고; 10196 bp 단편을 아가로스 겔, 이어서 겔 추출 키트 (미국 캘리포니아주 발렌시아 소재의 퀴아젠) 상에서 정제하였다. 단리된 DNA 단편을 DNA 폴리머라아제 I, 큰 (클레나우) 단편 (미국 매사추세츠주 입스위치 소재의 뉴 잉글랜드 바이오랩스)으로 처리하고, 그 후, T4 DNA 라이가아제로 자기 라이게이션시켰다. 그 후, 생성된 플라스미드를 NgoMIV 및 BamHI (미국 매사추세츠주 입스위치 소재의 뉴

잉글랜드 바이오랩스)으로 절단하고, 7533 bp 단편을 아가로스 겔, 이어서 겔 추출 키트 (미국 캘리포니아주 발렌시아 소재의 쿼아젠) 상에서 정제하였다. 단리된 DNA 단편을 DNA 폴리머라아제 I, 큰 (클레나우) 단편 (미국 매사추세츠주 입스위치 소재의 뉴 잉글랜드 바이오랩스)으로 처리하고, 그 후, T4 DNA 라이가아제로 자기 라이게이션시켰다. 플라스미드를 단리하고, 서열 분석에 의해 적당한 결실에 대하여 점검하였다. 올바른 플라스미드 단리체를 pYZ067DkivDDilvD로 표기하였다.

[0602] pYZ090 (서열 번호 195)으로부터 유도된 pK9G9.OLE1p.ilvD (서열 번호 773)는, DHAD의 발현을 위한, 효모 OLE1 프로모터 (뉴클레오티드 5986-5387)로부터 발현되고 FBA1 종결 서열 (뉴클레오티드 3632-3320) 이 뒤따르는 스트렙토코커스 뮤탄스 유래의 ilvD 유전자의 코딩 영역 (뉴클레오티드 위치 5377-3641)을 갖는 키메라 유전자, 및 KARI 발현을 위한, 효모 ILV5 프로모터 (뉴클레오티드 427-1620)로부터 발현되고 ILV5 종결 서열 (뉴클레오티드 2685-3307)이 뒤따르는 엔에어로스티페스 카카에 유래의 ilvC 유전자의 변이체 K9G9 (핵산 및 아미노산 서열 번호 774 및 647)의 코딩 영역 (뉴클레오티드 1628-2659)을 갖는 키메라 유전자를 함유하도록 제작하였다. 플라스미드의 제작은 하기와 같았다. 효모 FBA1 프로모터로부터 발현되고 FBA1 종결 서열이 뒤따르는 스트렙토코커스 뮤탄스 유래의 ilvD 유전자의 코딩 영역을 갖는 플라스미드 pYZ067로부터의 키메라 유전자를, 제한 효소 NgoMIV 및 BamHI을 이용한 절단 후에 pYZ090 내로 라이게이션시켰다. CUP1 프로모터의 3' 말단으로부터의 280 bp 및 alsS 코딩 영역은, 제한 효소 SpeI 및 PacI으로 절단하고 생성된 큰 DNA 단편을 자기 라이게이션시킴으로써 상기 생성된 플라스미드로부터 결실시켰다. ilvD의 상류의 효모 FBA1 프로모터는 상기 생성된 플라스미드로부터, 제한 효소 NgoMIV 및 PmlI를 이용한 절단에 의해 제거하고, 프라이머 pOLE1-NgoMI (서열 번호 775) 및 pOLE1-PmlI (서열 번호 776)로 증폭시킨 효모 OLE1 프로모터로 대체하였다. 락토코커스 락티스 유래의 ilvC 유전자의 코딩 영역은, 제한 효소 PmeI 및 SfiI로 절단함으로써 상기 생성된 플라스미드로부터 결실시키고, 이어서 큰 DNA 단편을 겔 정제하였다. 엔에어로스티페스 카카에 유래의 변이형 K9G9 ilvC 유전자의 코딩 영역 (서열 번호 777)을 PmeI 및 SfiI을 이용하여 pLH701 (서열 번호 778)로부터 절단해 내고, 겔 정제하였다. 2개의 DNA 단편을 라이게이션하여 pK9G9.OLE1p.ilvD를 생성하였다.

[0603] 실시예 11

[0604] PNY2240 및 PNY2242의 제작

[0605] 주 PNY2240을 플라스미드 pLH702 (서열 번호 181) 및 pBP915 (서열 번호 182)를 이용한 형질전환 후 PNY2211로부터 유도하였다. 형질전환체를 히스티딘 또는 우라실을 포함하지 않는 합성 완전 배지 (탄소 공급원으로서 1% 에탄올) 상에 플레이팅하였다. 형질전환체들을 탄소 공급원으로서 0.05% 에탄올 및 2% 글루코스를 함유하는 동일 배지에 패치하였다. 3개의 패치를 이용하여 액체 배지 (우라실이 제외된, 탄소 공급원으로서 0.3% 글루코스 및 0.3% 에탄올을 포함하는 합성 완전 배지)에 접종하였다. 아이소부탄올 생성을 시험하기 위하여, 우라실이 제외된, 탄소 공급원으로서 2% 글루코스 및 0.05% 에탄올을 함유하는 합성 완전 배지 내로 액체 배양물을 계대 배양하였는데, 상기 배지는 BME 비타민 믹스 (시그마 카탈로그 번호 B6891)를 또한 함유하였다. 배양물을 진탕하면서 (인포르스 멀티트론(Infors Multitron) 진탕기에서 250 rpm) 30℃에서 밀봉 세럼 바이알에서 (15 ml 바이알 중 10 ml의 배지) 인큐베이션하였다. 48시간 후, 배양 배지를 여과하고 (스핀-엑스(Spin-X) 컬럼), HPLC로 분석하였다 (이는 전체적으로 본 명세서에 참고로 포함된 미국 특허 공개 제2007/0092957호에 기재된 바와 같음). 하나의 클론을 PNY2240으로 표기하였다.

[0606] 주 PNY2242를 플라스미드 pLH702 (서열 번호 181) 및 pYZ067ΔkivDΔhADH (본 명세서에서 상기에 설명됨)를 이용한 형질전환 후 PNY2238로부터 유도하였다. 형질전환체를 히스티딘 또는 우라실을 포함하지 않는 합성 완전 배지 (탄소 공급원으로서 1% 에탄올) 상에 플레이팅하였다. 대신, 탄소 공급원으로서 0.05% 에탄올 및 2% 글루코스를 함유하는 동일 배지에 형질전환체들을 패치하였다. 3개의 패치를 상기에 설명한 바와 같이 아이소부탄올 생성에 대하여 검사하였다. 상기 3개의 패치 전부는 글루코스 소비 및 아이소부탄올 생성 면에서 유사하게 작동하였다. 하나의 클론을 PNY2242로 표기하고, 발효 조건 하에서 추가로 특성화하였으며, 이는 본 명세서에서 하기에 설명한 바와 같았다.

[0607] 실시예 12

[0608] 사카로마이세스 세레비지에 주 BP1064 (PNY1503)의 제작

[0609] 주 BP1064를 CEN.PK 113-7D (CBS 8340; CBS (Centraalbureau voor Schimmelcultures) 진균류 생물다양성 센터, 네덜란드)로부터 유도하였으며, 하기의 유전자의 결실을 함유한다: URA3, HIS3, PDC1, PDC5, PDC6, 및 GPD2. BP1064를 플라스미드 pYZ090 (서열 번호 195) 및 pLH468 (서열 번호 139)로 형질전환시켜 주 NGCI-070

(BP1083; PNY1504)을 생성하였다.

[0610] 전체 코딩 서열이 완전히 제거되는 결실을 형질전환체의 선발을 위한 G418 내성 마커 또는 URA3 유전자 중 어느 하나, 및 표적 유전자의 상류 및 하류의 상동성 영역을 함유하는 PCR 단편과의 상동 재조합에 의해 생성하였다. loxP 부위가 측면에 위치하는 G418 내성 마커를 Cre 리코비나아제를 이용하여 제거하였다 (pRS423::PGAL1-cre; 서열 번호 271). URA3 유전자는 상동 재조합에 의해 제거하여 스카리스 결실을 생성하거나, 또는 loxP 부위가 측면에 위치할 경우, Cre 리코비나아제를 사용하여 제거하였다.

[0611] URA3 결실

[0612] 내인성 URA3 코딩 영역을 결실시키기 위하여, ura3::loxP-kanMX-loxP 카세트를 pLA54 주형 DNA (서열 번호 386)로부터 PCR 증폭시켰다. pLA54는 케이. 락티스TEF1 프로모터 및 kanMX 마커를 함유하며, loxP 부위가 측면에 위치하여 Cre 리코비나아제를 사용한 재조합 및 마커의 제거를 가능하게 한다. PCR을 퓨전 DNA 폴리머라아제 및 프라이머 BK505 및 BK506 (서열 번호 294 및 295)을 사용하여 행하였다. 각각의 프라이머의 URA3 부분은 URA3 프로모터의 상류의 5' 영역 및 코딩 영역의 하류의 3' 영역으로부터 유래되어, loxP-kanMX-loxP 마커의 통합이 URA3 코딩 영역의 대체를 야기하게 하였다. PCR 생성물로 표준 유전자 기술 (문헌[Methods in Yeast Genetics, 2005, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, pp. 201-202])을 사용하여 CEN.PK 113-7D를 형질전환시키고, 형질전환체를 30°C에서 G418 (100 µg/ml)을 함유하는 YPD 상에서 선발하였다. 형질전환체를 프라이머 LA468 및 LA492 (서열 번호 296 및 297)를 사용한 PCR에 의해 스크리닝하여 올바른 통합을 검증하고, CEN.PK 113-7D Δura3::kanMX로 표기하였다.

[0613] HIS3 결실

[0614] 스카리스 HIS3 결실을 위한 PCR 카세트에 있어서의 4개의 단편을 퓨전 하이 피델리티 PCR 마스터 믹스 (미국 매사추세츠주 입스위치 소재의 뉴 잉글랜드 바이오랩스)를 사용하고, 젠트라 퓨어진 이스트/백트 키트 (미국 캘리포니아주 발렌시아 소재의 퀴아젠)로 제조된 주형으로서의 CEN.PK 113-7D 게놈 DNA를 사용하여 증폭시켰다. HIS3 단편 A를 프라이머 oBP452 (서열 번호 298) 및 프라이머 oBP453 (서열 번호 299) - HIS3 단편 B의 5' 말단에 대하여 상동성을 갖는 5' 테일을 함유함 - 을 이용하여 증폭시켰다. HIS3 단편 B를, HIS3 단편 A의 3' 말단에 대하여 상동성을 갖는 5' 테일을 함유하는 프라이머 oBP454 (서열 번호 300), 및 HIS3 단편 U의 5' 말단에 대하여 상동성을 갖는 5' 테일을 함유하는 프라이머 oBP455 (서열 번호 301)를 사용하여 증폭시켰다. HIS3 단편 U를, HIS3 단편 B의 3' 말단에 대하여 상동성을 갖는 5' 테일을 함유하는 프라이머 oBP456 (서열 번호 302), 및 HIS3 단편 C의 5' 말단에 대하여 상동성을 갖는 5' 테일을 함유하는 프라이머 oBP457 (서열 번호 303)을 사용하여 증폭시켰다. HIS3 단편 C를, HIS3 단편 U의 3' 말단에 대하여 상동성을 갖는 5' 테일을 함유하는 프라이머 oBP458 (서열 번호 304), 및 프라이머 oBP459 (서열 번호 305)를 사용하여 증폭시켰다. PCR 생성물을 PCR 정제 키트 (퀴아젠)를 사용하여 정제하였다. HIS3 단편 AB는 HIS3 단편 A 및 HIS3 단편 B를 혼합하고 프라이머 oBP452 (서열 번호 298) 및 oBP455 (서열 번호 301)를 사용하여 증폭시킴으로써 중첩 PCR에 의해 생성하였다. HIS3 단편 UC는 HIS3 단편 U 및 HIS3 단편 C를 혼합하고 프라이머 oBP456 (서열 번호 302) 및 oBP459 (서열 번호 305)를 사용하여 증폭시킴으로써 중첩 PCR에 의해 생성하였다. 생성된 PCR 생성물을 아가로스 겔, 이어서 겔 추출 키트 (퀴아젠)에서 정제하였다. HIS3 ABUC 카세트는 HIS3 단편 AB 및 HIS3 단편 UC를 혼합하고 프라이머 oBP452 (서열 번호 298) 및 oBP459 (서열 번호 305)를 사용하여 증폭시킴으로써 중첩 PCR에 의해 생성하였다. PCR 생성물을 PCR 정제 키트 (퀴아젠)로 정제하였다.

[0615] CEN.PK 113-7D Δura3::kanMX의 적격성 세포를 만들고, 프로즌-이지 이스트 트랜스포메이션 II 키트 (미국 캘리포니아주 오렌지 소재의 자이모 리서치)를 사용하여 HIS3 ABUC PCR 카세트로 형질전환시켰다. 형질전환 혼합물을 30°C에서 2% 글루코스가 보충된 우라실이 결여된 합성 완전 배지 상에 플레이트하였다. his3 녹아웃이 있는 형질전환체를 젠트라 퓨어진 이스트/백트 키트 (퀴아젠)로 제조된 게놈 DNA를 사용하여, 프라이머 oBP460 (서열 번호 306) 및 oBP461 (서열 번호 307)을 사용한 PCR에 의해 스크리닝하였다. 올바른 형질전환체를 주 CEN.PK 113-7D Δura3::kanMX Δhis3::URA3으로서 선발하였다.

[0616] Δura3 부위로부터의 KanMX 마커 제거 및 Δhis3 부위로부터의 URA3 마커 제거

[0617] KanMX 마커는 프로즌-이지 이스트 트랜스포메이션 II 키트 (자이모 리서치)를 사용하여 CEN.PK 113-7D Δura3::kanMX Δhis3::URA3을 pRS423::PGAL1-cre (서열 번호 271)로 형질전환시키고, 30°C에서 히스티딘 및 우라실이 결여된 그리고 2% 글루코스가 보충된 합성 완전 배지 상에 플레이트함으로써 제거하였다. 형질전환체를 30°C에서 대략 6시간 동안 1% 갈락토스가 보충된 YP에서 성장시켜 Cre 리코비나아제 및 KanMX 마커 절제를 유도

하고, 회복을 위하여 30℃에서 YPD (2% 글루코스) 플레이트 상에 플레이팅하였다. 단리체를 YPD에서 하룻밤 성장시키고, 30℃에서 5-플루오로-오로트산 (0.1%)을 함유하는 합성 완전 배지 상에 플레이팅하여, URA3 마커가 상실된 단리체를 선발하였다. 5-FOA 내성 단리체를 pRS423::PGAL1-cre 플라스미드의 제거를 위하여 YPD에서 성장시켜 YPD 상에 플레이팅하였다. 단리체를 YPD+G418 플레이트, 우라실이 결여된 합성 완전 배지 플레이트, 및 히스티딘이 결여된 합성 완전 배지 플레이트 상에서의 성장을 분석함으로써 pRS423::PGAL1-cre 플라스미드, URA3 마커, 및 KanMX 마커의 상실에 대하여 점검하였다. G418에 대하여 민감하고 우라실 및 히스티딘에 대하여 영양요구성인 올바른 단리체를 주 CEN.PK 113-7D Δura3::loxP Δhis3으로서 선발하고, BP857로 표기하였다. 결실 및 마커 제거는, 젠트라 퓨어진 이스트/백트 키트 (퀴아젠)를 이용하여 제조한 게놈 DNA를 사용하여, Δura3의 경우 프라이머 oBP450 (서열 번호 308) 및 oBP451 (서열 번호 309)을 사용하여 그리고 Δhis3의 경우 프라이머 oBP460 (서열 번호 306) 및 oBP461 (서열 번호 307)을 사용하여 PCR 및 서열결정에 의해 확인하였다.

[0618] PDC6 결실

[0619] 스카리스 PDC6 결실을 위한 PCR 카세트를 위한 4개의 단편을 퓨전 하이 피델리티 PCR 마스터 믹스 (뉴 잉글랜드 바이오랩스) 및 젠트라 퓨어진 이스트/백트 키트 (퀴아젠)로 제조된 주형으로서의 CEN.PK 113-7D 게놈 DNA를 사용하여 증폭시켰다. PDC6 단편 A를, 프라이머 oBP440 (서열 번호 310) 및 프라이머 oBP441 (서열 번호 311) - PDC6 단편 B의 5' 말단에 대하여 상동성을 갖는 5' 테일을 함유함 - 을 사용하여 증폭시켰다. PDC6 단편 B를, PDC6 단편 A의 3' 말단에 대하여 상동성을 갖는 5' 테일을 함유하는 프라이머 oBP442 (서열 번호 312), 및 PDC6 단편 U의 5' 말단에 대하여 상동성을 갖는 5' 테일을 함유하는 프라이머 oBP443 (서열 번호 313)을 사용하여 증폭시켰다. PDC6 단편 U를, PDC6 단편 B의 3' 말단에 대하여 상동성을 갖는 5' 테일을 함유하는 프라이머 oBP444 (서열 번호 314), 및 PDC6 단편 C의 5' 말단에 대하여 상동성을 갖는 5' 테일을 함유하는 프라이머 oBP445 (서열 번호 315)를 사용하여 증폭시켰다. PDC6 단편 C를, PDC6 단편 U의 3' 말단에 대하여 상동성을 갖는 5' 말단을 함유하는 프라이머 oBP446 (서열 번호 316), 및 프라이머 oBP447 (서열 번호 317)을 사용하여 증폭시켰다. PCR 생성물을 PCR 정제 키트 (퀴아젠)를 사용하여 정제하였다. PDC6 단편 AB는 PDC6 단편 A 및 PDC6 단편 B를 혼합하고 프라이머 oBP440 (서열 번호 310) 및 oBP443 (서열 번호 313)을 사용하여 증폭시킴으로써 중첩 PCR에 의해 생성하였다. PDC6 단편 UC는 PDC6 단편 U 및 PDC6 단편 C를 혼합하고 프라이머 oBP444 (서열 번호 314) 및 oBP447 (서열 번호 317)을 사용하여 증폭시킴으로써 중첩 PCR에 의해 생성하였다. 생성된 PCR 생성물을 아가로스 겔, 이어서 겔 추출 키트 (퀴아젠)에서 정제하였다. PDC6 ABUC 카세트는 PDC6 단편 AB 및 PDC6 단편 UC를 혼합하고 프라이머 oBP440 (서열 번호 310) 및 oBP447 (서열 번호 317)을 사용하여 증폭시킴으로써 중첩 PCR에 의해 생성하였다. PCR 생성물을 PCR 정제 키트(퀴아젠)로 정제하였다.

[0620] CEN.PK 113-7D Δura3::loxP Δhis3의 적격성 세포를 만들고, 프로즌-이지 이스트 트랜스포메이션 II 키트 (자이모 리서치)를 사용하여 PDC6 ABUC PCR 카세트로 형질전환시켰다. 형질전환 혼합물을 30℃에서 2% 글루코스가 보충된 우라실이 결여된 합성 완전 배지 상에 플레이팅하였다. pdc6 녹아웃이 있는 형질전환체를 젠트라 퓨어진 이스트/백트 키트 (퀴아젠)로 제조된 게놈 DNA를 사용하여, 프라이머 oBP448 (서열 번호 318) 및 oBP449 (서열 번호 319)를 사용한 PCR에 의해 스크리닝하였다. 올바른 형질전환체를 주 CEN.PK 113-7D Δura3::loxP Δhis3 Δpdc6::URA3으로 선발하였다.

[0621] CEN.PK 113-7D Δura3::loxP Δhis3 Δpdc6::URA3을 YPD에서 하룻밤 성장시키고, 30℃에서 5-플루오로-오로트산 (0.1%)을 함유하는 합성 완전 배지 상에 플레이팅하여, URA3 마커가 상실된 단리체를 선발하였다. 결실 및 마커 제거는 젠트라 퓨어진 이스트/백트 키트 (퀴아젠)로 제조된 게놈 DNA를 사용하여, 프라이머 oBP448 (서열 번호 318) 및 oBP449 (서열 번호 319)를 이용한 PCR 및 서열결정에 의해 확인하였다. 상기 단리체 유래의 PDC6 유전자의 부재는 PDC6의 코딩 서열에 대하여 특이적인 프라이머, oBP554 (서열 번호 320) 및 oBP555 (서열 번호 321)를 사용한 음성 PCR 결과에 의해 입증하였다. 올바른 단리체를 주 CEN.PK 113-7D Δura3::loxP Δhis3 Δpdc6으로 선발하고, BP891로서 표기하였다.

[0622] PDC1 결실 ilvDSm 통합

[0623] PDC1 유전자를 결실시키고, 스트렙토코커스 뮤탄스 ATCC #700610으로부터의 ilvD 코딩 영역으로 대체하였다. PDC1 결실-ilvDSm 통합을 위한 PCR 카세트를 위한, A 단편에 이은, 스트렙토코커스 뮤탄스 유래의 ilvD 코딩 영역을 퓨전 하이 피델리티 PCR 마스터 믹스 (뉴 잉글랜드 바이오랩스) 및 젠트라 퓨어진 이스트/백트 키트 (퀴아젠)로 제조된 주형으로서의 NYLA83 게놈 DNA를 사용하여 증폭시켰다. NYLA83은 전체적으로 본 명세서에 참고로 포함된 미국 특허 공개 제2009/0305363호에 기재된 PDC1 결실-ilvDSm 통합을 지닌 주이다. PDC1 단편 A-ilvDSm (서열 번호 322)을 프라이머 oBP513 (서열 번호 326) 및 프라이머 oBP515 (서열 번호 327) - PDC1 단편

B의 5' 말단에 대하여 상동성을 갖는 5' 테일을 함유함 - 를 사용하여 증폭시켰다. PDC1 결실-ilvDSm 통합을 위한 PCR 카세트를 위한 B, U, 및 C 단편을 퓨전 하이 피델리티 PCR 마스터 믹스 (뉴 잉글랜드 바이오랩스)를 사용하여 그리고 젠트라 퓨어진 이스트/백트 키트 (퀴아젠)를 이용하여 제조한 주형으로서의 CEN.PK 113-7D 게놈 DNA를 사용하여 증폭시켰다. PDC1 단편 B를, PDC1 단편 A-ilvDSm의 3' 말단에 대하여 상동성을 갖는 5' 테일을 함유하는 프라이머 oBP516 (서열 번호 328), 및 PDC1 단편 U의 5' 말단에 대하여 상동성을 갖는 5' 테일을 함유하는 프라이머 oBP517 (서열 번호 329)을 사용하여 증폭시켰다. PDC1 단편 U를, PDC1 단편 B의 3' 말단에 대하여 상동성을 갖는 5' 테일을 함유하는 프라이머 oBP518 (서열 번호 330), 및 PDC1 단편 C의 5' 말단에 대하여 상동성을 갖는 5' 테일을 함유하는 프라이머 oBP519 (서열 번호 331)를 사용하여 증폭시켰다. PDC1 단편 C를, PDC1 단편 U의 3' 말단에 대하여 상동성을 갖는 5' 테일을 함유하는 프라이머 oBP520 (서열 번호 332), 및 프라이머 oBP521 (서열 번호 333)을 사용하여 증폭시켰다. PCR 생성물을 PCR 정제 키트 (퀴아젠)를 사용하여 정제하였다. PDC1 단편 A-ilvDSm-B는 PDC1 단편 A-ilvDSm 및 PDC1 단편 B를 혼합하고 프라이머 oBP513 (서열 번호 326) 및 oBP517 (서열 번호 329)을 사용하여 증폭시킴으로써 중첩 PCR에 의해 생성하였다. PDC1 단편 UC는 PDC1 단편 U 및 PDC1 단편 C를 혼합하고 프라이머 oBP518 (서열 번호 330) 및 oBP521 (서열 번호 333)을 사용하여 증폭시킴으로써 중첩 PCR에 의해 생성하였다. 생성된 PCR 생성물을 아가로스 겔, 이어서 겔 추출 키트 (퀴아젠)에서 정제하였다. PDC1 A-ilvDSm-BUC 카세트 (서열 번호 323)는 PDC1 단편 A-ilvDSm-B 및 PDC1 단편 UC와 혼합하고 프라이머 oBP513 (서열 번호 326) 및 oBP521 (서열 번호 333)을 사용하여 증폭시킴으로써 중첩 PCR에 의해 생성하였다. PCR 생성물을 PCR 정제 키트 (퀴아젠)로 정제하였다.

[0624] CEN.PK 113-7D Δ ura3::loxP Δ his3 Δ pdc6의 적격성 세포를 만들고, 프로즌-이지 이스트 트랜스포메이션 II 키트 (자이모 리서치)를 사용하여 PDC1 A-ilvDSm-BUC PCR 카세트로 형질전환시켰다. 형질전환 혼합물을 30°C에서 2% 글루코스가 보충된 우라실이 결여된 합성 완전 배지 상에 플레이팅하였다. pdc1 녹아웃 ilvDSm 통합이 있는 형질전환체를 젠트라 퓨어진 이스트/백트 키트 (퀴아젠)로 제조된 게놈 DNA를 사용하여 프라이머 oBP511 (서열 번호 336) 및 oBP512 (서열 번호 337)를 사용한 PCR에 의해 스크리닝하였다. 상기 단리체 유래의 PDC1 유전자의 부재는 PDC1의 코딩 서열에 대하여 특이적인 프라이머, oBP550 (서열 번호 338) 및 oBP551 (서열 번호 339)를 사용한 음성 PCR 결과에 의해 입증하였다. 올바른 형질전환체를 주 CEN.PK 113-7D Δ ura3::loxP Δ his3 Δ pdc6 Δ pdc1::ilvDSm-URA3으로 선발하였다.

[0625] CEN.PK 113-7D Δ ura3::loxP Δ his3 Δ pdc6 Δ pdc1::ilvDSm-URA3을 YPD에서 하룻밤 성장시키고, 30°C에서 5-플루오로-오로트산 (0.1%)을 함유하는 합성 완전 배지 상에 플레이팅하여, URA3 마커가 상실된 단리체를 선발하였다. PDC1의 결실, ilvDSm의 통합 및 마커 제거를 젠트라 퓨어진 이스트/백트 키트 (퀴아젠)로 제조된 게놈 DNA를 사용하여 프라이머 oBP511 (서열 번호 336) 및 oBP512 (서열 번호 337)를 사용하여 PCR 및 서열결정에 의해 확인하였다. 올바른 단리체를 주 CEN.PK 113-7D Δ ura3::loxP Δ his3 Δ pdc6 Δ pdc1::ilvDSm으로서 선발하고, BP907로서 표기하였다.

[0626] PDC5 결실 sadB 통합

[0627] PDC5 유전자를 결실시키고, 아크로모박터 자일로숙시단스 유래의 sadB 코딩 영역으로 대체하였다 (상기 sadB 유전자는 전체적으로 본 명세서에 참고로 포함된 미국 특허 공개 제2009/0269823호에 기재되어 있다). PDC5 결실-sadB 통합을 위한 PCR 카세트의 절편을 먼저 플라스미드 pUC19-URA3MCS 내로 클로닝하였다.

[0628] pUC19-URA3MCS는 pUC19를 기반으로 하며, 다중 클로닝 부위 (MCS) 내에 위치하는, 사카로마이세스 세레비지에 유래의 URA3 유전자의 서열을 함유한다. pUC19는 pMB1 레플리콘 및 에스케리키아 콜라이에서의 복제 및 선발을 위한 베타-락타마아제를 코딩하는 유전자를 함유한다. URA3의 코딩 서열에 더하여, 이 유전자의 상류 및 하류로부터의 서열을 효모에서의 URA3 유전자의 발현을 위하여 포함시켰다. 백터는 클로닝 목적을 위해 사용될 수 있으며, 효모 통합 백터로서 사용될 수 있다.

[0629] 사카로마이세스 세레비지에 CEN.PK 113-7D 게놈 DNA로부터의 URA3 코딩 영역의 250 bp 상류 및 150 bp 하류와 함께 URA3 코딩 영역을 포함하는 DNA를 퓨전 하이 피델리티 PCR 마스터 믹스 (뉴 잉글랜드 바이오랩스)를 사용하여 BamHI, AscI, PmeI 및 FseI 제한효소 부위를 함유하는 프라이머 oBP438 (서열 번호 334) 및 XbaI, PacI 및 NotI 제한효소 부위를 함유하는 oBP439 (서열 번호 335)로 증폭시켰다. 게놈 DNA를 젠트라 퓨어진 이스트/백트 키트 (퀴아젠)를 사용하여 제조하였다. PCR 생성물 및 pUC19 (서열 번호 325)를 BamHI 및 XbaI을 이용한 절단 후 T4 DNA 라이가아제로 라이게이션시켜 백터 pUC19-URA3MCS를 생성하였다. 상기 백터를 프라이머 oBP264 (서열 번호 342) 및 oBP265 (서열 번호 343)를 이용한 PCR 및 서열결정에 의해 확인하였다.

[0630] sadB 및 PDC5 단편 B의 코딩 서열을 pUC19-URA3MCS 내로 클로닝하여 PDC5 A-sadB-BUC PCR 카세트의 sadB-BU

부분을 생성하였다. sadB의 코딩 서열을, 주형으로서 pLH468-sadB (서열 번호 359)를 사용하여, AscI 제한효소 부위를 함유하는 프라이머 oBP530 (서열 번호 344), 및 PDC5 단편 B의 5' 말단에 대하여 상동성을 갖는 5' 테일을 함유하는 프라이머 oBP531 (서열 번호 345)을 이용하여 증폭시켰다. PDC5 단편 B를, sadB의 3' 말단에 대하여 상동성을 갖는 5' 테일을 함유하는 프라이머 oBP532 (서열 번호 346), 및 PmeI 제한효소 부위를 함유하는 프라이머 oBP533 (서열 번호 347)을 사용하여 증폭시켰다. PCR 생성물을 PCR 정제 키트 (퀴아젠)를 사용하여 정제하였다. sadB-PDC5 단편 B는 sadB 및 PDC5 단편 B의 PCR 생성물을 혼합하고, 프라이머 oBP530 (서열 번호 344) 및 oBP533 (서열 번호 347)을 사용하여 증폭시킴으로써 중첩 PCR에 의해 생성하였다. 생성된 PCR 생성물을 AscI 및 PmeI으로 절단하고, T4 DNA 라이가아제를 사용하여 적절한 효소를 이용한 절단 후에 pUC19-URA3MCS의 상응하는 부위 내로 라이게이션시켰다. 생성된 플라스미드는, 프라이머 oBP536 (서열 번호 348) 및 oBP546 (서열 번호 349) - PDC5 단편 C의 5' 말단에 대하여 상동성을 갖는 5' 테일을 함유함 - 를 사용한 sadB-단편 B-단편 U의 증폭을 위한 주형으로 사용하였다. PDC5 단편 C를, PDC5 sadB-단편 B-단편 U의 3' 말단에 대하여 상동성을 갖는 5' 테일을 함유하는 프라이머 oBP547 (서열 번호 350), 및 프라이머 oBP539 (서열 번호 351)를 사용하여 증폭시켰다. PCR 생성물을 PCR 정제 키트 (퀴아젠)를 사용하여 정제하였다. PDC5 sadB-단편 B-단편 U-단편 C는 PDC5 sadB-단편 B-단편 U 및 PDC5 단편 C를 혼합하고 프라이머 oBP536 (서열 번호 348) 및 oBP539 (서열 번호 351)를 사용하여 증폭시킴으로써 중첩 PCR에 의해 생성하였다. 생성된 PCR 생성물을 아가로스 겔에 이어서 겔 추출 키트 (퀴아젠)에서 정제하였다. PDC5 A-sadB-BUC 카세트 (서열 번호 324)를, 천연 PDC5 코딩 서열의 바로 상류의 50 뉴클레오티드에 대하여 상동성을 갖는 5' 테일을 함유하는 프라이머 oBP542 (서열 번호 352), 및 oBP539 (서열 번호 351)를 사용하여 PDC5 sadB-단편 B-단편 U-단편 C를 증폭시킴으로써 생성하였다. PCR 생성물을 PCR 정제 키트 (퀴아젠)로 정제하였다.

[0631] CEN.PK 113-7D Δura3::loxP Δhis3 Δpdc6 Δpdc1::ilvDSm의 적격성 세포를 만들고, 프로존-이지 이스트 트랜스포메이션 II 키트 (자이모 리서치)를 사용하여 PDC5 A-sadB-BUC PCR 카세트에 형질전환시켰다. 형질전환 혼합물을 30℃에서 1% 에탄올을 보충한 (글루코스 없음) 우라실이 결여된 합성 완전 배지 상에 플레이팅하였다. pdc5 녹아웃 sadB 통합이 있는 형질전환체를 젠트라 퓨어인 이스트/백트 키트 (퀴아젠)로 제조된 게놈 DNA를 사용하여 프라이머 oBP540 (서열 번호 353) 및 oBP541 (서열 번호 354)을 사용한 PCR에 의해 스크리닝하였다. 상기 단리체 유래의 PDC5 유전자의 부재는 PDC5의 코딩 서열에 대하여 특이적인 프라이머, oBP552 (서열 번호 355) 및 oBP553 (서열 번호 356)을 사용한 음성 PCR 결과에 의해 입증하였다. 올바른 형질전환체를 주 CEN.PK 113-7D Δura3::loxP Δhis3 Δpdc6 Δpdc1::ilvDSm Δpdc5::sadB-URA3으로 선발하였다.

[0632] CEN.PK 113-7D Δura3::loxP Δhis3 Δpdc6 Δpdc1::ilvDSm Δpdc5::sadB-URA3을 YPE (1% 에탄올)에서 하룻밤 성장시키고, 에탄올이 보충된 (글루코스 없음) 그리고 5-플루오로-오로트산 (0.1%)을 함유하는 합성 완전 배지에서 30℃에서 플레이팅하여, URA3 마커가 상실된 단리주를 선발하였다. PDC5의 결실, sadB의 통합, 및 마커 제거를 젠트라 퓨어인 이스트/백트 키트 (퀴아젠)로 제조된 게놈 DNA를 사용하여 프라이머 oBP540 (서열 번호 353) 및 oBP541 (서열 번호 354)을 사용하여 PCR에 의해 확인하였다. 올바른 단리체를 주 CEN.PK 113-7D Δura3::loxP Δhis3 Δpdc6 Δpdc1::ilvDSm Δpdc5::sadB로서 선발하고, BP913으로 표기하였다.

[0633] GPD2 결실

[0634] 내인성 GPD2 코딩 영역을 결실시키기 위하여, gpd2::loxP-URA3-loxP 카세트 (서열 번호 361)는 loxP-URA3-loxP PCR (서열 번호 360)을 주형 DNA로서 사용하여 PCR 증폭시켰다. loxP-URA3-loxP는 loxP 리코비나아제 부위가 측면에 위치하는 ATCC # 77107로부터의 URA3 마커를 함유한다. PCR을 퓨전 DNA 폴리머라아제 및 프라이머 LA512 및 LA513 (서열 번호 340 및 341)을 사용하여 행하였다. 각각의 프라이머의 GPD2 부분은 GPD2 코딩 영역의 상류의 5' 영역 및 상기 코딩 영역의 하류의 3' 영역으로부터 유래되어, loxP-URA3-loxP 마커의 통합이 GPD2 코딩 영역의 대체를 야기하게 하였다. PCR 생성물로 BP913을 형질전환시키고, 형질전환체를 1% 에탄올이 보충되고 (글루코스 없음) 우라실이 결여된 합성 완전 배지 상에서 선발하였다. 형질전환체를 스크리닝하여, 프라이머 oBP582 및 AA270 (서열 번호 357 및 358)을 사용한 PCR에 의해 올바른 통합을 검증하였다.

[0635] URA3 마커는 pRS423::PGAL1-cre (서열 번호 271)를 이용한 형질전환에 의해 그리고 30℃에서 1% 에탄올이 보충된, 히스티딘이 결여된 합성 완전 배지 상에 플레이팅함으로써 재생시켰다. 형질전환체를 1% 에탄올이 보충된 그리고 5-플루오로-오로트산 (0.1%)을 함유하는 합성 완전 배지 상에 도말하고, 30℃에서 인큐베이션하여 URA3 마커를 상실한 단리체를 선발하였다. 5-FOA 내성 단리체를 pRS423::PGAL1-cre 플라스미드의 제거를 위하여 YPE (1% 에탄올)에서 성장시켰다. 결실 및 마커 제거를 프라이머 oBP582 (서열 번호 357) 및 oBP591 (서열 번호 362)을 사용한 PCR에 의해 확인하였다. 올바른 단리체를 주 CEN.PK 113-7D Δura3::loxP Δhis3 Δpdc6 Δ

pdc1::ilvDSm Δpdc5::sadB Δgpd2::loxP로서 선발하고, BP1064 (PNY1503)로 표기하였다.

- [0636] 실시예 13
- [0637] PNY2204 및 아이소부탄올 경로 플라스미드의 제작
- [0638] 이 실시예의 목적은 아세토락테이트 신타아제를 코딩하는 유전자를 염색체 XII 내의 PDC1 코딩 서열과 TRX1 코딩 서열 사이의 자연 발생 유전자간 영역 내로 통합시키는 것을 가능하게 하기 위한 벡터의 제작을 설명하는 것이다. 이 벡터의 사용에서 생기는 주를 또한 설명한다.
- [0639] 통합 벡터 pUC19-kan::pdc1::FBA-alsS::TRX1의 제작
- [0640] FBA-alsS-CYCt 카세트는 1.7kb BbvCI/PacI 단편을 pRS426::GPD::alsS::CYC (전체적으로 본 명세서에 참고로 포함된 미국 특허 제7,851,188호에 개시됨)로부터 pRS426::FBA::ILV5::CYC (전체적으로 본 명세서에 참고로 포함된 미국 특허 제7,851,188호)로 이동시킴으로써 제작하였는데, 이는 이전에 ILV5 유전자가 유리되도록 BbvCI/PacI로 절단하였다. 라이게이션 반응물로 이. 콜라이 TOP10 세포를 형질전환시키고, 형질전환체를 프라이머 N98SeqF1 (서열 번호 363) 및 N99SeqR2 (서열 번호 365)를 사용하여 PCR에 의해 스크리닝하였다. FBA-alsS-CYCt 카세트는 AfIII 부위 (클레나우 단편을 사용하여 라이게이션에 적합한 말단을 만들)에서의 pUC19-URA3::ilvD-TRX1 내로의 클로닝 (클론 "B")을 위하여 BglIII 및 NotI을 사용하여 상기 벡터로부터 단리하였다. 벡터 내에서 둘 모두의 배향으로 alsS 카세트를 함유하는 형질전환체를 수득하였으며, 구성(configuration) "A"의 경우 프라이머 N98SeqF4 (서열 번호 364) 및 N1111 (서열 번호 366)을 사용하여 그리고 구성 "B"의 경우 N98SeqF4 (서열 번호 364) 및 N1110 (서열 번호 367)을 사용하여 PCR에 의해 확인하였다. 그 후, URA3 유전자 (1.2 kb NotI/NaeI 단편)를 제거하고 제네티신 카세트를 부가함으로써 제네티신 선발가능 버전의 "A" 구성 벡터를 만들었다. 클레나우 단편을 사용하여 라이게이션에 적합한 모든 말단을 만들고, 형질전환체를 프라이머 BK468 (서열 번호 368) 및 N160SeqF5 (서열 번호 210)를 사용하여 이전의 URA3 마커와 동일한 배향의 제네티신 내성 유전자를 갖는 클론이 선발되도록 PCR로 스크리닝하였다. 생성된 클론을 pUC19-kan::pdc1::FBA-alsS::TRX1 (클론 A)(서열 번호 387)로 칭하였다.
- [0641] alsS 통합 주 및 아이소부탄올-생산 유도체의 제작
- [0642] 상기에 기재된 pUC19-kan::pdc1::FBA-alsS 통합 벡터를 *PmeI*로 선형화하고 PNY1507 (실시예 8)을 형질전환시켰다. *PmeI*은 클로닝된 pdc1-TRX1 유전자간 영역 내에서 상기 벡터를 절단하며, 따라서 그 위치에서의 표적화된 통합에 이르게 된다 (문헌[Rodney Rothstein, Methods in Enzymology, 1991, volume 194, pp. 281-301]). 형질전환체를 YPE + 50 μg/ml G418 상에서 선발하였다. 패치된 형질전환체를 프라이머 N160SeqF5 (서열 번호 210) 및 oBP512 (서열 번호 337)를 사용하여 통합 사건에 대하여 PCR에 의해 스크리닝하였다. 2개의 형질전환체를 아이소부탄올을 만드는 그 주의 능력의 평가에 의해 간접적으로 아세토락테이트 신타아제 기능에 대하여 시험하였다. 이를 행하기 위하여, 추가의 아이소부탄올 경로 유전자를 이. 콜라이-효모 서틀 벡터 (pYZ090 ΔalsS 및 pBP915, 하기에 기재됨) 상에 공급하였다. 하나의 클론을 PNY2205로 표기하였다. 플라스미드 무함유 모 주(parent strain)를 PNY2204 (MATa ura3Δ::loxP his3Δ pdc6Δ pdc1Δ::P[PDC1]-DHAD|ilvD_Sm-PDC1t-pUC19-loxP-kanMX-loxP-P[FBA1]-ALS|alsS_Bs-CYC1t pdc5Δ::P[PDC5]-ADH|sadB_Ax-PDC5t gpd2Δ::loxP fra2Δ adh1Δ::UAS(PGK1)P[FBA1]-kivD_L1(y)-ADH1t)로 표기하였다.
- [0643] 아이소부탄올 경로 플라스미드 (pYZ090 ΔalsS 및 pBP915)
- [0644] pYZ090 (서열 번호 195)을 *SpeI* 및 *NotI*으로 절단하여 대부분의 CUP1 프로모터와 전부의 *alsS* 코딩 서열 및 CYC 종결 서열을 제거하였다. 그 후, 상기 벡터를 클레나우 단편을 이용하여 처리한 후 자기 라이게이션시키고, 이. 콜라이 Stb13 세포를 형질전환시켜 암피실린 내성에 대하여 선발하였다. 상기 DNA 영역의 제거는 프라이머 N191 (서열 번호 370)을 사용하여 PCR에 의해 라이게이션 접합부를 가로지르는 DNA 서열결정에 의해 2개의 독립적인 클론에 대하여 확인하였다. 생성된 플라스미드를 pYZ090 ΔalsS (서열 번호 371)로 명명하였다. pLH468 플라스미드를 효모에서의 DHAD, KivD 및 HADH의 발현용으로 제작하였다. pBP915 (서열 번호 182)는 *kivD* 유전자와, *kivD*의 상류의 *TDH3* 프로모터의 957 염기쌍을 결실시킴으로써 pLH468 (서열 번호 139)로부터 제작하였다. pLH468을 *SwaI*로 절단하고, 큰 단편 (12896 bp)을 아가로스 겔, 이어서 겔 추출 키트 (미국 캘리포니아주 발렌시아 소재의 퀴아젠) 상에서 정제하였다. DNA의 단리된 단편을 T4 DNA 라이가아제를 이용하여 자기 라이게이션시키고, 전기적격성 TOP10 에스케리키아 콜라이 (미국 캘리포니아주 칼스배드 소재의 인비트로젠)를 형질전환시켰다. 형질전환체로부터의 플라스미드를 단리하고, *SwaI* 제한 효소를 이용한 제한효소 분석에 의해 적당한 결실에 대하여 점검하였다. 또한 단리체들을 프라이머 oBP556 (서열 번호 372) 및

oBP561 (서열 번호 373)을 사용하여 결실 부위를 가로질러서 서열결정하였다. 적당한 결실을 갖는 클론을 pBP915 (pLH468 Δ kivD)(서열 번호 182)로 표기하였다.

[0645] pYZ090은 pHR81 (미국 버지니아주 머내서스 소재의 ATCC #87541) 골격을 기반으로 한다. pYZ090 을 ALS의 발현을 위한, 효모 CUP1 프로모터 (뉴클레오티드 2-449)로부터 발현되며 CYC1 종결 서열 (뉴클레오티드 2181-2430)이 뒤따르는 바실러스 서브틸리스 유래의 alsS 유전자의 코딩 영역 (뉴클레오티드 위치 457-2172)을 갖는 키메라 유전자, 및 KARI의 발현을 위한, 효모 ILV5 프로모터 (2433-3626)로부터 발현되며 ILV5 종결 서열 (뉴클레오티드 4682-5304)이 뒤따르는 락토코커스 락티스 유래의 ilvC 유전자의 코딩 영역 (뉴클레오티드 3634-4656)을 갖는 키메라 유전자를 함유하도록 제작하였다.

[0646] 실시예 14.

[0647] 아이소부탄을 생성-PNY1910 및 PNY2242

[0648] 방법:

[0649] 접종 배지의 제조

[0650] 1 L의 접종 배지는 하기를 함유하였다: 아미노산을 포함하지 않는 6.7 g의 효모 질소 베이스(Yeast Nitrogen Base; 디프코(Difco) 0919-15-3); 히스티딘, 류신, 트립토판 및 우라실을 포함하지 않는 2.8 g의 효모 합성 드롭아웃 배지 보충제(Yeast Synthetic Drop-out Medium Supplement; 시그마 Y2001); 20 mL의 1% (w/v) L-류신; 4 mL의 1% (w/v) L-트립토판; 3 g의 에탄올; 10 g의 글루코스.

[0651] 균정 발효 배지의 제조

[0652] 접종 후 브로쓰의 부피는 800 mL이었으며, 이때 최종 조성은 리터당 하기와 같았다: 5 g의 황산암모늄, 2.8 g의 제1인산칼륨, 1.9 g의 황산마그네슘 7수화물, 0.2 mL의 소포제 (시그마 DF204), 히스티딘, 류신, 트립토판 및 우라실을 포함하지 않는 효모 합성 드롭아웃 배지 보충제 (시그마 Y2001), 16 mg의 L-류신, 4 mg의 L-트립토판, 6 mL의 비타민 혼합물 (1 L의 물 중, 50 mg의 바이오틴, 1 g의 판토텐산칼슘, 1 g의 니코틴산, 25 g의 미오-이노시톨, 1 g의 티아민 클로라이드 염산염, 1 g의 피리독솔 염산염, 0.2 g의 p-아미노벤조산), 6 mL의 미량 미네랄 용액 (1 L의 물 중, 15 g의 EDTA, 4.5 g의 황산아연 7수화물, 0.8 g의 염화망간 무수물, 0.3 g의 염화코발트 6수화물, 0.3 g의 황산구리 5수화물, 0.4 g의 다이소듐 몰리브덴 무수물, 4.5 g의 염화칼슘 2수화물, 3 g의 황산철 7수화물, 1 g의 붕산, 0.1 g의 요오드화칼륨), 30 mg의 티아민 HCl, 30 mg의 니코틴산. pH를 2 N KOH를 이용하여 5.2로 조정하고, 글루코스를 10 g/L로 첨가하였다.

[0653] 접종물의 준비

[0654] 전체 바이알 배양물 (대략 1 mL)을 10 mL의 접종 배지 내에 피펫팅함으로써 직접적으로 냉동 바이알로부터 125 mL 진탕 플라스크에 접종하였다. 상기 플라스크를 260 rpm 및 30°C에서 인큐베이션하였다. 상기 주를 약 1.0의 OD까지 하룻밤 성장시켰다. $\lambda = 600$ nm에서의 OD를 베크만(Beckman) 분광 광도계 (미국 소재의 베크만)에서 측정하였다.

[0655] 생물반응기의 실험 설계

[0656] 0.8 L를 작업 부피로 하여 1 L 바이오스탯(Biostat) B DCU3 발효기 (미국 소재의 사르토리우스(Sartorius))에서 발효를 실시하였다. 오프가스(off-gas) 조성물을 프리마(Prima) DB 질량 분광계 (미국 소재의 서모 일렉트론 코포레이션(Thermo Electron Corp.))에 의해 모니터링하였다. 전체 발효 전체에 걸쳐서 온도를 30°C로 유지하고, pH를 2 N KOH를 이용하여 5.2로 제어하였다. 80 mL의 접종물을 접종한 직후, dO를 교반에 의해 30%로 제어하고, pH를 5.25로 제어하고, 통기량을 0.2 L/min으로 제어하였다. 일단 대략 3의 OD가 도달되면, 가스를 혐기성 배양을 위하여 N₂로 스위칭하였다. 발효 전체에 걸쳐, 글루코스를 50% (w/w) 용액의 수동 첨가에 의해 과량 (5-20 g/L)으로 유지하였다.

[0657] 배양 실험의 분석 방법

[0658] $\lambda = 600$ nm에서의 OD는 잘 혼합된 브로쓰 샘플을 큐벳 (CS500 브이더블유알 인터내셔널(VWR International), 독일 소재) 내로 피펫팅함으로써 분광 광도계에서 측정하였다. 샘플의 바이오매스 농도가 분광 광도계의 선형 흡광도 범위 (전형적으로 0.000 내지 0.300의 OD 값)를 초과하면, 샘플을 0.9% NaCl 용액으로 희석시켜 선형 범위 내의 값을 생성하였다.

[0659] 배양 상청액 중 글루코스, 아이소부탄올, 및 다른 발효 부산물의 측정은 굴절률 (RI) 및 다이오드 어레이 (210 nm) 검출기를 갖춘, 바이오-라드 아미넥스(Bio-Rad Aminex) HPX-87H 컬럼 (미국 소재의 바이오-라드)을 사용한 HPLC에 의해 실시하였다. 0.01 N의 H_2SO_4 를 이동상으로 사용하여 40℃의 컬럼 온도 및 0.6 mL/min의 유량으로 크로마토그래피 분리를 성취하였다. 아이소부탄올 체류 시간은 이들 조건 하에서 32.2분이다. 오프가스 샘플 중 아이소부탄올의 농도를 질량-분광계로 측정하였다.

[0660] 결과

[0661] 광학 밀도 (OD)로 측정되는 최대 바이오매스 농도, 아이소부탄올 생성량의 체적비, 최종 아이소부탄올 역가, 및 글루코스에 대한 아이소부탄올 수율이 하기 표에 제시되어 있다. 주 PNY2242는 주 PNY1910보다 더 높은 역가 및 더 빠른 속도를 가졌으며, 더욱 높은 비속도(specific rate) 및 비역가(specific titer)로 아이소부탄올을 생성하였다. 비속도가 도 5에 예시되어 있다. 배양 상청액 중 DHIV + DHMB의 측정은 PNY2242 주와 비교하여 PNY1910에서 3배 더 높았다 (도 6). 글루코스에 대한 글리세롤, 피루브산, BDO, DHIV+DHMB*, αKIV, 및 아이소부티르산의 수율이 도 7에 예시되어 있다.

[0662] *HPLC법에 의해 분석한 DHIV는 DHIV 및 DHMB 둘 모두를 포함한다.

[0663] [표 19]

주	최대 OD600	속도 (g/L/h)	역가 (g/L)	수율 (g/g)
PNY1910	5.0	0.16	10.9	0.25
PNY2242	5.0	0.23	16.1	0.27

[0664]

[0665] 실시예 15.

[0666] K9G9의 오류 발생이 쉬운 PCR 라이브러리의 제작

[0667] K9G9의 오류 발생이 쉬운 PCR을 수행하여 라이브러리를 생성하였으며, 상기 라이브러리는 NADH에 비하여 NADPH에 대한 Km 값이 증가된 변이체에 대하여 스크리닝될 수 있다. K9G9의 돌연변이 유발 PCR을 진모르프(GeneMorph)(등록상표) II 이지클론 도메인 돌연변이 유발 키트(EZClone Domain Mutagenesis Kit) (카탈로그 번호 200552; 미국 캘리포니아주 라 졸라 소재의 애질런트 테크놀로지스, 스트라타진 제품부(Agilent Technologies, Stratagene Products Division))를 이용하여 수행하였다. 프라이머 K9G9_EZ_F1 (AAA CAT GGA AGA ATG TAA GAT GGC; 서열 번호 390) 및 K9G9_EZ_R1 (TCA GTT GTT AAT CAA CTT GTC TTC G; 서열 번호 391)은 인테그레이티드 디엔에이 테크놀로지스, 인크. (미국 아이오와주 코랄빌 소재)에 의해 상업적으로 합성되었다. 프라이머, 주형, 및 ddH₂O 이외의, 여기서 사용한 시약들은 상기에 나타난 키트와 함께 공급되었다. 돌연변이 유발 PCR 혼합물은 4 μl의 pHR81-PIIv5-KARI-K9.G9 (서열 번호 392) (770 ng/μg), 1.25 μl의 각각의 프라이머 (스톡 1 μl당 100 ng), 5 μl의 10× 뮤타자임(Mutazyme) II 반응 완충제, 1 μl의 40 mM dNTP 믹스, 1.5 μl의 뮤타자임 II DNA 폴리머라아제, 및 36 μl의 ddH₂O로 이루어졌다. 하기 조건을 PCR 반응에 이용하였다: 출발 온도는 95℃, 2.0분이었으며, 이어서 30회의 가열/냉각 사이클이었다. 각각의 사이클은 95℃, 30초; 48℃, 30초, 및 72℃, 2.0분으로 이루어졌다. 상기 온도 사이클링의 완료시에, 샘플을 72℃에서 10.0분 더 유지하고, 그 후 4℃에서 샘플 회수를 기다리면서 유지하였다. 반응 생성물을 아가로스 겔 전기영동 (1% 아가로스, 1X TBE 완충제)을 통하여 주형으로부터 분리하고, 스트라타프렘(StrataPrep)(등록상표) DNA 겔 추출 키트 (카탈로그 번호 400766, 미국 캘리포니아주 라 졸라 소재의 애질런트 테크놀로지스, 스트라타진 제품부)를 사용하여 회수하였으며, 이는 제조업자가 권고한 바와 같았다.

[0668] 단리된 반응 생성물을 메가프라이머로서 이용하여 상기에 나타난 키트의 "이지클론 반응"에서의 유전자 라이브러리를 생성하였다. 메가프라이머, 주형 및 ddH₂O 이외의, 여기서 사용한 시약들은 상기에 나타난 키트와 함께 공급되었다. 반응물은 25 μl의 2x 이지클론 효소 믹스, 4 μl의 메가프라이머 (125 ng/μl), pBAD.KARI 벡터 중 K9G9 (25 ng/μl) 2 μl, 3 μl의 이지클론 용액, 및 16 μl의 ddH₂O로 이루어졌다. 하기 조건을 반응에 이용하였다: 출발 온도는 95℃, 1.0분이었으며, 이어서 30회의 가열/냉각 사이클이었다. 각각의 사이클은 95℃, 50초; 60℃, 50초; 및 68℃, 10.0분으로 이루어졌다. 상기 온도 사이클링의 완료시에, 샘플들을 72℃에서 10.0분 더 유지하고, 그 후 4℃에서 샘플 회수를 기다리면서 유지하였다. 1 μl의 Dpn I (10 U/μl)을 첨가

하고, 혼합물을 37℃에서 4시간 동안 인큐베이션하였다.

[0669] 그 후, 4 μ l의 Dpn I 절단 "이지클론 반응" 생성물로 50 μ l의 XL10-골드(등록상표) 초적격성(Ultracompetent) 이. 콜라이 세포 (진모르프(등록상표) II 이지클론 도메인 돌연변이 유발 키트 내에 제공됨)를 형질전환시켰으며, 이는 제조업자가 권고한 바와 같았다. 형질전환체를 LB 배지 및 100 μ g/ml의 암피실린(카탈로그 번호 L1004, 미국 캘리포니아주 홀리스터 소재의 테크노바 인크.(Teknova Inc.))을 함유하는 한천 배지 상에 도말하고, 37℃에서 하룻밤 인큐베이션하고, 4℃에서 보관하였다. 총 10개의 형질전환체에 대하여 형질전환당 4 μ l의 Dpn I 절단 "이지클론 반응" 생성물 및 50 μ l의 세포를 이용하여 이들 단계를 반복하였다. XL-골드에서의 생성된 라이브러리를 M9 염을 함유하는 용액을 이용하여 한천 플레이트에서 긁어내고, 합하고, LB 배지 및 100 μ g/ml의 암피실린을 함유하는 배지 내에 희석시키고, 37℃에서 하룻밤 인큐베이션하였다. 상기 라이브러리 DNA를 제조업자에 의해 제공된 프로토콜에 따라 퀴아프랩 스핀 미니프랩 키트(QIAprep Spin Miniprep Kit; 카탈로그 번호 2706; 미국 캘리포니아주 발렌시아 소재의 퀴아젠)를 이용하여 세포로부터 단리하였다. 그 후, 증폭된 라이브러리를 사용하여, 바이오라드 진 펄서 II (미국 캘리포니아주 허큘리스 소재의 바이오라드 레보리토리즈 인크.)를 이용하여 전기적격성 이. 콜라이 Bw25113 (Δ ilvC) 주를 형질전환시켰다. 형질전환된 클론을 LB 배지 및 100 μ g/ml의 암피실린 (#101320-154, 미국 캘리포니아주 홀리스터 소재의 테크노바 인크.)을 함유하는 한천 배지 상에 도말하고, 37℃에서 하룻밤 인큐베이션하였다. 실시예 16에서 설명한 바와 같이 클론을 고 처리량 스크리닝에 이용하였다.

[0670] 실시예 16.

[0671] NADH 활성의 NADP⁺ 저해의 감소에 대한 스크리닝을 통한, NADPH에 대한 K_M이 증가된 K9G9 변이체의 확인

[0672] 실시예 15에서 설명한 K9G9 라이브러리를 NADH-의존성 KARI 활성의 NADP⁺ 저해가 감소된 변이체에 대하여 스크리닝하였다. NADH에 의한 활성의 NADP⁺ 저해가 감소된 K9G9 변이체는 NADH에 대한 K_M에 대한 NADPH에 대한 K_M의 비의 증가를 잠재적으로 나타낼 수 있다. NADH에 대한 K_M에 비하여 NADPH에 대한 K_M을 증가시키기 위한 특별한 목적에서, 스크린으로부터의 히트(hit)들을 부분적으로 정제하고, 동력학적 분석을 수행하여 NADH 및 NADPH에서의 V_{max} 및 K_M 파라미터를 결정하였다.

[0673] K9G9 유전자 라이브러리의 고 처리량 스크리닝 분석

[0674] 돌연변이 KARI 효소의 유전자 라이브러리의 고 처리량 스크리닝을 본 명세서에 설명된 바와 같이 수행하였다: 554.4 $\times$ g/L의 글리세롤, 68 mM의 (NH₄)₂SO₄, 4 mM의 MgSO₄, 17 mM의 시트르산나트륨, 132 mM의 KH₂PO₄, 36 mM의 K₂HPO₄를 함유하는 10x 냉동 배지를 분자적으로 순수한 물로 제조하고, 여과 살균하였다. 냉동 배지는 10x 냉동 배지를 LB 배지로 희석시킴으로써 준비하였다. 분취량 (200 μ l)의 1x 냉동 배지를 96웰 아카이브 플레이트(카탈로그 번호: 3370, 미국 뉴욕주 코닝 소재의 코닝 인크.)의 각각의 웰에 사용하였다.

[0675] LB 한천 배지로부터의 클론을 선발하고, 냉동 배지를 함유하는 96웰 아카이브 플레이트 내에 접종하고, 진탕 없이 37℃에서 하룻밤 성장시켰다. 그 후, 상기 아카이브 플레이트를 -80℃에서 보관하였다. 미국 특허 제 8,129,162호에 설명된 이. 콜라이 주 Bw25113(Δ ilvC)를 pBAD-HisB (인비트로젠)로 형질전환시킨 것을 항상 음성 대조군으로 사용하였다. 상기 라이브러리의 양성 대조군은 이. 콜라이 주 Bw25113(Δ ilvC) 내의 K9G9-KARI였으며, 이는 미국 특허 제 8,129,162호에 기재된 바와 같았다.

[0676] 아카이브 플레이트로부터의 클론을 깊은 96웰 플레이트 내에 접종하였다. 각각의 웰은 해동된 아카이브 플레이트로부터의 세포 3.0 μ l, 유도제로서 0.02%(w/v) 아라비노스 및 암피실린 100 μ g/ml을 함유하는 LB 배지 200 μ l를 함유하였다. 세포를 진탕시키면서 (900 rpm) 80% 습도를 이용하여 37℃에서 하룻밤 성장시키고, 원심분리 (3750 rpm, 25℃에서 5분) (에펜도르프 원심분리기, 미국 뉴욕주 웨스트베리 소재의 브린크만 인스트루먼츠, 인크.)에 의해 수확하고, 세포 펠렛을 이후의 분석을 위하여 -20℃에서 보관하였다.

[0677] 분석 기질, (R,S)-아세토라테이트를 올라마우프 및 솔로스에 의해 설명된 바와 같이 합성하였다 (문헌 [Aulabaugh and Schloss, Biochemistry, 29: 2824-2830, 1990]). 이 분석법에서 사용한 모든 다른 화학물질은 시그마로부터 구매하였다. KARI에 의한 아세토라테이트의 α , β -다이하이드록시아이소발레레이트로의 효소적 전환에 이어, 플레이트 판독기 (사파이어(Sapphire) 2, 스위스 만네도르프 소재의 테칸(Tecan))을 이용하여 340 nm에서 반응물로부터의 보조 인자, NADH의 산화를 측정하였다. NADH에 대하여 6220 M⁻¹cm⁻¹의 몰 흡광 계수를

이용하여 활성을 계산하였다.

[0678] 깊은 웰의 플레이트 중의 냉동 세포 펠렛 및 버그버스터 (독일 다름스타트 소재의 노바겐(Novagen) 71456)를 실온에서 30분 동안 동시에 가온하였다. 75 μ l의 50% 버그버스터 (v/v, 물 중)를 30분간의 가온 후 각각의 웰에 첨가하고, 세포를 플레이트 진탕기를 이용하여 현탁시켰다. 세포 펠렛/50% 버그 버스터 현탁물을 포함하는 플레이트를 실온에서 30분 동안 인큐베이션하였다. 세포 용해물을 75 μ L의 d.d수(d.d water)로 희석시켜 0.5X 용해물을 생성하였다. 희석시킨 세포 무함유 추출물의 분석은 2.5 mM의 NADP+를 포함하거나 또는 포함하지 않는, 2.4 mM의 (R/S)-아세토락테이트, 100 mM의 HEPES pH 6.8, 100 mM의 KCl, 10 mM의 MgCl₂, 150 μ M의 NADH, 12.5 μ L의 0.5 X 세포 용해물을 함유하는 완충제에서 30℃에서 수행하였다.

[0679] NADH KARI 활성의 NADP+ 저해가 감소된 K9G9 변이체의 확인

[0680] NADP+의 부재 하에서의 측정된 NADH 산화 속도에 대한 NADP+의 존재 하에서의 측정된 NADH 산화 속도의 비를 각각의 변이체 및 양성 대조 웰 (플레이트당 2개)에 대하여 계산하였다. 양성 대조 웰 (총 104개) 전부에 있어서의 비들의 평균 및 표준 편차를 계산하였다.

[0681] 변이체 웰은, NADP+의 부재 하에서의 속도가 0.1 OD/hr보다 크고 상기 속도 비가 0.45 초과 (양성 대조군의 평균보다 더 높은 3개의 표준 편차), 그리고 1 미만일 경우 초기 히트를 함유하는 것으로 간주하였다. 총 521개의 히트를 4607가지의 잠재적인 변이체의 풀로부터 확인하였다. 이들 초기 히트들을 결합시켜 추가 분석용의 더욱 작은 라이브러리를 형성하였다.

[0682] 초기 라이브러리 히트의 이차 스크리닝

[0683] 결합된 히트 라이브러리를 생물학적 삼중체로 성장시키고, 세포 무함유 추출물을 제조하고, 상기에 기재된 바와 같이 분석하였다. 그 후, 속도 비들을 상기와 같이 변이체 및 양성 대조군에 대하여 계산하였다. 상술된 동력학적 분석용으로 선발된 최종 히트는 하기 기준을 충족시켰다: NADP+의 부재 하에서의 속도는 0.6 OD/hr 초과였으며, 속도 비는 0.51 초과, 그리고 1 미만이었으며, 3개의 생물학적 레플리케이트 중 적어도 2개는 이 기준을 통과하였다. 17개의 히트를 동력학적 분석용으로 확인하였으며, 100 μ g/mL의 암피실린이 첨가된 LB 플레이트 상에 도말하였다.

[0684] K9G9 변이체의 서열 분석

[0685] 이차 HTS 스크리닝으로부터 확인된 17가지의 변이체의 DNA 서열결정은 프라이머 pBAD-For (ATGCCATAGCATTTTATCC; 서열 번호 393) 및 pBAD-Rev (CTGATTAACTCTGTATCAGGCT; 서열 번호 394)를 이용하여 템플리파이(TempliPhi)TM (지이 헬스케어(GE Healthcare))를 사용함으로써 달성하였다.

[0686] [표 20]

K9G9 에 있어서의 아미노산 치환

변이체	서열	아미노산 치환
878 C1	873	어떠한 것도 확인되지 않음
879 A7	874	K90M
879 C2	875	H37Q
880 A11	876	A182T, P320Q
880 B4	877	K57E
880 D11	878	K90M, A174V
881 A2	879	K90M, I133V, K282T
881 G3	880	Y53F, E74G
881 G9	881	K90E
882 B12	882	H118R
882 C10	883	G31S, R61S, C121Y, D129N, G183D
882 C7	933	E54G
882 F9	934	K90E, Q160H
882 G6	935	G55A
882 G12	936	V142L, S285Y
883 C4	937	A170V
883 G9	938	L197M, K310M

[0687]

[0688] 부분 정제된 변이체 단백질의 동력학적 분석

[0689] 미국 특허 제8,129,162호에 기재된 이. 콜라이 주 Bw25113 (Δ ilvC)을 사용하여 17가지의 변이체 및 양성 대조군 K9G9를 발현시켰다. 주는 125 mL의 배플형, 통기식 여과 뚜껑 플라스크에서 진탕하면서 100 μ g/mL의 암피실린을 함유하는 LB 브로쓰 (#46-060-CM, 미국 버지니아주 매너서스 소재의 메디아테크(Mediatech)) 10 mL에서

37℃에서 8시간 동안 성장시켰다. 200 μ L의 이 배양물을 사용하여 100 μ g/mL의 암피실린 및 0.2% (w/v)의 아라비노스가 첨가된 100 mL LB 브로쓰에 접종하였다. 이들 배양물을 500 mL의 배플형, 통기식 여과 뚜껑 플라스틱에서 진탕하면서 37℃에서 16 내지 18시간 동안 성장시켰다. 세포를 하나의 20 mL 및 2개의 40 mL 분취물로 수확하고, 상청액을 가만히 따라내고, 펠렛을 -80℃에서 냉동시켰다.

[0690] 단백질을 부분 정제하기 위하여, 상기 20 mL 세포 배양물 수확물에 상응하는 세포 펠렛을 해동시키고, 1 mL 버그 버스터 마스터 믹스 (노바젠 71456, 독일 다름스타트 소재)에 재현탁시켰다. 세포 현탁물을 실온에서 15분 동안 인큐베이션하고, 이어서 60℃에서 15분간 인큐베이션하여 열불안정성 단백질을 변성시켰다. 세포 잔사 및 변성된 단백질을 4℃에서 30분 동안 원심분리함으로써 펠렛화하였다. K9G9 및 변이체를 포함하는 열안정성 세포질 단백질을 함유하는 상청액을 회수하고, 4℃에서 보관하였다.

[0691] 열안정성 세포질 단백질 분획의 전체 단백질량을 쿠마시 플러스(Coomassie Plus) (미국 일리노이주 록포드 소재의 서모 사이언티픽(Thermo Scientific) #23238)를 이용하여 브래드포드 분석법(Bradford Assay)에 의해 측정하였다. BSA를 표준물로 이용하였다. 캐리(Cary) 300 분광 광도계 (미국 델라웨어주 월밍턴 소재의 애질런트 테크놀로지스(Agilent Technologies))를 사용하여 595 nm에서 흡광도를 측정함으로써 단백질의 농도를 측정하였다.

[0692] NADH 및 NADPH에 대한 V_{max} 및 K_m 값을 결정하기 위하여, 부분 정제된 단백질을 NADH (0, 16.4, 32.8, 65.7, 98.5, 164.3 및 246.5 μ M) 및 NADPH (0, 12.8, 25.6, 51.2, 76.8 및 128 μ M)의 다양한 농도에서 분석하였다. 분석은 30℃에서 100 mM의 HEPES (pH 6.8), 10 mM의 $MgCl_2$, 100 mM의 KCl 및 4.8 mM의 R/S-아세토락테이트에서 행하였다. 0.1 내지 0.35 mg/mL의 전체 단백질을 상기 분석물에 첨가하였다. S-아세토락테이트의 DHIV로의 전환 속도는 캐리 300 분광 광도계 (미국 델라웨어주 월밍턴 소재의 애질런트 테크놀로지스)를 사용하여 340 nm에서 NAD(P)H의 산화를 모니터링함으로써 측정하였다. $6220 M^{-1}cm^{-1}$ 의 몰 흡광 계수를 이용하여 활성을 계산하였다. V_{max} 및 K_m 값을 보조 인자 농도에 대한 비활성 (U/mg)의 도시에 의해 계산하였으며, 그 데이터를 칼레이다그래프(Kaleidagraph) 소프트웨어 (미국 펜실베이니아주 리딩 소재의 시너지(Synergy))를 사용하여 미하엘리스-멘텐 등식에 피팅시켰다.

[0693] [표 21]

NAD(P)H 소비량 분석을 통하여 결정한, 부분적으로 정제된 K9G9 변이체에 있어서의 동력학적 값

변이체	NADPH의 V_{max} (U/mg)	NADPH의 K_m (μ M)	NADPH의 V_{max}/K_m (L/min*mg)	NADH의 V_{max} (U/mg)	NADH의 K_m (μ M)	NADH의 V_{max}/K_m (L/min*mg)
K9G9	1.53	45.5	0.034	1.09	67.4	0.016
878 C1	0.75	42.2	0.018	0.62	107.8	0.006
879 A7	2.51	546	0.005	1.44	263	0.006
879 C2	1.27	103	0.012	1.23	187	0.007
880 A11	0.72	86.9	0.008	0.51	117	0.004
880 B4	1.23	233	0.005	1.14	133	0.009
880 D11	1.38	130	0.011	1.50	232	0.006
881 A2	0.88	93.5	0.009	1.13	166.8	0.007
881 G3	0.69	99.2	0.007	0.69	61.8	0.011
881 G9	1.03	158	0.007	0.96	310	0.003
882 B12	0.87	30.3	0.029	0.49	78.9	0.006
882 C10	0.71	34.1	0.021	0.56	97.9	0.006
882 C7	1.62	45.3	0.036	0.96	75.6	0.013
882 F9	1.39	256	0.005	1.19	335	0.004
882 G6	0.95	47.4	0.020	0.74	98.7	0.007
882 G12	1.06	63.5	0.017	0.75	81.2	0.009
883 C4	1.26	46.8	0.027	0.67	83.9	0.008
883 G9	1.26	38	0.033	1.01	71.9	0.014

[0694]

[0695] 실시예 17.

[0696] 부위 지정 돌연변이 유발을 통한 K9 KARI 변이체의 수동 재조합

[0697] K9G9 유도체 K9JB4 및 K9JG3 (실시예 16에서 각각 880 B4 및 881 G3으로 확인됨)의 부위 지정 돌연변이 유발을 수행하여 실시예들에 기재되어 있는 다른 아미노산 변화를 포함시켰다. 초기 단계는 N87P 치환을 추가하는 것이었으며, 이는 실시예 5에 설명되어 있다. 돌연변이를 프라이머 N87PC1 (CTGACATCATTATGATCTTGATCCAGATGAAAAGCAGGCTACCATGTAC; 서열 번호 395) 및 N87PC1r (GTACATGGTAGCCTGCTTTTCATCTGGGATCAAGATCATAATGATGTCAG; 서열 번호 396)을 이용하여 KARI 유전자 내로 도입하

였으며, 이는 퀵체인지(QuikChange)(등록상표) II 부위 지정 돌연변이 유발 키트 (카탈로그 번호 200523; 미국 캘리포니아주 라 줄라 소재의 애질런트 테크놀로지스, 스트라타진 제품부)를 이용하였다. 프라이머, 주형 및 ddH₂O를 제외하고는, 여기서 사용한 모든 시약은 상기에 나타난 키트와 함께 공급되었다. 프라이머들은 인테그레이티드 디엔에이 테크놀로지스, 인크. (미국 아이오와주 코랄빌 소재)에 의해 상업적으로 합성되었다. 주형은 이. 콜라이 벡터 내의 K9 KARI 변이체 (pBAD.KARI)였다. K9JB4의 돌연변이 유발에 있어서, 반응 혼합물은 1 μ l의 K9JB4 (50 ng/ μ l), 1 μ l의 각각의 프라이머 (150 ng/ μ l), 5 μ l의 10 $\times$ 반응 완충제, 1 μ l의 dNTP 믹스, 1 μ l의 Pfu 울트라(Ultra) HF DNA 폴리머라아제, 및 40 μ l의 ddH₂O를 함유하였다. K9JG3 반응 혼합물에 있어서, 1 μ l의 K9JB4 (50 ng/ μ l)를 1 μ l의 K9JG3 (50 ng/ μ l)으로 치환하였다. 하기 조건을 둘 모두의 반응에 이용하였다: 출발 온도는 95 $^{\circ}$ C, 30초였으며, 이어서 16회의 가열/냉각 사이클이었다. 각각의 사이클은 95 $^{\circ}$ C, 30초; 55 $^{\circ}$ C, 30초; 및 68 $^{\circ}$ C, 5.0분으로 이루어졌다. 상기 온도 사이클링의 완료시에, 샘플들을 4 $^{\circ}$ C에서 샘플 회수를 기다리면서 유지하였다. 1 μ l의 Dpn I (10 U/ μ l)을 각각의 반응물에 첨가하고, 혼합물을 37 $^{\circ}$ C에서 1시간 동안 인큐베이션하였다.

[0698] 제조업자의 지시에 따라 2 μ l의 각각의 돌연변이 유발 반응물로 원 샷(One Shot)(등록상표) TOP10 화학적 적격성 이. 콜라이 (인비트로젠, 카탈로그 번호 C404003)를 형질전환시켰다. 형질전환체를 LB 배지 및 100 μ g/ml의 암피실린 (카탈로그 번호 # L1004, 미국 캘리포니아주 홀리스터 소재의 테크노바 인크.)을 함유하는 한천 배지 상에 도말하고, 37 $^{\circ}$ C에서 하룻밤 인큐베이션하였다. 그 후, 다수의 형질전환체를 템플리파이™ (지이 헬스케어) 기반 DNA 서열결정용으로 선발하였는데, 상기 서열결정은 프라이머 pBAD-For (ATGCCATAGCATTTTATCC; 서열 번호 393) 및 pBAD-Rev (CTGATTAACTGTATCAGGCT; 서열 번호 394)를 이용하였다. KARI 서열이 확인된 형질전환체를 100 μ g/ml의 암피실린을 함유하는 LB 배지 내에 접종하고, 225 rpm에서 진탕하면서 33 $^{\circ}$ C에서 인큐베이션하였다. 플라스미드 DNA를 제조업자에 의해 제공된 프로토콜에 따라 쿼이프렙 스핀 미니프렙 키트 (카탈로그 번호 2706; 미국 캘리포니아주 발렌시아 소재의 쿼이젠)를 이용하여 세포로부터 단리하였다. 생성된 클론 K9JB4P 및 K9JG3P는 각각 K9JB4 및 K9JG3으로부터 유도되었다.

[0699] 변형을 이용하여, 상기에 기재된 바와 같이 추가의 부위 지정 돌연변이 유발을 수행하였다.

[0700] 변이체 K9JA1을 K9JG3P로부터 유도하였으며, 이는 프라이머 oK57E1 (GGTTTATTCGAAGGTGCGGAGGAGTGGAAAAGAGCTG; 서열 번호 397) 및 oK57E1r (CAGCTCTTTTCCACTCTCCGCACCTTCGAATAAACC; 서열 번호 398)을 이용하였다. 돌연변이 유발 반응물은 1 μ l의 K9JG3P (50 ng/ μ l), 1 μ l의 각각의 프라이머 (150 ng/ μ l), 5 μ l의 10 $\times$ 반응 완충제, 1 μ l의 dNTP 믹스, 1 μ l의 Pfu울트라 HF DNA 폴리머라아제 및 40 μ l의 ddH₂O를 함유하였다. 이. 콜라이 형질전환체에 있어서의 액체 배양물을 33 $^{\circ}$ C 대신 37 $^{\circ}$ C에서 인큐베이션하였다.

[0701] 변이체 K9SB2를 K9JB4P로부터 유도하였으며, 이는 프라이머 oY53F1 (GTAACGTTATCATTTGGTTTATACGAAGGTGCGGAGGAG; 서열 번호 399) 및 oY53F1r (CTCCTCCGCACCTTCGAATAAACCAATGATAACGTTAC; 서열 번호 400)을 이용하였다. 돌연변이 유발 반응물은 1 μ l의 K9JB4P (50 ng/ μ l), 1 μ l의 각각의 프라이머 (150 ng/ μ l), 5 μ l의 10 $\times$ 반응 완충제, 1 μ l의 dNTP 믹스, 1 μ l의 Pfu울트라 HF DNA 폴리머라아제 및 40 μ l의 ddH₂O를 함유하였다. 이. 콜라이 형질전환체에 있어서의 액체 배양물을 33 $^{\circ}$ C 대신 37 $^{\circ}$ C에서 인큐베이션하였다.

[0702] 변이체 K9SB2-K90L을 K9SB2로부터 유도하였으며, 이는 프라이머 oK90L1 (GATCTTGATCCAGATGAATTGCAGGCTACCATGTACAAAAAC; 서열 번호 401) 및 oK90L1r (GTT TTT GTA CAT GGT AGC CTG CAA TTC ATC TGG GAT CAA GAT C; 서열 번호 402)을 이용하였다. 돌연변이 유발 반응물은 2.5 μ l의 K9SB2 (50 ng/ μ l), 1 μ l의 각각의 프라이머 (150 ng/ μ l), 5 μ l의 10 $\times$ 반응 완충제, 1 μ l의 dNTP 믹스, 1 μ l의 Pfu울트라 HF DNA 폴리머라아제, 및 38.5 μ l의 ddH₂O를 함유하였다. 가열/냉각 사이클에 있어서, 55 $^{\circ}$ C, 30초의 단계를 1분이 되도록 증가시켰다. 이. 콜라이 형질전환체에 있어서의 액체 배양물을 33 $^{\circ}$ C 대신 37 $^{\circ}$ C에서 인큐베이션하였다.

[0703] 변이체 K9SB2-K90M을 K9SB2로부터 유도하였으며, 이는 프라이머 oK90M1 (CTTGATCCAGATGAAATGCAGGCTACCATGTACAAAAAC; 서열 번호 403) 및 oK90M1r (GTT TTT GTA CAT GGT AGC CTG CAT TTC ATC TGG GAT CAA G; 서열 번호 404)을 이용하였다. 돌연변이 유발 반응물은 2.5 μ l의 K9SB2 (50 ng/ μ l), 1 μ l의 각각의 프라이머 (150 ng/ μ l), 5 μ l의 10 $\times$ 반응 완충제, 1 μ l의 dNTP 믹스, 1 μ l의 Pfu울트라 HF DNA 폴리머라아제, 및 38.5 μ l의 ddH₂O를 함유하였다. 가열/냉각 사이클에 있어서, 55 $^{\circ}$ C, 30초의 단계를 1분이 되도록 증가시켰다. 이. 콜라이 형질전환체에 있어서의 액체 배양물을 33 $^{\circ}$ C 대신 37 $^{\circ}$ C에서 인큐베이션하

였다.

[표 22]

K9G9 변이체 및 조합의 아미노산 치환

변이체	아미노산 서열 번호	핵산 서열 번호	아미노산 치환
K9JB4	417	418	S56A, K57E, S58E
K9JB4P	419	420	S56A, K57E, S58E, N87P
K9JG3	421	422	Y53F, S56A, S58E, E74G
K9JG3P	423	424	Y53F, S56A, S58E, E74G, N87P
K9JA1	425	426	Y53F, S56A, K57E, S58E, E74G, N87P
K9SB2	427	428	Y53F, S56A, K57E, S58E, N87P
K9SB2-K90L	429	430	Y53F, S56A, K57E, S58E, N87P, K90L
K9SB2-K90M	431	432	Y53F, S56A, K57E, S58E, N87P, K90M

실시예 18

K_M NADH에 대한 K_M NADPH의 비가 증가된 정제된 K9G9 유도체의 동력학적 특성

K9G9 및 변이체들을 미국 특허 제8,129,162호에 기재된 바와 같이 이. 콜라이 주 Bw25113 ($\Delta ilvC$)에서 과다발 현시켰으며, 보조 인자 친화성 및 최대 속도의 더욱 정확한 결정을 취득하기 위하여 정제하였다.

발현 및 특성화에 있어서, 미국 특허 제8,129,162호에 개시된 이. 콜라이 Bw25113 ($\Delta ilvC$)의 전기적격성 주를 이. 콜라이 플라스미드 (pBAD.KARI)를 사용하여 형질전환시켰으며, 이는 바이오라드 진 펄서 II (미국 캘리포니아주 허큘리스 소재의 바이오-라드 래보러토리즈 인크.)를 이용하였다. 형질전환된 클론을 LB 배지 및 100 μ g/ml의 암피실린 (#101320-154, 미국 캘리포니아주 홀리스터 소재의 테크노바 인크.)을 함유하는 한천 배지 상에 도말하고, 37°C에서 하룻밤 인큐베이션하였다. 각각의 주에 있어서 단일한 형질전환체를 100 μ g/mL의 암피실린을 포함하는 LB 플레이트 상에 도말하였다. 이들 플레이트 각각으로부터의 단일한 콜로니를 사용하여 100 μ g/mL의 암피실린을 포함하는 10 mL LB 브로쓰에 접종하였다. 이들 배양물을 통기식 여과 뚜껑을 갖춘 125 mL의 배플형 플라스크에서 진탕하면서 37°C에서 8시간 동안 성장시켰다. 200 μ L의 이 배양물을 사용하여, 100 μ g/mL의 암피실린 및 0.2% (w/v)의 아라비노스를 포함하는 LB 브로쓰를 함유하는 통기식 여과 뚜껑을 갖춘 2개의 500 mL 배플형 플라스크에 접종하였다. 발현 배양물을 진탕하면서 37°C에서 16 - 18시간 동안 성장시켰다. 세포를 원심분리를 통하여 40 mL 분취물로 수확하고; 상청액을 버리고, 세포 펠렛을 정제할 때까지 -80°C에서 냉동시켰다.

K9G9 및 모든 변이체를 동일 공정을 이용하여 정제하였다. 각각 40 mL의 세포 배양 분취물을 대표하는 2개의 세포 펠렛을 4 mL의 버그 버스터 마스터 믹스 (노바젠 71456, 독일 다름스타트 소재)에 재현탁시키고, 실온에서 15분 동안, 이어서 60°C에서 15분 동안 인큐베이션하였다. 변성된 단백질 및 세포 잔사를 4°C에서 30분 동안 7,000 rpm에서 원심분리함으로써 펠렛화하였다. 상청액을 가만히 따라내어 남겨두고, 아크로디스크(Acrodisc) 0.2 μ m 시린지 필터 (PN4192, 미국 미시간주 앤 아버 소재의 폴(Pall))를 통하여 여과하였다. K9G9를 지이 헬스케어 하이로드(HiLoad) 26/60 수퍼덱스(Superdex) 200 겔 여과 컬럼 (17-1071-01, 영국 버킹엄셔 소재)을 이용하여 여과 열처리 세포 무함유 추출물로부터 정제하였다. 상기 컬럼을 2.0 mL/min의 유량으로 50 mM HEPES (pH 7.5) 5 mM $MgCl_2$ 완충제를 이용하여 0.2 CV 평형을 이용하여 사전 평형시킨 후 단백질 로딩을 하였다. K9G9 및 변이체들을 2.0 mL/min의 유량으로 50 mM HEPES (pH 7.5) 5 mM $MgCl_2$ 완충제로 이루어진 1.5 CV 등용매 단계에 걸쳐 용출시켰다. 2.5 mL 부피의 분획을 사형 패턴으로 Frac-950 분획 수집기 (영국 버킹엄셔)를 사용하여 수집하였다. K9G9 및 변이체들 전부는 분획 D5 - E5 또는 D6 - E4 사이에서 용출되었다. 분획들을 15 mL 아미콘 울트라(Amicon Ultra) YM-30 스핀 필터 (UFC903008, 미국 매사추세츠주 빌레리카 소재의 밀리포어 (Millipore))를 사용하여 폴딩하고, 10 mL의 100 mM HEPES (pH 6.8) 및 10 mM $MgCl_2$ 완충제를 이용하여 세척하였다. 여과액을 버리고, 정제된 단백질을 100 mM HEPES (pH 6.8) 및 10 mM $MgCl_2$ 를 함유하는 1 mL 완충제를 사용하여 상기 막으로부터 용출시켰다.

NADH 및 NADPH의 V_{max} 및 K_M 값을 결정하기 위하여, 정제된 단백질을 NAD(P)H의 다양한 농도 (0 내지 1000 μ M)에서 분석하였는데, 이는 NAD(P)H 재생 시스템과 커플링되었다. 분석을 30°C에서, 100 mM MOPS, pH 6.8, 10 mM $MgCl_2$, 1 mM EDTA, 5 mM (R/S)-아세토락테이트, 1mM 글루코스-6-포스페이트, 3 mU/ μ L의 글루코스-6-포스페

이트 데하이드로게나아제를 함유하는 완충제에서 행하였다. 상기 반응물을 3배 부피의 0.1% 포름산을 이용하여 10분 후에 켄칭하였다. DHIV 농도를 LC-MS를 사용하여 측정하였다. S-아세토락테이트의 DHIV로의 전환 속도는 고정된 시점에서의 DHIV의 생성량을 측정함으로써 결정하였다. V_{max} 및 K_m 값을 보조 인자 농도에 대한 비활성 (U/mg)의 도시에 의해 계산하였으며, 그 데이터를 미하엘리스-멘텐 등식에 피팅시켰다. 아세토락테이트 K_m 값 (NADH의 고정된 농도에서)의 측정에 의하면, 보조 인자 K_m 결정에 이용한 고정된 아세토락테이트 농도는 포화되었음이 나타났다.

[표 23]

DHIV 형성량 분석을 통하여 결정된, 정제된 K9G9 변이체에 있어서의 동력학적 값

변이체	NADPH의 V_{max} (U/mg)	NADPH의 K_m (μ M)	NADPH의 V_{max}/K_m (L/min*mg)	NADH의 V_{max} (U/mg)	NADH의 K_m (μ M)	NADH의 V_{max}/K_m (L/min*mg)
K9G9	2.2	24.1	0.091	1.9	78.2	0.024
K9JB4	2.7	249	0.011	3.4	115	0.030
K9JB4P	2	83.2	0.024	2.9	34.1	0.085
K9G3	3.1	113	0.027	2.8	106	0.026
K9G3P	1.8	33.6	0.054	2.1	18.1	0.116
K9JA1	2.6	63.4	0.041	3.4	14	0.243
K9SB2	1.7	44.8	0.038	1.8	11.6	0.155
K9SB2-K90L	2.1	173	0.012	2.4	28.6	0.084
K9SB2-K90M	1.8	245	0.007	2.2	41.3	0.053

실시예 19.

K_m NADH에 대한 K_m NADPH의 비가 증가된 K9G9 유도체의 아이소부탄을 생성

K9JB4, K9JB4P, K9JG3, K9JG3P, K9JA1, 및 K9SB2를 위한 효모 발현 플라스미드는 이. 콜라이 벡터 (pBAD.KARI) 유래의 변이형 KARI 유전자를 PmeI 및 SfiI 부위에서 pHR81-PIIv5-KARI-K9.G9 내로 서브클로닝함으로써 만들었다. pHR81-PIIv5-KARI-K9.G9 및 pHR81-PIIv5-KARI-K9.D3 (서열 번호 181)과 함께인 생성된 플라스미드들을 효모에서의 아이소부탄을 생성 및 부산물 형성에 대하여 분석하였다. 효모 경로 주를 PNY2259 (MATa ura3 $\Delta::loxP$ his3 Δ pdc6 Δ pdc1 $\Delta::P[PDC1]-DHAD|ilvD_Sm-PDC1t-P[FBA1]-ALS|alsS_Bs-CYC1t$ pdc5 $\Delta::P[PDC5]-ADH|sadB_Ax-PDC5t$ gpd2 $\Delta::loxP$ fra2 $\Delta::P[PDC1]-ADH|adh_H1-ADH1t$ adh1 $\Delta::UAS(PGK1)P[FBA1]-kivD_Lg(y)-ADH1t$ yprc15 $\Delta::P[PDC5]-ADH|adh_H1-ADH1t$ ymr226c Δ ald6 $\Delta::loxP$; 실시예 22) 숙주에서 만들었으며, 이는 경로 플라스미드 #1로서의 상기 KARI 벡터, 및 경로 플라스미드 #2로서의 pBP915 (pRS423-PFBA1-DHAD-PGPM1-hADH1; 서열 번호 182)의 동시 형질전환에 의한 것이었다. 형질전환된 세포를 히스티딘 또는 우라실을 포함하지 않는 합성 배지 (탄소 공급원으로서 1% 에탄올) 상에 플레이팅하였다. 3개의 형질전환체를 동일 배지의 신선한 플레이트로 옮겼다. 상기 형질전환체들을 세럼 바이알에서 혐기성 조건 하에서 아이소부탄을 생성에 대하여 검사하였다.

SE-Ura-His 플레이트 상의, 형질전환으로부터의 효모 콜로니는 3-5일 후에 나타났다. 각각의 변이체로부터의 3개의 콜로니를 신선한 SE-Ura-His 플레이트 상에 패치하고, 30°C에서 3일 동안 인큐베이션하였다.

성장 배지 및 절차

2가지 유형의 배지, 즉 호기성 전배양 배지 및 혐기성 배양 배지를 효모 주의 성장 절차 동안 사용하였다. 모든 화학물질은 달리 나타내지 않으면 시그마 (미국 미주리주 세인트 루이스 소재)로부터 획득하였다.

호기성 전배양 배지 (SE-Ura-His): 아미노산을 포함하지 않는 6.7 g/L의 효모 질소 베이스 (디프코, 291940, 미국 메릴랜드주 스팅스), 히스티딘, 류신, 트립토판 및 우라실을 포함하지 않는 1.4 g/L의 효모 합성 드롭아웃 배지 보충제, 0.2% 에탄올, 0.2% 글루코스, 0.01% (w/v) 류신 및 0.002% (w/v) 트립토판.

혐기성 배양 배지 (SEG-Ura-His): 50 mM의 MES (pH 5.5, 아미노산을 포함하지 않는 6.7 g/L의 효모 질소 베이스 (디프코, 291940, 미국 메릴랜드주 스팅스), 히스티딘, 류신, 트립토판 및 우라실을 포함하지 않는 1.4 g/L의 효모 합성 드롭아웃 배지 보충제, 0.1% 에탄올, 3% 글루코스, 0.01% 류신, 0.002% 트립토판, 30 mg/L의 니코틴산, 30 mg/L의 티아민 및 10 mg/L의 에르고스테롤 - 50/50 (v/v)의 트윈(Tween) /에탄올 용액 중에 만들어짐.

패치된 세포를 0.2% 글루코스 및 0.2% 에탄올을 포함하는 25 mL SEG-Ura,HIS 배지 내에 접종하고, 대략 1.5 내지 2의 표적 OD₆₀₀ 값을 성취할 때까지, 진탕하면서 30°C에서 대략 48시간 동안 뚜껑을 닫은 상태로 점진적 산소

제한 조건 하에 성장시켰다. OD₆₀₀ 값을 기록하였다. 세포를 원심분리를 통하여 펠렛화하고, 상청액을 버렸다. 세포 펠렛을 코이 언에어로빅 백(Coy Anaerobic Bag; 그래스 레이크(Grass Lake), 미국 미시간주) 내로 옮겼으며, 여기서 펠렛을 1.0 mL 혐기성 성장 배지 (SEG-Ura-His)에 재현탁시켰다. 재현탁된 세포 펠렛을 사용하여, 50 mL 세럼 병 (휘톤(Wheaton), 223748, 미국 뉴저지주 밀빌) 내의 30 mL SEG-Ura-His 배지에 0.2의 표적 초기 OD₆₀₀ 값으로 접종하였다. 모든 혐기성 배지, 세럼 바이알, 스토퍼 및 크립프를 24시간 이상 동안 혐기성 백에서 탈기시킨 후 접종하였다. 세럼 병을 스토퍼로 처리하고, 크립핑하고, 혐기성 백에서 옮기고, 240 rpm에서 진탕하면서 30℃에서 성장시켰다. 혐기성 배양물을 24 내지 72시간 동안 성장시켰으며, 이때 표적 OD₆₀₀ 값은 1.2 이상이었다. 추가의 혐기성 성장 단계에서는 접종물로서 이전의 혐기성 배양 단계로부터의 세포를 이용하였다. 3개의 형질전환체를 각각의 변이체에 대하여 평가하였다.

[0723] K9G9 KARI 변이체를 갖는 효모 주의 HPLC 분석

[0724] 샘플을 혐기성 성장 기간의 마지막에 OD₆₀₀ 값을 수득하기 위하여 그리고 HPLC 분석을 위하여 취하였다. HPLC 분석을 워터스 2695 분리 유닛, 2996 포토다이오드 어레이 검출기, 및 2414 굴절률 검출기 (미국 매사추세츠주 밀포드 소재의 워터스) - 쇼텍스 슈가(Shodex Sugar) SH-G 예비 컬럼 및 쇼텍스 슈가 SH1011 분리 컬럼 (쇼텍스, 제이엠 사이언스(Shodex, JM Science), 미국 뉴욕주 그랜드 아일랜드 소재)을 갖춤 - 를 사용하여 수행하였다. 0.5 mL/min의 유량을 이용하여 0.01 N 황산에서 등용매 용출에 의해 화합물들을 분리하였다. 워터스 엠파워 프로(Empower Pro) 소프트웨어를 이용하여 크로마토그램을 분석하였다.

[0725] 글리세롤, 아이소부탄올에 대한 몰 수율, 및 글리세롤/아이소부탄올 비를 결정하였다. 평균 및 표준 편차를 각각의 변이체에 있어서의 3회 분석에 의해 계산하였다. 그 후 스튜던트 t-검정(Student's t-test)을 이용하여 상기 값들의 차이가 K9D3 대조군 값으로부터 통계적으로 유의한지를 결정하였다. 새로운 변이체에 있어서, NADH에 대한 K_M에 비하여 NADPH에 대한 K_M 값이 증가하는 것은 예상되는 결과인, 감소된 NADPH 이용성이다. 하기 표 및 도 9에 보고된 결과는, K_M NADH에 대한 K_M NADPH의 비가 증가된 새로운 변이체들은 K9D3 및 K9G9에 비하여 글리세롤에 대한 아이소부탄올의 더 높은 비를 나타냄을 나타낸다. K9SB2는 K9D3과 비교하여 아이소부탄올 역가의 35% 증가를 보였다.

[0726] [표 24]

K9G9 변이체의 동력학적 데이터 및 아이소부탄올 데이터

	변이체	K _M (NADPH)/ K _M (NADH)	아이소부탄올/ 글리세롤 비	아이소부탄올 몰 수율	아이소부탄올 역가 (mM)
실험 1	K9D3	0.24	1.67 ± 0.02	0.581 ± 0.007	33.9 ± 1.8
	K9JB4	2.2	2.10 ± 0.06	0.603 ± 0.006	35.3 ± 1.3
	K9JG3	1.1	2.07 ± 0.06	0.598 ± 0.004	39.1 ± 1.6
실험 2	K9D3	0.24	1.75 ± 0.06	0.586 ± 0.010	63.7 ± 0.6
	K9JA1	4.5	2.24 ± 0.04	0.611 ± 0.002	78.2 ± 5.4
	K9SB2	3.9	2.21 ± 0.10	0.608 ± 0.007	77.2 ± 1.4
실험 3	K9G9	0.31	2.10 ± 0.03	0.624 ± 0.011	54.2 ± 1.5
	K9JG3P	1.9	2.56 ± 0.08	0.652 ± 0.009	66.3 ± 4.0
	K9JB4P	2.4	2.54 ± 0.07	0.654 ± 0.006	61.5 ± 3.4

[0727]

[0728] 실시예 20.

[0729] K9SB2의 오류 발생이 쉬운 PCR의 라이브러리의 제작

[0730] K9SB2의 오류 발생이 쉬운 PCR의 라이브러리를 하기와 같이 변형시켜 K9G9 라이브러리와 유사한 방식으로 제작하였다. 돌연변이 유발 PCR 혼합물은 pBAD.KARI 벡터 내의 K9SB2 9.5 μl (190 ng/ μl), 1.25 μl의 프라이머 K9G9_EZ_F1 (100 ng/ μl), 1.25 μl의 프라이머 K9G9_EZ_R1 (100 ng/ μl), 5 μl의 10× 뮤타자임 II 반응 완충제, 1 μl의 40 mM dNTP 믹스, 1.5 μl의 뮤타자임 II DNA 폴리머라아제 및 30.5 μl의 ddH₂O로 이루어졌다. "이식론 반응물"은 25 μl의 2× 이식론 효소 믹스, 3 μl의 메가프라이머 (K9SB2 돌연변이 유발

PCR 생성물, 190 ng/ μ l), 2.6 μ l의 K9SB2 주형 DNA (19 ng/ μ l), 3 μ l의 이질클론 용액 1, 및 16 μ l의 ddH₂O를 함유하였다. Dpn I 단계에 있어서, 상기 혼합물을 37°C에서 3시간 동안 인큐베이션하였다. 실시예 21에서 설명한 바와 같이 클론을 고 처리량 스크리닝에 이용하였다.

[0731] 실시예 21.

[0732] 증가된 NADH 대 NADPH 활성 비를 바탕으로 한, Km NADPH에 대한 Km NADPH의 추가의 증가된 비를 갖는 K9SB2 변이체의 스크리닝

[0733] 실시예 20에서 설명한 K9SB2 라이브러리를 NADPH 친화성이 감소된 변이체에 대하여 스크리닝하였다. NADH에 대한 Km에 비하여 NADPH에 대한 Km을 증가시키기 위한 특별한 목적에서, 스크린으로부터의 히트들을 부분적으로 정제하고, 동력학적 분석을 수행하여 NADH 및 NADPH에서의 V_{max} 및 K_m 파라미터를 결정하였다.

[0734] K9SB2 유전자 라이브러리의 고 처리량 스크리닝 분석

[0735] 하기를 제외하고서, 실시예 16에서 설명한 바와 같이 HTS를 사용하여 변이체를 스크리닝하였다. 분석 완충제는 2.4 mM의 (R/S)-아세트락테이트, 100 mM의 HEPES pH 6.8, 10 mM의 MgCl₂, 150 μ M의 NADH 또는 100 μ M의 NADPH 및 12.5 μ L의 0.5 X 세포 용해물로 이루어졌다.

[0736] 150 μ M NADH의 측정된 산화 속도에 대한 100 μ M NADPH의 측정된 산화 속도의 비를 각각의 변이체 및 양성 대조 웰 (플레이트당 2개)에 대하여 계산하였다. 변이체 웰은, 상기 NADH 속도가 0.6 OD/hr보다 크고 상기 속도 비(NADPH/NADH)가 0.37보다 작을 경우 (양성 대조군의 평균보다 더 낮은 3개의 표준 편차) 초기 히트를 함유하는 것으로 간주하였다. 총 218개의 히트를 4947가지의 잠재적인 변이체의 풀로부터 확인하였다. 이들 초기 히트들을 결합시켜 추가 분석용의 더욱 작은 라이브러리를 형성하였다.

[0737] 결합된 초기 히트 라이브러리를 생물학적 삼중체로 성장시키고, 세포 무함유 추출물을 제조하고, 상기에 기재된 바와 같이 분석하였다. 그 후, 속도 비들을 상기와 같이 변이체 및 양성 대조군에 대하여 계산하였다. 상술된 동력학적 분석용으로 선발된 최종 히트는 하기 기준을 충족시켰다: NADPH/NADH 속도 비는 0.45 미만이었으며, NADH 속도는 0.6 OD/hr 초과였고, 3개의 생물학적 레플리케이트 중 적어도 2개는 이 기준을 통과하였다. 107가지의 변이체를 확인하였다.

[0738] 또한 데이터를 분석하여 NADH 보조 인자에 의한 S-아세트락테이트의 DHIV로의 전환 속도가 더 높은 변이체들을 확인하였다. NADH 산화의 평균 속도 및 표준 편차를 모든 양성 대조군에 대하여 계산하였다. 변이체는 NADH 산화 속도가 양성 대조군의 속도 (2.524 OD/hr)의 속도보다 3 이상의 표준 편차만큼 더 높을 경우 잠재적 히트로 간주하였다. 68가지의 변이체를 확인하였으며, 서열 분석에 의해 17가지가 적어도 하나의 아미노산 치환을 가짐을 결정하였다. 치환 T93A 및 T93I 각각은 2회 나타났으며, 변이체 2017 B12 및 D6을 추가의 분석용으로 선발하였다.

[0739] 이차 HTS 스크리닝으로부터 확인된 107가지의 변이체의 DNA 서열결정은 프라이머 pBAD-For (ATGCCATAGCATTTTATCC; 서열 번호 393) 및 pBAD-Rev (CTGATTTAATCTGTATCAGGCT; 서열 번호 394)를 이용하여 템플리파이™ (지이 헬스케어)를 사용하여 달성하였다. 105개의 서열은 모 서열과 달랐으며, 아미노산 치환들은 하기 두 표 중 첫 번째 표에 열거되어 있다.

[0740] NADH 속도 스크리닝으로부터 확인된 68가지의 변이체의 DNA 서열결정은 프라이머 pBAD-For (ATGCCATAGCATTTTATCC; 서열 번호 393) 및 pBAD-Rev (CTGATTTAATCTGTATCAGGCT; 서열 번호 394)를 이용하여 템플리파이™ (지이 헬스케어)를 사용하여 달성하였다. 17개의 서열은 야생형과 달랐으며, 반복적으로 나타난 2가지의 치환의 아미노산 치환이 하기 두 번째 표에 열거되어 있다.

[0741] [표 25]

K9SB2 변이체의 아미노산 치환

변이체	서열	아미노산 치환
K9SB2	427	Y53F, S56A, K57E, S58E, N87P
2011 A2	433	Y53F, G55D, S56A, K57E, S58E, N87P
2011 A3	434	Y53F, S56A, K57E, S58E, N87P, M94I
2011 A5	435	Y53F, S56A, K57E, S58E, M83I, N87P, L185M, E217D
2011 A7	436	Y53F, S56A, K57E, S58E, N87P, D98V
2011 A9	437	Y53F, S56A, K57E, S58E, F67I, N87P
2011 A11	438	Y53F, S56A, K57E, S58E, N87P, M94T, K126E, T273A
2011 B1	439	Y53F, S56A, K57E, S58E, N87P, M94I, A279T
2011 B2	440	Y53F, S56A, K57E, S58E, N87P, Q91L
2011 B3	441	Y53L, S56A, K57E, S58E, N87P
2011 B4	442	Y53F, S56A, K57E, S58E, N87P, P135T
2011 B7	443	Y53F, S56A, K57E, S58E, N87P, L185M
2011 B8	444	Y53F, S56A, K57E, S58E, N87P, C233S, F296Y
2011 B10	445	Y53F, S56A, K57E, S58E, N87P, A303D
2011 C1	446	Y53F, S56A, K57E, S58E, E63K, N87P, G251D, K294R
2011 C3	448	Y53F, S56A, K57E, S58E, A72V, N87P, N102Y, F189I, Y245H
2011 C6	449	Y53F, S56A, K57E, S58E, I84F, N87P
2011 C7	450	E13V, Y53F, S56A, K57E, S58E, M94I, N87P, T141I
2011 C8	451	Y53F, S56A, K57E, S58E, A72V, N87P, N102Y, F189I, Y245H
2011 C9	447	Y53F, S56A, K57E, S58E, N87P, E194D
2011 C10	452	Y53F, A56G, K57E, S58E, K60N, N87P
2011 C12	453	Y53F, S56A, K57E, S58E, I84L, N87P, N97T
2011 D1	454	L39M, Y53F, S56A, K57E, S58E, E68G, N87P
2011 D2	455	Y53F, S56A, K57E, S58E, N87P, M94I, V307I
2011 D3	456	Y53F, S56A, K57E, S58E, F67I, N87P
2011 D4	457	Y53F, S56A, K57E, S58E, N87P, P135L, A202V
2011 D5	458	Y53F, S56A, K57E, S58E, N87P, G164S, G199A
2011 D6	500	Y53F, S56A, K57E, S58E, N87P, S247C
2011 D8	459	Y53F, S56A, K57E, S58E, N87P, N116I
2011 D9	460	Y53F, S56A, K57E, S58E, N87P, K90M
2011 D11	461	Y53F, S56A, K57E, S58E, N87P, M94L, T259I
2011 D12	462	Y53F, S56A, K57E, S58E, M83K, N87P
2011 E3	463	Y53F, S56A, K57E, S58E, N87P, I122V, L297W
2011 E4	464	Y53F, S56A, K57E, S58E, N87P, A112S, Q160R
2011 E8	465	Y53F, S56A, K57E, S58E, N87P, V142I, P320L
2011 E11	466	Y53F, S56A, K57E, S58E, N87P, Q91L
2011 F2	467	Y53F, S56V, K57E, S58E, N87P, A210T
2011 F4	468	Y53F, S56A, K57E, S58E, N87P, F189I
2011 F6	469	Y53F, S56G, K57E, S58E, K60N, N87P
2011 F9	470	A41V, Y53F, S56A, K57E, S58E, N87P, S305P
2011 F10	471	Y53F, S56A, K57E, S58E, L85M, N87P

[0742]

2011 G1	472	E13V, Y53F, S56A, K57E, S58E, M94I, N87P, T141I
2011 G3	473	H35Q, Y53F, S56A, K57E, S58E, N87P
2011 G4	474	Y53F, S56A, K57E, S58E, I84F, N87P
2011 G8	475	A26T, Y53F, S56A, K57E, S58E, N87P, K90M
2011 G9	476	Y53F, S56T, K57E, S58E, N87P
2011 G10	477	Y53F, S56A, K57E, S58E, I84F, N87P
2011 H1	478	Y53F, S56A, K57E, S58E, N87P, M94I, V156L
2011 H5	479	Y53F, S56A, K57E, S58E, I84L, N87P
2011 H7	480	Y53F, S56T, K57E, S58E, N87P
2011 H9	481	Y53F, G55C, S56A, K57E, S58E, N87P
2012 A2	482	Y53F, S56A, K57E, S58E, R61G, I86V, N87P
2012 A7	483	Y53F, S56A, K57E, S58E, T71S, A76V, N87P
2012 A8	484	Y53F, S56A, K57E, S58E, N87P, M212T
2012 A9	485	H35Q, Y53F, S56A, K57E, S58E, A72T, N87P
2012 A10	486	A36T, Y53F, S56A, K57E, S58E, N87P
2012 A11	487	Y53F, S56A, K57E, S58E, N87P, S247T
2012 B3	488	Y53F, S56A, K57E, S58E, N87P, P135S
2012 B4	489	Y53F, S56A, K57E, S58E, W59R, N87P, K278E
2012 B5	490	Y53F, S56V, K57E, S58E, N87P, I234V
2012 B6	491	Y30H, Y53F, S56A, K57E, S58E, N87P
2012 B9	492	I50M, Y53F, S56A, K57E, S58E, N87P
2012 C2	493	Y53L, S56A, K57E, S58E, N87P
2012 C3	494	Y53F, S56A, K57E, S58E, N87P, F115V, T191S, V208I, C209W, F292I
	495	Y53F, S56A, K57E, S58E, N87P, M94I
2012 C6	496	Y53F, S56A, K57E, S58E, F67L, N87P
2012 C8	497	Y53F, S56A, K57E, S58E, N87P, M94I, M169T
2012 C10	498	Y53F, S56A, K57E, S58E, F67I, N87P, T276I
2012 D1	499	Y53F, S56A, K57E, S58E, I84F, N87P, M132T
2012 D8	501	Y53F, S56A, K57E, S58E, N87P, P135S
2012 D11	502	K8N, Y53F, S56A, K57E, S58E, N87P, K90M, T141I
2012 E5	503	Y53F, S56A, K57E, S58E, I84L, N87P
2012 E9	504	Y53F, S56A, K57E, S58E, N87P, V142I, T191S, C233S
2012 F1	505	Y53F, S56A, K57E, S58E, N87P, H235Y
2012 F2	506	Y53F, S56V, K57E, S58E, N87P, V232D
2012 F3	507	Y53F, S56A, K57E, S58E, N87P, K90M, V142I, T187S
2012 F4	508	Y53F, S56A, K57E, S58E, N87P, M94I, G149D
2012 F7	509	E13V, Y53F, S56A, K57E, S58E, N87P, M94I, T141I
2012 F10	510	Y53F, S56A, K57E, S58E, Q65H, N87P, F189I
2012 F12	511	Y53F, S56V, K57E, S58E, N87P
2012 G3	512	Y53F, S56T, K57E, S58E, N87P, R190S
2012 G4	513	Y53F, S56A, K57E, S58E, N87P, N102S, V142I
2012 G5	514	Y53F, S56A, K57E, S58E, I84L, N87P
2012 G8	515	Y53F, S56A, K57E, S58E, K77N, N87P, A92V
2012 G9	516	Y53F, S56A, K57E, S58E, N87P, M94I, V307I
2012 G10	517	Y53F, S56A, K57E, S58E, N87P, T195I
2012 G12	518	Y53F, S56A, K57E, S58E, N87P, F309I
2012 H1	519	Y53F, S56A, K57E, S58E, N87P, K90T, A180S
2012 H3	520	Y53F, S56A, K57E, S58E, W59C, N87P

[0743]

2012 H7	521	Y53F, S56A, K57E, S58E, N87P, M94I, A202T
2012 H9	522	H35N, Y53F, S56A, K57E, S58E, N87P
2012 H11	523	Y53F, S56A, K57E, S58E, A72V, N87P, L211M, I240M
2013 A2	524	Y53F, S56T, K57E, S58E, N87P, Q288H
2013 A4	525	Y53F, S56A, K57E, S58E, L85M, N87P
2013 A5	526	L52S, Y53F, S56A, K57E, S58E, N87P
2013 B2	527	A36T, Y53F, S56A, K57E, S58E, N87P, V203I
2013 B5	528	Y53F, S56A, K57E, S58E, N87P, P135T
2013 B7	529	I9M, Y53F, S56A, K57E, S58E, N87P, K90E
2013 B8	530	Y53F, G55C, S56A, K57E, S58E, N87P
2013 B9	531	Y53F, S56V, K57E, S58E, N87P
2013 B11	532	A38V, Y53F, S56A, K57E, S58E, N87P
2013 C1	533	Y53L, S56A, K57E, S58E, N87P, M237I
2013 C6	534	K23M, Y53F, S56A, K57E, S58E, N87P, E194D
2013 C8	535	Y53F, S56A, K57E, S58E, N87P, P135T
2013 C12	536	Y53F, S56A, K57E, S58E, A72V, N87P, T93S, A176V, H235Y
	537	Y53F, S56A, K57E, S58E, N87P, M94R, K310M

[0744]

[0745] [표 26]

K9SB2 변이체의 아미노산 치환

변이체	아미노산 서열 번호; 핵산 서열 번호	아미노산 치환
K9SB2	427	Y53F, S56A, K57E, S58E, N87P
2017 B12	639; 640	Y53F, S56A, K57E, S58E, N87P, T93I
2017 D6	641; 642	Y53F, S56A, K57E, S58E, N87P, T93A

[0746]

[0747] 부분 정제된 K9SB2 변이체 단백질의 동력학적 분석

[0748] 미국 특허 제8,129,162호에 기재된 이. 콜라이 주 *Bw25113* ($\Delta ilvC$)를 사용하여 이차 HTS 스크리닝으로부터의 107가지의 변이체 및 양성 대조군 K9SB2를 발현시켰다. 아카이브 플레이트로부터의 클론을 깊은 96웰 플레이트 내에 접종하였다. 각각의 웰은 해동된 아카이브 플레이트로부터의 세포 3.0 μ l, 유도제로서 0.02%(w/v) 아라비노스 및 암피실린 100 μ g/ml을 함유하는 LB 배지 200 μ l를 함유하였다. 세포를 진탕시키면서 (900 rpm) 80% 습도에서 37℃에서 하룻밤 성장시키고, 원심분리 (4000 rpm, 4℃에서 7 min) (75004251, 미국 일리노이주 록포드 소재의 서모 사이언티픽)에 의해 수확하고, 세포 펠렛을 이후의 분석을 위하여 -80℃에서 보관하였다.

[0749] 깊은 웰의 플레이트 중의 냉동 세포 펠렛을 실온에서 30분 동안 동시에 해동시켰다. 75 μ l의 50% 버그버스터 (노바겐 71456, 독일 다름슈타트) (v/v, 물 중)를 각각에 첨가하고, 세포를 플레이트 진탕기를 이용하여 현탁시켰다. 50% 버그 버스터 중 세포 현탁물을 실온에서 30분 동안 인큐베이션하고, 그 후 60℃에서 15분간 인큐베이션하였다. 세포 잔사 및 변성된 열불안정성 단백질을 원심분리 (4000 rpm, 4℃에서 15 min) (75004251, 미국 일리노이주 록포드 소재의 서모 사이언티픽)에 의해 펠렛화하고, 75 μ L의 상청액을 평평한 바닥의 96웰 플레이트 (코닝, 3370, 미국 뉴욕주 코닝 소재)로 옮기고, 75 μ L의 100 mM HEPES (pH 6.8), 100 mM KCl, 10 mM $MgCl_2$ 로 2배 희석시켰다.

[0750] 전체 단백질을 쿠마시 플러스 (서모 사이언티픽, #23238, 미국 일리노이주 록포드)를 이용하여 브래드포드 분석법을 사용함으로써 결정하였다. BSA를 표준물로 이용하였다. 캐리 300 분광 광도계 (미국 델라웨어주 월밍턴 소재의 애질런트 테크놀로지스)를 사용하여 595 nm에서 흡광도를 측정함으로써 단백질의 농도를 측정하였다.

[0751] NADH 및 NADPH에 대한 V_{max} 및 K_m 값을 결정하기 위하여, 부분 정제된 단백질을 NADH (20, 30, 40, 60, 80, 120, 200 및 300 μ M) 및 NADPH (60, 80, 120, 200, 300 및 400 μ M)의 다양한 농도에서 분석하였다. 분석은 30℃에서 100 mM의 HEPES (pH 6.8), 10 mM의 $MgCl_2$, 100 mM의 KCl 및 4.8 mM의 R/S-아세토락테이트에서 행하였다. 0.005 내지 0.015 mg/mL의 전체 단백질을 상기 분석물에 첨가하였다. S-아세토락테이트의 DHIV로의 전환 속도는 스펙트라맥스(Spectramax) 384 플러스 플레이트 판독기 (미국 캘리포니아주 서니베일 소재의 몰레큘러 디바이스즈(Molecular Devices))를 사용하여 340 nm에서 NAD(P)H의 산화를 모니터링함으로써 측정하였다. 6220 $M^{-1}cm^{-1}$ 의 몰 흡광 계수를 이용하여 활성을 계산하였다. V_{max} 및 K_m 값을 보조 인자 농도에 대한 비활성 (U/mg)의 도시에 의해 계산하였으며, 그 데이터를 칼레이다그래프 소프트웨어 (미국 펜실베이니아주 리딩 소재의 시너지)를 사용하여 미하엘리스-멘텐 등식에 피팅시켰다.

[0752] [표 27]

NAD(P)H 소비량 분석을 통하여 결정한, 부분적으로 정제된 K9SB2 변이체에 있어서의 동력학적 값

변이체	NADPH의 V_{max} (U/mg)	NADPH의 K_m (μ M)	NADPH의 V_{max}/K_m (L/min*mg)	NADH의 V_{max} (U/mg)	NADH의 K_m (μ M)	NADH의 V_{max}/K_m (L/min*mg)
K9SB2	1.79	153	0.012	2.10	87.4	0.024
2011 A2	1.48	897	0.002	1.94	71.2	0.027
2011 A3	1.43	371	0.004	1.33	44.1	0.030
2011 A5	0.93	109	0.009	0.23	17.4	0.017
2011 A7	2.27	334	0.007	1.70	69.2	0.025
2011 A9	2.09	266	0.008	0.40	n/d	n/a
2011 A11	2.21	294	0.008	1.25	25.8	0.048
2011 B1	1.79	421	0.004	64.5	55171	0.001
2011 B2	0.33	254	0.001	0.02	264	0.007
2011 B3	1.87	505	0.004	0.68	225	0.003
2011 B4	0.36	294	0.001	0.26	171	0.002
2011 B7	n/d	1.03	n/a	0.48	25.5	0.019
2011 B8	0.83	109.5	0.008	0.61	28.7	0.021
2011 B10	0.88	171	0.005	0.46	7.53	0.061
2011 C1	1.06	404	0.003	1.51	191	0.008
2011 C3	1.08	844	0.001	2.53	500	0.005
2011 C6	1.19	388	0.003	1.90	189	0.010
2011 C7	0.71	946	0.001	1.95	457	0.004
2011 C8	1.73	1546	0.001	3.73	750	0.005
2011 C9	1.02	123	0.008	2.29	177	0.013
2011 C10	1.02	656	0.002	3.84	899	0.004
2011 C12	2.74	244	0.011	3.08	99	0.031
2011 D1	4.68	501	0.009	3.54	80.4	0.044
2011 D2	2.34	547	0.004	2.18	77.3	0.028
2011 D3	0.05	306	0.0002	0.05	44.3	0.001
2011 D4	0.47	857	0.001	0.44	91.6	0.005
2011 D5	0.75	550	0.001	0.42	57.1	0.007
2011 D6	0.70	200	0.004	0.52	25.11	0.021
2011 D8	0.04	214	0.0002	0.04	38.9	0.001
2011 D9	0.18	407	0.0004	0.16	50.2	0.003
2011 D11	0.78	185	0.004	0.61	15.0	0.041
2011 D12	0.74	190	0.004	0.80	39.0	0.021
2011 E3	0.77	163	0.005	1.54	128	0.012
2011 E4	1.59	270	0.006	3.78	234	0.016
2011 E8	0.91	435	0.002	2.16	252	0.009
2011 E11	5.56	6466	0.001	3.06	511	0.006
2011 F2	0.39	692	0.001	1.79	136	0.013
2011 F4	2.07	242	0.009	2.00	68.3	0.029
2011 F6	1.19	946	0.001	1.35	231	0.006
2011 F9	0.57	269	0.002	0.45	27.1	0.017

[0753]

2011 F10	1.43	390	0.004	1.55	79.8	0.019
2011 G1	1.31	1533	0.001	0.50	40.1	0.013
2011 G3	0.61	1003	0.001	1.00	283	0.004
2011 G4	1.00	316	0.003	2.13	174	0.012
2011 G8	0.90	482	0.002	1.22	106	0.012
2011 G9	0.30	530	0.001	3.48	549	0.006
2011 G10	0.93	358	0.003	2.13	238	0.009
2011 H1	0.96	218	0.004	1.57	114	0.014
2011 H5	1.17	81.2	0.014	1.85	86.5	0.022
2011 H7	0.20	435	0.001	1.25	173	0.007
2011 H9	1.30	741	0.002	1.55	177	0.009
2012 A2	1.71	264	0.007	1.58	55.4	0.029
2012 A7	1.98	215	0.009	1.87	67.3	0.028
2012 A8	1.19	91.8	0.013	1.24	22.7	0.055
2012 A9	0.44	481	0.001	0.34	38.1	0.009
2012 A10	1.21	340	0.004	1.31	66.7	0.020
2012 A11	1.99	342	0.006	1.37	35.9	0.038
2012 B3	0.88	1214	0.001	0.42	63.8	0.007
2012 B4	4.34	1593	0.003	1.66	95.1	0.018
2012 B5	4.88	7389	0.001	1.19	85.2	0.014
2012 B6	3.72	428	0.009	2.23	51.3	0.044
2012 B9	1.17	523	0.002	0.87	63.7	0.014
2012 C2	4.43	923	0.005	3.16	189	0.017
2012 C3	1.40	203	0.007	1.80	68.9	0.026
2012 C5	1.73	348	0.005	3.66	268	0.014
2012 C6	2.18	234	0.009	2.90	103	0.028
2012 C8	1.53	394	0.004	2.63	194	0.014
2012 C10	1.12	286	0.004	1.41	78.9	0.018
2012 D1	0.72	599	0.001	0.58	64.5	0.009
2012 D8	1.17	1528	0.001	0.47	73.6	0.006
2012 D11	0.43	334	0.001	0.52	27.7	0.019
2012 E5	2.13	257	0.008	4.77	313	0.015
2012 E9	1.07	1326	0.001	n/d	n/d	n/a
2012 F1	1.70	272	0.006	1.48	62.5	0.024
2012 F2	0.39	925	0.0004	0.79	97.8	0.008
2012 F3	1.40	2213	0.0006	0.94	142	0.007
2012 F4	2.70	719	0.004	1.50	72.9	0.021
2012 F7	0.86	840	0.001	0.77	117	0.007
2012 F10	1.92	170	0.011	1.75	27.4	0.064
2012 F12	0.90	1582	0.0006	1.75	117	0.015
2012 G3	1.47	1003	0.002	1.07	127	0.008
2012 G4	0.73	615	0.001	0.63	81.6	0.007
2012 G5	1.92	240	0.008	2.00	83.5	0.024
2012 G8	1.17	315	0.004	0.99	58.7	0.017
2012 G9	2.41	717	0.003	1.37	91.5	0.015
2012 G10	1.06	400	0.003	0.71	39.8	0.018
2012 G12	1.58	147	0.011	2.00	70.0	0.029
2012 H1	1.49	195	0.008	1.74	68.9	0.025
2012 H3	14.98	7389	0.002	1.45	99.0	0.015

[0754]

2012 H7	1.14	246	0.005	1.30	76.2	0.017
2012 H9	0.37	210	0.002	n/d	n/d	n/a
2012 H11	0.65	162	0.004	0.62	32.0	0.019
2013 A2	0.58	285	0.002	0.64	71.0	0.009
2013 A4	0.63	188	0.003	0.86	81.5	0.011
2013 A5	0.61	886	0.001	0.88	210	0.004
2013 B2	0.62	282	0.002	0.71	70.4	0.010
2013 B5	6.68	7389	0.001	0.083	150	0.006
2013 B7	1.22	433	0.003	1.12	79.0	0.014
2013 B8	0.42	90.7	0.005	1.27	191	0.007
2013 B9	5.31	13970	0.0004	1.38	217	0.006
2013 B11	0.48	212	0.002	0.60	63.2	0.010
2013 C1	0.49	149	0.003	0.68	64.8	0.0105
2013 C6	0.54	163	0.003	0.36	24.4	0.015
2013 C8	2.87	3752	0.001	0.70	188	0.004
2013 C12	0.75	495	0.002	0.79	115	0.007
2013 D1	1.31	1608	0.001	0.87	188	0.005

[0755]

[0756]

실시예 22:

[0757]

주 PNY2259의 제작

[0758]

이 실시예의 목적은 *adh1Δ* 유전자좌에서 PNY2238에서 *kivD_L1(y)*의 염색체 카피를 *kivD_Lg(y)*로 대체하는 데

사용되는 제작물들의 조립을 설명하는 것이다.

- [0759] 결실/통합은 형질전환체의 선발을 위한 *URA3* 유전자 및 표적 영역의 상류 및 하류의 상동성 영역을 함유하는 PCR 생성물을 이용하여 상동 재조합에 의해 생성하였다. *URA3* 유전자를 상동 재조합에 의해 제거하여 스카리스 결실/통합을 생성하였다. *kivD_Lg(y)*를 통합시키기 위한 플라스미드는 사카로마이세스 세레비지에의 *ADH1* 유전자와 내로 UAS(PGK1)P[FBA1]-*kivD_Ll(y)*를 통합시키기 위하여 제작한 플라스미드로부터 유도하였다. *ADH1* 유전자와 내로의 UAS(PGK1)P[FBA1]-*kivD_Ll(y)*의 통합에 사용하는 플라스미드의 제작을 하기에 설명한다. 플라스미드들을 pUC19-URA3MCS에서 제작하였다.
- [0760] *ADH1* 결실/UAS(PGK1)P[FBA1]-*kivD_Ll(y)* 통합 플라스미드의 제작
- [0761] 사카로마이세스 세레비지에에서의 발현용으로 코돈 최적화된, 락토코커스 락티스 유래의 *kivD* 코딩 영역, *kivD_Ll(y)*를 주형으로서 pLH468 (서열 번호 139)을 사용하여, *PmeI* 제한효소 부위를 함유하는 프라이머 oBP562 (서열 번호 197), 및 *ADH1* 단편 B의 5' 말단에 대하여 상동성을 갖는 5' 테일을 함유하는 프라이머 oBP563 (서열 번호 198)을 이용하여 증폭시켰다. *ADH1* 단편 B를 사카로마이세스 세레비지에 CEN.PK 113-7D 게놈 DNA로부터, *kivD_Ll(y)*의 3' 말단에 대하여 상동성을 갖는 5' 테일을 함유하는 프라이머 oBP564 (서열 번호 199), 및 *FseI* 제한효소 부위를 함유하는 프라이머 oBP565 (서열 번호 200)를 이용하여 증폭시켰다. PCR 생성물을 PCR 정제 키트 (미국 캘리포니아주 발렌시아 소재의 퀴아젠)를 사용하여 정제하였다. *kivD_Ll(y)*-*ADH1* 단편 B는 *kivD_Ll(y)* 및 *ADH1* 단편 B의 PCR 생성물을 혼합하고, 프라이머 oBP562 (서열 번호 197) 및 oBP565 (서열 번호 200)를 사용하여 증폭시킴으로써 중첩 PCR에 의해 생성하였다. 생성된 PCR 생성물을 *PmeI* 및 *FseI*으로 절단하고, T4 DNA 라이가아제를 사용하여 적절한 효소를 이용한 절단 후에 pUC19-URA3MCS의 상응하는 부위 내로 라이게이션시켰다. *ADH1* 단편 A를, *SacI* 제한효소 부위를 함유하는 프라이머 oBP505 (서열 번호 201), 및 *AscI* 제한효소 부위를 함유하는 프라이머 oBP506 (서열 번호 202)을 사용하여 게놈 DNA로부터 증폭시켰다. *ADH1* 단편 A의 PCR 생성물을 *SacI* 및 *AscI*으로 절단하고, T4 DNA 라이가아제를 사용하여 *kivD_Ll(y)*-*ADH1* 단편 B를 함유하는 플라스미드의 상응하는 부위 내로 라이게이션시켰다. *ADH1* 단편 C를, *PacI* 제한효소 부위를 함유하는 프라이머 oBP507 (서열 번호 203), 및 *SalI* 제한효소 부위를 함유하는 프라이머 oBP508 (서열 번호 204)를 사용하여 게놈 DNA로부터 증폭시켰다. *ADH1* 단편 C의 PCR 생성물을 *PacI* 및 *SalI*로 절단하고, T4 DNA 라이가아제를 사용하여 *ADH1* 단편 A-*kivD_Ll(y)*-*ADH1* 단편 B를 함유하는 플라스미드의 상응하는 부위 내로 라이게이션시켰다. 하이브리드 프로모터 UAS(PGK1)-P_{FBA1} (서열 번호 406)를 벡터 pRS316- UAS(PGK1)-P_{FBA1}-GUS로부터, *AscI* 제한효소 부위를 함유하는 프라이머 oBP674 (서열 번호 205), 및 *PmeI* 제한효소 부위를 함유하는 프라이머 oBP675 (서열 번호 206)를 사용하여 증폭시켰다. UAS(PGK1)-P_{FBA1} PCR 생성물을 *AscI* 및 *PmeI*로 절단하고, T4 DNA 라이가아제를 사용하여 *kivD_Ll(y)*-*ADH1* 단편 ABC를 함유하는 플라스미드의 상응하는 부위 내로 라이게이션시켜 pBP1181을 생성하였다.
- [0762] pBP1716 및 pBP1719의 제작
- [0763] *kivD_Ll(y)*를 *ADH1* 결실/UAS(PGK1)P[FBA1]-*kivD_Ll(y)* 통합 플라스미드 pBP1181로부터 제거하였다. 플라스미드를 *PmeI* 및 *FseI*로 절단하고, 큰 DNA 단편을 아가로스 겔, 이어서 겔 추출 키트 (퀴아젠) 상에서 정제하였다. *ADH1* 단편 B를, *PmeI* 제한효소 부위를 함유하는 프라이머 oBP821 (서열 번호 407), 및 *FseI* 제한효소 부위를 함유하는 프라이머 oBP484 (서열 번호 408)를 사용하여 pBP1181로부터 증폭시켰다. *ADH1* 단편 B의 PCR 생성물을 *PmeI* 및 *FseI*로 절단하고, T4 DNA 라이가아제를 사용하여, 겔 정제된 큰 DNA 단편의 상응하는 부위 내로 라이게이션시켰다. *kivD_Ll(y)*의 3' 500 bp에 상응하는 PCR 단편을 PNY1528에 있어서 *kivD_Ll(y)*의 표적화된 결실을 위하여 생성된 벡터 내로 클로닝하였다. 상기 단편을, *NotI* 제한효소 부위를 함유하는 프라이머 oBP822 (서열 번호 409), 및 *PacI* 제한효소 부위를 함유하는 oBP823 (서열 번호 410)을 사용하여 pBP1181로부터 증폭시켰다. 상기 단편을 *NotI* 및 *PacI*으로 절단하고, 적절한 제한 효소를 이용한 절단 후 *kivD_Ll(y)* 결실을 갖는 상기 플라스미드 내의 *URA3*의 하류의 상응하는 부위 내로 T4 DNA 라이가아제를 사용하여 라이게이션시켰다. 생성된 플라스미드를 pBP1716으로 표기하였다.
- [0764] 사카로마이세스 세레비지에에서의 발현용으로 코돈 최적화된 리스테리아 그라이이 유래의 *kivD* 코딩 영역 (서열 번호 411), *kivD_Lg(y)*는 DNA2.0 (미국 캘리포니아주 멘로 파크 소재)에 의해 합성되었다. *kivD_Lg(y)*를, *PmeI* 제한효소 부위를 함유하는 프라이머 oBP828 (서열 번호 412), 및 *PmeI* 제한효소 부위를 함유하는 oBP829 (서열 번호 413)를 사용하여 증폭시켰다. 생성된 PCR 생성물을 *PmeI*로 절단하고, 적절한 효소를 이용한 절단 후에 T4 DNA 라이가아제를 사용하여 pBP1716 내의 상응하는 부위 내로 라이게이션시켰다. 클로닝된 유전자의 배향을 프라이머 FBAp-F (서열 번호 414) 및 oBP829 (서열 번호 413)를 사용하여 PCR에 의해 점검하였다.

kivD_Lg(y)를 올바른 배향으로 포함하는 단리체를 pBP1719로 표기하였다.

[0765] 주 PNY2259의 제작

[0766] kivD_Ll(y) 결실/kivD_Lg(y) 통합 카세트를 pBP1719로부터 프라이머 oBP505 (서열 번호 201) 및 oBP823 (서열 번호 410)을 사용하여 증폭시켰다. PNY2238의 적격성 세포를 만들고, 프로즌-이지 이스트 트랜스포메이션 II 키트 (자이모 리서치; 미국 캘리포니아주 오렌지 소재)를 상기 PCR 생성물로 형질전환시켰다. 형질전환 혼합물을 30℃에서 1% 에탄올을 보충한 우라실이 결여된 합성 완전 배지 상에 플레이팅하였다. 형질전환 주를 PCR (점프스타트(JumpStart)(TM) REDTaq (c) 레디믹스(ReadyMix)(TM))에 의해 프라이머 Ura3-end F (서열 번호 222) 및 HY-50 (서열 번호 415)을 사용하여 스크리닝하였다. 형질전환체를 YPE (1% 에탄올)에서 성장시키고, 1% EtOH가 보충되고 5-플루오로-오로트산 (0.1%)을 함유하는 합성 완전 배지 상에서 30℃에서 플레이팅하여, UR43 마커가 상실된 단리주를 선발하였다. kivD_Ll(y)의 결실 및 kivD_Lg(y)의 통합을 프라이머 HY-50 및 oBP834 (서열 번호 416)을 사용하여 PCR에 의해 확인하였다. 하나의 올바른 단리체는 동일 유전자좌에 kivD_Lg(y)를 함유하였으며, 이는 PNY2238에서 kivD_Ll(y)와 동일한 프로부터로부터 발현되었는데, 이를 PNY2259로 표기하였다.

[0767] 실시예 23

[0768] NADH에 대하여 보조 인자 선호성을 갖는 변이체들을 확인하기 위한 2 부위-포화 유전자 라이브러리의 제작

[0769] 실시예 4에서, 축퇴성 코돈 NNK를 갖는 프라이머를 사용하였다 (N은 모든 4개의 뉴클레오타이드 A, C G 및 T를 나타내는 반면, K는 G 및 T를 나타냄). 이 실시예에서, K9 KARI의 위치 53, 56 및 58에 대하여 A, C, D, E, F, G, H, I, L, M, N, P, Q, V, W, 또는 Y의 각각의 개별 아미노산 변화를 코딩하는 프라이머들을 함유하는 프라이머 혼합물들을 이용하였으며, S, T, K, 및 R로의 치환은 이들 위치에 바람직하지 못하기 때문에 배제하였다. K9 KARI의 3개의 NADPH 포스페이트 결합 부위 (53, 56 및 58)를 표적화하는 포화 라이브러리의 크기는 (실시예 4에서와 같이 NNK 축퇴성 암호 프라이머를 사용하여 32*32*32 또는 32,768가지의 변이체와 비교하여) 4,096이다.

[0770] 하나의 라이브러리 제작 방법은 위치 58로부터 출발하였다. 먼저 동일한 위치를 표적화하는 모든 프라이머들을 혼합함으로써 프라이머 혼합물을 만들었다 (예를 들어, 프라이머 혼합물, K9_53f는 위치 53을 표적화하는 동일한 물의 모든 16가지의 정방향 프라이머들 (하기 표에 열거됨)을 혼합함으로써 만들었다). 이와 유사하게, K9_56f 및 K9_58f를 제조하였다. 공유된 역방향 프라이머는 K9_191G_112210r (서열 번호 174): GGTTCAGTTTCGCCTCTGAAGGTAGTTTC (이 실시예에서 SR로 칭함)이다. 위치 58에서의 돌연변이를 먼저 PCR을 통하여 AS6F1 내로 도입하였다. 돌연변이 유발 절차는 실시예 4에서 설명한 것과 유사하다. 간략하게는, K9_58f 및 SR을 포스포릴화하였다. 그 후, 포스포릴화된 프라이머를 직접적으로 사용하여 위치 58에서의 돌연변이를 AS6F1 내로 도입하였으며, 이는 유에스비 체인지_잇(USB Change_It) 키트 (미국 오하이오주 클레블랜드 소재의 유에스비 코퍼레이션(USB Corporation), #78480)를 사용하였다. 주형을 Dpn I을 이용하여 제거하였다. 클린업된 PCR 생성물 (자이모 DNA 클린 앤드 콘센트레이터-5(Zymo DNA Clean & Concentrator-5); 미국 캘리포니아주 어바인 소재의 자이모 리서치 코퍼레이션(Zymo Research Corporation), 카탈로그 번호 D4003)로 KOBW-3a 세포를 형질전환시켰다. 37℃ 인큐베이터에서 LB 한천 플레이트 상에서 하룻밤 성장시킨 후, 모든 세포를 수집하고, 쿼아프랩 스핀 미니프랩 키트 (미국 캘리포니아주 발렌시아 소재의 쿼아젠 인크., 카탈로그 번호 27106)를 사용하여 DNA를 추출하였다.

[0771] 그 후, 추출된 DNA를 주형으로 사용하여, 위치 58의 돌연변이 유발과 동일한 K9_56f 및 SR을 이용하여 위치 56에서 돌연변이를 도입하였다. 마지막으로, 위치 53에서의 돌연변이를 유사하게 도입하였다. 모든 3개의 위치 (53, 56 및 58)에서의 돌연변이를 AS6F1 내로 도입한 후, 실시예 4에서 설명한 것과 동일하게 새로운 라이브러리를 스크리닝하였으며, 일부 선발된 돌연변이체는 하기 표에 열거되어 있다.

[0772] 다른 방법은 위치 53에서 시작하였다. 프라이머 혼합물 K9_56r 및 K9_58r을 하기 표에 열거된 프라이머들을 사용하여 유사하게 제조하였다. 공유된 정방향 프라이머는 pBAD_266f: CTCTCTACTGTTTCTCCATACCCG (서열 번호 634; 이 실시예에서 SF로 칭해짐)이다. 위치 53에서의 돌연변이를 먼저 PCR을 통하여 AS6F1 내로 도입하였다. 상기 돌연변이 유발 절차는 AS6F1을 주형으로 사용하고 K9_53f 및 SR을 두 PCR 프라이머로 사용하는 상기에 설명한 것과 유사하다. 생성된 돌연변이된 DNA (위치 53에서)를 주형으로 사용하고 K9_56r 및 SF를 돌연변이 유발 프라이머로 사용하여 위치 56에 돌연변이를 도입하였다. 마지막으로, 위치 58에서의 돌연변이를 K9_58r 및 SF를 사용하여 유사하게 도입하였다. 모든 3개의 위치 (53, 56 및 58)에서의 돌연변이를 AS6F1 내로 도입한 후,

상기에서 설명한 것과 동일하게 새로운 라이브러리를 스크리닝하였으며, 일부 선발된 돌연변이체는 하기 표에 열거하였다.

[표 28]

정방향 돌연변이 프라이머

K9-KARJ의 표적화된 위치(들)	프라이머
53	K9_53F_031411f GTTATCATCGGATTATTGGAAGGA (서열 번호 544) K9_53L_031411f GTTATCATCGGATTATTGGAAGGA (서열 번호 545) K9_53Y_031411f GTTATCATCGGATTATATGAAGGA (서열 번호 546) K9_53C_031411f GTTATCATCGGATTATGTGAAGGA (서열 번호 547) K9_53W_031411f GTTATCATCGGATTATGGGAAGGA (서열 번호 548) K9_53P_031411f GTTATCATCGGATTACCAGAAGGA (서열 번호 549) K9_53H_031411f GTTATCATCGGATTACATGAAGGA (서열 번호 550) K9_53Q_031411f GTTATCATCGGATTACAAGAAGGA (서열 번호 551) K9_53I_031411f GTTATCATCGGATTAATTGAAGGA (서열 번호 552) K9_53M_031411f GTTATCATCGGATTAATTGAAGGA (서열 번호 553) K9_53N_031411f GTTATCATCGGATTAATGAAGGA (서열 번호 554) K9_53V_031411f GTTATCATCGGATTAGTTGAAGGA (서열 번호 555) K9_53A_031411f GTTATCATCGGATTAGCTGAAGGA (서열 번호 556) K9_53D_031411f GTTATCATCGGATTAGATGAAGGA (서열 번호 557) K9_53E_031411f GTTATCATCGGATTAGAAGAAGGA (서열 번호 558) K9_53G_031411f GTTATCATCGGATTAGGTGAAGGA (서열 번호 559)
56	K9_56F_031411f GGATTACCTGAAGGATTCAAA (서열 번호 560) K9_56L_031411f GGATTACCTGAAGGATTGAAA (서열 번호 561) K9_56Y_031411f GGATTACCTGAAGGATATAAA (서열 번호 562) K9_56C_031411f GGATTACCTGAAGGATGTAAA (서열 번호 563) K9_56W_031411f GGATTACCTGAAGGATGGAAA (서열 번호 564) K9_56P_031411f GGATTACCTGAAGGACCAAAA (서열 번호 565) K9_56H_031411f GGATTACCTGAAGGACATAAA (서열 번호 566) K9_56Q_031411f GGATTACCTGAAGGACAAAAA (서열 번호 567) K9_56I_031411f GGATTACCTGAAGGAATTAAA (서열 번호 568) K9_56M_031411f GGATTACCTGAAGGAATGAAA (서열 번호 569) K9_56N_031411f GGATTACCTGAAGGAAATAAA (서열 번호 570) K9_56V_031411f GGATTACCTGAAGGAGTAAA (서열 번호 571) K9_56A_031411f GGATTACCTGAAGGAGCTAAA (서열 번호 572) K9_56D_031411f GGATTACCTGAAGGAGATAAA (서열 번호 573) K9_56E_031411f GGATTACCTGAAGGAGAAAAA (서열 번호 574) K9_56G_031411f GGATTACCTGAAGGAGTAAA (서열 번호 575)
58	K9_58F_051611f GATTACCTGAAGGATTTAAATTCTGGAAGAGAGC (서열 번호 576) K9_58L_051611f GATTACCTGAAGGATTTAAATTGTGGAAGAGAGC (서열 번호 577) K9_58Y_051611f GATTACCTGAAGGATTTAAATATTGGAAGAGAGC (서열 번호 578)

578)	K9_58C_051611f	GATTACCTGAAGGATTTAAATGTTGGAAGAGAGC	(서열 번호 579)
579)	K9_58W_051611f	GATTACCTGAAGGATTTAAATGGTGAAGAGAGC	(서열 번호 580)
580)	K9_58P_051611f	GATTACCTGAAGGATTTAAACCATGGAAGAGAGC	(서열 번호 581)
581)	K9_58H_051611f	GATTACCTGAAGGATTTAAACATTGGAAGAGAGC	(서열 번호 582)
582)	K9_58Q_051611f	GATTACCTGAAGGATTTAAACAATGGAAGAGAGC	(서열 번호 583)
583)	K9_58I_051611f	GATTACCTGAAGGATTTAAATTTGGAAGAGAGC	(서열 번호 584)
584)	K9_58M_051611f	GATTACCTGAAGGATTTAAATGTGGAAGAGAGC	(서열 번호 585)
585)	K9_58N_051611f	GATTACCTGAAGGATTTAAAAATGGAAGAGAGC	(서열 번호 586)
586)	K9_58V_051611f	GATTACCTGAAGGATTTAAAGTTTGAAGAGAGC	(서열 번호 587)
587)	K9_58A_051611f	GATTACCTGAAGGATTTAAAGCTTGAAGAGAGC	(서열 번호 588)
588)	K9_58D_051611f	GATTACCTGAAGGATTTAAAGATTGAAGAGAGC	(서열 번호 589)
589)	K9_58E_051611f	GATTACCTGAAGGATTTAAAGAATGGAAGAGAGC	(서열 번호 590)
590)	K9_58G_051611f	GATTACCTGAAGGATTTAAAGGTTGGAAGAGAGC	(서열 번호 591)

[0775]

[0776]

[표 29]

정방향 돌연변이 프라이머

K9-KARI의 표적화된 위치(들)	프라이머
56	K9_56F_071211r GCTCTCTTCCATGGTTTGAATCCTTC (서열 번호 592) K9_56L_071211r GCTCTCTTCCATGGTTTCAATCCTTC (서열 번호 593) K9_56Y_071211r GCTCTCTTCCATGGTTTATATCCTTC (서열 번호 594) K9_56C_071211r GCTCTCTTCCATGGTTTACATCCTTC (서열 번호 595) K9_56W_071211r GCTCTCTTCCATGGTTTCCATCCTTC (서열 번호 596) K9_56P_071211r GCTCTCTTCCATGGTTTGGTCCTTC (서열 번호 597) K9_56H_071211r GCTCTCTTCCATGGTTTATGTCCTTC (서열 번호 598) K9_56Q_071211r GCTCTCTTCCATGGTTTGTCTCCTTC (서열 번호 599) K9_56I_071211r GCTCTCTTCCATGGTTTAAATCCTTC (서열 번호 600) K9_56M_071211r GCTCTCTTCCATGGTTTCATTCCTTC (서열 번호 601) K9_56N_071211r GCTCTCTTCCATGGTTTATTTCCCTTC (서열 번호 602) K9_56V_071211r GCTCTCTTCCATGGTTTAATCCTTC (서열 번호 603) K9_56A_071211r GCTCTCTTCCATGGTTTAGCTCCTTC (서열 번호 604) K9_56D_071211r GCTCTCTTCCATGGTTTATCTCCTTC (서열 번호 605) K9_56E_071211r GCTCTCTTCCATGGTTTCTCTCCTTC (서열 번호 606) K9_56G_071211r GCTCTCTTCCATGGTTTACCTCCTTC (서열 번호 607)
58	K9_58F_071211r GTTCTTCTGCTCTCTCCAGAATTT (서열 번호 608) K9_58L_071211r GTTCTTCTGCTCTCTCCACAATTT (서열 번호 609) K9_58Y_071211r GTTCTTCTGCTCTCTCCAATATTT (서열 번호 610) K9_58C_071211r GTTCTTCTGCTCTCTCCAACATTT (서열 번호 611) K9_58W_071211r GTTCTTCTGCTCTCTCCACCATTT (서열 번호 612) K9_58P_071211r GTTCTTCTGCTCTCTCCATGGTTT (서열 번호 613) K9_58H_071211r GTTCTTCTGCTCTCTCCAATGTTT (서열 번호 614) K9_58Q_071211r GTTCTTCTGCTCTCTCCATTGTTT (서열 번호 615) K9_58I_071211r GTTCTTCTGCTCTCTCCAAATTTT (서열 번호 616) K9_58M_071211r GTTCTTCTGCTCTCTCCACATTTT (서열 번호 617) K9_58N_071211r GTTCTTCTGCTCTCTCCAATTTT (서열 번호 618) K9_58V_071211r GTTCTTCTGCTCTCTCCAAACTTT (서열 번호 619) K9_58A_071211r GTTCTTCTGCTCTCTCCAAGCTTT (서열 번호 620) K9_58D_071211r GTTCTTCTGCTCTCTCCAATCTTT (서열 번호 621) K9_58E_071211r GTTCTTCTGCTCTCTCCATTCTTT (서열 번호 622) K9_58G_071211r GTTCTTCTGCTCTCTCCAACCTTT (서열 번호 623)

[0777]

[0778] [표 30]

측정된 K_M 값을 갖는 일부 돌연변이체의 목록

돌연변이체	서열 번호 (핵산, 아미노산)	돌연변이	K_M (μM) (NADH)	K_M (μM) (NADPH)
K9 WT	26, 27	--	326	0.2
BI7D12	625, 624	Y53Q, S56V, S58D , I86V, N87P, T131M, T191G	39	196
BI10F1	627, 626	Y53E, S56V, S58D , I86V, N87P, T131M, T191G	74	573
BJ6G6	629, 628	Y53P, S56D, S58Q , I86V, N87P, T131M, T191G	298	672
BJ7D6	631, 630	Y53P, S56V, S58E , I86V, N87P, T131M, T191G	37	236
BJ7F7	633, 632	Y53A, S56D, S58Q , I86V, N87P, T131M, T191G	269	762

[0779]

[0780] 실시예 24

[0781] ald6 Δ 주 및 아이소부탄올-생산 유도체의 제작

[0782] 락토바실러스 플란타룸 유래의 xpk1 및 eutD 유전자를 위한 발현 카세트를 함유하는 pRS426::GPD-xpk1+ADH-eutD 유래의 5.3 kb (BglII/EcoRV) DNA 단편 (서열 번호 383)을, 상기 실시예 9에 설명된 pUC19::ald6D::loxP-URA3-loxP 벡터의 SnaBI 부위에서 ALD6 측면 위치 서열들 사이에 부가하였다. 라이게이션 반응물로 이. 콜라이 Stb13 세포를 형질전환시키고, 이를 LB Amp 플레이트 상에서 인큐베이션하여 형질전환체를 선발하였다. xpk1-eutD 카세트의 삽입을 PCR (프라이머)에 의해 확인하였다. 양성 클론 (pUC19:: Δ ald6::URA3::xpkS)을 수득하였다.

[0783] 상기에 설명한 벡터를 AhdI로 선형화하고, 2.0 mL의 YPE (효모 추출물, 펩톤 - 1% 에탄올을 포함 -)배지에서의 2.5시간의 추가의 파생 인큐베이션을 포함하는, 제조업자의 프로토콜에 대한 변경을 이용하여 자이모 리서치 프로즌-이지 효모 형질전환 키트 (카탈로그 번호 T2001)를 사용하여 제조한 PNY1507 (본 명세서에 설명됨) 세포를 형질전환시켰다. 탄소 공급원으로서 1% 에탄올이 제공된, 우라실이 제외된 합성 완전 배지 상에 플레이트링함으로써 형질전환체를 수득하였다. 패치된 형질전환체를 PCR에 의해 스크리닝하여 결실/통합을 확인하였는데, 이는 프라이머 N1090 및 N1213 (서열 번호 779 및 242)을 이용하였다. 히스티딘 마커 선발을 이용하여, Cre 리코비나아제를 지닌 플라스미드 (pRS423::GAL1p-Cre; 서열 번호 271)로 상기 주를 형질전환시켰다. 형질전환체를 0.5% 갈락토스로 보충된 YPE 상에서 계대하였다. 콜로니를 5-FOA에 대한 내성 (URA3 마커의 상실)에 대하여 그리고 히스티딘 영양요구성 (Cre 플라스미드의 상실)에 대하여 스크리닝하였다. 측면에 위치하는 loxP 부위를 통한 URA3 유전자의 적당한 제거는 프라이머 N1212 및 N1214 (서열 번호 241 및 281)를 이용하여 PCR에 의해 확인하였다. 마지막으로, alsS 통합 플라스미드 (서열 번호 780)로 이 주를 형질전환시켰으며, 이는 포함된 제네티신 선발 마커를 이용한 것이었다. 통합체를 프라이머 N160SeqF5 및 oBP512 (서열 번호 388 및 337)를 사용하여 확인하였다.

[0784] 플라스미드 pYZ090 Δ alsS 및 pBP915 (서열 번호 371 및 182)로 상기 주를 형질전환시켰으며, 이는 아세트산리튬 형질전환법 (문헌["Methods in Yeast Genetics" 2005. Amberg, Burke and Strathern]에서 프로토콜 #2)에 의한 것이었다. 탄소 공급원으로서 에탄올을 포함하는, 히스티딘 및 우라실이 제외된 합성 완전 배지 상에 플레이트링함으로써 형질전환체를 선발하였다. 형질전환체를 패치시키고, 그 후 2% 글루코스 및 0.05% 에탄올을 포함하는, 히스티딘 및 우라실이 제외된 합성 완전 배지 상에 재패치하였다. 6개의 클론을 성장 및 아이소부탄올 생성에 대하여 평가하였다. 이들 중 하나를 PNY2216으로 표기하였다.

[0785] 실시예 25

[0786] 에스. 세레비지에 주 PNY2211로부터의 YMR226c 결실 (PNY2248의 제작)

[0787] 에스. 세레비지에 주 BY4743 ymr226c Δ ::KanMX4 (ATCC 4020812)에서 이용가능한 PCR 증폭 선형 KanMX4-기반 결실 카세트를 사용하여 상동 재조합에 의해 유전자 YMR226c를 에스. 세레비지에 주 PNY2211 (실시예 9에서 설명함)로부터 결실시켰다. 정방향 및 역방향 PCR 프라이머 N1237 (서열 번호 784) 및 N1238 (서열 번호 785)은 염색체 XIII로부터 2,051 bp ymr226c Δ ::KanMX4 결실 카세트를 증폭시켰다. 상기 PCR 생성물은 ymr226c Δ ::KanMX4 결실 카세트 측면에 위치하는, 각각 253 및 217 bp의 상류 및 하류 서열을 함유하였으며, 이는 주

PNY2211에 있어서 천연 YMR226c 유전자좌의 측면에 위치하는 서열들에 대하여 100%의 상동성을 갖는다. 재조합 및 유전자 교환이 효과적으로 측면 상동 서열들에서 일어나서 YMR226c 유전자를 결실시키고 ymr226c Δ ::KanMX4 결실 카세트를 통합시킨다.

[0788] 이전에 문헌[*Methods in Yeast Genetics*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, pp. 201-202 (2005)]에 기재된 아세트산리튬법을 이용하여 적격성으로 만든 주 PNY2211을 대략 2.0 μ g의 PCR 증폭 생성물로 형질전환시키고, 형질전환 믹스를 YPE + 제네티신 (50 μ g/mL) 상에 플레이팅하고, 통합된 ymr226c Δ ::KanMX4 카세트를 포함하는 세포의 선발을 위하여 30℃에서 인큐베이션하였다. 형질전환체를 PCR에 의해 ymr226c Δ ::KanMX4에 대하여 스크리닝하였으며, 이는 밖으로 향하는 5' KanMX4 결실 카세트-특이적 내재적 프라이머 N1240 (서열 번호 786)가 측면에 위치하는, 안으로 향하는 염색체 특이적 프라이머 N1239 (서열 번호 243)와 쌍을 형성한 것, 및 밖으로 향하는 3' KanMX4 결실 카세트-특이적 프라이머 N1241 (서열 번호 787)가 측면에 위치하는, 안으로 향하는 염색체-특이적 프라이머 N1242 (서열 번호 244)와 쌍을 형성한 것을 이용한다. 양성 PNY2211 ymr226c Δ ::KanMX4 클론들을 수득하고, 이들 중 하나를 PNY2248로 표기하였다.

[0789] 실시예 26

[0790] YMR226c 녀아웃에서 DHMB 수율이 감소된 아이소부탄올의 생성

[0791] PNY2211 ymr226c Δ ::KanMX4 형질전환체 및 비결실 대조군 (천연 YMR226c를 포함하는 PNY2211)은, 먼저 아이소부탄올 경로-함유 플라스미드 pYZ090 Δ alsS (서열 번호 371) 및 pBP915 (서열 번호 182)를 문헌[*Methods in Yeast Genetics*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY (2005)]에 기재된 퀵 앤드 더티(Quick and Dirty) 아세트산리튬 형질전환법에 의해 동시에 도입함으로써 글루코스 배지에서의 부탄올 생성에 대하여 시험하였다. 플라스미드 선발은 에탄올을 함유하는 선발 플레이트 (1.0% 에탄올을 포함하는, -his -ura 합성 완전 배지) 상에서의 히스티딘 및 우라실 영양요구성을 바탕으로 하였다. 3 내지 5일 후에, 가장 강한 성장을 나타내는 몇몇 형질전환체를 SD 2.0% 글루코스 + 0.05% 에탄올 -his -ura 상에 패치시킴으로써 글루코스 배지에 적응시키고, 300℃에서 48 내지 72시간 인큐베이션하였다. 가장 강한 성장을 나타내는 3개의 스트리크(streak)를 사용하여, 125 mL의 통기성 플라스크 내의 SD 0.2% 글루코스 + 0.2% 에탄올 -his -ura 중에 10 mL 시드 배양물(seed culture)을 접종하고, 30℃, 250 rpm에서 대략 24시간 동안 성장시켰다. 그 후, 세포를 125 mL의 타이트하게 뚜껑이 닫히는 플라스크 내의 2% 글루코스 + 0.05% 에탄올 -his -ura를 포함하는 합성 완전 배지 내에 계대배양하고, 30℃에서 48시간 인큐베이션하였다. 접종 후 그리고 48시간의 인큐베이션 후 수집한 배양 상청액을 HPLC로 분석하여 아이소부탄올의 생성량을 결정하고 LC/MS로 분석하여 DHMB를 정량화하였다. 대조 주는 글루코스 1 몰당 0.03 내지 0.07 몰의 물 수율로 DHMB를 생성하는 것으로 관찰되었다. DHMB에 상응하는 피크는 ymr226c Δ 주들의 배양 상청액에서 관찰되지 않았으며, 상기 주들 중 하나를 PNY2249로 표기하였다.

[0792] 실시예 27

[0793] 효모 녀아웃 라이브러리를 사용한 아세토라테이트 리덕타아제 (ALR) 활성 효소를 코딩하는 유전자의 확인

[0794] 오픈 바이오시스템즈(등록상표) (미국 매사추세츠주 월섬 소재의 서모 피셔 사이언티픽의 부)로부터 입수가 가능한, 주 BY4743으로부터 유래된 >6000가지의 효모 주의 녀아웃 ("KO") 컬렉션으로부터, 95가지의 후보 데하이드로게나아제 유전자 녀아웃 주를 선택하였다. 녀아웃 주의 출발 배양물(Starter culture)을 풍부 배지 YPD 상에서 96웰의 깊은 웰 플레이트 (코스타 3960, 미국 뉴욕주 코닝 소재의 코닝, 인크.; 또는 이와 유사한 것)에서 성장시키고, 대략 0.3의 출발 OD 600 nm에서, 0.67%의 효모 질소 베이스, 0.1%의 카사미노산, 2%의 글루코스, 및 0.1 M의 K⁺MES - pH 5.5 - 를 함유하는 배지에서 계대배양하였다. 샘플들을 DHMB 및 DHIV 측정용으로 5일의 기간에 걸쳐 취하였다. DHIV와, DHMB의 2가지의 이성체를 분리하고, 아틀란티스(Atlantis) T3 (파트 번호 186003539) 컬럼을 사용하여, 워터스 (미국 매사추세츠주 밀포드 소재) 어퀴티(Acquity)TQD 시스템에서 액체 크로마토그래피-질량 분광법 ("LC/MS")에 의해 정량화하였다. 컬럼을 30℃에서 유지하였으며, 유량은 0.5 mL/분이었다. A 이동상은 수 중의 0.1% 포름산이고, B 이동상은 아세토니트릴 중의 0.1% 포름산이었다. 각각의 런은 99% A, 1분; 1분에 걸친 25% B로의 선형 구배; 이어서 99% A, 1분으로 이루어졌다. 컬럼 유출물은, 워터스 ACQ-TQD (s/n QBA688) 질량 스펙트럼(mass spec) 검출기에 의해 콘 전압을 32.5 V로 하여, m/z = 133 (negative ESI)에서의 피크에 대하여 모니터링하였다. 소위 "패스트 DHMB"는 전형적으로 1.10분에 나왔으며, 이어서 1.2분에 DHIV가 나왔고, "슬로우" DHMB는 1.75분에 나왔다. 기준선 분리를 수득하고, 1 M 수성 스톱으로부터 만든 표준 용액의 분석을 참조하여, DHIV에 대한 피크 면적을 μ M DHIV 농도로 전환하였다. 이들 측정은 DHMB 수준의 변화의 대부분이 처음 48-60시간에 일어남을 나타냈으며, 따라서 후속 실험에서 단일

샘플을 대략 이 시점에 수집하였다. 이 실험에서, 패스트 DHMB는 슬로우 DHMB보다 훨씬 더 높은 수준으로 발견되었는데, 슬로우 DHMB는 항상 검출가능한 것은 아니었다. 대부분의 배양물에서 패스트 DHMB에 대한 DHIV의 비는 대략 3이었지만, YMR226C 유전자가 결여된 주는 매우 낮은 수준의 패스트 DHMB, 및 보통의 DHIV를 일관적으로 나타내서, DHIV/패스트 DHMB 비는 약 100이었다. 이는 YMR226Cp가 이 배경에서 주요 ALR임을 시사하였다.

[0795] 이 배경에서 YMR226Cp가 주요 ALR임을 확인하기 위하여, ALR 및 KARI의 시험관 내 수준을 *ymr226c* 결실 주 (미국 버지니아주 매너서스 소재의 아메리칸 타입 컬처 컬렉션(ATCC), ATCC #4020812) 및 그의 모 주, BY4743 (ATCC #201390; 미국 버지니아주 매너서스 소재의 아메리칸 타입 컬처 컬렉션)에서 검사하였다. 6 ml YPD를 함유하는 50 ml 튜브에 YPD 한천 플레이트로부터의 것을 접종하였으며, 하룻밤 성장시켰다 (30°C, 250 rpm). 세포를 펠렛화하고, 물에서 1회 세척하고, 효모 프로테아제 저해제 콕테일 (스위스 바젤 소재의 로슈(Roche), 카탈로그 번호 11836170001, 판매사가 지시하는 대로 사용함, 10 ml의 완충제당 1개의 정제)을 함유하는 1 ml 효모 세포질 완충제 (문헌[Van Eunen *et al.* *FEBS Journal* 277: 749-760 (2010)])에 재현탁시켰다. 톨루엔 (0.02 ml, 미국 뉴저지주 페어 론 소재의 피셔 사이언티픽)을 첨가하고, 투과화를 위하여 볼텍스 지니(Vortex Genie) 2 진탕기 (미국 뉴욕주 보헤미아 소재의 사이언티픽 인더스트리즈(Scientific Industries), 모델 G-560)에서 10분 동안 최고 속도에서 튜브를 진탕시켰다. 상기 튜브들을 30°C의 수조 내에 두고, 기질을 하기 최종 농도가 되도록 첨가하였다: 9.4 mM이 되도록 하는 (S)-아세토락타이트 (하기에서 실시예 29에서 설명한 바와 같이 효소에 의해 제조함), NADPH (미국 미주리주 세인트 루이스 소재의 시그마-알드리치) 0.2 mM + NAD(P)H-재생 시스템 - 대략 10 mM의 글루코스-6-포스페이트 및 2.5 U/ml의 류코노스톡 메센테로이데스 글루코스-6-포스페이트 데하이드로게나아제 (미국 미주리주 세인트 루이스 소재의 시그마, 카탈로그 번호 G8404)로 이루어짐 - . 시기 적절한 간격으로, 분취물 (0.15 ml)을 2% 포름산의 0.15 ml 분취물에 첨가하여 반응을 종결시켰다. 그 후, 샘플들을 상기에 기재된 바와 같이 LC/MS에 의해 DHIV와, DHMB의 둘 모두의 이성체에 대하여 분석하였으며; 단지 패스트 DHMB 및 DHIV가 관찰되었다. 상기 두 주에서의 상기 두 효소의 비활성을 표 31에 나타낸다.

[0796] [표 31]

KARI 및 ALR 효소 활성

주	KARI	ALR
BY4743	단백질 1 mg당 1.7 mU	20 mU/mg
YMR226C 결실 주	단백질 1 mg당 2.2 mU	0.1 mU/mg

[0797] 이 데이터는 YMR226C 유전자 생성물이 ALR 활성의 >99%를 차지하였음을 시사한다.

[0798] 실시예 28

[0799] 효모 과다발현 라이브러리를 사용한 아세토락타아제 리덕타아제(ALR) 활성 효소를 코딩하는 유전자의 확인

[0800]

[0801] 유도성 프로모터의 제어 하에, 공지된 효모 유전자 + C-말단 태그를 지닌 플라스미드를 각각이 포함하는, Y258의 >5000개의 형질전환체의 "효모 ORF" 컬렉션 (오픈 바이오시스템즈(등록상표), 서모 피셔 사이언티픽의 부, 미국 매사추세츠주 월섬 소재)으로부터, 데하이드로게나아제 활성과 결부된 유전자를 함유하는 플라스미드를 포함하는 96가지의 주를 판매사 (오픈 바이오시스템즈(등록상표))가 권고하는 성장 및 유도 프로토콜의 적응에 의해 96웰 형식으로 성장시켰다. 상기 세포를 펠렛화하고, 상기에 기재된 바와 같이 톨루엔으로 투과성화하고, 진한 기질 믹스를 첨가하여 실시예 27에서와 같은 최종 농도를 제공하였다. 그 후, 시기 적절한 샘플들을 취하고, DHIV와, DHMB의 둘 모두의 이성체에 대하여 분석하였다. ALR/KARI의 비들을 계산하고, 비교하였다. 상기 비가 상승된 주들은 ALR 활성의 과다생성을 위한 후보였다. 패스트 DHMB 및 DHIV 형성의 상대적인 속도에 대한 데이터를 미니탭(Minitab)(등록상표) (미국 위싱턴주 레드몬드 소재의 마이크로소프트 인크.(Microsoft Inc.)) 박스플롯(boxplot)으로 디스플레이할 때, 상기 비들 중 절반은 대략 9-13 사이에 들어가며 대부분의 나머지는 3과 19 내에 들어갔다. 아웃라이어(outlier)으로 확인되는 예외는 YER081W, YIL074C, YMR226C, YBR006W, 및 YOR375C를 포함하였으며, 이들의 ALR/KARI 비는 22와 40 사이에 들어갔다. 슬로우 DHMB 및 DHIV 형성의 상대적인 속도의 유사한 분석에서, 상기 비들 중 절반은 9와 11 사이에 들어가지만, YMR226C, YPL275W, YER081W, 및 YOL059W는 외부에 있는 것으로 나타났으며, 이때 비는 13과 25 사이였다. 따라서, YMR226C 및 YER081W의 과다발현은 둘 모두의 DHMB의 합성을 증가시켰다. 게다가, YIL074C, YBR006W, 및 YOR375C는 패스트 DHMB 합성을 증가시켰으며, YPL275W 및 YOL059W는 슬로우 DHMB 합성을 증가시켰다. 과다발현에 있어서 확인된 유전자의 게놈 DNA 서열 (이는 인트론을 포함할 수 있음) 및 ORF 번역 서열이 표 6에 제공되어 있다.

[0802] 실시예 29

- [0803] DHMB에 의한 KARI의 저해
- [0804] (S)-아세트락테이트의 효소적 생성
- [0805] (S)-아세트락테이트를 DHMB 합성을 위한 출발 물질로서 사용하였다. (S)-아세트락테이트를 하기와 같이 효소에 의해 만들었다. IPTG 제어 하에, 클램프엘라 BudB (이전에 미국 특허 제7,851,188호에서 개시되었으며, 상기 미국 특허는 전체적으로 본 명세서에 참고로 포함됨; 상기 특허의 실시예 9 참조)를 발현하도록 변형된 이. 콜라이 TOP10 주 (미국 캘리포니아주 칼스배드 소재의 인비트로젠)를 효소 공급원으로 사용하였다. 이것을 200-1000 ml 배양 부피로 성장시켰다. 예를 들어, 200 ml을 0.5 L 코니칼(conical) 플라스크에서 0.1 mg/ml의 암피실린 (미국 미주리주 세인트 루이스 소재의 시그마)을 함유하는 루리아 브로쓰(Luria Broth) (미국 버지니아주 매너서스 소재의 메디아테크(Mediatech))에서 성장시켰으며, 이를 250 rpm에서 37℃에서 진탕하였다. 대략 0.4의 OD 600에서, 아이소프로필티오갈락토시드 (미국 미주리주 세인트 루이스 소재의 시그마)를 0.4 mM이 되도록 첨가하고, 성장을 2시간 동안 계속 시킨 후 세포를 원심분리에 의해 수집하여 대략 1 g의 습윤 중량의 세포를 생성하였다. 이와 마찬가지로, 부분 정제를 대략 0.5 내지 5 g의 습윤 세포의 규모로 행하였다. 예를 들어, 대략 0.5 g의 세포를 25 mM Na-MES - pH 6 - 를 함유하는 2.5 ml 완충제에 현탁시키고, 0℃에서 초음파 처리에 의해 파쇄하고, 원심분리에 의해 청정시켰다. 조 추출물에 0.1 mM 티아민 파이로포스페이트, 10 mM MgCl₂, 및 1 mM EDTA (전부 미국 미주리주 세인트 루이스 소재의 시그마로부터의 것임)를 보충하였다. 다음, 0.07 ml의 10% (w/v) 수성 스트렙토마이신 설페이트 (미국 미주리주 세인트 루이스 소재의 시그마)를 첨가하고, 샘플을 56℃ 수조에서 20분 동안 가열하였다. 이것을 원심분리에 의해 청정시키고, 황산암모늄을 50%의 포화도로 첨가하였다. 상기 혼합물을 원심분리하고, 펠렛을 0.5 ml의 25 mM Na-MES - pH 6.2 - 에서 형성(bring up)하고, 추가의 특성화 없이 사용하였다. 아세트락테이트 신타아제를 다양한 규모로 또한 행하였다. 대량 제조를 하기와 같이 행하였다: 5.5 g의 피루브산나트륨을 대략 45 ml이 되도록 25 mM Na-MES - pH 6.2 - 에 용해시키고, 10 mM MgCl₂, 1 mM 티아민 파이로포스페이트, 1 mM EDTA (전부 미국 미주리주 세인트 루이스 소재의 시그마로부터의 것), 25 mM의 아세트산나트륨 (미국 뉴저지주 페어 론 소재의 피셔 사이언티픽), 및 0.25 ml의 BudB 제제를 보충하였다. 상기 혼합물을 실온에서 pH 미터(meter) 아래에서 교반시켰다. 반응이 진행됨에 따라, CO₂가 방출되었고, pH가 상승하였다. 피루브산 (미국 매사추세츠주 워드 힐 소재의 알파(Alfa))을 연동 펌프를 통하여 서서히 첨가하여 pH를 6 내지 7로 유지하였다. pH가 상승함에 따라, 효소 반응이 느려지지만, 이것을 6 미만으로 강하시킬 경우, 아세트락탄의 탈카르복실화가 문제시된다. 반응이 완료되었을 때, 상기 혼합물을 -80℃에 보관하였다.
- [0806] DHMB의 합성
- [0807] DHMB를 (S)-아세트락테이트로부터 화학적으로 합성하였다. 대략 8의 pH의, 대략 0.8 M의 3 ml의 조 아세트락테이트 제제를 1.2 당량의 NaBH₄ (미국 위스콘신주 밀워키 소재의 알드리치 케미칼 컴퍼니(Aldrich Chemical Co))로 처리하였다. 상기 반응물을 실온에서 하룻밤 정치시킨 후 둘로 나누고, 이동상으로서 물을 사용하여 바이오겔(Biogel) P-2 (미국 캘리포니아주 허큘리스 소재의 바이오-라드)의 60 cm x 1 cm 직경의 컬럼에서 두 부분으로 탈염시켰다. 혼합 DHMB를 함유하는 분획들을 회전 증발에 의해 농축시키고, 황산을 이용하여 pH 2.2로 조정하였다.
- [0808] DHMB의 부분입체 이성체들을, 215 nm에서의 UV 검출을 이용하여, 워터스 아틀란티스 T3 (5 μm, 4.6 x 150 mm)가 0.3 mL/min의 유량으로 pH 2.5의 0.2% 수성 포름산에서 실온에서 실행되는 상태에서, HPLC 시스템 (LKB 2249 펌프 및 구매 제어기 (LKB, 현재, 영국 첼폰트 세인트 자일스 소재의 제너럴 일렉트릭(General Electric)의 부임) 및 휴렛-패커드(Hewlett-Packard) (현재, 애글린트(Agilent), 미국 캘리포니아주 산타 클라라 소재) 1040A UV/vis 검출기로 이루어짐)을 사용하여 분리하였다. "페스트" DHMB는 8.1분에서 용출되고, "슬로우" DHMB는 13.7분에서 용출되었다. DHIV는 존재하지 않았다. 풀링한 분획들을 거의 건조되게 하고, 톨루엔을 이용하여 동시 증발시켜 잔류 포름산을 제거하였다. 그 후, 잔사를 물에 용해시키고, 트라이에틸아민 (미국 뉴저지주 페어 론 소재의 피셔)으로 염기성으로 만들었다.
- [0809] DHMB의 농도 측정 및 절대 구조
- [0810] 정제된 DHMB 용액의 농도를 하기와 같이 측정하였다. 상기 농도를 NaBH₄ 환원에서 사용된 아세트락테이트의 mmol을 바탕으로 하여 개산하였다. DHMB의 부분에, 공지된 양의 벤조산나트륨 (고체 벤조산 (ACS 등급, 미국 뉴저지주 페어 론 소재의 피셔 사이언티픽)을 수성 NaOH에 용해시킴으로써 제조함)을 첨가하여 (대략적으로) 등물량의 2성분 혼합물을 제공하였다. 또한 DHIV의 유사한 샘플을 주문 합성 (미국 뉴욕주 알바니 소재의 알바니

몰레큘러 리서치(Albany Molecular Research))을 통하여 수득한 고체 나트륨 염으로부터 제조하였다. 상기 샘플들을 D₂O (미국 위스콘신주 밀워키 소재의 알드리치)를 이용하여 수회 동시 증발시키고, D₂O에 재용해시켰다. 적분된 양성자 NMR 스펙트럼을 수득하였으며, 이를 이용하여 벤조에이트에 대한 DHIV 또는 DHMB의 몰비를 결정하였다. DHMB의 상기 NMR 스펙트럼을 CDCl₃ 중 자유 산에 대한 문헌의 스펙트럼 (문헌[Kaneko *et al.*, *Phytochemistry* 39: 115-120 (1995)])과 비교해 보면 페스트 DHMB가 에리트로 이성체(erythro isomer)임이 나타났다. 효소에 의해 합성된 아세토라테이트는 C-2에서 (S) 배열을 갖기 때문에, 페스트 DHMB는 2S, 3S 배열을 갖는다. 슬로우 DHMB는 트레오(threo) 2S, 3R 배열을 갖는다.

[0811] 또한 NMR 샘플의 회석은, 별도로 제조된 벤조산 용액을 표준물로서 사용하여 LC/MS에 의해 분석하였다. 벤조산, DHIV, 및 DHMB의 두 이성체를 아틀란티스 T3 (파트 번호 186003539) 컬럼을 사용하여, 워터스 (미국 매사추세츠주 밀포드 소재) 어퀴티TQD 시스템에서 LC/MS에 의해 분리하고 정량화하였으며, 이는 상기에 기재된 바와 같았다. 벤조산은 m/z=121 (음의 ESI)에서 검출되었으며, 2.05분에서 나타났다. 상기 혼합물 중 벤조에이트의 농도는 예상치의 실험적 불확실성 내에 있었다. 또한 상기 실험에 의하면, DHMB의 어느 하나의 이성체는, 상기 기기의 응답 범위 전체에 걸쳐, LC/MS에서 DHIV의 약 80%의 민감성 (즉, 주입된 nmol당 관찰된 MS 피크 면적)을 가짐이 밝혀졌다. 따라서, DHIV 표준물을 사용하여 세포 추출물에서 또는 효소 반응물에서 발견되는 DHMB를 정량화할 경우, 겔보기 DHMB 농도에 1.25를 곱할 필요가 있다.

[0812] DHMB에 의한 KARI의 저해의 측정

[0813] 정제된, 락토코커스 락티스 유래의 유전자에 의해 코딩되는 KARI (서열 번호 864), JEA1로 공지된 슈도모나스 플루오레센스 KARI의 유도체 (서열 번호 799; 전체적으로 본 명세서에 참고로 포함된 미국 특허 공개 제 2010/0197519호), 또는 K9D3으로 공지된 엔에어로스티페스 카카에 KARI의 변이체 (서열 번호 788)를, 30℃로 설정된, TCC240A 온도 제어 유닛을 갖춘 시마즈(Shimadzu; 일본 교토 소재) UV160U 기기에서의 분광 광도적 분석에 있어서 DHMB 저해에 대한 그의 민감성에 대하여 시험하였다. 완충제는 10 mM MgCl₂ 및 1 mM EDTA를 함유하는 0.1 M K+ Hepes - pH 6.8 - 었다. NADPH는 0.2 mM로 존재하였으며, 라세믹 아세토라테이트는 3 mM 또는 0.725 mM의 (S) 이성체로 존재하였다. 페스트 DHMB 또는 슬로우 DHMB 중 어느 하나의 존재 및 부재 하에서 NADPH 산화 속도를 측정하였다. 각각의 샘플의 V_{max}를 미하엘리스-멘텐 등식을 이용하여, 관찰된 속도 및 공지된 아세토라테이트 K_M으로부터 계산하였다. 부피 K_i를 아세토라테이트에 대비한 경쟁적 저해에 대하여 변경한 미하엘리스-멘텐 등식을 사용하여 DHMB의 존재 하에서 각각 측정한 것에 대하여 계산하였다 (미하엘리스-멘텐 등식에서 K_M 항에 (1+[I]/K_i))를 곱하고, 상기 등식을 K_i에 대하여 푼다. 그 결과는 NMR 실험의 완료시에 mM로 전환시켰으며, 이를 표 32에 나타낸다.

[0814] [표 32]

DHMB 이성체에 의한 KARI 저해에 있어서의 K_i 값

주	페스트 DHMB	슬로우 DHMB
JEA1	0.23 mM	0.23 mM
K9D3	0.3 mM	0.2 mM
앨. 락티스	2.8 mM	2.3 mM

[0815]

[0816] 실시예 30

[0817] DHMB에 의한 DHAD의 저해

[0818] 스타필로코커스 뮤탄스 유래의 정제된 다이하이드록시산 데하이드라타아제 (DHAD)를, 문헌[Szamosi *et al.*, *Plant Phys.* 101: 999-1004 (1993)]에 기재된 비색 분석법의 변경을 이용하여 DHMB에 의한 다이하이드록시아이소발레레이트 (DHIV)의 2-케토아이소발레레이트 (2-KIV)로의 전환의 저해에 대하여 시험하였다. 상기 분석은 30℃로 유지한 히팅 블록 내에 둔 2 mL 에펜도르프 튜브에서 일어났다. 분석 혼합물은 최종 부피가 0.8 mL이었으며, 이는 100 mM Hepes-KOH 완충제, pH 6.8, 10 mM MgCl₂, 0.5-10 mM DHIV, 0-40 mM DHMB, 및 18 µg DHAD를 함유하였다. 10x의 진한 기질 스톱을 첨가함으로써 분석을 시작하였다. 샘플을 0.1 및 30분에 꺼내고 (0.35 mL), 제2 에펜도르프 튜브에서 0.05%의 2,4-다이니트로페닐하이드라진 (알드리치)을 포함하는 0.35 mL의 0.1 N HCl 내로 혼합하여 넣음으로써 반응을 종결시켰다. 실온에서 30분 동안 인큐베이션한 후, 0.35 mL의 4 N NaOH를 상기 혼합물에 첨가하고, 혼합하고, 15,000xG에서 2분 동안 원심분리기 (벤크만-쿨터 마이크로퓨지(Beckman-Coulter Microfuge) 18)에서 원심분리하였다. 그 후, 540 nm에서의 상기 용액의 흡광도를 캐리 300 바이오

UV-Vis 분광 광도계 (바리안(Varian))를 사용하여 1 cm 경로 길이 큐벳에서 측정하였다. 인증 2-KIV (플루카 (Fluka))를 사용한 표준 곡선을 바탕으로 하면, 540 nm에서의 1 OD 흡광도가 0.28 mM 2-KIV에 의해 생성된다. 2-KIV 형성 속도를 패스트 DHMB 또는 슬로우 DHMB 중 어느 하나의 존재 및 부재 하에서 측정하였다. 둘 모두의 DHMB 형태는 DHIV의 경쟁적 저해제와 같이 거동하였다. 이들의 저해 상수 (K_i)를 단순한 경쟁적 저해에 대한 미하엘리스-멘텐 등식으로부터 계산하였다: $v = S \cdot V_{\max} / (S + K_M \cdot (1 + I/K_i))$ 상기 식에서, v 는 측정된 2-KIV 형성 속도이며, S 는 초기 DHIV 농도이고, $V_{\max}$ 는 10 mM DHIV에서 그리고 DHMB가 없는 상태에서 관찰된 속도로부터 계산된 최대 속도이며, K_M 은 0.5 mM의 이전에 측정된 상수이고, I 는 DHMB의 농도이다. DHMB의 패스트 및 슬로우 이성체는 계산된 저해 상수가 각각 7 mM 및 5 mM였다.

[0819] 실시예 31

[0820] YMR226C 상동체의 확인

[0821] 사카로마이세스 세레비자이의 YMR226C 유전자의 상동체를, 젠뱅크 비중복 뉴클레오티드 데이터베이스의 BLAST 검색 (blast.ncbi.nlm.nih.gov/Blast.cgi), 사카로마이세스 게놈 데이터베이스의 진균류 게놈 BLAST 검색 도구 (www.yeastgenome.org/cgi-bin/blast-fungal.pl), 및 제노레부레스 프로젝트(Genolevures Project)의 BLAST 도구 (genolevures.org/blast.html#)에 의해 찾았다. YMR226C에 대하여 높은 서열 동일성을 나타내는 18가지의 효모 종으로부터의 독특한 서열을 확인하였으며, 이들 유전자에 있어서의 전 ORF를 관련 데이터베이스에서의 기입된 기록으로부터 회복하였다. 이들 ORF에 의해 코딩되는 폴리펩티드 서열을 벡터 NTI (미국 캘리포니아주 칼스배드 소재의 인비트로젠)의 번역 특징(Translation feature)에 의해 결정하였다. 폴리뉴클레오티드 및 폴리펩티드 서열을 하기 표 33에 나타낸다. 효모 종, 뉴클레오티드 데이터베이스 등록 번호, 및 DNA 및 단백질 서열이 상기 표에 주어졌다. 에스. 클루이베리 서열은 주어진 등록 번호로 제노레부레스 데이터베이스에 있으며; 다른 것들은 젠뱅크에 있다. 서열들 사이의 동일성 (%)을 표 34에 나타낸다.

[0822] 18개의 ORF를 얼라인엑스를 사용하여 정렬하였다 (벡터 NTI; 뉴로스포라 크라사 유래의 추정 NADP⁺-의존성 데하이드로게나아제를 코딩하는 유전자 (XM_957621, YMR226C 뉴클레오티드 서열을 이용한 젠뱅크 BLAST 검색에서 확인됨)를 외집단으로 사용하였다). 생성된 계통수를 도 11에 나타내며, 서열 정렬을 도 12에 나타낸다.

[0823] YMR226C에 대한 이들 상동체의 서열 동일성은 최소 55% (야로위아 리포라이티카 및 쉬조사카로마이세스 폼베)로부터 최대 90% (에스. 파라독수스)까지의 범위이다. 또한, BLAST 검색에 의하면 484 염기쌍에 걸쳐 92% 서열 동일성을 갖는 에스. 파스토리아누스(*S. pastorianus*) 유래의 cNDA (등록 번호: CJ997537)가 나타났지만, 이 종은 에스. 바야누스 (이의 YMR226C 상동체는 에스. 세레비자이 서열에 대하여 82% 동일성을 나타냄) 사이의 하이브리드이기 때문에, 그리고 단지 ORF 부분 서열이 입수가가능하기 때문에, 이 서열을 비교에 포함시키지 않았다. 표준 실험실 주 S288C 유래의 YMR226C 서열을 에스. 세레비자이의 12가지의 다른 주로부터의 서열과 비교했을 때, 단지 4개의 단일-뉴클레오티드 다형성이 발견되며 (서열 동일성: 99.5%), 이는 이것이 상기 종에서 고도로 보존된 유전자임을 나타내는 것이었다.

[0824] [표 33]

YMR226C 호모 상동체

종	등록 번호 (데이터베이스 등록일)	핵산 서열 번호	아미노산 서열 번호
사카로마이세스 파라독수스	AABY01000127 (3/7/2011)	698	699
사카로마이세스 바야누스	AACA01000631 (3/7/2011)	700	701
사카로마이세스 카스텔리이	AACF01000116 (3/7/2011)	702	703
사카로마이세스 미카테	AACH01000019 (3/7/2011)	704	705
아쉬비아 고시피이	AE016819 (3/7/2011)	706	707
칸디다 글라브라타	CR380959 (3/7/2011)	708	709
테바료마이세스 한세나이	CR382139 (3/7/2011)	710	711
세페르소마이세스 스티피티스 (이전에는 피키아 스티피티스)	XM_001387479 (3/7/2011)	712	713
메이에로자이마 구일리에르몬디이 (이전에는 피키아 구일리에르몬디이)	XM_001482184 (3/7/2011)	714	715
반데르발토자이마 폴리스포라 (이전에는 클루이베로마이세스 폴리스포루스)	XM_001645671 (3/4/2011)	716	717
칸디다 두블리니엔시스	XM_002419771 (3/7/2011)	718	719
자이코사카로마이세 스 로옥시이	XM_002494574 (3/7/2011)	720	721
라칸세아 서모톨레란스 (이전에는 클루이베로마이세스 서모톨레란스)	XM_002553230 (3/4/2011)	722	723
클루이베로마이세스 락티스	XM_451902 (3/4/2011)	724	725
사카로마이세스 클루이베리	SAKL0H04730 (3/7/2011)	726	727
야로위아 리포라이트카	XM_501554 (3/8/2011)	728	729
위조사카로마이세스 폼베	NM_001018495 (3/8/2011)	730	731

[0825]

[0826] [표 34]

YMR226C 상동체 동일성 (%)

종	Sm	Sb	Sca	Ag	DH	SS	Mg	Cd	Cg	Vp	Sk	Kl	Lt	Zr	Sce	Sp	Yl	NC
Spa	88	82	70	64	62	62	58	57	67	68	68	69	68	68	90	55	55	56
Sm		82	70	64	60	62	58	56	67	69	68	70	68	69	86	57	56	57
Sb			71	63	59	62	58	53	67	66	68	70	69	67	82	56	56	58
Sca				60	62	61	60	59	65	69	69	71	64	70	69	57	53	54
Ag					56	60	57	54	59	61	62	62	62	62	63	54	55	55
DH						64	62	61	61	63	62	61	59	63	62	57	57	53
SS							68	64	61	62	62	64	62	63	62	56	58	58
Mg								60	57	58	60	60	59	62	59	57	57	56
Cd									57	62	59	60	54	60	58	57	53	49
Cg										69	70	68	67	67	66	55	56	55
Vp											71	72	67	70	71	58	52	51
Sk												77	71	72	69	53	54	54
Kl													71	72	71	56	52	54
Lt														69	69	53	60	58
Zr															69	58	55	55
Sce																55	55	56
Sp																	58	60
Yl																		61
NC																		

[0827]

[0828] 표 34의 키(key): 사카로마이세스 파라독수스 ("Spa"); 사카로마이세스 미카태 ("Sm"); 사카로마이세스 바야누스 ("Sb"); 사카로마이세스 카스텔리이 ("Sca"); 아쉬비아 고시피이 ("Ag"); 테바료마이세스 한세니이 ("Dh"); 쉐페르소마이세스 스티피티스 ("Ss"); 메이에로자이마 구일리에르몽디이 ("Mg"); 칸디다 두블리니엔시스 ("Cd"); 칸디다 글라브라타 ("Cg"); 반테르발토자이마 폴리스포라 ("Vp"); 사카로마이세스 클루이베리 ("Sk"); 클루이베로마이세스 락티스 ("Kl"); 라칸세아 서모틀레란스 ("Lt"); 자이고사카로마이세스 로옥시이 ("Zr"); 사카로마이세스 세레비지에 ("Sce"); 쉬조사카로마이세스 폼베 ("Spo"); 야로위아 리포라이티카 ("Yl"); 뉴로스포르라 크라사 ("Nc")

[0829] 실시예 32(가공)

[0830] 효소적 분석을 이용한, 아이소부티르알데히드의 아이소부티르산으로의 전환 능력에 대한 에스. 세레비지에 유래의 알데히드 데하이드로게나아제의 스크리닝

[0831] 이 실시예는 에스. 세레비지에로부터의 어떠한 내인성 알데히드 데하이드로게나아제가 아이소부티르알데히드를 아이소부티르산으로 효소에 의해 전환시킬 수 있는지를 결정하는 방법을 보여 준다.

[0832] 공지된 알데히드 데하이드로게나아제 효소 내에 개별 파괴를 함유하는 에스. 세레비지에 주를 다음과 같이 ATCC로부터 획득한다: BY4741 $\Delta ald2::kanMX4$ (ATCC #4000753); BY4741 $\Delta ald3::kanMX4$ (ATCC #4000752); BY4741 $\Delta ald4::kanMX4$ (ATCC #4001671); BY4741 $\Delta ald5::kanMX4$ (ATCC #4000213); 및 BY4741 $\Delta ald6::kanMX4$ (ATCC #4002767).

[0833] 상기 결실 주들은 먼저 30℃에서 5 ml YPD 배지를 함유하는 튜브 내에서 하룻밤 성장시킨다. 5 ml의 하룻밤 배양물을 500 ml 플라스크 내의 100 ml의 배지 내로 옮기고, 220 rpm에서 진탕하면서 30℃에서 인큐베이션한다. 상기 배양물은 이것이 600 nm에서 1 내지 2 O.D.에 도달할 때 수확한다. 샘플을 10 mL의 20 mM 트리스(pH 7.5)로 세척하고, 그 후 1 mL의 동일한 트리스 완충제에서 재현탁시킨다. 샘플을 0.1 mm 실리카(라이싱 매트릭스 B, MP 바이오메디칼스(Lysing Matrix B, MP biomedical))가 함유된 2.0 mL 튜브 내로 옮겼다. 그 후, 상기 세포를 비드-비터(bead-beater) (BI0101)에서 파쇄한다. 마이크로퓨지에서 4℃에서 30분 동안 13,000 rpm으로 원심분리함으로써 상청액을 수득하였다. 전형적으로는, 0.06 내지 0.1 mg의 조 추출 단백질을 단일 분석에서 사용한다. 조 추출물 중 단백질을 쿠마시 염색(Coomassie stain)을 사용하는 브래드포드 분석법으로 측정하였다.

[0834] 알데히드 데하이드로게나아제 활성을 시그마-알드리치에 의해 그리고 보스티안(Bostian) 및 베츠(Betts)에 의해(문헌[Bostian, K.A. and Betts, G.F. (1978) Biochemical Journal 173, 773-786]) 주어진 프로토콜에 따라 측정한다. 상기 결실 주로부터의 조 추출물 및 구매가능한 알데히드 데하이드로게나아제를 이 방법을 이용하여 시험한다.

[0835] 대안적인 분석법은 밀봉 유리 GC 바이알에서 대략 0.1 mg의 조 추출 단백질에 아이소부티르알데히드를 1 내지

30 mM의 농도로 첨가하고 이를 30℃에서 30분 동안 인큐베이션하는 것이다. 그 후, 상기 추출물을 0.22 μM 스피린 필터 (코닝, 카탈로그 번호 8169)를 통하여 3000 rpm에서 3분 동안 원심분리하고, 여과액을 GC-MS 분석을 위하여 GC 바이알로 옮긴다. 아이소부티르알데히드 및 아이소부티르산이 검출된다.

[0836] 상기 GC 방법은 검출을 위한 5975 질량 분광계, 및 DB-1701 컬럼 (30 m x 0.25 mm ID, 0.25 μM 필름)을 갖춘 애질런트 7890 GC (미국 캘리포니아주 산타 클라라 소재의 애질런트로부터의 것)를 이용한다. 캐리어 가스는 1.0 mL/min의 일정 유량의 헬륨이며; 주입기 분할비(injector split)는 250℃에서 1:10이며; 오븐 온도는 40℃, 1분; 10℃/min으로 40℃에서 120℃까지; 그리고 30℃/min으로 120℃에서 240℃까지이다. MS 검출은 아이소부티르알데히드 및 아이소부티르산의 확인 및 정량화를 위해 전체 스캔 모드(full scan mode)로 사용한다. 보정된 표준 곡선을 다음의 화합물에 대하여 생성하였다: 아이소부티르알데히드, 아이소부티르산, 및 아이소부탄올.

[0837] 실시예 33

[0838] 에스. 세레비지에에서의 아이소부탄올 경로 유전자 발현을 위한 발현 벡터의 제작

[0839] pLH475-JEA1 제작

[0840] 효모에서의 ALS 및 KARI의 발현을 위해 pLH475-JEA1 플라스미드 (서열 번호 419)를 제작하였다. pLH475-JEA1은 하기 키메라 유전자: 1) CUP1 프로모터 (서열 번호 789), 바실러스 서브틸리스 유래의 아세트락테이트 신타아제 코딩 영역 (AlsS; 서열 번호 790; 단백질 서열 번호 791) 및 CYC1 종결 서열 2 (서열 번호 792); 2) ILV5 프로모터 (서열 번호 793), Pf5.IlvC-JEA1 코딩 영역 및 ILV5 종결 서열 (서열 번호 794); 및 3) FBA1 프로모터 (서열 번호 795), 에스. 세레비지에 KARI 코딩 영역 (ILV5; 서열 번호 796; 단백질 서열 번호 797) 및 CYC1 종결 서열 (서열 번호 798)을 함유하는 pHR81 벡터 (ATCC #87541)였다.

[0841] Pf5.IlvC-JEA1 코딩 영역은, 본 명세서에 전체적으로 참고로 포함된 미국 특허 공개 제2009/0163376호 및 제2010/0197519호에 기재된, 슈도모나스 플루오레센스로부터 유래하는 KARI를 코딩하나 돌연변이를 함유하는 서열이다. Pf5.IlvC-JEA1 코딩된 KARI (핵산 및 아미노산 서열 번호 799 및 800)는, 천연 슈도모나스 플루오레센스 KARI와 비교하여, 아미노산 변화를 갖는다.

[0842] 발현 벡터 pLH468

[0843] 효모에서의 DHAD, KivD, 및 HADH의 발현을 위해 pLH468 플라스미드 (서열 번호139)를 제작하였다.

[0844] 사카로마이세스 세레비지에에서의 발현을 위해 최적화된 코돈에 기초하여 DNA2.0에 의해 엘. 락티스 케토아이소발레레이트 데카르복실라아제 (KivD) 및 말 간 알코올 데하이드로게나아제 (HADH)의 코딩 영역을 합성하고 (각각 서열 번호 801 및 802), 플라스미드 pKivDy-DNA2.0 및 pHadhy-DNA2.0에 제공하였다. 코딩된 단백질은 각각 서열 번호 803 및 804의 서열이다. KivD 및 HADH에 대한 개개의 발현 벡터를 제작하였다. 본 명세서에 참고로 포함된 미국 특허 출원 제2008/0182308호, 실시예 17에 기재된 pLH467 (pRS426::P_{TDH3}-kivDy-TDH3t)를 조립하기 위하여, 벡터 pNY8 (pRS426.GPD-ald-GPDt로도 명명됨)을 AscI 및 SfiI 효소로 절단하였고, 따라서 GPD 프로모터 및 ald 코딩 영역을 잘라내었다. 5' 프라이머 OT1068 및 3' 프라이머 OT1067 (서열 번호 806 및 807)를 사용하여, pNY8로부터의 TDH3 프로모터 단편 (서열 번호 805)을 PCR 증폭시켜 5' 말단에서 AscI 부위 및 3' 말단에서 SpeI 부위를 부가하였다. AscI/SfiI 절단된 pNY8 벡터 단편을 AscI 및 SpeI로 절단된 TDH3 프로모터 PCR 생성물과 라이게이션하였고, 코돈 최적화된 kivD 코딩 영역을 함유하는 SpeI-SfiI 단편을 벡터 pKivD-DNA2.0로부터 단리하였다. 3중 라이게이션은 벡터 pLH467 (pRS426::PTDH3-kivDy-TDH3t)를 생성하였다. pLH467 (서열 번호 808)을 제한효소 지도화 및 서열결정에 의해 검증하였다.

[0845] pLH435 (pRS425::PGPM1-Hadhy-ADH1t)를, 본 명세서에 참고로 포함되는 미국 특허 공개 제2009/0305363호, 실시예 3에 기재된, 벡터 pRS425::GPM-sadB (서열 번호 809)로부터 유도하였다. pRS425::GPM-sadB는 GPM1 프로모터 (서열 번호 810), 아크로모박터 자일로속시단스의 부탄올 데하이드로게나아제로부터의 코딩 영역 (sadB; DNA 서열 번호 811; 단백질 서열 번호 812), 및 ADH1 종결 서열 (서열 번호 444)를 함유하는 키메라 유전자를 갖는 pRS425 벡터 (ATCC #77106)이다. pRS425::GPMp-sadB는 sadB 코딩 영역의 5' 및 3' 말단에 각각 BbvI 및 PacI 부위를 함유한다. 프라이머 OT1074 및 OT1075 (서열 번호 813 및 814)를 사용하여 부위 지정 돌연변이 유발에 의해 NheI 부위를 sadB 코딩 영역의 5' 말단에서 부가하여, 벡터 pRS425-GPMp-sadB-NheI을 생성하였고, 이를 서열결정에 의해 검증하였다. pRS425::P_{GPM1}-sadB-NheI을 NheI 및 PacI로 절단하여 sadB 코딩 영역을 빼내고, 벡터 pHadhy-DNA2.0으로부터의 코돈 최적화된 HADH 코딩 영역을 함유하는 NheI-PacI 단편과 라이게이션시켜,

pLH435 (서열 번호 815)를 생성하였다.

[0846] 단일 벡터 내에서 KivD 및 HADH 발현 카세트를 조합하기 위해, 효모 벡터 pRS411 (ATCC # 87474)을 *SacI* 및 *NotI*로 절단하고, P_{GPM1}-*Hadhy-ADH1t* 카세트를 함유하는 pLH435로부터의 *SalI-NotI* 단편과 함께 P_{TDH3}-*kivDy-TDH3t* 카세트를 함유하는 pLH467로부터의 *SacI-SalI* 단편과 삼중 라이게이션 반응으로 라이게이션시켰다. 이로써 벡터 pRS411::P_{TDH3}-*kivDy*-P_{GPM1}-*Hadhy* (pLH441, 서열 번호 816)를 수득하였고, 이를 제한효소 지도화에 의해 검증하였다.

[0847] 더 낮은 아이소부탄을 경로에서 모든 세 유전자, *ilvD*, *kivDy* 및 *Hadhy*를 위한 동시 발현 벡터를 생성하기 위하여, pRS423 FBA *ilvD*(Strep) (서열 번호 817)를, *IlvD* 유전자의 공급원으로서 사용하였다. 이 서열 벡터는 이. 콜라이에서의 유지를 위한 F1 복제 기점 (뉴클레오티드 1423 내지 1879) 및 효모에서의 복제를 위한 2 마이크로 기점 (뉴클레오티드 8082 내지 9426)을 함유한다. 벡터는 FBA1 프로모터 (nt 2111 내지 3108; 서열 번호 795) 및 FBA1 종결 서열 (nt 4861 내지 5860; 서열 번호 818)을 갖는다. 또한, 이는 효모에서의 선발을 위한 His 마커 (뉴클레오티드 504 내지 1163), 및 이. 콜라이에서의 선발을 위한 암피실린 내성 마커(뉴클레오티드 7092 내지 7949)를 갖는다. 스트렙토코커스 뮤탄스 UA159 (ATCC #700610) 유래의 *ilvD* 코딩 영역 (뉴클레오티드 3116 내지 4828; 서열 번호 819; 단백질 서열 번호 820)은 FBA1 프로모터와 FBA1 종결 서열 사이에 있어서 발현을 위한 키메라 유전자를 형성한다. 또한, *ilvD* 코딩 영역(뉴클레오티드 4829 내지 4849)에 융합되는 루미오 태그가 존재한다.

[0848] 제1 단계는 pRS423 FBA *ilvD*(Strep) (pRS423-FBA(SpeI)-*IlvD*(스트렙토코커스 뮤탄스)-루미오로도 칭해짐)를 *SacI* 및 *SacII* (*SacII* 부위는 T4 DNA 폴리머라아제를 사용하여 불린트 말단화됨)를 이용하여 선형화하여, 총 길이가 9,482 bp인 벡터를 제공하는 것이었다. 제2 단계는 *SacI* 및 *KpnI* (*KpnI* 부위는 T4 DNA 폴리머라아제를 사용하여 불린트 말단화됨)를 사용하여 pLH441로부터 *kivDy-hADHy* 카세트를 단리하는 것이었으며, 이는 6,063 bp 단편을 제공한다. 이 단편을 pRS423-FBA(SpeI)-*IlvD*(스트렙토코커스 뮤탄스)-루미오로부터의 9,482 bp 벡터 단편과 라이게이션하였다. 이는 벡터 pLH468 (pRS423::P_{FBA1}-*ilvD*(Strep)루미오-FBA1t-P_{TDH3}-*kivDy-TDH3t*-P_{GPM1}-*hadhy-ADH1t*)를 생성하였으며, 이를 제한효소 지도화 및 서열결정에 의해 확인하였다.

[0849] 실시예 34

[0850] 피루베이트 데카르복실라아제 및 헥소키나아제 2에서의 변형을 포함하는 에스. 세레비지에 숙주 주의 제작

[0851] 이 실시예는 에스. 세레비지에의 내인성 PDC1, PDC5, 및 PDC6 유전자의 삽입-불활성화를 설명한다. PDC1, PDC5, 및 PDC6 유전자는 피루베이트 데카르복실라아제의 3가지 주요 아이소자임을 코딩한다. 글루코스의 포스포릴화 및 전사 억제를 초래하는 헥소키나아제 2가 또한 불활성화된다. 생성된 PDC/HXK2 불활성화 주를, 발현 벡터 pLH475-JEA1 및 pLH468을 위한 숙주로서 사용하였다.

[0852] pdc6:: P_{GPM1}-sadB 통합 카세트의 제작 및 PDC6 결실:

[0853] pRS425::GPM-sadB (참고로 포함된 미국 특허 공개 제2009/0305363호에 기재됨)로부터의 GPM-sadB-ADH1t 절편 (서열 번호 821)을 pUC19-URA3r로부터의 URA3r 유전자에 연결하여, pdc6::P_{GPM1}-sadB-ADH1t-URA3r 통합 카세트를 제조하였다. pUC19-URA3r (서열 번호 822)은 75 bp 상동 반복 서열이 측면에 위치하는 pRS426 (ATCC # 7710 7)로부터의 URA3 마커를 함유하여, 생체 내 상동 재조합 및 URA3 마커의 제거를 가능하게 한다. 2개의 DNA 절편을 SOE PCR (문헌[Horton *et al.* (1989) *Gene* 77:61-68]에 기재된 바와 같음)에 의해, 주형으로서 pRS425::GPM-sadB 및 pUC19-URA3r 플라스미드 DNA를 사용하여, 퓨전 DNA 폴리머라아제 (미국 매사추세츠주 비벌리 소재의 뉴 잉글랜드 바이오랩스 인크.; 카탈로그 번호 F-540S) 및 프라이머 114117-11A 내지 114117-11D (서열 번호: 823, 824, 825, 826), 및 114117-13A 및 114117-13B (서열 번호: 827 및 828)에 의해 연결하였다. SOE PCR (114117-13A 및 114117-13B)을 위한 외측 프라이머는 각각 PDC6 프로모터 및 종결 서열의 상류 및 하류의 영역에 상동성인 5' 및 3'의 약 50 bp 영역을 포함하였다. 표준 유전자 기술 (문헌[*Methods in Yeast Genetics*, 2005, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, pp. 201-202])을 사용하여, BY4700 (ATCC # 200866)을 완성된 카세트 PCR 단편으로 형질전환시키고, 형질전환체를 30°C에서, 2% 글루코스를 보충하고 우라실이 결여된 합성 완전 배지 상에서 유지하였다. 형질전환체를, 프라이머 112590-34G 및 112590-34H (서열 번호: 829 및 830), 및 112590-34F 및 112590-49E (서열 번호: 831 및 832)를 사용하여 PCR에 의해 스크리닝하여, PDC6 코딩 영역의 결실을 갖는 PDC6 유전자좌에서의 통합을 검증하였다. URA3r 마커를 표준 프로토콜에 따라 30°C에서, 2% 글루코스 및 5-FOA가 보충된 합성 완전 배지 상에 플레이트링하여 재생하였다. 5-

FOA 플레이트로부터 SD-URA 배지 상으로 콜로니를 패칭함으로써 성장의 부재를 검증하여 마커 제거를 확인하였다. 생성된 확인된 주는 다음의 유전자형을 갖는다: BY4700 *pdv6::P_{GPM1}-sadB-ADH1t*.

[0854] *pdv1::P_{PDC1}-ilvD* 통합 카세트의 제작 및 PDC1 결실:

[0855] 주형으로서 *pLH468* 및 *pUC19-URA3r* 플라스미드 DNA를 사용하여, SOE PCR (문헌[Horton et al. (1989) *Gene* 77:61-68]에 기재된 바와 같음)에 의해, 퓨전 DNA 폴리머라아제 (미국 매사추세츠주 비벌리 소재의 뉴 잉글랜드 바이오랩스 인크.; 카탈로그 번호 F-540S) 및 프라이머 114117-27A 내지 114117-27D (서열 번호: 823, 824, 825, 826)로, *pLH468*로부터의 *ilvD-FBA1t* 절편을 *pUC19-URA3r*로부터의 *URA3r* 유전자에 연결하여, *pdv1::PPDC1-ilvD-FBA1t-URA3r* 통합 카세트 (서열 번호: 833)를 제조하였다.

[0856] SOE PCR (114117-27A 및 114117-27D)을 위한 외측 프라이머는 *PDC1* 프로모터의 하류 및 *PDC1* 코딩 서열의 하류의 영역에 상동성인 5' 및 3'의 약 50 bp 영역을 포함하였다. 표준 유전자 기술 (문헌[Methods in Yeast Genetics, 2005, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, pp. 201-202])을 사용하여, BY4700 *pdv6::P_{GPM1}-sadB-ADH1t*를 완성된 카세트 PCR 단편으로 형질전환시키고 형질전환체를 30°C에서 2% 글루코스가 보충되고 우라실이 결여된 합성 완전 배지 상에 유지하였다. 프라이머 114117-36D 및 135 (서열 번호 834 및 835), 및 프라이머 112590-49E 및 112590-30F (서열 번호 832 및 836)을 사용하여 PCR에 의해 형질전환체를 스크리닝하여, *PDC1* 코딩 서열의 결실을 갖는 *PDC1* 유전자좌에서의 통합을 검증하였다. *URA3r* 마커를 표준 프로토콜에 따라 30°C에서, 2% 글루코스 및 5-FOA가 보충된 합성 완전 배지 상에 플레이팅하여 재생하였다. 5-FOA 플레이트로부터 SD-URA 배지 상으로 콜로니를 패칭함으로써 성장의 부재를 입증하여 마커 제거를 확인하였다. 생성된 확인된 주 "NYLA67"는 다음의 유전자형을 갖는다: BY4700 *pdv6::P_{GPM1}-sadB-ADH1t pdv1::P_{PDC1}-ilvD-FBA1t*.

[0857] *HIS3* 결실

[0858] 내인성 *HIS3* 코딩 영역을 결실시키기 위하여, *URA3r2* 주형 DNA (서열 번호 837)로부터 *his3::URA3r2* 카세트를 PCR-증폭시켰다. *URA3r2*은 500 bp 상동 반복 서열이 측면에 위치하는 pRS426 (ATCC # 77107)으로부터의 *URA3* 마커를 함유하여, 생체 내 상동 재조합 및 *URA3* 마커의 제거를 가능하게 한다. 퓨전 DNA 폴리머라아제 및 프라이머 114117-45A 및 114117-45B (서열 번호: 838 및 839)를 사용하여 PCR을 행하였고, 이는 약 2.3 kb PCR 생성물을 생성하였다. 각각의 프라이머의 *HIS3* 부분은 *HIS3* 프로모터의 상류의 5' 영역 및 코딩 영역의 하류의 3' 영역으로부터 유래하여 *URA3r2* 마커의 통합이 *HIS3* 코딩 영역의 대체를 야기하도록 하였다. 표준 유전자 기술 (문헌[Methods in Yeast Genetics, 2005, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, pp. 201-202])을 사용하여 PCR 생성물로 NYLA67을 형질전환시키고, 30°C에서, 2% 글루코스가 보충되고 우라실이 결여된 합성 완전 배지 상에서 형질전환체를 선발하였다. 형질전환체를 스크리닝하여, 30°C에서, 2% 글루코스가 보충되고 히스티딘이 결여된 합성 완전 배지 상으로의 형질전환체의 레플리카(replica) 플레이팅에 의해 올바른 통합을 검증하였다. 표준 프로토콜에 따라 30°C에서 2% 글루코스 및 5-FOA가 보충된 합성 완전 배지 상에 플레이팅하여 *URA3r* 마커를 재생하였다. 5-FOA 플레이트로부터 SD -URA 배지로 콜로니를 패칭하여, 성장이 없음을 검증함으로써, 마커 제거를 확인하였다. NYLA73으로 칭해지는, 생성된 확인된 주는 다음의 유전자형을 갖는다: BY4700 *pdv6::P_{GPM1}-sadB-ADH1t pdv1::P_{PDC1}-ilvD-FBA1t Δhis3*.

[0859] *HXK2*의 결실:

[0860] 퓨전 DNA 폴리머라아제 및 프라이머 384 및 385 (서열 번호: 840 및 841)를 사용하여, *hvk2::URA3r2* 주형 (상기에 기재됨)으로부터 *URA3r* 카세트를 PCR-증폭시켰고, 이는 약 2.3 kb PCR 생성물을 생성하였다. 각각의 프라이머의 *HXK2* 부분은 *HXK2* 프로모터의 상류의 5' 영역 및 코딩 영역의 하류의 3' 영역으로부터 유래하여, *URA3r2* 마커의 통합이 *HXK2* 코딩 영역의 대체를 야기하도록 하였다. 표준 유전자 기술 (문헌[Methods in Yeast Genetics, 2005, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, pp. 201-202])을 사용하여 PCR 생성물로 NYLA73을 형질전환시키고, 30°C에서 2% 글루코스가 보충되고 우라실이 결여된 합성 완전 배지 상에서 형질전환체를 선발하였다. 형질전환체를 PCR에 의해 스크리닝하여, 프라이머 N869 및 N871 (서열 번호: 842 및 843)을 사용하여 *HXK2* 코딩 영역이 대체된 *HXK2* 유전자좌에서의 올바른 통합을 검증하였다. *URA3r2* 마커는 표준 프로토콜에 따라 30°C에서 2% 글루코스 및 5-FOA가 보충된 합성 완전 배지 상에 플레이팅함으로써 재생하였다. 5-FOA 플레이트로부터의 콜로니를 SD -URA 배지 상에 패치하여 성장의 부재를 검증함으로써 마커의 제거를 확인하였고, PCR에 의해 프라이머 N946 및 N947 (서열 번호: 844 및 845)를 사용하여 올바른 마커 제거

를 검증하였다. NYLA83으로 명명된 생성된 확인된 주는 다음의 유전자형을 갖는다: BY4700 *pd6::P_{GPM1}-sadB-ADH1t pdc1::P_{PDC1}-ilvD-FBA1t Δhis3 Δhvk2*.

[0861] pd5::kanMX 통합 카세트의 제작 및 PDC5 결실

[0862] 퓨전 DNA 폴리머라아제 및 프라이머 PDC5::KanMXF 및 PDC5::KanMXR (서열 번호: 846 및 847)를 사용하여 주 YLR134W 염색체 DNA (ATCC No. 4034091)로부터 *pd5::kanMX4* 카세트를 PCR-증폭시켰고, 이는 약 2.2 kb PCR 생성물을 생성하였다. 각각의 프라이머의 PDC5 부분은 PDC5 프로모터의 상류의 5' 영역 및 코딩 영역의 하류의 3' 영역으로부터 유래하여, kanMX4 마커의 통합이 PDC5 코딩 영역의 대체를 야기하도록 하였다. PCR 생성물로 NYLA83을 형질전환시키고, 형질전환체를 상기와 같이 선발하고 스크리닝하였다. NYLA84로 명명한, 확인된 올바른 형질전환체는 다음의 유전자형을 갖는다: BY4700 *pd6::P_{GPM1}-sadB-ADH1t pdc1::P_{PDC1}-ilvD-FBA1t Δhis3 Δhvk2 pd5::kanMX4*.

[0863] 표준 유전자 기술 (문헌[*Methods in Yeast Genetics*, 2005, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY])을 사용하여, 플라스미드 벡터 pLH468 및 pLH475-JEA1를 주 NYLA84 (BY4700 *pd6::P_{GPM1}-sadB-ADH1t pdc1::P_{PDC1}-ilvD-FBA1t Δhis3 Δhvk2 pd5::kanMX4*) 내로 동시에 형질전환시키고, 생성된 주를 30℃에서 1% 에탄올이 보충되고 히스티딘 및 우라실 결여된 합성 완전 배지 상에 유지하였다.

[0864] 실시예 35 (가공)

[0865] 피루베이트 데카르복실라아제, 헥소키나아제 2, 및 알데히드 데하이드로게나아제에서의 변형을 포함하는 에스. 세레비지에 숙주 주의 제작.

[0866] 이 실시예는 아이소부티르산의 형성을 무효화하거나 또는 감소시키는 하나 이상의 알데히드 데하이드로게나아제의 불활성화를 설명한다. *ALD2*의 파괴를 예로서 사용하나, 본 방법은 임의의 알데히드 데하이드로게나아제의 파괴에 사용할 수 있다. 생성된 NYLA84-유도된 주는 발현 벡터 pLH475-JEA1 및 pLH468을 위한 숙주로서 사용한다.

[0867] ALD2의 결실:

[0868] 퓨전 DNA 폴리머라아제 및 프라이머 T001 및 T002 (서열 번호: 848 및 849)를 사용하여 *ald2::URA3r2* 주형 (상기에 기재됨)으로부터 *URA3r* 카세트를 PCR-증폭시켰고, 이는 약 2.3 kb PCR 생성물을 생성하였다. 각각의 프라이머의 *ALD2* 부분은 *ALD2* 유전자의 5' 서열 및 3' 서열로부터 유래하여, *URA3r2* 마커의 통합이 *ALD2* 코딩 영역의 대체를 야기하도록 한다. 표준 유전자 기술(문헌[*Methods in Yeast Genetics*, 2005, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, pp. 201-202])을 사용하여 PCR 생성물로 NYLA84를 형질전환시키고, 30℃에서 1% 에탄올이 보충된 우라실이 결여된 합성 완전 배지 상에서 형질전환체를 선발한다. 형질전환체를 PCR에 의해 스크리닝하여, 프라이머 T003 및 T004 (서열 번호: 850 및 851)를 사용하여, *ALD2* 코딩 영역이 대체된 *ALD2* 유전자좌에서의 올바른 통합을 검증한다. *URA3r2* 마커는 표준 프로토콜에 따라 30℃에서 1% 에탄올 및 5-FOA이 보충된 합성 완전 배지 상에 플레이트링함으로써 재생한다. 5-FOA 플레이트로부터 콜로니를 SE -URA 배지 상에 패치하여 성장이 없음을 검증함으로써 마커 제거를 확인하고, PCR에 의해 프라이머 T004 및 T005 (서열 번호: 851 및 852)를 사용하여 올바른 마커 제거를 검증한다. 생성된 확인된 주는 NYLA84 *Δald2*로 명명하며, 다음의 유전자형을 갖는다: BY4700 *pd6::P_{GPM1}-sadB-ADH1t pdc1::P_{PDC1}-ilvD-FBA1t Δhis3 Δhvk2 Δald2*.

[0869] 실시예 36

[0870] NYLA84에서 재조합 에스. 세레비지에에서의 아이소부탄올의 생성

[0871] HPLC 방법

[0872] 글루코스 및 발효 부산물 조성물의 분석은 당업계에 잘 알려져 있다. 예를 들어, 고성능 액체 크로마토그래피 (HPLC) 방법에서는 굴절률(RI) 검출기와 함께, 쇼텍스 SH-G 가드 컬럼을 동반하는 쇼텍스 SH-1011 컬럼(둘 모두 미국 매사추세츠주 밀포드 소재의 워터스 코퍼레이션으로부터 입수가능)을 사용한다. 0.01 M H₂SO₄를 이동상으로 사용하여 50℃의 컬럼 온도 및 0.5 mL/min의 유량으로 크로마토그래피 분리를 달성한다. 아이소부탄올 체류 시간은 47.6분이다.

- [0873] NYLA84에서 재조합 에스. 세레비지에에서의 아이소부탄올의 생성A84
- [0874] 이 실시예의 목적은 효모 주 NYLA84에서의 아이소부탄올의 생성을 설명하는 것이다. 효모 주는 피루베이트 데카르복실라아제의 3개의 아이소자임을 코딩하는 유전자, PDC1, PDC5, 및 PDC6의 결실, 및 헥소키나아제 2를 코딩하는 HXK2의 결실을 포함한다. 상기 주는 또한 AlsS (아세토락테이트 신타아제), KARI (케토산 리덕토아이스머라아제), DHAD (다이하이드록시산 데하이드라타아제), KivD (케토아이스발레이트 데카르복실라아제), SadB (2차 알코올 데하이드로게나아제), 및 HADH (말 간 알코올 데하이드로게나아제)의 이중성 발현을 위한 제작물을 함유한다.
- [0875] 주 제작
- [0876] 표준 PEG/아세트산리튬-매개 형질전환 방법에 의해, 플라스미드 pLH468 및 pLH475-JEA1를 NYLA84 내로 도입하였다. 형질전환체를 글루코스, 히스티딘 및 우라실 이 결여된 합성 완전 배지 상에서 선발하였다. 에탄올 (1% v/v)을 탄소 공급원으로서 사용하였다. 3일 후에, 2% 글루코스 및 0.5% 에탄올 둘 모두를 탄소 공급원으로서 보충한 히스티딘 및 우라실 이 결여된 합성 완전 배지에 형질전환체를 패치하였다. 45% 글리세롤을 사용하여 로그기(log-phase) 배양물을 15% (w/v)의 최종 글리세롤 농도로 희석하여 냉동 바이알을 제조하였다.
- [0877] 아이소부탄올의 생성
- [0878] 냉동 바이알을 사용하여, 2% 글루코스 및 0.5% 에탄올 둘 모두를 탄소 공급원으로서 보충한 히스티딘 및 우라실 이 결여된 합성 완전 배지 80 ml를 접종하였다.
- [0879] 배양 조건:
- [0880] 1 리터 발효기를 준비하고 0.4 L 물로 살균하였다. 냉각 후에, 필터 살균된 배지를 첨가하여, 접종 후 800 mL 중의 하기 최종 농도를 제공하였다:
- [0881] 배지 (최종 농도):
- [0882] 6.7 g/L의, 아미노산을 포함하지 않는 효모 질소 베이스 (디프코)
- [0883] 2.8 g/L의, 히스티딘, 류신, 트립토판 및 우라실을 포함하지 않는 효모 합성 드롭아웃 배지 보충제 (시그마 Y2001 Y2001)
- [0884] 20 mL/L의 1% (w/v) L-류신
- [0885] 4 mL/L의 1% (w/v) L-트립토판
- [0886] 1 mL/L의 에르고스테롤/트윈/에탄올 용액
- [0887] 0.2 mL/L의 시그마 DF204F204
- [0888] 10 g/L의 글루코스
- [0889] KOH를 사용한 pH 5.5, 30% dO, 온도 30°C, 및 기류 0.2 SLPM (또는, 0.25 vvm)로 제어하도록 발효기를 설정하였다. 접종시에, 기류를 초기에는 0.01 SLPM로 설정하였고, 그 후에 일단 성장이 확립되면 0.2 SLPM로 증가시켰다. 글루코스는 수동 첨가에 의해 내내 5-15 g/L로 유지하였다.
- [0890] 공기를 0.25 vvm로 발효기에 계속적으로 공급하였다. 지속적인 통기로, 발효조의 수성상으로부터 아이소부탄올의 상당한 스트리핑이 유발된다. 스트리핑에 기인한 아이소부탄올의 손실을 정량화하기 위하여, 발효조를 빠져나오는 오프-가스를 질량 분석기 (프리마 dB 질량 분석기, 서모 일렉트론 코포레이션 (미국 위스콘신주 매디슨 소재))로 직접 보내, 가스 스트림 중의 아이소부탄올의 양을 정량화하였다. 74 또는 42의 전하에 대한 질량의 비(mass to charge ratio)에서의 아이소부탄올 피크를 계속 모니터링하여, 가스 스트림 중의 아이소부탄올의 양을 정량화하였다.
- [0891] 수성상 중의 글루코스 및 유기산을 HPLC를 사용하여 발효 중에 모니터링하였다. 또한 글루코스를 글루코스 분석기 (와이에스아이 인크.(YSI, Inc.) (미국 오하이오주 옐로우 스프링스 소재))를 사용하여 빠르게 모니터링한다. 발효기로부터 수성상을 주기적으로 빼낸 후에, 수성상 중 아이소부탄올을, 상기의 본 명세서의 HPLC 방법에 기재된 바와 같이 HPLC에 의해 정량화하였다. 스트리핑으로 인한 아이소부탄올 손실에 대해 보정한 유효 역가는 7.5 g/L였다. 아이소부탄올의 역가는 1.28 g/L였다 (도 14).

- [0892] 실시예 37 (가공)
- [0893] NYLA84 $\Delta ald2$ 에서 재조합 에스. 세레비지에에서의 아이소부탄올의 생성.
- [0894] 이 실시예의 목적은 효모 주 NYLA84에서의 아이소부탄올의 생성을 설명하는 것이다. *ALD2*의 파괴를 예로서 사용하나, 본 실시예는 임의의 알데히드 탈하이드로게나아제의 파괴에 사용할 수 있다. NYLA84 $\Delta ald2$ 효모 주는 피루베이트 데카르복실라아제의 3개의 아이소자임을 코딩하는 유전자, *PDC1*, *PDC5*, 및 *PDC6*의 결실, 및 헥소키나아제 2를 코딩하는 *HXX2*의 결실, 및 알데히드 탈하이드로게나아제를 코딩하는 *ALD2*의 결실을 포함한다. 주는 또한 AlsS (아세트락테이트 신타아제), KARI (케토산 리덕토아이소머라아제), DHAD (다이하이드록시산 탈하이드라타아제), KivD (케토아이소발레이트 데카르복실라아제), 및 SadB (2차 알코올 탈하이드로게나아제)의 이중성 발현을 위한 제작물을 함유한다.
- [0895] 주 제작
- [0896] 표준 PEG/아세트산리튬-매개 형질전환 방법에 의해, 플라스미드 pLH468 및 pLH475-JEA1를 NYLA84 $\Delta ald2$ 내로 도입한다. 형질전환체를 글루코스, 히스티딘 및 우라실로 결여된 합성 완전 배지 상에서 선발한다. 에탄올 (1% v/v)을 탄소 공급원으로서 사용한다. 3일 후에, 2% 글루코스 및 0.5% 에탄올 둘 모두를 탄소 공급원으로서 보충한 히스티딘 및 우라실이 결여된 합성 완전 배지에 형질전환체를 패치한다. 45% 글리세롤을 사용하여 로그기 배양물을 15% (w/v)의 최종 글리세롤 농도로 희석하여 냉동 바이알을 제조한다.
- [0897] 아이소부탄올의 생성
- [0898] 냉동 바이알을 사용하여, 0.25% 글루코스 및 0.5% 에탄올 둘 모두를 탄소 공급원으로서 보충한 히스티딘 및 우라실이 결여된 합성 완전 배지 80 ml를 접종한다. 1 리터 발효기를 준비하고 0.4 L 물로 살균한다. 냉각 후에, 필터 여과된 배지를 첨가하여, 접종 후 800 mL 중에서 하기 최종 농도를 제공한다:
- [0899] 배지 (최종 농도):
- [0900] 6.7 g/L의, 아미노산을 포함하지 않는 효모 질소 베이스 (디프코)
- [0901] 2.8 g/L의, 히스티딘, 류신, 트립토판 및 우라실을 포함하지 않는 효모 합성 드롭아웃 배지 보충제 (시그마 Y2001)
- [0902] 20 mL/L의 1% (w/v) L-류신
- [0903] 4 mL/L의 1% (w/v) L-트립토판
- [0904] 1 mL/L의 에르고스테롤/트윈/에탄올 용액
- [0905] 0.2 mL/L의 시그마 DF204
- [0906] 10 g/L의 글루코스
- [0907] KOH를 사용한 pH 5.5, 30% dO, 온도 30°C, 및 기류 0.2 SLPM (또는, 0.25 vvm)로 제어하도록 발효기를 설정한다. 접종시에, 기류를 초기에는 0.01 SLPM로 설정하고, 그 후에 일단 성장이 확립되면 0.2 SLPM로 증가시킨다. 글루코스는 수동 첨가에 의해 내내 5-15 g/L로 유지한다.
- [0908] 스트립핑에 기인한 아이소부탄올의 손실을 정량화하기 위하여, 발효조를 빠져나오는 오프-가스를 질량 분석기 (프리마 dB 질량 분석기, 미국 위스콘신주 매디슨 소재의 서모 일렉트론 코포레이션)로 직접 보내, 가스 스트림 중의 아이소부탄올의 양을 정량화한다. 74 또는 42의 전하에 대한 질량의 비에서의 아이소부탄올 피크를 계속 모니터링하여, 가스 스트림 중의 아이소부탄올의 양을 정량화한다. 수성상 중의 글루코스 및 유기산을 HPLC를 사용하여 발효 중에 모니터링한다. 또한 글루코스를 글루코오스 분석기 (와이에스아이 인크.(미국 오하이오주 옐로우 스프링스 소재))를 사용하여 빠르게 모니터링한다. 발효기로부터 수성상을 주기적으로 빼낸 후에, 수성상 중 아이소부탄올 및 아이소부티르산을 HPLC에 의해 정량화한다.
- [0909] 실시예 38
- [0910] 아이소부티르산 생성의 분석
- [0911] 0.05% 에탄올이 보충된, 히스티딘 및 우라실을 제외한 합성 완료 배지 내로 주를 접종하였다. 일단 배양물이 정지기에 도달하였으면, 신선한 배지 내로 계대 배양하였다 (출발 OD = 0.2). PNY2205 (실시예 13)의 경우에는, 배지를 0.05% 에탄올로 보충하여, 포스포케톨라아제 경로를 갖지 않는 PDC KO 효모에서 관찰되는 C2

요건을 충족시켰다. PNY2209 (실시예 13)의 경우에는, 세포를 에탄올을 포함하는 배지 및 에탄올을 포함하지 않는 배지 내로 계대 배양하였다. ald6Δ 클론 (PNY2216 및 동계(isogenic) 클론, 실시예 24)의 경우에는, 에탄올을 포함하지 않는 배지를 사용하였다. 스크루 캡 진탕 플라스크에서 배양물을 성장시켰다 (125 ml 플라스크 내에 20 ml의 배지). 접종 직후에 배양물 상청액을 수집하였고 48시간 후에 다시 수집하였다. 이들 배양 상청액을 HPLC (본 명세서에 참고로 포함된 미국 특허 공개 제20070092957호)에 의해 분석하여 글루코스 소비량 및 아이소부티르산 농도를 결정하였다.

[0912] [표 35]

ald6Δ 주에서의 아이소부티르산의 물 수율

주	아이소부티르산 물 수율 (소비된 글루코스 1 몰당 물)
PNY2205*	0.09
PNY2209*	0.07
PNY2209	0.09
PNY2216 및 5가지의 다른 동계(isogenic) 클론	0.03

*는 배양 배지에 0.05% 에탄올을 보충하였음을 나타냄

[0913]

[0914] 실시예 39

[0915] 신규 주에서의 아이소부탄올 생성의 증가 및 부산물 생성의 감소

[0916] 이 실시예의 목적은 남아 있는 아이소부탄올 경로 유전자들을 공급하도록 플라스미드 pK9G9.0LE1p.ilvD를 사용하여 새롭게 제작된 주를 이용한 소규모 배양 실험을 설명하는 것이다.

[0917] 신규 숙주 주 PNY1528, PNY2243, PNY2237 및 PNY2238을, 플라스미드 pK9G9.0LE1p.ilvD를 사용하여 형질전환시키고 아이소부탄올 생성에 대해 시험하였다. 탄소 공급원으로서 1% 에탄올을 갖는, 우라실이 제외된 합성 완전 배지 상에서 선발함으로써 형질전환체를 수득하였다. 탄소 공급원으로서 2% 글루코스 및 0.05% 에탄올을 갖는, 우라실이 제외된 합성 완전 배지 상에 형질전환체를 패치하였다. 패치를 이용하여 액체 배지 (우라실이 제외된, 탄소 공급원으로서 0.3% 글루코스 및 0.3% 에탄올을 포함하는 합성 완전 배지)에 접종하였다. 아이소부탄올 생성을 시험하기 위하여, 우라실이 제외된, 탄소 공급원으로서 2% 글루코스 및 0.05% 에탄올을 함유하는 합성 완전 배지 내로 액체 배양물을 계대 배양하였는데, 상기 배지는 BME 비타민 믹스 (시그마 카탈로그 번호 B6891)를 함유하였다. 배양물을 진탕하면서 (인포르스 멀티트론 진탕기에서 250 rpm) 30℃에서 밀봉 세럼 바이알에서 (15 ml 바이알 중 10 ml의 배지) 인큐베이션하였다. 48시간 후에, 배양 배지를 여과하고 (스핀-엑스 컬럼), HPLC로 분석하였다 (이는 미국 특허 공개 제20070092957호에 앞서 기재된 바와 같음). 경로 부산물 아이소부티레이트의 물 수율은 ald6Δ를 지닌 주에서 50%만큼 감소하였다. PNY2238을 기반으로 한 주는 글루코스 소비량 및 아이소부탄올 역가 (하기 표에 나타난 클론 K 결과)가 더 높은 것으로 나타났다.

[0918] [표 36]

ald6Δ 주에서의 아이소부티르산의 물 수율

주	관련 표현형	아이소부티르산 물 수율 (소비된 글루코스 1 몰당 물)
PNY2205*	ALD6+, C2-의존성	0.09
PNY2209*	ALD6+, C2-독립성	0.07
PNY2209	ALD6+, C2-독립성	0.09
PNY2216 및 5가지의 다른 동계 클론	ALD6-, C2-독립성	0.03

*는 배양 배지에 0.05% 에탄올을 보충하였음을 나타냄

[0919]

[0920] 모든 주는 플라스미드 pK9G9.0LE1p.ilvD를 함유하였다. PNY2237- 및 PNY2238-유도된 주의 경우, 제공된 데이터는 2개의 생물학적 레플리케이트의 평균이다.

[0921] 실시예 40

[0922] PNY2259에서 생성된 변이체의 아이소부탄올 생성의 스크리닝 및 부위 포화 유전자 라이브러리의 제작

[0923] 위치 309에서의 모든 20개의 개별 아미노산 변화를 코딩하는 프라이머를 함유하는 역방향 프라이머 혼합물 (본 실시예에서 K9_309r로 칭해짐)(표 37) 및 정방향 프라이머 K9_131T_080211f: ggacttgatgtcactatgatc (본 실시

예에서 K9_131Tf로 칭해짐)를 이용하여, 309의 K9 KARI의 위치를 표적화하는 단일-부위 포화 라이브러리를 생성하였다. 변이체 K9SB2를 함유하는 플라스미드 (pHR81-ILV5p.K9SB2.TEF1(M4)p.ilvD (서열 번호 930, "pLH744"로도 지칭됨).

[0924] 간략하게는, 보통의 PCR을 통해 메가프라이머를 제조하였다. 메가프라이머 PCR 혼합물은 45 μ l의 슈퍼믹스 (SuperMix; 미국 캘리포니아주 칼스배드 소재의 인비트로젠, #10572063), 2.0 μ l의 K9_131Tf (20 μ M), 2.0 μ l의 K9_309r (20 μ M) 및 1.0 μ l의 주형 (50 ng/ μ l)으로 이루어졌다. 메가프라이머를 제조하기 위한 PCR 프로그램은 출발 온도가 95 $^{\circ}$ C, 1.0분이었으며, 이어서, 35회의 가열/냉각 사이클이었다. 각각의 사이클은 95 $^{\circ}$ C, 20초; 55 $^{\circ}$ C, 20초; 및 72 $^{\circ}$ C, 1.0분으로 이루어졌다. 이어서, 실시예 5에 나타난 것과 동일한 절차를 사용하여 K9SB2 내로 돌연변이를 도입하는 데 메가프라이머를 사용하였다. 단일-부위 포화 라이브러리로부터의 클론을 서열결정하였다.

[0925] pLH744와 함께 생성된 독특한 변이체를 효모에서의 아이소부탄올 생성 및 부산물 형성에 대해 분석하였다 (각각의 돌연변이체에 대해 3회). KARI 벡터로 형질전환시켜, PNY2259 (MATa ura3 Δ ::loxP his3 Δ pdc6 Δ pdc1 Δ ::P[PDC1]-DHAD|ilvD_Sm-PDC1t-P[FBA1]-ALS|alsS_Bs-CYC1t pdc5 Δ ::P[PDC5]-ADH|sadB_Ax-PDC5t gpd2 Δ ::loxP fra2 Δ ::P[PDC1]-ADH|adh_H1-ADH1t adh1 Δ ::UAS(PGK1)P[FBA1]-kivD_Lg(y)-ADH1t yprc Δ 15 Δ ::P[PDC5]-ADH|adh_H1-ADH1t ymr226c Δ ald6 Δ ::loxP; 실시예 22) 숙주에서 효모 경로 주를 제조하였다. 형질전환된 세포를 히스티딘 또는 우라실을 포함하지 않는 합성 배지 (탄소 공급원으로서 1% 에탄올) 상에 플레이팅하였다. 3가지 형질전환체를 동일한 배지의 신선한 플레이트로 옮겼다. 48웰 플레이트 (미국 캘리포니아주 유니온 시티 소재의 악시젠(Axygen), 카탈로그 번호 391-05-061)에서 혐기성 조건 하에 아이소부탄올 생성에 대해 형질전환체를 시험하였다.

[0926] SE-Ura 플레이트 상의, 형질전환으로부터의 효모 콜로니는 5-7일 후에 나타났다. 각각의 변이체로부터 3개 이상의 콜로니를 신선한 SE-Ura 플레이트 상으로 패치하고, 30 $^{\circ}$ C에서 3일 동안 인큐베이션하였다.

[0927] 성장 배지 및 절차

[0928] 효모 주의 성장 절차 동안 2가지 유형의 배지, 즉 호기성 예비 배양물 배지 및 혐기성 배양 배지를 사용하였다. 달리 언급되지 않는다면 모든 화학물질은 시그마 (미국 미주리주 세인트 루이스 소재)로부터 입수하였다.

[0929] 호기성 예비 배양 배지 (SE-Ura): 아미노산을 포함하지 않는 6.7 g/L의 효모 질소 베이스 (디프코, 291940, 미국 메릴랜드주 스팅크스 소재), 히스티딘, 류신, 트립토판 및 우라실을 포함하지 않는 1.4 g/L의 효모 합성 드롭아웃 배지 보충제, 0.2% 에탄올, 0.2% 글루코스, 0.01% w/v 류신, 0.002% w/v 히스티딘, 및 0.002% w/v 트립토판.

[0930] 혐기성 배양 배지 (SEG-Ura-His): 50/50 v/v 트윈 /에탄올 용액 중에 제조된, 50 mM MES (pH 5.5, 아미노산을 포함하지 않는 6.7 g/L 효모 질소 베이스 (디프코, 291940, 미국 메릴랜드주 스팅크스 소재), 히스티딘, 류신, 트립토판 및 우라실을 포함하지 않는 1.4 g/L 효모 합성 드롭아웃 배지 보충제, 0.1% 에탄올, 3% 글루코스, 0.01% 류신, 0.002% w/v 히스티딘, 0.002% 트립토판, 30 mg/L 니코틴산, 30 mg/L 티아민 및 10 mg/L 에르고스테롤.

[0931] 패치된 세포를 48웰 플레이트 내에 접종하였다. 각각의 웰은 1.5 ml의 호기성 배지를 함유한다. 플레이트를 투과성 포일로 덮고, 하룻밤 진탕하면서 30 $^{\circ}$ C에서 성장시켰다. 이어서, 세포 밀도 (OD₆₀₀)를 측정하였다. 각각의 웰에 대해 최종 OD₆₀₀이 0.2인 1.5 ml의 세포 현탁물 (혐기성 배지 중)을 제조하기 위한 세포의 양을 계산하고, 혐기성 배지를 사용하여 1.5 ml의 세포 현탁물을 제조하고, 각각의 웰에 첨가하였다. 제조업자의 프로토콜 (미국 미시간주 그레이스 레이크 소재의 코이 래보라토리 프로덕츠 인크.(Coy Laboratory Products Inc.))에 따라 혐기성 챔버를 사용하여 48웰 플레이트 내의 산소를 제거하였다. 이어서 2일 동안 진탕하면서 30 $^{\circ}$ C에서 세포를 성장시켰다. 이어서, 세포 밀도 (OD₆₀₀)를 측정하였다. 아이소부탄올 생성 측정을 위해, 최상으로 성장된 돌연변이체를 48웰 플레이트에서 동일하게 2일 동안 성장시켰다. 2일의 혐기성 성장 후에, 세포 밀도 (OD₆₀₀)를 측정하였다. 세포를 4,000 g에서 5분 동안 원심분리하고, LC/MS를 사용한 아이소부탄올 측정을 위해 상청액을 수집하였다.

[0932] K9 KARI 돌연변이체를 갖는 효모 주의 LC/MS 분석

[0933] 혐기성 성장 기간이 끝났을 때, LCMS 분석을 위한 샘플을 취하였다. 워터스 어퀴티 UPLC HSS T3 분리 컬럼 (미국 매사추세츠주 밀포드 소재의 워터스)을 갖는, 워터스 어퀴티 UPLC 분리 유닛 및 어퀴티 TQD 트리플 쿼드 질

량 분석기 (미국 매사추세츠주 밀포드 소재의 워터스)를 사용하여 LCMS 분석을 수행하였다. 0.5 mL/min의 유량에서, 99% 수성으로 출발하여 99% 유기로 끝나는, 물 (+ 0.1% 포름산) 및 아세트니트릴 (+ 0.1% 포름산)의 역상구배를 사용하여 화합물들을 분리하였다. 워터스 매스링스(Waters Masslynx) 4.1 소프트웨어 (미국 매사추세츠주 밀포드 소재의 워터스)를 사용하여 크로마토그램을 분석하였다. 표준물의 3회 반복 분석의 보정 곡선을 사용하는 워터스 콰린스(Waters Quanlynx) 소프트웨어 (미국 매사추세츠주 밀포드 소재의 워터스)를 사용하여 아이소부탄올에 대한 마이크로 몰 수율을 계산하였다.

[표 37]

역방향 프라이머

K9-KARI의 표적화된 위치(들)	프라이머
309	K9_309I_111711r: CTTTCTCATAGCCTTAATGTGGAC (서열 번호 884) K9_309L_111711r: CTTTCTCATAGCCTTAAGTGGAC (서열 번호 885) K9_309Y_111711r: CTTTCTCATAGCCTTATAGTGGAC (서열 번호 886) K9_309C_111711r: CTTTCTCATAGCCTTACAGTGGAC (서열 번호 887) K9_309W_111711r: CTTTCTCATAGCCTTCCAGTGGAC (서열 번호 888) K9_309P_111711r: CTTTCTCATAGCCTTGGGTGGAC (서열 번호 889) K9_309H_111711r: CTTTCTCATAGCCTTATGGTGGAC (서열 번호 890) K9_309Q_111711r: CTTTCTCATAGCCTTTGGTGGAC (서열 번호 891) K9_309M_111711r: CTTTCTCATAGCCTTCATGTGGAC (서열 번호 892) K9_309N_111711r: CTTTCTCATAGCCTTATGTGGAC (서열 번호 893) K9_309V_111711r: CTTTCTCATAGCCTTAACGTGGAC (서열 번호 894) K9_309A_111711r: CTTTCTCATAGCCTTAGCGTGGAC (서열 번호 895) K9_309D_111711r: CTTTCTCATAGCCTTATCGTGGAC (서열 번호 896) K9_309E_111711r: CTTTCTCATAGCCTTTTCGTGGAC (서열 번호 897) K9_309G_111711r: CTTTCTCATAGCCTTACCGTGGAC (서열 번호 898) K9_309S_111711r: CTTTCTCATAGCCTTAGAGTGGAC (서열 번호 899) K9_309T_111711r: CTTTCTCATAGCCTTAGTGTGGAC (서열 번호 900) K9_309R_111711r: CTTTCTCATAGCCTTCTGTGGAC (서열 번호 901) K9_309K_111711r: CTTTCTCATAGCCTTCTGTGGAC (서열 번호 902)

[표 38]

2 일간의 혐기성 성장 후 OD₆₀₀ 및 아이소부탄올 역가 (mM)를 갖는 일부 변이체의 목록

변이체	아미노산 서열 번호	반복	OD ₆₀₀	아이소부탄올 역가 (mM)
K9C4	927	1	0.5502	76.99
		2	0.6578	84.05
		3	0.7301	98.91
K9C8	928	1	0.6887	116.57
		2	0.5309	78.77
		3	0.6859	102.49
K9SB2	427	1	0.6314	88.39
		2	0.5977	81.18
		3	0.2325	44.60

실시예 41

K9SB2의 절단된 버전인, K9SB2_SH (K9SB2_short)의 제작

처음 5개의 N-말단 아미노산 (MEECK)이 결여된 K9SB2의 버전을 코딩하는 유전자를, 퓨전(등록상표) 하이-피델리티 PCR 키트 (카탈로그 번호 E0553L, 뉴 잉글랜드 바이오랩스)를 사용하여 PCR에 의해 제조하였다. 프라이머 Kshort1 (AAACCGGTTTAAACAGTATGGCTAAGATTACTACCAAGAAGACTG; 서열 번호 903) 및 YGrev (TTCTGTTTTATCAGACCGCTTC; 서열 번호 904)는 인테그레이티드 디엔에이 테크놀로지스, 인크.(미국 아이오와주 코랄빌 소재)에 의해 합성되었다. 프라이머, dNTP 믹스 (카탈로그 번호 18427-013, 미국 캘리포니아주 칼스배드 소재의 인비트로젠), 주형, 및 ddH₂O 이외의, 여기서 사용한 시약들은 상기에 나타난 키트와 함께 공급되었다. PCR 혼합물은 1 µl의, pBAD.KARI 플라스미드 (20 ng/µl; 실시예 17에 기재된 플라스미드 제제; 서열 번호 428) 중 K9SB2, 2 µl의 각각의 프라이머 (20 µM 스톡), 10 µl의 5× 퓨전 HF 완충제, 2 µl의 5 mM dNTP 믹

스, 0.5 μ l의 폴리머라아제, 및 28 μ l의 ddH₂O로 이루어졌다. 하기 조건을 PCR 반응에 사용하였다: 출발 온도는 98℃, 2분이었으며, 이어서, 30회의 가열/냉각 사이클이었다. 각각의 사이클은 98℃, 10초; 48℃, 30초; 및 72℃, 1.5분으로 이루어졌다. 상기 온도 사이클링의 완료시에, 샘플을 72℃에서 10분 더 유지하고, 그 후 4℃에서 샘플 회수를 기다리면서 유지하였다. 반응 생성물을 아가로스 겔 전기영동 (1% 아가로스, 1X TBE 완충제)을 통하여 주형으로부터 분리하고, 일러스트라(Illustra) GFX™ PCR DNA 및 겔 밴드 정제 키트(Gel Band Purification Kit) (카탈로그 번호 28-9034-70, 미국 뉴저지주 피스카타웨이 소재의 지이 헬스케어 사이언시즈(GE Healthcare Sciences))를 사용하여 회수하였으며, 이는 제조업자가 권고한 바와 같았다. 정제된 PCR 생성물을 PmeI 및 SfiI로 절단하고, pBAD-ps-JEA1 (서열 번호 905)의 상응하는 부위 내로 라이게이션하고, 생성된 제작물, pBAD.KARI 내의 K9SB2_SH (서열 번호 929)의 서열을 확인하였다.

실시예 42

K9SB2 및 기타 KARI 효소의 연구를 위한 효모 발현 플라스미드의 생성

pHR81-ILV5p-K9SB2-TEF(M4)-IlvD (pLH744; 서열 번호 930)의 제작:

K9SB2 유전자를 PmeI 및 SfiI 절단에 의해 pHR81-ILV5p-K9SB2로부터 잘라내고, PmeI 및 SfiI 부위에서 pHR81-ILV5p-K9D3-TEF(M4)-IlvD (pBP2090)에 라이게이션시켰다. 라이게이션된 벡터를 pHR81-ILV5p-K9SB2-TEF(M4)-IlvD (pLH744)로 칭하였고, 이러한 벡터를 서열결정에 의해 확인하였다.

K9SB2_SH_DHAD (서열 번호 931)의 제작:

(실시예 41에 기재된) 절단된 K9SB2_SH PCR 생성물을 pLH744의 PmeI 및 SfiI 부위 내로 라이게이션하고, 서열결정에 의해 확인하였다.

K9SB2_SH_81 (서열 번호 932)의 제작

(실시예 41에 기재된) 절단된 K9SB2_SH PCR 생성물을 pHR81-PIlv5-KARI-K9.G9의 PmeI 및 SfiI 부위 내로 라이게이션하고, 서열결정에 의해 확인하였다.

K9SB2 유도체의 효모 발현을 위한 플라스미드의 제작

플라스미드 pLH744 및 플라스미드 pHR81-PIlv5-KARI-K9.G9의 PmeI 및 SfiI 부위 내로, pBAD.KARI 플라스미드로부터의 상응하는 KARI 유전자를 서브클로닝하여, (실시예 17 및 실시예 21에 기재된) 표 39에 열거된 변이체에 대한 효모 발현 플라스미드를 생성하였다. 제작물을 서열결정에 의해 확인하였다.

[표 39]

효모 발현 플라스미드 내로 서브클로닝된 K9SB2 유도체

K9SB2 유도체	다른 명칭	아미노산 서열 번호	핵산 서열 번호
K9SB2-K90M	K9_David	431	432
K9SB2-G55D	K9_Eliza	433	946
K9SB2-Q91L	K9_Frank	440	947
K9SB2-A303D	K9_Grace	445	948
K9SB2-M94-V307I	K9_Ingrid	455	949
K9SB2-F67I	K9_Jarvis	437	950
K9SB2-A56G-K90N	K9_Kelly	452	951
K9SB2-G55C	K9_Norman	481	952
K9SB2-P135S	K9_Ophelia	488	953
K9SB2-F53L	K9_Pat	441	954
K9SB2-Q94I	K9_Quentin	495	955
K9SB2-F67L	K9_Ralph	496	956
K9SB2-K8N-K90M-T141I	K9_Sophia	502	957
K9SB2-F13V-M94I-T141I	K9_Tiberius	509	958
K9SB2-A56V	K9_Ursula	511	959
K9SB2-I84L	K9_Victor	514	960
K9SB2-W59C	K9_Watson	520	961
K9SB2-T93A	K9_Xavier	641	642

K9SB2 유도체의 절단된 버전의 효모 발현을 위한 플라스미드의 제작

[0954] 변형을 이용하여, 실시예 41에 기재된 바와 같이 N-말단 절단된 버전의 K9SB2 유도체를 제조하였다. 프라이머 Kshort1을, 인테그레이티드 디엔에이 테크놀로지스, 인크. (미국 아이오와주 코랄빌 소재)에 의해 합성된 5'-포스포릴화 프라이머 KPSH1 (AAACAGTATG GCT AAG ATT TAC TAC CAA GAA GAC TG; 서열 번호 906)로 대체하였다. PCR에서의 K9SB2를 표 39에 열거된 변이체 (pBAD.KARI 플라스미드)의 풀링된 혼합물로 대체하였다. 정제된 PCR 생성물을 SfiI로 절단하고 pHR81-PIIv5-KARI-K9.G9의 PmeI 및 SfiI 부위 내로 라이게이션하였다. 템플리파이™ (지이 헬스케어) 및 프라이머 pHR81-F (ACACCCAGTATTTCCCTTTC) 및 pHR81- Rev (CTA GTG TAC AGA TGT ATG TCG G) (서열 번호: 924 및 925)을 사용한 DNA 서열결정을 수행하여 각각의 절단된 유도체를 확인하였다. 확인된 변이체가 표 40에 나타나있다. 절단된 버전에 대한 코딩 서열을 후속적으로 플라스미드 pLH744의 PmeI 및 SfiI 부위 내로 서브클로닝하고 서열결정에 의해 확인하였다.

[0955] [표 40]

K9SB2 유도체의 절단된 버전

변이체	아미노산 서열 번호	핵산 서열 번호
K9_David_SH	196	263
K9_Eliza_SH	266	907
K9_Frank_SH	267	908
K9_Grace_SH	389	909
K9_Ingrid_SH	405	910
K9_Jarvis_SH	781	911
K9_Kelly_SH	782	912
K9_Norman_SH	783	913
K9_Ophelia_SH	835	914
K9_Pat_SH	853	915
K9_Quentin_SH	854	916
K9_Ralph_SH	855	917
K9_Ursula_SH	856	918
K9_Watson_SH	857	858
K9_Xavier_SH	859	919

[0956]

[0957] K9 Zeke_SH (K9SB2-K90M-T93A) 및 K9 Annabel_SH (K9SB2-K90M-T93I)의 제작

[0958] 퀵체인지(등록상표) 라이트닝(Lightning) 부위 지정 돌연변이 유발 키트(카탈로그 #210518; 미국 캘리포니아주 라 줄라 소재의 애질런트 테크놀로지스, 스트라타진 제품부)를 사용하여 부위 지정 돌연변이를 통해 K9_David_SH로부터 K9_Zeke_SH (서열 번호 860, 단백질 서열 번호 861) 및 K9_Annabel_SH (서열 번호 862, 단백질 서열 번호 863)를 유도하였다. 프라이머, 주형 및 ddH₂O를 제외하고는, 여기서 사용한 모든 시약은 상기에 나타난 키트와 함께 공급되었다. 프라이머는 인테그레이티드 디엔에이 테크놀로지스, 인크. (미국 아이오와주 코랄빌 소재)에 의해 합성되었다.

[0959] K9_Zeke_SH의 경우, 사용된 프라이머는 oMT93A (GATCCAGATGAAATGCAGGCTGCCATGTACAAAACGACATCG; 서열 번호 920) 및 oMT93Arev (CGATGTCGTTTTGTACATGGCAGCTGCATTTCATCTGGGATC; 서열 번호 921)였다. 반응 혼합물은 1 μl의, pHR81-PIIv5-KARI-K9.G9-유도된 벡터 (20 ng/μl) 중 K9_David_SH, 1 μl의 각각의 프라이머 (150 ng/ul), 5 μl의 10× 반응 완충제, 1 μl의 dNTP 믹스, 1.5 μl 퀵솔루션(QuikSolution), 1 ul의 퀵체인지 라이트닝 효소, 및 38.5 μl의 ddH₂O를 함유하였다. K9_Annabel_SH의 경우에는, 프라이머를 oMT93I (GATCCAGATGAAATGCAGGCTATCATGTACAAAACGACATCG; 서열 번호 922) 및 oMT93Irev (CGATGTCGTTTTGTACATGATAGCTGCATTTCATCTGGGATC; 서열 번호 923)으로 대체하였다.

[0960] 하기 조건을 둘 모두의 반응에 이용하였다: 출발 온도는 95℃, 2분이었으며, 이어서 18회의 가열/냉각 사이클이었다. 각각의 사이클은 95℃, 20초; 60℃, 10초; 및 68℃, 6분으로 이루어졌다. 상기 온도 사이클링의 완료시에, 샘플들을 4℃에서 샘플 회수를 기다리면서 유지하였다. 2 μl의 Dpn I을 각각의 반응물에 첨가하고, 혼합물을 37℃에서 1시간 동안 인큐베이션하였다. 제조업자의 지시에 따라 2 μl의 각각의 돌연변이 유발 반응물로 된 샷(등록상표) Stb13™ 화학적 적격성 이. 콜라이 (인비트로젠, 카탈로그 번호 C7373-03)를 형질전환시켰다. 형질전환체를 LB 배지 및 100 μg/ml의 암피실린 (카탈로그 번호 L1004, 미국 캘리포니아주 홀리스터 소재의 테크노바 인크.)을 함유하는 환천 배지 상에 도말하고, 37℃에서 하룻밤 인큐베이션하였다. 그 후, 다수의 형질전환체를 템플리파이™ DNA 서열결정™ (지이 헬스케어)을 위해 선발하였는데, 상기 서열결정은 프라이머 pHR81-F (ACACCCAGTATTTCCCTTTC; 서열 번호 924) 및 pHR81- Rev (CTA GTG TAC AGA TGT ATG TCG G; 서열 번호 925)를 이용하였다. 확인된 서열을 갖는 변이체 (pHR81-PIIv5-KARI-K9.G9 유도된 벡터 내의 K9_Zeke_SH 및 K9_Annabel_SH)를 pLH744 (서열 번호 930)의 PmeI 및 SfiI 부위 내로 서브클로닝하였다.

- [0961] 실시예 43
- [0962] 48웰 플레이트에서 혐기성 조건 하에 성장된 K9SB2 및 유도체의 아이소부탄을 생성
- [0963] 실시예 42에 기재된 바와 같이 pH744로부터 유도된 벡터를 사용하여 제조된 K9SB2, K9SB2_SH, 및 K9SB2-T93A를 위한 효모 발현 플라스미드를 사용하여, 48웰 플레이트에서 혐기성 조건 하에 성장된 효모에서의 아이소부탄을 생성을 평가하였다. KARI 변이체에 대한 코딩 서열을 함유하는 플라스미드로 형질전환시키고 우라실을 포함하지 않는 합성 배지 (탄소 공급원으로서 1% 에탄올) 상에 플레이팅하여, 숙주 PNY2259 (MATa ura3Δ::loxP his3Δ pdc6Δ pdc1Δ::P[PDC1]-DHAD|ilvD_Sm-PDC1t-P[FBA1]-ALS|alsS_Bs-CYC1t pdc5Δ::P[PDC5]-ADH|sadB_Ax-PDC5t gpd2Δ::loxP fra2Δ::P[PDC1]-ADH|adh_H1-ADH1t adh1Δ::UAS(PGK1)P[FBA1]-kivD_Lg(y)-ADH1t yprcΔ15Δ::P[PDC5]-ADH|adh_H1-ADH1t ymr226cΔ ald6Δ::loxP)에서 아이소부탄을 생성 주를 제조하였다. SE-Ura 플레이트 상의, 형질전환으로부터의 효모 콜로니는 30℃에서의 3-5일의 인큐베이션 후에 나타났다. 각각의 변이체로부터 3개 이상의 콜로니를 신선한 SE-Ura 플레이트 상에 패치하고 30℃에서 인큐베이션하였다.
- [0964] 효모 배양 조건:
- [0965] 호기성 배양 배지: 2 g/l 에탄올을 갖는 SE-Ura 배지.
- [0966] 혐기성 배양 배지: 10 mg/l 에르고스테롤, 50 mM MES 완충제 (pH 5.5), 30 mg/l 티아민, 및 30 mg/l 니코틴산 이 보충된, 30 g/l 글루코스 및 1 g/l 에탄올을 갖는 SEG-Ura.
- [0967] 48웰 플레이트: 악시젠 카탈로그 번호 P-5ML-48-C-S, 5 ml/웰 총 부피, 배양물 부피: 1.5 ml/웰.
- [0968] 플레이트를, 호기성 배양을 위한 투과성 접착 필름 (VWR ; 카탈로그 번호 60941-086)으로 덮었다. 플레이트를 30℃에서 225 rpm으로 진탕하였다. 혐기성 배양을 위해, 혐기성 챔버 내에서의 2시간 동안 평형화에 의해, 통기성 필름으로 덮은 새로 접종된 플레이트에서 산소를 제거하였다. 이어서, 플레이트 커버를 접착성 알루미늄 커버 (VWR ; 카탈로그 번호 89049-034)로 교환하고 각각의 플레이트를, 기밀 플라스틱 박스 (미츠비시 가스 케미칼 아메리카, 인크(Mitsubishi Gas Chemical America, Inc); 미국 뉴욕주 뉴욕 소재; 카탈로그 50-25)에, 신선한 산소 스캐빈저 팩 (미츠비시 가스 케미칼 아메리카, 인크; 미국 뉴욕주 뉴욕 소재; 카탈로그 10-01)과 함께 넣었다. 전체 조립체 (밀봉된 기밀 플라스틱 박스 내의 플레이트(들) 및 산소 스캐빈저 팩)를 혐기성 챔버로부터 꺼내고 30℃에서 225 rpm로 진탕하였다.
- [0969] 실험 프로토콜
- [0970] SE-Ura 한천 플레이트 상의 단일 효모 콜로니를 신선한 SE-Ura 한천 플레이트 상에 도말하고, 세포의 농후한 패치가 성장할 때까지 30℃에서 인큐베이션하였다. 초기 호기성 배양을 위해 이러한 세포의 루프를 48웰 플레이트 내의 액체 예비 배양물에 접종하였다. 하룻밤 진탕한 후에, 각 웰의 0.15 ml를 평평한 바닥의 96웰 플레이트 내로 옮기고, 분광광도계 디바이스 (미국 캘리포니아주 서니베일 소재) 플레이트 판독기를 사용하여 600 nm에서 각 웰의 흡광도를 측정함으로써, 각각의 배양 웰의 OD₆₀₀을 측정하였다. 실험적으로 결정된 보정선에 기초한 선형 형질전환을 이러한 플레이트 판독기-측정된 광학 밀도에 적용하여 이를 큐벳-기반 분광 광도계에 대한 비견되는 흡광도 값으로 전환하였다.
- [0971] 각각의 호기성 예비 배양물 웰의 계산된 부분을, 1.5 ml의 SEG -Ura 배지를 갖는 신선한 48웰 플레이트의 상응하는 웰에 접종하여, 0.2의 초기 OD₆₀₀ (큐벳 분광 광도계 흡광도 유닛)을 달성하였다. 신선한 플레이트를 접종하는 과정에서, 호기성 예비 배양물 플레이트를 원심분리하여, 각각의 웰로부터 상청액을 제거하고, 각각의 웰 내의 세포를 신선한 SEG -Ura 배지에 재현탁시켰다. 이러한 혐기성 배양 플레이트를 3일 동안 진탕하였다. 배양물 상청액 내의 아이소부탄을 농도를 HPLC로 측정하였다 (표 41).

[0972] [표 41]

제 1 혐기성 계대 사이클에서 도달한 아이소부탄올 역가

변이체	웰의 번호	평균 아이소부탄올 역가 (mM)	아이소부탄올 역가 (mM)의 표준 편차
K9SB2	16	35.2	11.9
K9SB2_SH	8	67.4	12.2
K9SB2-T93A	8	40.8	9.1

[0973]

[0974] 하기와 같이 제1 혐기성 배양으로부터 후속 혐기성 배양을 시작하였다. 각각의 혐기성 배양물 웰의 계산된 부분을, 1.5 ml의 SEG -Ura 배지를 갖는 신선한 48웰 플레이트의 상응하는 웰에 접종하여, 0.2의 초기 OD₆₀₀ (큐벳 분광 광도계 흡광도 단위)을 달성하였다. 신선한 플레이트를 접종하는 과정에서, 하나의 배양으로부터 다음 배양으로의 대사산물의 이월을 최소화하기 위하여, 성장 플레이트를 원심분리하여, 각각의 웰로부터 상청액을 제거하고, 각각의 웰 내의 세포를 신선한 SEG -Ura 배지에 재현탁시켰다. 후속 (제2 사이클) 혐기성 배양 플레이트를 2일 동안 진탕하였다. 배양물 상청액 내의 아이소부탄올 농도를 HPLC로 측정하였다 (표 42).

[0975] [표 42]

제 2 혐기성 계대 사이클에서 도달한 아이소부탄올 역가

변이체	웰의 번호	평균 아이소부탄올 역가 (mM)	아이소부탄올 역가 (mM)의 표준 편차
K9SB2	16	67.6	10.8
K9SB2_SH	8	85.7	9.2
K9SB2-T93A	8	76.3	16.8

[0976]

[0977] 실시예 44

[0978] K9G9, K9SB2, 및 K9SB2_SH의 동력학적 특성

[0979] K9G9, K9SB2, K9SB2_SH를 이. 콜라이 주 Bw25113 ($\Delta ilvC$) 내의 pBAD.KARI 플라스미드를 통해 과다발현시키고 동력학적 파라미터의 상세한 측정을 위해 정제하였다.

[0980] 하기 변형을 가지고, 실시예 18에 기재된 바와 같이 발현, 정제 및 보조 인자 동력학적 분석을 수행하였다. 발현 배양물을 125 mL 통기식 캡 플라스크(vented cap flask) 내의, 100 µg/mL 암피실린 및 0.2% (w/v) 아라비노스를 포함하는 20 mL의 LB 중에서 성장시켰다. 발현 배지에 1/10 부피의 8시간 배양물을 사용하여 접종하였다. 발현 배지를 K9 SB2에 대해서는 18시간 동안, 그리고 K9SB2_SH에 대해서는 24시간 동안 성장시켰다.

[0981] [표 43]

NADH 및 NADPH 에 의해 반응들을 비교하는 동력학적 파라미터

변이체	서열 번호	NADPH 의 V _{max} (U/mg)	NADPH 의 K _m (µM)	NADPH 의 V _{max} /K _m (L/min*mg)	NAD H 의 V _{max} (U/mg)	NAD H 의 K _m (µM)	NADH 의 V _{max} /K _m (L/min*mg)
K9G9	644	2.2	24.1	0.091	1.9	78.2	0.024
K9SB2	427	1.7	44.8	0.038	1.8	11.6	0.155
K9SB2 _SH	637	1.7	109.9	0.015	1.7	13.3	0.128

[0982]

[0983] 실시예 45

[0984] K9SB2 및 유도체의 아이소부탄올 생성

[0985] 하기와 같이 변형시켜, 하기 실시예 42에 기재된 바와 같이, 48웰 플레이트에서 아이소부탄올 생성 분석을 수행하였다. 호기성 및 혐기성 배양 배지는 실시예 19에 기재된 것과 동일하였으나, 0.01% w/v 히스티딘을 첨가하였다. 캐리 300 분광 광도계 (미국 텔라웨어주 윌밍턴 소재의 애질런트 테크놀로지스)를 사용하여 호기성 예비

배양물의 OD₆₀₀ 값을 측정하였다. M-20 로터(rotor)가 구비된 허레우스 멀티퓨그(Heraeus Multifug) X1R (미국 매사추세츠주 윌섬 소재의 서모 사이언티픽)을 사용하여 호기성 예비 배양물 세포를 펠렛화하고 소비한 배양 배지를 폐기하였다. 셀 펠렛을 갖는 플레이트를 코이 혐기성 백(Coy Anaerobic Bag; 미국 미시간주 그래스 레이크 소재)으로 옮기고 100 μ L의 혐기성 배양 배지를 각각의 펠렛에 첨가하였다. 혐기성 배양 배지를 24시간 이상 동안 탈기시킨 후에 48웰 플레이트에 1.5 mL 분취물을 넣었으며; 이러한 과정은 혐기성 백 안에서 수행하였다. 적절한 부피의 세포 재현탁물을 갖는 혐기성 배양 플레이트를 인큐베이션하고 접착성 알루미늄 포일로 덮었다. 플레이트를 MCG 2.5L 언에어로팩(AnaeroPack) 시스템 (일본 소재의 MCG)에 넣었다. 박스를 밀봉하고 혐기성 백으로부터 꺼내어, 220 rpm에서 진탕하면서 30°C 인큐베이터에 80시간 동안 넣어 두었다. 변이체 당 3개의 형질전환체를 분석하였고, K9D3 및 K92B2에 대해 6가지 형질전환을 분석하였다 (표 44).

[표 44]

아이소부탄올 역가

변이체	서열 번호	아이소부탄올 역가 (mM)
K9D3	645	76.4 $\pm$ 9.7
K9SB2	427	93.8 $\pm$ 3.3
K9_Frank	440	27.8 $\pm$ 0.5
K9_Grace	445	89.4 $\pm$ 15.5
K9_Ingrid	455	88.3 $\pm$ 5.4
K9_Jarvis	437	91.8 $\pm$ 5.8
K9_Kelly	452	40.9 $\pm$ 4.5
K9_Norman	481	66.2 $\pm$ 10.6
K9_Ophelia	488	28.3 $\pm$ 5.3
K9_Pat	441	77.1 $\pm$ 25.6
K9_Quentin	495	80.1 $\pm$ 10.8
K9_Ralph	496	82.1 $\pm$ 21.3
K9_Sophia	502	25.3 $\pm$ 13.4
K9_Tiberius	509	11.4 $\pm$ 10.4
K9_Ursula	511	57.3 $\pm$ 26.2
K9_Victor	514	93.0 $\pm$ 11.4

하기 변형을 이용하여, 실시예 19에 기재된 바와 같이 변이체를 선발하기 위하여 세럼 바이알에서 아이소부탄올 생성 분석을 수행하였다. 실시예 42에 기재된 바와 같이 제조된 pLH744로부터 유도된 플라스미드로부터 효모에서 KARI 변이체를 발현시켰다. 호기성 예비 배양물 및 혐기성 성장 배지 둘 모두에 0.01% w/v의 최종 농도로 히스티딘을 첨가하였다. 변이체당 3 또는 4개의 형질전환체를 분석하였다 (표 45, 표 46, 및 표 47).

[표 45]

아이소부탄올 역가 실험 1

변이체	서열 번호	세대	시간	아이소부탄올 역가 (mM)	아이소부탄올/글리세롤	
K9SB2	427	1	75	28.2 $\pm$ 6.85	2.11 $\pm$ 0.16	
K9SB2_SH	637	1	75	87.9 $\pm$ 1.04	2.49 $\pm$ 0.04	
K9SB2	427	2	48	85.0 $\pm$ 4.35	2.67 $\pm$ 0.06	
K9SB2_SH	637	2	48	97.4 $\pm$ 4.79	2.49 $\pm$ 0.07	

[0091] [표 46]

아이소부탄올 역가 실험 2

변이체	서열 번호	계대	시간	아이소부탄올 역가 (mM)	아이소부탄올/ 글리세롤
K9SB2	427	1	44	30.2 ± 4.1	2.30 ± 0.15
K9SB2_SH	637	1	44	40.3 ± 5.6	2.41 ± 0.07
K9_David	431	1	44	36.2 ± 3.5	2.40 ± 0.10
K9_David_SH	196	1	44	41.9 ± 6.4	2.37 ± 0.04
K9_Grace	445	1	44	40.2 ± 5.5	2.30 ± 0.05
K9_Pat	441	1	44	29.0 ± 3.1	2.35 ± 0.10
K9SB2	427	2	46	59.2 ± 0.6	2.61 ± 0.03
K9SB2_SH	637	2	46	74.8 ± 3.0	2.58 ± 0.03
K9_David	431	2	46	63.8 ± 0.6	2.65 ± 0.06
K9_David_SH	196	2	46	76.7 ± 1.5	2.53 ± 0.04
K9_Grace	445	2	46	62.3 ± 3.1	2.52 ± 0.07
K9_Pat	441	2	46	50.1 ± 2.1	2.48 ± 0.07

[0092]

[0093] [표 47]

아이소부탄올 역가

변이체	서열 번호	계대	시간	아이소부탄올 역가 (mM)
K9SB2	427	2	50	66.2 ± 1.8
Annabel SH	862	2	50	74.4 ± 2.4
Zeke SH	860	2	50	76.8 ± 8.3

[0094]

HMMSR 0 [2.2g]
NAME Functionally Verified KARs
LENG 354

ALPH Amino

MAP yes

ild -n Functionally Verified KARs exp-KARl.hmm exp-KARl_mod.ain
ibrate exp-KARl.hmm
lec 8 17.34.51 2008
-1000 -1000 -8455 -4 -8455 -4
55
558 85 338 -294 453 -1158 197 249 902 -1085 -142 -21 -313 45 531 201 384 -1998 -644
2708 0.110102

File format version: a unique identifier for this save file format.
Name of the profile HMM.
Model length: the number of match states in the model.
Symbol alphabet. This determines the symbol alphabet and the size of the symbol emission probability distributions. Bmino, the alphabet size is set to 20 and the symbol alphabet to "ACDEFGHIKLMNPQRSTVWY" (alphabetic order).
Map annotation flag. If set to yes, each line of data for the match state/consensus column in the main section of the file is followed by an extra number. This number gives the index of the alignment column that the match state was made from. This information provides a "map" of the match states (1..M) onto the columns of the alignment (1..alen). It is used for quickly aligning the model back to the original alignment, e.g. when using hmmbuild -mapall.
Command line for every HMMER command that modifies the save file. This one means that hmmbuild (default parameters) was applied to generate the save file.
Command line for every HMMER command that modifies the save file. This one means that hmmscalibrate (default parameters) was applied to the save profile.
Sequence number: the number of sequences the HMM was trained on.
Creation date. When was the save file generated.
Eight "special" transitions for controlling parts of the algorithm-specific parts of the Plan7 model. The null probability used to convert these back to model probabilities is 1.0. The order of the eight fields is N->B, N->N, E->C, E->J, C->T, C->C, J->B, J->J.
The transition probability distribution for the null model (single G state).
The extreme value distribution parameters μ and λ respectively, both floating point values. These values are set when the model is calibrated with hmmscalibrate. They are used to determine E-values of bit scores.

	C	D	E	F	G	H	I	K	L	M	N	P	Q	R	S	T	V	W	Y	Position in alignment	
m->m	m->n	m->d	n->m	n->n	d->m	d->d	b->m	m->e													
1(O)	-650	-1366	-1463	-44	-1453	-1166	-219	-1455	321	-1417	-911	-227	-1496	3263	122	-643	-684	-1239	-1542	-1030	7100%
-	-145	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-38	-5840	-6882	-894	-1115	-701	-1378	-650													
2(M)	-4231	-3929	-5216	-5402	-3438	-4370	-4528	-3232	-5113	-2613	5320	-5052	-4790	-4977	-4823	-4692	-4459	-3629	-4103	-4017	7200%
-	-147	-501	232	42	-382	397	104	-625	209	-467	-722	276	396	44	95	361	121	-368	-296	-251	
-	-3303	-3318	-325	-3473	-136	-701	-1378														
3(F)	-1308	-1104	-2227	-2120	3516	-2093	-244	-196	-1891	64	66	-1626	-2278	-1503	-1798	-1617	-1350	-389	305	1335	8600%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-38	-5840	-6882	-894	-1115	-701	-1378	-1060													
4(A)	1616	-1744	1125	33	-2015	-1540	-262	-1686	937	-1765	-911	-252	-1658	154	-383	-488	640	-3	-2038	-1421	8700%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-901	-7402	-1125	-894	-1115	-2352	-314														
5(C)	-346	2578	1084	-712	2092	-1540	-384	-167	-624	-482	125	-731	-1705	-451	-883	-631	-338	-50	-774	-133	8800%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-721	275	394	45	96	359	118	-369	-295	-249	
-	-1009	-1006	-7567	-131	-3527	-1916	-444														
6(S)	800	-586	-1937	-1415	-821	-1740	-954	1279	-1204	-584	19	-1258	-1964	-1013	-1358	1715	-476	1117	-1320	-938	9000%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-17	-6953	-7995	-894	-1115	-146	-3378														
7(K)	-956	-2411	-803	501	-2743	-1919	-558	-2483	2435	-2420	-1502	57	-2010	1146	458	829	224	-2040	-2577	-1913	9100%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378														
8(V)	-2472	-2010	-5089	-4702	-2534	-4799	-4391	2241	-4574	-151	-1318	-4442	-4600	-4417	-4628	-4080	-82	3023	-3952	-3510	9200%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378														

[0095]

9(Y)	-4673	-3663	-5210	-5505	2423	-5069	-1332	-3424	-5065	-392	-2838	-3726	-4920	-3635	-4458	-4313	-4533	-3643	-581	4349	9300%
.	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	.	.	.	.	.	.	.	.	.	.	.	.	.	.
10(Y)	-2170	-2625	-2489	-2097	-1555	-2986	-1481	-2628	906	-2674	-2098	-2051	-3206	-1513	-1078	-2758	1039	-2435	-2009	4185	9400%
.	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	.	.	.	.	.	.	.	.	.	.	.	.	.	.
11(D)	-2498	-4412	3500	1042	-4581	-2437	-1765	-4500	733	-4361	-3682	515	-2061	-1429	-2799	-2158	-2558	-3974	-4550	-3541	9500%
.	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	.	.	.	.	.	.	.	.	.	.	.	.	.	.
12(K)	11	-2371	348	819	-2692	-535	-527	-2443	2294	-2387	-1461	590	-1960	-68	904	-67	-837	-1993	-2554	-1871	9600%
.	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	.	.	.	.	.	.	.	.	.	.	.	.	.	.
13(D)	-2663	-4633	3700	580	-4789	-2487	-1872	-4738	731	-4578	-3963	1073	-3046	-1551	2987	-2292	-2742	-4201	-4759	-3709	9700%
.	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	.	.	.	.	.	.	.	.	.	.	.	.	.	.
14(C)	2503	3193	-4266	-3818	-2010	-3276	-2896	762	-3517	-1437	-1051	-3233	-3509	-3212	-3411	-2499	-1792	1507	-2796	-2431	9800%
.	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	.	.	.	.	.	.	.	.	.	.	.	.	.	.
15(D)	-1363	-2905	2748	542	-3202	-2072	-520	-2977	290	-2912	-2023	1270	-2294	-489	-1186	53	1116	-2518	-3086	-2349	9900%
.	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	.	.	.	.	.	.	.	.	.	.	.	.	.	.
16(L)	-1268	-1113	-3338	-540	-1057	-2827	-1716	569	-2409	2299	-236	-2381	-2862	-2089	-2316	-232	-1213	1306	-1645	-1304	10000%
.	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	.	.	.	.	.	.	.	.	.	.	.	.	.	.
17(S)	-1350	-2877	588	1045	-3189	-496	-520	-2963	-628	-2901	-2011	1860	-2289	-489	-1184	2139	190	-2503	-3077	-2343	10100%
.	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-2336	-8139	-325	-894	-1115	-701	-1378	.	.	.	.	.	.	.	.	.	.	.	.	.	.
18(G)	-454	-832	-968	-1110	-2112	3143	-1211	-2091	-1317	-2264	-1691	-878	-1499	-1202	-1421	-646	-774	-1550	-1916	-1919	10200%
.	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-38	-5840	-6882	-894	-1115	-3098	-179	.	.	.	.	.	.	.	.	.	.	.	.	.	.
19(H)	-898	-1313	-545	-482	-320	-1336	4297	-1552	-169	-1483	-1035	-579	-1675	-363	-322	-834	-851	-1354	-725	107	10300%
.	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-38	-5840	-6882	-894	-1115	-3098	-179	.	.	.	.	.	.	.	.	.	.	.	.	.	.
20(D)	-872	-1812	3234	432	-2215	-967	-433	-2172	-569	-2269	-1704	99	-1453	-184	-1141	-728	-873	-1814	-2146	-1646	10400%
.	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-38	-5840	-6882	-894	-1115	-3098	-179	.	.	.	.	.	.	.	.	.	.	.	.	.	.
21(E)	-766	-1695	521	2831	-2050	-1029	-293	-1804	-118	-1918	-1331	69	-1441	-4	-527	-653	-814	-1512	-1988	-1505	10500%
.	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-38	-5840	-6882	-894	-1115	-3098	-179	.	.	.	.	.	.	.	.	.	.	.	.	.	.
22(Y)	-1337	-1229	-1681	-1596	1268	-1957	121	-918	-1294	-769	-585	-1229	-2163	-1111	-1301	-1443	-1359	-932	592	3032	10600%
.	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-38	-5840	-6882	-894	-1115	-109	-3775	.	.	.	.	.	.	.	.	.	.	.	.	.	.
23(O)	-2294	-1931	-4749	-4227	-1724	-4227	-3320	2306	-3952	1993	-634	-3878	94	-3538	-3812	-3411	-2247	1576	-2891	-2629	10700%
.	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	.	.	.	.	.	.	.	.	.	.	.	.	.	.

9(Y)	-4673	-3663	-5210	-5505	2423	-5069	-1332	-3424	-5065	-392	-2838	-3726	-4920	-3835	-4458	-4313	-4533	-3643	-581	4349	9300%
.	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	.	.	.	.	.	.	.	.	.	.	.	.	.	.
10(Y)	-2170	-2625	-2489	-2097	-1555	-2986	-1481	-2628	906	-2674	-2098	-2051	-3206	-1513	-1078	-2258	1039	-2435	-2009	4185	9400%
.	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	.	.	.	.	.	.	.	.	.	.	.	.	.	.
11(D)	-2498	-4412	3500	1042	-4581	-2437	-1765	-4500	733	-4361	-3682	515	-2061	-1429	-2799	-2158	-2558	-3974	-4550	-3541	9500%
.	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	.	.	.	.	.	.	.	.	.	.	.	.	.	.
12(K)	11	-2371	348	819	-2692	-535	-527	-2443	2294	-2387	-1461	590	-1960	-68	904	-67	-837	-1993	-2554	-1871	9600%
.	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	.	.	.	.	.	.	.	.	.	.	.	.	.	.
13(D)	-2663	-4633	3700	580	-4789	-2487	-1872	-4738	731	-4578	-3963	1073	-3046	-1551	2987	-2292	-2742	-4201	-4759	-3709	9700%
.	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	.	.	.	.	.	.	.	.	.	.	.	.	.	.
14(C)	2503	3193	-4266	-3818	-2010	-3276	-2896	762	-3517	-1437	-1051	-3233	-3509	-3212	-3411	-2499	-1792	1507	-2796	-2431	9800%
.	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	.	.	.	.	.	.	.	.	.	.	.	.	.	.
15(D)	-1363	-2905	2748	542	-3202	-2072	-520	-2977	290	-2912	-2023	1270	-2294	-489	-1186	53	1116	-2518	-3086	-2349	9900%
.	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	.	.	.	.	.	.	.	.	.	.	.	.	.	.
16(L)	-1268	-1113	-3338	-540	-1057	-2827	-1716	569	-2409	2299	-236	-2381	-2862	-2089	-2316	-232	-1213	1306	-1645	-1304	10000%
.	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	.	.	.	.	.	.	.	.	.	.	.	.	.	.
17(S)	-1350	-2877	588	1045	-3189	-496	-520	-2963	-628	-2901	-2011	1860	-2289	-489	-1184	2139	190	-2503	-3077	-2343	10100%
.	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-2336	-8139	-325	-894	-1115	-701	-1378	.	.	.	.	.	.	.	.	.	.	.	.	.	.
18(G)	-454	-832	-968	-1110	-2112	3143	-1211	-2091	-1317	-2264	-1691	-878	-1499	-1202	-1421	-646	-774	-1550	-1916	-1919	10200%
.	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-38	-5840	-6882	-694	-1115	-3098	-179	.	.	.	.	.	.	.	.	.	.	.	.	.	.
19(W)	-898	-1313	-545	-483	-320	-1336	4297	-1552	-160	-1493	-1035	-579	-1675	-363	-322	-834	-951	-1354	-725	107	10300%
.	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-38	-5840	-6882	-694	-1115	-3098	-179	.	.	.	.	.	.	.	.	.	.	.	.	.	.
20(D)	-873	-1812	3234	432	-2215	-3967	-433	-2172	-569	-2269	-1704	99	-1453	-184	-1141	-728	-973	-1814	-2146	-1646	10400%
.	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-38	-5840	-6882	-694	-1115	-3098	-179	.	.	.	.	.	.	.	.	.	.	.	.	.	.
21(E)	-166	-1695	521	2831	-2050	-1029	-293	-1804	-118	-1919	-1331	69	-1441	-4	-527	-653	-814	-1512	-1988	-1505	10500%
.	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-38	-5840	-6882	-694	-1115	-3098	-179	.	.	.	.	.	.	.	.	.	.	.	.	.	.
22(Y)	-1337	-1228	-1681	-1596	1268	-1957	121	-918	-1294	-769	-585	-1229	-2163	-1111	-1301	-1443	-1359	-932	522	3932	10600%
.	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-38	-5840	-6882	-694	-1115	-3098	-179	.	.	.	.	.	.	.	.	.	.	.	.	.	.
23(I)	-2264	-1931	-4749	-4227	-1734	-4277	-3300	2308	-3952	1960	-434	-3878	94	-3538	-3812	-3411	-2247	1578	-2891	-2629	10700%
.	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-38	-5840	-6882	-694	-1115	-3098	-179	.	.	.	.	.	.	.	.	.	.	.	.	.	.

38(G)	677	-2128	-3838	-4171	-4647	3536	-3816	-4506	-4340	-4749	-3857	-3009	-3149	-3871	-4137	-1784	-2005	-3297	-4725	-4735	13700%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-	-8	-8139	-9181	-894	-1115	-701	-1378	-	-	-	-	-	-	-	-	-	-	-	-	
39(H)	-2667	-3375	-2682	-2114	-3744	-3201	4738	-3782	-445	-3553	-2886	-2112	866	-1265	1506	-2614	-2557	-3469	-3282	-2908	13800%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-	-8	-8139	-9181	-894	-1115	-701	-1378	-	-	-	-	-	-	-	-	-	-	-	-	
40(A)	3631	-2768	-4492	-4815	-4888	-2992	-4271	-4781	-4818	-5025	-4365	-3727	-3728	-4477	-4545	-2567	-2762	-3852	-4724	-4942	13900%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-	-8	-8139	-9181	-894	-1115	-701	-1378	-	-	-	-	-	-	-	-	-	-	-	-	
41(H)	-3103	-3404	-2950	-2573	-783	-3679	4549	-3407	-1372	-3071	-2715	-2454	-3764	2546	-1428	-2990	-2976	-3308	2269	-295	14000%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-	-8	-8139	-9181	-894	-1115	-701	-1378	-	-	-	-	-	-	-	-	-	-	-	-	
42(A)	3357	-1795	-4134	-4277	-4057	-2118	-3548	-3549	-4035	-4024	-3192	-2817	-2900	-3608	-3823	217	-1660	-276	-4363	-4211	14100%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-	-8	-8139	-9181	-894	-1115	-701	-1378	-	-	-	-	-	-	-	-	-	-	-	-	
43(O)	-1061	1950	-2044	-1475	-1236	2372	-1154	-789	-1218	1062	1123	743	-2448	2895	-1441	-1392	-1005	-693	-1678	-1278	14200%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-	-8	-8139	-9181	-894	-1115	-701	-1378	-	-	-	-	-	-	-	-	-	-	-	-	
44(N)	-4000	-4117	-3369	-3749	-5073	-3911	-4123	-6022	-4503	-5797	-5419	4397	-4479	-4255	-4592	-4115	-4312	-5371	-4650	-4731	14300%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-	-8	-8139	-9181	-894	-1115	-701	-1378	-	-	-	-	-	-	-	-	-	-	-	-	
45(L)	-4414	-3800	-5638	-5628	-2290	-4980	-4628	-1886	-5423	3316	-1236	-5514	-4997	-4750	-5002	-5379	-4399	-2629	-3665	-3690	14400%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-	-8	-8139	-9181	-894	-1115	-701	-1378	-	-	-	-	-	-	-	-	-	-	-	-	
46(R)	-1731	-3015	275	-931	-3487	-2518	-973	-3116	2321	-2955	-2123	224	-2603	256	2808	-1596	-1613	-2730	-2995	-2515	14500%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-	-8	-8139	-9181	-894	-1115	-701	-1378	-	-	-	-	-	-	-	-	-	-	-	-	
47(D)	-2896	-4843	3855	944	-5037	-2600	-2082	-5082	-2528	-4903	-4373	-1209	-3196	-1786	-3536	-2501	-3007	-4517	-5004	-3956	14600%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-	-8	-8139	-9181	-894	-1115	-701	-1378	-	-	-	-	-	-	-	-	-	-	-	-	
48(S)	-1536	-2212	-2363	-2679	-4293	-2279	-3062	-4365	-3331	-4524	-3676	288	-3026	-2967	-3497	3508	-1962	-3259	-4477	-4066	14700%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-	-8	-8139	-9181	-894	-1115	-701	-1378	-	-	-	-	-	-	-	-	-	-	-	-	
49(G)	-2521	-3968	1232	-911	-4849	3373	-2126	-4854	-2535	-4752	-4136	-53	-3115	-1836	-3440	-2284	-2716	-4157	-4880	-3914	14800%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-	-8	-8139	-9181	-894	-1115	-701	-1378	-	-	-	-	-	-	-	-	-	-	-	-	
50(V)	-2767	-2324	-5232	-4770	396	-4827	-3784	-36	-4546	848	-611	-4472	-4518	-3980	-4367	-4081	-2716	3323	-3037	-2660	15500%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-	-8	-8139	-9181	-894	-1115	-701	-1378	-	-	-	-	-	-	-	-	-	-	-	-	
51(D)	-1684	-3285	2735	2014	-3554	-2196	-1177	-3350	92	-3279	-2427	692	-2505	-770	-1595	-1483	-1669	332	-3460	-2676	15700%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-	-8	-8139	-9181	-894	-1115	-701	-1378	-	-	-	-	-	-	-	-	-	-	-	-	
52(V)	-3122	-2888	-5092	-5160	-3522	-4180	-4687	-905	-5060	-2626	-2570	-4662	-4579	-4940	-4923	-4013	-3297	3796	-4414	-4190	15800%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-	-8	-8139	-9181	-894	-1115	-701	-1378	-	-	-	-	-	-	-	-	-	-	-	-	

	-8	-8139	-9181	-894	-1115	-701	-1378	*													
53(V)	369	366	-3075	-2462	-683	-2557	-1420	1415	378	-757	-117	-2098	-2610	-1809	-2037	-1630	1166	2145	-1385	-343	15900%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*													
54(V)	-2624	-2122	-5302	-4991	-2772	-5108	-5139	2623	-4948	-1533	-1475	-4794	-4871	-4894	-5106	-4488	-2620	3088	-4511	-3996	16000%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*													
55(G)	929	-2107	-3852	-4182	-4633	3492	-3809	-4486	-4335	-4732	-3835	-2997	-3132	-3863	-4127	-1761	-1982	-3275	-4720	-4725	16100%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*													
56(L)	-3427	-2938	-5791	-5325	-1449	-5374	-4410	-543	-5063	3041	-255	-5207	-4820	-4126	-4691	-4757	-3351	883	-3184	-3234	16200%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*													
57(R)	-3040	-3724	-3266	82	-4620	-3470	-1396	-3905	804	-3529	-2874	-2133	-3439	-978	3800	-2894	-2709	-3662	-3353	-3267	16300%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*													
58(K)	31	-2412	803	1532	-2743	-1920	-559	-2483	1772	-2421	-1503	-556	1229	727	1079	-566	-893	-2041	-2579	-1915	16400%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*													
59(G)	-2671	-4661	1614	587	-4832	3103	-1901	-4803	-2269	-4648	-4047	421	-3049	-1587	-3230	-2297	-2766	-4245	-4850	-3752	16500%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*													
60(S)	-1499	-2308	-1932	-1859	-4006	1604	-2121	-3754	1362	-3793	-2945	-1833	-2827	-1794	-1902	2738	-1771	-2970	-3910	-3479	16600%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-247	-8139	-2699	-494	-1115	-701	-1378	*													
61(K)	1362	-2232	-619	-98	-2567	-427	-435	-2309	1599	-2265	-1349	1101	-1861	886	-512	833	-740	-1868	-2441	-1767	16700%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-9	-1904	-4943	-894	-1115	-344	-2238	*													
62(S)	-1288	-1904	-3742	-4011	-4384	-2155	-5593	-4209	-3996	-4479	-3573	-2789	-2948	-3606	-3832	3517	228	-3028	-4600	-4451	16800%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*													
63(W)	726	-873	-3261	-2634	1926	-2567	-1425	660	-2252	-701	-68	-2174	-2617	-1698	18	-1648	-972	983	4091	-958	16900%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*													
64(E)	-1527	-2404	121	1636	-2722	-1878	-556	-2474	1241	-2419	-1497	350	-1985	-100	-659	96	70	-2025	-2589	-1903	17000%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*													
65(K)	-8	-2242	-899	770	-2502	-1963	609	-2192	2589	22	-1353	-631	-2052	692	-617	-889	-906	-361	-2455	-1830	17100%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*													
66(A)	3631	-2768	-4492	-4814	-4888	-2992	-4271	-4781	-4818	-5025	-4365	-3727	-3728	-4477	-4545	-2567	-2762	-3852	-4724	-4942	17200%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*													

[1001]

96(A)	1981	-2315	-809	-268	-2645	-531	-579	-2374	232	-2350	-1445	-567	1217	711	445	447	-874	-1951	-2540	-1883	20700%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-	-8	-8139	-9181	-894	-1115	-701	-1378	*												
97(D)	491	-2251	1394	1381	-2671	-1854	1062	-2421	1010	-2367	-1440	-489	-1947	1017	362	-760	250	-623	-2535	-1852	20800%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-	-8	-8139	-9181	-894	-1115	-701	-1378	*												
98(V)	-2039	-1706	-4456	-3939	-1846	-3939	-3049	1986	-3656	1460	-826	804	-3870	-3351	-3565	-3105	-2000	2330	-2796	-2442	20900%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-	-8	-8139	-9181	-894	-1115	-701	-1378	*												
99(Y)	-4840	-3766	-5230	-5581	1898	-5109	-1300	-3727	-5135	-3041	-3132	-3723	-4964	-3861	-4501	-4357	-4690	-3683	3325	4377	21000%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-	-8	-8139	-9181	-894	-1115	-701	-1378	*												
100(E)	-163	-2353	-734	1681	-2674	-1859	888	-2422	1668	-792	-1443	777	-1952	890	286	-766	238	-1975	-2536	-1856	21100%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-257	-8139	-2649	-894	-1115	-701	-1378	*													
101(E)	1017	-2763	862	2042	-3060	-1913	-775	-2838	-493	-2773	-1886	1956	-2143	-136	-1056	265	-1163	-2377	-2948	-2207	21200%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-	-5	-7891	-6833	-894	-1115	-338	-2261	*												
102(E)	944	2422	863	2138	-2740	-436	-567	-2493	894	-2437	-1515	-518	-1994	1767	-673	105	-885	-1023	-2605	-1917	21300%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-	-8	-8139	-9181	-894	-1115	-701	-1378	*												
103(I)	-2660	-2156	-5316	-4965	-2520	-5119	-4900	3165	-4894	297	-1251	-4775	-4828	-4705	-4975	-4470	-2642	2240	-4202	-3814	21400%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-	-8	-8139	-9181	-894	-1115	-701	-1378	*												
104(E)	1068	-2341	-760	2003	-628	-1887	876	-2380	1240	-2347	-1436	-529	-1983	881	-618	-804	-855	-1954	-2530	-1862	21500%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-	-8	-8139	-9181	-894	-1115	-701	-1378	*												
105(P)	-343	-3144	1561	442	-3538	-469	-1216	-3329	-1038	-3274	-2420	-848	2974	812	-1635	469	-1644	-2849	-3462	-2693	21600%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-	-8	-8139	-9181	-894	-1115	-701	-1378	*												
106(N)	-1173	-2375	-814	627	-2376	-2071	1767	-2279	-479	-2336	-1509	3151	-2218	-415	-957	-1093	-1120	-138	-2486	647	21700%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-	-8	-8139	-9181	-894	-1115	-701	-1378	*												
107(M)	-3415	-2850	-5626	-5252	-1352	-5488	-4282	1361	-5022	2621	2728	-5181	-4778	-4005	-4613	-4776	-3292	69	-3071	-3194	21800%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-	-8	-8139	-9181	-894	-1115	-701	-1378	*												
108(K)	-1941	-3058	-1997	-1232	-3650	-2740	-1025	-3210	3059	-3010	-2208	499	-2766	1457	1261	-1817	90	-2858	-3002	-2622	21900%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-	-8	-8139	-9181	-894	-1115	-701	-1378	*												
109(P)	1129	-2428	-740	964	-2747	-1913	589	-2491	1139	-2440	-1525	-552	1941	1446	-655	-480	-913	-2050	-2610	-1935	22000%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-	-8	-8139	-9181	-894	-1115	-701	-1378	*												
110(G)	-2276	-2907	-2347	-2709	-4832	3554	-3345	-5005	-3678	-5053	-4315	1193	-3507	-3243	-3937	-2418	-2674	-3974	-4703	-4521	22100%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	

	-	-8139	-9181	-894	-1115	-701	-1378	*													
111(A)	1730	-2349	958	-198	-2661	-1868	-535	-2405	927	-2362	-1444	414	-1966	788	-830	790	-840	-303	-2540	-1863	22200%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*													
112(T)	1350	-1148	-14	-1461	-1155	-2314	-1111	758	-1275	-1024	1167	-1475	-2388	-1111	-1501	334	1843	354	-1581	-1182	22300%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*													
113(L)	-3333	-2796	-5806	-5293	-1506	-5535	-4502	1096	-5103	2935	-282	-5232	-4857	-4172	-4762	-4864	-3236	506	-3264	-3351	22400%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*													
114(A)	1769	-1525	158	-857	-1603	148	-891	-1181	-752	187	-712	-1040	-2228	660	-1135	-1111	-913	1305	-1913	-1442	22500%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*													
115(F)	-4110	-3437	-5436	-5431	4216	-5143	-2159	-1742	-5074	563	-1124	-4290	-4871	-3987	-4561	-4547	-4016	-2374	-1356	-292	22600%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*													
116(A)	3091	-1829	-3998	-4219	-4413	119	-3637	-4216	-4134	-4469	-3523	-2798	-2896	-3656	-3927	1514	-1679	-2983	-4632	-4539	22700%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*													
117(H)	-5197	-4538	-4720	-5009	-4036	-4506	5435	-6314	-4911	-5786	-5667	-4864	-4960	-5011	-4732	-5391	-5395	-4022	-4063	-3641	22800%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*													
118(G)	-4432	-4203	-5692	-5462	-5893	3834	-5028	-6627	-5765	-6297	-5970	-5141	-4804	-5546	-5385	-4727	-4815	-5862	-4924	-5849	22900%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*													
119(F)	-4944	-3387	-5534	-5444	4093	-5246	-2370	-1514	-5107	1089	-868	-4443	-4880	-3998	-4592	-4639	-3934	-2200	-1536	-523	23000%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*													
120(N)	865	-1899	-2020	-1781	-2956	-2135	-1925	-2602	-1809	3	-2052	3468	-2633	-1676	-2141	413	-1437	-2139	-3194	-2737	23100%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*													
121(I)	-2673	-2169	-5324	-4969	-2477	-5123	-4876	3293	-4893	358	-1211	-4780	-4824	-4681	-4961	-4472	-2653	-1969	-4158	-3791	23200%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*													
122(H)	-3381	-3705	-3197	-3491	-4166	638	5216	-5496	-3798	-5304	-4811	-3481	-4185	-3479	-3879	-3508	-3702	-4793	-4170	-3759	23300%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*													
123(Y)	-4816	-3757	-5210	-5549	3410	-5097	2153	-3719	-5105	-3041	-3127	-3715	-4955	-3851	-4483	-4344	-4669	-3870	-547	3677	23400%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*													
124(G)	-1065	-2519	948	-272	-2820	1844	-938	-2566	972	-284	-1622	1553	-2090	-229	-802	-938	-1011	-2133	-2708	-2021	23500%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*													

[1005]

154(V)	129	-1901	-989	821	-2060	-1969	654	-1704	498	-52	-1037	-703	-2057	695	-796	443	-344	1871	-2192	-1626	26500%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-	-8139	-9181	-894	-1115	-701	-1378	*													
155(O)	576	-2355	344	1159	-2675	-1856	-515	-508	1502	-2370	-1444	571	-1949	1878	419	-764	-822	-1978	-2538	-1856	26600%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-	-8139	-9181	-894	-1115	-701	-1378	*													
156(G)	3239	-3889	516	2361	-5355	3646	-3337	-5629	-3818	-5498	-4951	-2619	-3905	-3187	-4377	-3211	-3532	-4837	-4895	-4826	26700%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-295	-250	
-	-155	-5318	-9181	-2159	-366	-701	-1378	*													
157(G)	753	-2518	-789	488	-2848	2300	-672	-2942	596	-2529	-1627	481	-2112	-224	471	-962	-1024	-2149	-2694	-2033	27300%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-	-8139	-9181	-894	-1115	-701	-1378	*													
158(G)	52	2212	-3792	-4133	-4698	3627	-3843	-4580	-4356	-4812	-3937	-3058	-3216	-3901	-4170	-1874	-2095	-3384	-4734	-4756	27400%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-	-8139	-9181	-894	-1115	-701	-1378	*													
159(V)	-2483	-2030	-5123	-4769	-2667	-4797	-4593	2349	-4661	-1545	-1424	-4502	-4648	-4554	-4752	-4115	823	2986	-4159	-3678	27500%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-	-8139	-9181	-894	-1115	-701	-1378	*													
160(P)	-2541	-3139	-2413	-2753	-4726	-2991	-3342	-5055	-3527	-5058	-4393	1199	4031	-3244	-3757	-2665	-2911	-4148	-4583	-4362	27600%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-	-8139	-9181	-894	-1115	-701	-1378	*													
161(C)	1577	3078	1357	-656	-2654	-219	891	-2359	-617	-2434	-1578	-891	-2199	-545	-1093	872	-1043	-1946	-2701	-2102	27700%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-	-8139	-9181	-894	-1115	-701	-1378	*													
162(L)	-2149	-2404	-3995	-3997	-2053	-3121	-3283	-1687	-3689	3041	-1206	-3360	-3626	-3433	-3567	469	-2414	-1973	-3145	-2755	27800%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-	-8139	-9181	-894	-1115	-701	-1378	*													
163(I)	-2527	-2052	-5072	-4613	-2047	-4674	-3911	2658	-4413	117	841	-4323	-4439	-4023	-4320	-3916	-2489	2176	-3342	-3040	27900%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-	-8139	-9181	-894	-1115	-701	-1378	*													
164(A)	3631	-2768	-4492	-4815	-4888	-2992	-4271	-4781	-4818	-5025	-4365	-3727	-3728	-4477	-4545	-2567	-2762	-3852	-4724	-4942	28000%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-	-8139	-9181	-894	-1115	-701	-1378	*													
165(V)	-2623	-2122	-5301	-4990	-2770	-5102	-2426	-4946	-1532	-1474	-4791	-4869	-4891	-5102	-4483	-2619	3200	-4506	-3991	28100%	
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-	-8139	-9181	-894	-1115	-701	-1378	*													
166(H)	-495	-2631	903	-2051	722	-3242	3753	-2386	-2056	-2342	-1863	-2047	-3330	-1815	-2362	-2318	-2233	-2272	981	3315	28200%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-	-8139	-9181	-894	-1115	-701	-1378	*													
167(Q)	-4589	-4392	-3927	-4146	-5099	-4221	-4099	-5973	-3840	-5564	-5304	-4230	-4693	4575	-3820	-4704	-4772	-5612	-4577	-4751	28300%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-	-8139	-9181	-894	-1115	-701	-1378	*													
168(D)	-2873	-4609	3943	-302	-4948	-2633	-2157	-5087	-2604	-4922	-4387	428	-3235	-1872	-3575	-2522	-3009	-4491	-4932	-3946	28400%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	

[1006]

	-	-8139	-9181	-894	-1115	-701	-1378	*													
169(A)	1776	-1612	-1274	138	-1698	-2092	-816	-1295	150	-1526	-780	-843	1212	-538	-1013	-1056	-894	1275	-1968	635	28500%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-	-8139	-9181	-894	-1115	-701	-1378	*													
170(S)	-1545	-2420	1061	-1518	-4049	-2206	-2206	-3839	-2264	-3938	-3103	-1627	-2814	-1909	-2758	2666	2313	-3045	-4143	-3550	28600%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-	-8139	-9181	-894	-1115	-701	-1378	*													
171(G)	-2999	-3461	-2978	-3207	-5161	3669	-3454	-5283	15	-5188	-4565	-3174	-3946	-3312	-3140	-3128	-3317	-4492	-4622	-4749	28700%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-	-8139	-9181	-894	-1115	-701	-1378	*													
172(N)	-887	-2349	-736	911	-2668	-1860	-518	-2415	711	-2363	-1440	1880	-1954	436	799	671	1230	-539	-2534	-1855	28800%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-	-8139	-9181	-894	-1115	-701	-1378	*													
173(A)	3631	-2768	-4492	-4815	-4888	-2992	-4271	-4781	-4818	-5025	-4365	-3727	-3728	-4477	-4545	-2567	-2762	-3852	-4724	-4942	28900%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-	-8139	-9181	-894	-1115	-701	-1378	*													
174(K)	-1368	-2597	-1358	-720	-2953	-2298	1933	-2603	1975	851	-1706	-958	-2360	859	1012	-1273	1094	-2245	-2690	-2165	29000%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-	-8139	-9181	-894	-1115	-701	-1378	*													
175(D)	-27	-2613	2320	1049	-2823	-1962	1973	-2684	544	-2624	-1712	1408	-2115	345	-869	-969	-1058	-2231	-2793	-2087	29100%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-	-8139	-9181	-894	-1115	-701	-1378	*													
176(V)	-1098	-938	-3279	-2658	-899	-2643	-1513	1265	1006	1388	-124	-2232	-2688	-1950	-2151	-1725	-311	1669	-1425	627	29200%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-	-8139	-9181	-894	-1115	-701	-1378	*													
177(A)	3342	-1826	-4064	-4295	-4368	111	-3651	-4129	-4159	-4412	-3484	-2820	-2903	-3689	-3929	-1469	409	2949	-4599	-4508	29300%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-	-8139	-9181	-894	-1115	-701	-1378	*													
178(L)	-4414	-3800	-5638	-5628	-2290	-4980	-4628	-1886	-5423	3316	-1236	-5514	-4997	-4750	-5002	-5379	-4399	-2629	-3665	-3690	29400%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-	-8139	-9181	-894	-1115	-701	-1378	*													
179(S)	2216	-1831	-3961	-4157	-4409	656	-3608	-4213	-4076	-4462	-3514	-2782	-2892	-3613	-3895	2686	-1675	-2982	-4623	-4523	29500%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-	-8139	-9181	-894	-1115	-701	-1378	*													
180(Y)	-3634	-3059	-4918	-4872	-6	-4597	-1435	223	-4437	-250	-1998	-3545	-4494	-3522	-4004	-3782	-3536	-2621	2928	4349	29600%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-	-8139	-9181	-894	-1115	-701	-1378	*													
181(A)	3391	-1860	-3998	-4279	-4411	2128	-3684	-4207	-4197	-4490	-3565	-2837	-2929	-3729	-3959	705	-1718	-3001	-4636	-4534	29700%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-	-8139	-9181	-894	-1115	-701	-1378	*													
182(K)	370	307	113	-351	-2280	-1925	-615	324	1888	-2040	-1194	-632	-2022	-205	-736	1541	14	-1628	-2330	-1720	29800%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-	-8139	-9181	-894	-1115	-701	-1378	*													

183(G)	2572	-2028	-3834	-4246	-4573	2752	-3783	-4406	-4316	-4661	-3751	-2950	-3070	-3837	-4092	-1670	-1898	-3191	-4701	-4686	29900%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-	-8139	-9181	-894	-1115	-701	-1378	*													
184(I)	-2178	-1808	-4630	-4153	-2094	-4190	-3417	3121	-3909	3111	-1023	698	-4099	-3656	-3864	-3388	-2148	1742	-3147	-2761	30000%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-	-8139	-9181	-894	-1115	-701	-1378	*													
185(G)	-4433	-4203	-5092	-5462	-5893	3834	-5028	-6627	-5765	-6297	-5970	-5141	-4804	-5546	-5385	-4727	-4815	-5862	-4924	-5849	30100%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-	-8139	-9181	-894	-1115	-701	-1378	*													
186(G)	1392	2751	-4353	-4536	-4308	2864	-3681	-4084	-4233	-4354	-3475	-2859	-2890	-3744	-3957	712	-1656	-2914	-4553	-4470	30200%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-	-8139	-9181	-894	-1115	-701	-1378	*													
187(G)	855	-1822	-3738	-3769	-4188	2507	-3358	-3667	-4196	-3283	-2667	986	-3302	-3821	-1441	2334	-2867	-4408	-4236	30300%	
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-	-8139	-9181	-894	-1115	-701	-1378	*													
188(R)	-3706	-3692	-4480	-3846	1391	-4057	-2273	-3795	-1906	-3355	-3181	-3458	-4263	-2675	3948	-3768	-3671	-3813	-2293	-1328	30400%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-	-8139	-9181	-894	-1115	-701	-1378	*													
189(A)	2844	-1670	3814	-3873	-4048	-1958	-3316	-3787	-3686	-4061	-3156	-2598	-2734	-3361	-3578	1088	1807	-2713	-4286	-4138	30500%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-	-8139	-9181	-894	-1115	-701	-1378	*													
190(G)	-4178	-3965	-4853	-5222	-5696	3828	-4823	-6386	-5533	-6087	-5741	-4896	-4608	-5312	-5178	-4461	-4560	-5613	-4754	-5633	30600%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-	-8139	-9181	-894	-1115	-701	-1378	*													
191(V)	-2496	-2036	-5139	-4788	-2680	-4825	-4637	-2522	-4686	-1548	-1431	-4527	-4668	-4584	-4783	-4148	629	2919	-4191	-3706	30700%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-	-8139	-9181	-894	-1115	-701	-1378	*													
192(I)	-2762	-2307	-5270	-4785	1172	-4884	-3992	3346	-4576	662	572	-4539	-4535	-4008	-4400	-4134	-2703	757	-3223	-3041	30800%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-	-8139	-9181	-894	-1115	-701	-1378	*													
193(E)	454	-3086	-801	3278	-4125	-2346	-1868	-3919	-1868	-3932	-3156	-1291	1208	-1533	-2450	-1842	-2097	-3305	-4121	-3395	30900%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-	-8139	-9181	-894	-1115	-701	-1378	*													
194(T)	-2138	-3109	-4599	-4810	-4910	-3305	-4346	-4865	-4769	-5072	-4551	-3987	-3998	-4580	-4545	-2099	4033	-4113	-4684	-4915	31000%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-	-8139	-9181	-894	-1115	-701	-1378	*													
195(T)	-1323	-1975	-2764	-2978	-4207	-3152	-3105	-4004	-3295	-4229	-3354	134	-2901	-2988	-3387	-1113	3742	-2969	-4405	-4119	31100%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-	-8139	-9181	-894	-1115	-701	-1378	*													
196(F)	-4110	-3437	-5436	-5431	4218	-5143	-2159	-1742	5074	563	-1124	-4290	-4871	-3987	-4561	-4547	-4616	-2374	-1358	-292	31200%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-	-8139	-9181	-894	-1115	-701	-1378	*													
197(K)	-111	-2844	-1470	1220	-3294	-2415	-4860	-2539	2446	-2798	-1945	-1056	-2475	767	2054	-1421	-1432	-2544	-2858	-2356	31300%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	

[1008]

	-	-8139	-9181	-894	-1115	-701	-1378	*													
198(E)	542	-3735	1715	2886	-3881	-2368	-1442	-3818	-1468	-3725	-2924	-899	-2713	1281	-2087	-1777	-2041	-3331	-3909	-3046	31400%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-	-8139	-9181	-894	-1115	-701	-1378	*													
199(E)	-4574	-4665	-2714	3919	-5655	-3995	-3886	-6219	-4238	-5898	-5604	-3415	-4513	-3838	-4570	-4456	-4726	-5786	-4878	-5197	31500%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-	-8139	-9181	-894	-1115	-701	-1378	*													
200(T)	-1211	1331	-3446	-2962	-1399	-2495	-1930	-869	-2610	-1374	-769	-2403	-2815	-2309	-2526	-1685	3305	195	-1933	1430	31600%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-	-8139	-9181	-894	-1115	-701	-1378	*													
201(E)	-1941	-3222	-921	3293	-3618	-2473	-1316	-3186	916	-3225	-2465	-1147	-2749	-923	-1064	-1790	-1902	-171	-1358	-2802	31700%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-	-8139	-9181	-894	-1115	-701	-1378	*													
202(I)	-1286	-1898	-3764	-4016	-4329	-2157	-3572	-4120	-3959	-4408	-3517	-2785	-2947	-3588	-3797	697	3756	-2989	-4554	-4399	31800%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-	-8139	-9181	-894	-1115	-701	-1378	*													
203(D)	-4580	-4701	-4174	-3014	-5700	-3967	-3905	-6376	-4478	-6024	-5744	-3355	-4501	-3870	-4926	-4440	-4750	-5894	-4922	-5231	31900%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-	-8139	-9181	-894	-1115	-701	-1378	*													
204(L)	-3705	-3122	-6060	-5527	-1359	-5814	-4569	1065	-5292	3069	-146	-5564	-4963	4163	-4828	-5215	-3571	-1279	-3159	-3298	32000%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-	-8139	-9181	-894	-1115	-701	-1378	*													
205(F)	-3777	-3220	-5271	-5259	4268	-4892	-2120	417	-4916	-1143	-1314	-4142	-4753	-3956	-4473	-4270	-3740	-1899	-1349	-269	32100%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-	-8139	-9181	-894	-1115	-701	-1378	*													
206(G)	-4433	-4203	-5092	-5462	-5893	3834	-5028	-6627	-5765	-6297	-5970	-5141	-4804	-5546	-5385	-4727	-4815	-5862	-4924	-5849	32200%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-	-8139	-9181	-894	-1115	-701	-1378	*													
207(E)	-4574	-4665	-2714	3918	-5655	-3995	-3886	-6219	-4238	-5898	-5604	-3415	-4513	-3838	-4570	-4456	-4726	-5786	-4878	-5197	32300%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-	-8139	-9181	-894	-1115	-701	-1378	*													
208(Q)	-3157	-3146	-3170	-2459	-4497	-3515	-1163	-4161	-443	-3809	-3189	-2392	-3620	4200	1284	-3063	-2944	-3900	-3556	-3420	32400%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-	-8139	-9181	-894	-1115	-701	-1378	*													
209(A)	-2673	-1334	-3318	-2853	-1740	371	-2073	483	-2577	-1549	-928	-2359	-2198	2295	-2567	-1629	191	932	-2245	-1899	32500%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-	-8139	-9181	-894	-1115	-701	-1378	*													
210(V)	-2620	-2129	-5293	-4983	-2756	-5076	-5100	-1877	-4932	-1522	-1466	-4777	-4855	-4870	-5082	-4456	-2619	3416	-4480	-3969	32600%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-	-8139	-9181	-894	-1115	-701	-1378	*													
211(L)	-4414	-3800	-5638	-5628	-2250	-4960	-4678	-1886	-5423	3316	-1236	-5514	-4997	-4750	-5002	-5379	-4399	-2629	-3665	-3690	32700%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-	-8139	-9181	-894	-1115	-701	-1378	*													

212(C)	-2243	-5044	-4840	-4449	-1998	-3905	-3598	-31	-4138	449	-430	-3902	-4040	-3778	-4010	-3184	-2306	1347	-3209	-2883	32800%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-	-	-8139	-9181	-894	-1115	-701	-1378	*												
213(G)	-4433	-4203	-5092	-5462	-5893	3834	-5028	-6627	-5765	-6297	-5970	-5141	-4804	-5546	-5385	-4727	-4815	-5862	-4924	-5849	32900%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-	-	-8139	-9181	-894	-1115	-701	-1378	*												
214(G)	677	-2128	-3838	-4171	-4647	3536	-3816	-4506	-4340	-4749	-3857	-3009	-3149	-3871	-4137	-1784	-2005	-3297	-4725	-4735	33000%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-	-	-8139	-9181	-894	-1115	-701	-1378	*												
215(V)	378	724	-3707	-3104	-1180	-2986	-1919	1210	-2734	1302	-359	-2627	-3014	-2382	-2566	-2089	1123	1949	-1773	-1423	33100%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-	-	-8139	-9181	-894	-1115	-701	-1378	*												
216(M)	-948	-1407	-1515	156	-1452	-2164	1677	-1030	-821	-1302	1976	-1113	-2245	-718	-1173	773	1715	1332	-1794	-1343	33200%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-	-	-8139	-9181	-894	-1115	-701	-1378	*												
217(E)	1367	-2528	725	2286	-2932	240	-791	-2681	328	-2645	-1744	-674	-2162	-351	-939	545	-1095	-2227	-2828	-2143	33300%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-	-	-8139	-9181	-894	-1115	-701	-1378	*												
218(L)	-3705	-3122	-4060	-5527	-1359	-5814	-4569	1065	-5292	3069	-146	-5564	-4963	-4163	-4828	-5215	-3571	-1279	-3159	-3298	33400%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-	-	-8139	-9181	-894	-1115	-701	-1378	*												
219(V)	-2600	-2108	-5251	-4894	-2568	-5025	-4783	2479	-4810	-1354	1358	-4683	-4772	-4654	-4895	-4362	-2584	3018	-4181	-3758	33500%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-	-	-8139	-9181	-894	-1115	-701	-1378	*												
220(K)	-1633	-2905	-1573	706	-3375	-2487	-900	-3003	2925	-2849	-2008	-1128	-2541	1714	784	-1509	-105	-2617	-2894	-2418	33600%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-	-	-8139	-9181	-894	-1115	-701	-1378	*												
221(A)	2352	2066	-2593	-2000	-947	-2434	-1271	-486	-32	-832	714	-1817	-2498	-1493	-1792	-1483	274	453	-1419	501	33700%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-	-	-8139	-9181	-894	-1115	-701	-1378	*												
222(G)	224	-1905	-3562	-3696	-3684	3361	-3297	-3220	-3625	81	-2886	-2733	-2977	-3326	-3545	-1606	-1763	-2574	-4068	-3810	33800%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-	-	-8139	-9181	-894	-1115	-701	-1378	*												
223(F)	-4781	-3756	-5207	-5542	4341	-5070	-1342	-3653	-5111	-2971	-3065	-3743	-4949	-3874	-4496	-4351	-4650	-3829	-591	1725	33900%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-	-	-8139	-9181	-894	-1115	-701	-1378	*												
224(E)	-2413	-4114	221	3465	-4392	-2485	-1689	-4248	-1608	-4112	-3396	-1094	-2951	-1336	871	-2119	-2441	-3763	-4239	-3395	34000%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-	-	-8139	-9181	-894	-1115	-701	-1378	*												
225(T)	-1461	-1864	-3139	-2645	-2659	-2483	-2136	-1734	-1646	-2359	-1761	-2298	-2936	-1995	920	-1748	3354	967	-2989	-2654	34100%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-	-	-8139	-9181	-894	-1115	-701	-1378	*												
226(L)	-3831	-3268	-5314	-5148	-673	-5068	-2476	-1443	-4706	3059	-789	-4359	-4756	-3864	-4314	-4462	-3729	-2115	-1672	1736	34200%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	

[1010]

	-	-8139	-9181	-894	-1115	-701	-1378	*													
227(V)	-1815	-1960	-4426	-4359	-2977	-3037	-3800	-439	-4098	-2145	-1909	-3451	-3600	-3879	-4011	-2397	2510	2699	-3974	-3688	34300%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-	-	-8139	-9181	-894	-1115	-701	-1378	*												
228(E)	-2863	-4790	1397	3563	-4990	-2584	-2061	-5021	-2476	-4848	-4798	-1204	-3182	-1761	-3454	-2476	-2968	-4462	-4951	-3920	34400%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-	-	-8139	-9181	-894	-1115	-701	-1378	*												
229(A)	2686	-1916	-1440	275	-2529	-292	-1240	-2176	-998	-2345	1183	-1184	-2355	-905	-1425	512	-1158	-1817	-2697	-2174	34500%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-	-	-8139	-9181	-894	-1115	-701	-1378	*												
230(G)	-4435	-4203	-5092	-5462	-5893	3834	-5028	-6627	-5765	-6297	-5970	-5141	-4804	-5546	-5385	-4727	-4815	-5862	-4924	-5849	34600%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-	-	-8139	-9181	-894	-1115	-701	-1378	*												
231(V)	-4099	-3483	-4921	-5048	-109	-4705	-1565	-2914	-4494	-2334	2010	-3723	-4707	-3735	-4111	-4065	-4068	-3172	-847	-4618	34700%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-	-	-8139	-9181	-894	-1115	-701	-1378	*												
232(Q)	1711	-2410	-712	934	-2738	-1925	-464	-2477	-171	-2433	-1524	-574	-2035	2086	345	902	-623	-2942	-2608	-1941	34800%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-	-	-8139	-9181	-894	-1115	-701	-1378	*												
233(P)	-3403	-4071	-1922	817	-5220	-3359	-3173	-5423	-3337	-5281	-4771	-2564	4045	-2989	-3769	-3320	-3624	-4817	-4763	-4636	34900%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-	-	-8139	-9181	-894	-1115	-701	-1378	*												
234(E)	-2870	-4786	1265	3587	-4950	-2600	-2068	-5028	-2483	-4852	-4303	-1212	-3188	-1768	-3458	-2484	-2974	-4467	-4950	-3925	35000%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-	-	-8139	-9181	-894	-1115	-701	-1378	*												
235(M)	-3089	-2618	-5526	-4978	-1443	-5128	-4045	653	-4739	1429	4269	-4803	-4610	3915	-4423	-4378	-2995	1140	-3054	-3074	35100%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-	-	-8139	-9181	-894	-1115	-701	-1378	*												
236(A)	3631	-2768	-4492	-4815	-4888	-2952	-4271	-4781	-4818	-5025	-4365	-3727	-3728	-4477	-4545	-2567	-2762	-3852	-4724	-4942	35200%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-	-	-8139	-9181	-894	-1115	-701	-1378	*												
237(Y)	-4797	-3764	-5203	-5543	1114	-5669	-1339	-3694	-5111	-3013	-3107	-3741	-4951	-3878	-4497	-4354	-4668	-3665	-588	-4743	35300%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-	-	-8139	-9181	-894	-1115	-701	-1378	*												
238(F)	-3828	-3605	-4146	-4086	-4792	-4207	-2060	-3492	774	-3071	-3005	-3556	-4434	-3287	-2952	-3668	-3856	-3571	-1593	-494	35400%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-	-	-8139	-9181	-894	-1115	-701	-1378	*												
239(E)	-2775	-4471	-511	3582	-4815	-2610	-2057	-4863	-2317	-4711	-4124	1234	3182	-1755	-3103	-2442	-2884	-4366	-4753	-3820	35500%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-	-	-8139	-9181	-894	-1115	-701	-1378	*												
240(C)	-1407	5023	-4397	-4323	-3916	-2468	-3398	-1251	-3952	-2540	-2082	-3044	-3133	-3588	-3744	-1803	1473	1125	-3677	-3390	35600%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-	-	-8139	-9181	-894	-1115	-701	-1378	*												

241(L)	-3375	-2847	-5795	-5233	-1386	-5455	-4298	708	-5010	2859	1340	-5155	-4777	-4028	-4622	-4755	-3254	814	-3163	-3213	35700%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*													
242(H)	-2515	-4224	-445	940	-4287	-2505	-4583	-4377	-1764	-4237	-3571	2007	-3009	-1439	-2361	-2209	-2569	-3893	-4267	-3353	35800%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*													
243(E)	-3177	2571	-2701	3711	-4851	-3438	-3479	-4769	-3558	-4932	-4408	-3081	-4005	-3370	-3802	-3269	-3451	-4260	-4554	-4524	35900%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*													
244(L)	-85	-1333	-3893	-3288	-1111	-3185	-2083	1066	-2910	2310	1961	-2823	-3170	-2501	-2721	-2289	-113	436	-1859	-1556	36000%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*													
245(K)	-2513	-3173	-2941	-2370	-4402	-3094	-1824	-3895	3666	-3734	-3068	-2271	-3377	-1447	-616	-2537	751	-3485	-3625	-3471	36100%
-	-149	-500	233	43	-381	400	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-155	-3318	-9181	-196	-2974	-701	-1378	*													
246(L)	-3571	-3023	-5954	-5375	-1327	-5646	-4390	-632	-5138	2962	1671	-5358	-4852	-4044	-4689	-4963	-3436	742	-3082	-3239	36300%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*													
247(I)	-2980	-2494	-5473	-5109	-1958	-5196	-4587	3728	-4915	267	-761	-4933	-4833	-4427	-4799	-4598	-2949	64	-3627	-3397	36400%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*													
248(V)	-1685	-1668	-4005	-3732	-2081	-3082	-2853	-227	-3402	-1488	1383	-3123	-3419	-3145	-3320	367	-1607	3332	-2874	-2504	36500%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*													
249(D)	-2963	-4562	3864	-1035	-4953	-2751	-2187	-4998	767	-4822	-4260	-1424	-3314	-1891	-3072	-2624	-3060	-4467	-4770	-3962	36600%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*													
250(L)	-2768	-2715	-4842	-4633	-1875	-3988	-3790	-1038	-4207	3056	-562	-4150	-4179	-3740	-3989	-3399	699	-1545	-3154	-3067	36700%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*													
251(M)	-2822	-2356	-5342	-4861	-1758	-4985	-4151	2587	-4863	173	4005	-4649	-4601	-4076	-4487	-4251	-2764	766	-3321	-3210	36800%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*													
252(Y)	-4562	-3630	-5142	-5401	1518	-4992	-1300	-3544	-4968	-2963	-2986	-3671	-4869	-3786	-4393	381	-4432	-3662	2413	4375	36900%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*													
253(E)	-1959	-3457	-5668	3135	-3841	-2347	-1410	-3622	-1165	-3547	-2751	-1000	-2718	1879	-1666	-1757	692	-3162	-3709	-2960	37000%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*													
254(G)	-347	-2818	-1213	201	-3253	2635	-921	-2921	1474	2822	-1972	-1002	-2459	-486	658	-1397	-1435	-2521	-2923	-2378	37100%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*													
255(G)	-4435	-4203	-5092	-5462	-5893	3834	-5028	-6627	-5765	-6297	-5970	-5141	-4804	-5546	-5385	4727	-4815	-5862	-4924	-5849	37200%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	

	-8	-8139	-9181	-894	-1115	-701	-1378	*													
256(I)	-2042	-1769	-4321	-3740	-1316	-3753	-2668	3134	-3389	1017	1999	194	-3653	-2558	-3221	-2884	-1980	-344	-2325	-2073	37300%
	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
	-8	-8139	-9181	-894	-1115	-701	-1378	*													
257(A)	1914	-1640	-1237	128	-1748	461	-793	-1355	-577	-49	-817	-905	-2155	-498	-993	1037	162	-1149	-2002	624	37400%
	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
	-8	-8139	-9181	-894	-1115	-701	-1378	*													
258(N)	365	-3609	1001	557	-4083	1196	-1518	-3930	-1535	-3838	-3055	3219	-2763	-1148	-2243	-1845	-2131	-3433	-4027	-3144	37500%
	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
	-8	-8139	-9181	-894	-1115	-701	-1378	*													
259(M)	3656	-3159	-5816	-5350	-1349	-5471	-4248	-812	-4928	948	4920	-5248	-4838	-4039	-4539	-4860	-3558	-1557	-3044	-3030	37600%
	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
	-8	-8139	-9181	-894	-1115	-701	-1378	*													
260(R)	-1614	-1943	-2260	-1663	-896	-2765	-1089	-1596	-1089	360	-1133	2239	-2814	-1215	2408	-1789	-1535	-1468	1995	1546	37700%
	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
	-8	-8139	-9181	-894	-1115	-701	-1378	*													
261(Y)	-1548	-2973	568	509	-2846	-2207	1172	-2986	441	-2941	-2110	1645	-2446	-659	-1264	-1389	-1509	-2582	-2924	3695	37800%
	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
	-8	-8139	-9181	-894	-1115	-701	-1378	*													
262(S)	279	-1844	-3877	-4131	-4448	136	-3634	-4260	-4132	-4511	-3561	-2782	-2903	-3648	-3936	3391	-1692	-3009	-4661	-4566	37900%
	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
	-8	-8139	-9181	-894	-1115	-701	-1378	*													
263(J)	-2653	-2149	-5311	-4951	-2538	-5114	-4905	2957	-4891	332	-1267	-4770	-4827	-4713	-4978	-4466	-2636	2521	-4218	-3821	38000%
	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
	-8	-8139	-9181	-894	-1115	-701	-1378	*													
264(S)	-2212	-2711	-4019	-4348	-4697	-2899	-4045	-4998	-4527	-5102	-4364	-3492	-3638	-4203	-4355	3681	-2664	-3902	-4616	-4605	38100%
	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
	-8	-8139	-9181	-894	-1115	-701	-1378	*													
265(N)	-2725	-4778	2906	990	-4896	-2486	-1922	-2885	-2320	-4718	-4140	3045	-3069	-1612	-3311	-2334	-2821	-4325	-4919	-3794	38200%
	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
	-8	-8139	-9181	-894	-1115	-701	-1378	*													
266(T)	-2061	-3396	-596	903	-4071	-2367	-1713	-3874	-1685	-3859	-3102	2125	-2850	-1369	-4277	-1897	3157	-3351	-4055	-3271	38300%
	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
	-8	-8139	-9181	-894	-1115	-701	-1378	*													
267(A)	3410	-2035	-3979	-4290	-4573	659	-3796	-4400	-4332	-4660	-3754	-2918	-3080	-3858	-4099	-1690	-1908	-3194	-4697	-4697	38400%
	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
	-8	-8139	-9181	-894	-1115	-701	-1378	*													
268(E)	-2118	-3486	-1036	2964	-3935	-2588	-1284	-3596	1878	-3417	-2636	-1209	-2827	1323	773	-1823	-2032	-3199	-3455	-2917	38500%
	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
	-8	-8139	-9181	-894	-1115	-701	-1378	*													
269(Y)	-4524	-3818	-5100	-5310	1910	-4972	-1299	-3522	634	-2951	-2965	-3649	-4847	-3741	-4299	-4119	-4391	-3637	2997	4211	38600%
	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
	-155	-8139	-3345	-894	-1115	-701	-1378	*													

[1015]

299(F)	-4347	-3577	-4619	543	3858	-4820	-1320	-3438	-4609	-2900	-2894	-3532	-4750	-3628	-4209	-4078	-4237	-3546	-601	2917	41600%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*													
300(A)	2627	-1603	-4068	-3628	-2047	-3165	-2823	1205	-3349	-1486	-1089	-3103	-3432	-3073	-3298	-2387	197	1100	-2812	-2447	41700%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*													
301(K)	-2354	-3363	-2464	163	-4038	-3041	-1188	-3501	2928	-3234	-2487	-1893	-3047	-758	2488	-458	-2162	-3184	-3159	-2895	41800%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*													
302(M)	893	-2361	740	1780	-2680	-537	-524	-2429	930	-2376	1895	-498	-1958	-66	722	775	-833	-1982	-2545	-1894	41900%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*													
303(W)	-2965	-2553	-4795	-4482	3048	-4315	-1779	-1426	-4093	92	-987	-3564	-4179	-3330	-3743	-3478	-2878	664	4754	-84	42000%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*													
304(I)	-2621	-2138	-5246	-4839	-2354	-4887	-4577	3052	-4747	-250	1943	-4641	-4713	-4488	-4769	-4301	-2587	2384	-3834	-3596	42100%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*													
305(L)	684	-1318	-1741	781	-10478	-2223	-1037	-827	-1042	1683	-533	-1287	-2325	-915	-1348	787	-153	84	-1758	-1359	42200%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*													
306(E)	364	-4165	621	3314	-4393	-2398	-1868	-4262	-1636	-4169	-3446	728	-2893	-1338	-2626	-2045	-2402	-3767	-4362	-3401	42300%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*													
307(N)	-939	879	-1235	481	-1738	1352	935	-1357	-549	-1572	-816	2186	-2155	-482	165	-1022	-880	30	1990	1446	42400%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*													
308(Q)	687	-2363	-273	511	-2721	-316	-544	-2465	1211	-2405	-1484	584	-1989	2135	1528	-412	-668	-2019	-2566	-1895	42500%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*													
309(A)	2050	-1857	-1081	-528	-2013	-2012	188	-1645	-385	-1313	-1011	486	-2103	-346	1857	-563	306	-261	-2169	-1624	42600%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*													
310(G)	-1848	-2469	-2088	-2292	1262	2944	-2358	-3563	-2904	-3628	-3017	2347	-3167	-2550	-3185	-1986	-2162	-3006	-2814	-1874	42700%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*													
311(Y)	475	1019	-1606	-1042	225	-2162	-935	-946	-891	-1226	1222	357	-2767	1002	1577	-1172	-888	-800	-1730	2446	42800%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*													
312(P)	-87	-2373	-1362	756	-3738	-2205	-2007	-3445	-1934	-3575	-2761	-1555	3598	-1697	-2362	-1566	286	-2803	-3832	-3277	42900%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-2336	-8139	-325	-894	-1115	-701	-1378	*													
313(K)	804	-1483	564	-230	-1920	-1335	-161	-1605	2889	-1630	-1021	349	-1569	232	698	-786	-759	-1358	-1637	-1317	43000%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	

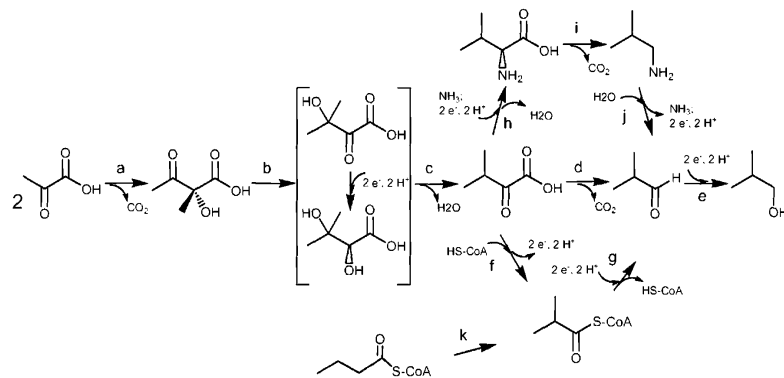
[1016]

	-38	-5840	-6882	-894	-1115	-3098	-179	*													
314(E)	756	-1695	521	2831	-2050	-1029	-293	-1804	-118	-1919	-1331	69	-1441	-4	-527	-653	-614	-1512	-1988	-1505	43100%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-109	-375	*													
315(T)	942	-2382	-739	1086	-2714	151	-581	-2459	-171	-2415	-1499	414	-2004	-128	839	1365	1730	-2017	-2592	-1915	43200%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*													
316(M)	-2196	-1920	-4499	-3891	1726	-3822	-2504	-645	-3523	1973	3030	-3442	-3673	-2938	-3257	-2944	-2114	326	-2014	1662	43300%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*													
317(H)	883	-2314	-747	517	-2618	-1863	1714	-647	1272	-2322	-1408	1011	472	-69	433	-772	1411	-299	-2507	-1836	43400%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*													
318(A)	2474	-2397	-816	-367	-2797	-273	-722	-2529	555	-2507	-1610	592	-2110	837	-805	-138	-1006	2092	2699	-2039	43500%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*													
319(M)	-154	-986	-2485	-337	1024	-375	-1232	325	-444	867	1235	-1752	-2474	-1419	1020	-1455	670	-411	831	535	43600%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*													
320(R)	-1311	-2432	-1349	-724	-2724	-2272	-799	-2361	613	-644	1079	976	597	-382	2908	-1246	-1219	-2044	-2579	-2061	43700%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*													
321(R)	897	-2364	-833	905	-2678	-1930	-568	-2405	1293	-2366	1485	-575	-2020	-117	2045	95	-893	-1984	-2543	-1892	43800%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*													
322(N)	505	-2300	-750	523	-2598	121	-525	594	485	95	-1395	1720	-1957	348	224	551	-821	-1910	-2497	-1828	43900%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*													
323(E)	444	-2266	-766	1488	-2551	-1871	-533	-2276	889	-2266	1474	1478	-1963	682	-629	-781	12	-279	-2472	476	44000%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*													
324(N)	1511	-1770	-726	-204	-2244	-1781	-426	282	601	2121	-1133	1769	-1827	1398	447	-748	-766	-1713	-1779	-1595	44100%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*													
325(N)	1053	-3109	-1756	1735	-3404	-2143	-1074	-3191	-846	-3124	-2254	2158	-2417	-657	-1430	-1361	197	-2727	-3303	-2540	44200%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*													
326(H)	2064	-3071	-1245	-1267	-3262	-2570	4711	-3611	-1060	-3528	-2812	-1479	-2961	1288	-1287	670	-2133	-3161	-3310	-2601	44300%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*													
327(O)	-891	-2294	332	948	-2585	-739	-537	-2316	-131	-138	-1391	-518	1404	2021	-634	-785	619	-561	-2495	-1830	44400%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*													

[1019]

도면

도면1



도면2

5	K9	(1)	MEECKMAKIYYQEDCNLSLLDGKTIIVIG Y GSQGHAA L NAKESGCNVII
	PF5	(1)	-----MKVPYDKDCDLSIIQGGKVAIIG Y GSQGHAA C NLKDSGVDVTV
10	K9	(51)	GLYEGSKSWKRAEEQ G FEVYTAAEAACKADIIMIL I NDEKQATMYKNDIE
	PF5	(45)	GLRKGSA T VAKAEAH G LKVTDVAAVAGADLVML T PDEFQSQLYKNEIE
15	K9	(101)	PNLEAGNMLMEAHGFNIHFGCIVPPKDVDTMIAPKGPHTVRSEYEEGK
	PF5	(95)	PNIKKGATLAFSHGF A IHYNQVVRADLDVIM I APKAPGHTVRSEFVKGG
20	K9	(151)	GVPCLVAVEQ D ATGKALDMALAY A LIGGARAGVLETTFTETETD L FG E
	PF5	(145)	GIPDLIAIYQ D ASGNAKNVALSYAA G VGGGRTGIIETTFKDETETD L FG E
25	K9	(201)	QAVLCGGVCALMQAGFETLVEAGYDPRNAYFECIHEMKLIVDLIYQSGFS
	PF5	(195)	QAVLCGGTVELV K AGFETLVEAGY A PEMAYFEC L HELKLIVDL M YEGG I A
30	K9	(251)	GMRYSISNTAEYGDYITGPKI I TEDTKKAMKILSDIQDGTFAKDFLVDM
	PF5	(245)	NMNYISISNNAEYGEYVTGPEVIN A ESRQAMRNALKRIQDGEYAKMFISEG
35	K9	(301)	SDAGSQVHF K AMRKLASEHPAEVVGEI R SLYSWSEDEK L INN---
	PF5	(295)	--ATGYPSMTAKRRN N AHGIEI I GEQLRSMPWIGANKIVDKAK N

도면3a

5	K9	(1)	MEECKMAKIYYQEDCNLSLLDGKTIIVIGYGSQGHAAHALN A KESG-CNVI
	K1	(1)	----MAATIWYEKDADLSVFEGKKVAVIGYGSQGHAAHALNLRDSG-VDVV
	K2	(1)	-----MAAQIWYENDGDLVLDGKKVAIIIGYGSQGHAAHALNLRDSG-VDVV
	K7	(1)	----MFKMYIEKDTDLNLLRGKKVAIIIGYGSQGHAAHALNLRDSG-VDVV
	S2	(1)	-QEELMKVYYDSADLGLIKSKKIALIIGYGSQGHAAHALNLRDSGVAEVA
10	K25	(1)	-----MVKVYYNEVTBN--LGDKTIALIIGYGSQGHAAHALNLRDTG-HNVV
	K26	(1)	----MAVQMEYEKDVKVPALDGKKIIVIGYGSQGHAAHALNLRDTG-HDVI
	엘. 락티스 (L. Lactis)	(1)	----MAVTMYIEDDVEVSALAGKQIIVIGYGSQGHAAHALNLRDSG-HNVI
15	K9	(50)	IGLYEGSKSWKRAEEQGFEVYTAAEAARKADIIMILI N DEKQATMYKNDI
	K1	(46)	VGLRPTSKSVEYAKEQGLEVQSVADATAEADVVMILLPDQYQAAYYKSEI
	K2	(46)	VGLRPTSKSVEYAKEQGLEVQSVPEAAAAEADIIMILAPDQYQRTIWANDI
	K7	(45)	VGLYNGSKSAKAAEAGLQVATVADAAKADLIMILLPDEKQAKIYNEEI
	S2	(50)	IALRPDSASVKAQDAGFKVLTNAEAAKWADILMILAPDEHQAAIYAE DL
20	K25	(43)	IGIRAG-KSAENAKNDGFTVLVSEAVQQAQNVVMILAPDEIQGELYEKEI
	K26	(46)	IGVRPG-KSFDKAKEDGFDITYVAEATKLADVIMILAPDEIQGELYEAEI
	엘. 락티스 (L. Lactis)	(46)	IGVRHG-KSFDKAKEDGFETFEVGEAVAKADVIMVLAPDEIQQSIYEEDI
25	K9	(100)	EPNLEAGNMLMFAHGFNIHFGCIVPKDQDV T MIAPKPGHTRVSEYEEG
	K1	(96)	EPNLKPGAALAFHGFNIHYGYIKPSEDDHVFVMVAPKGPCHIVRREYTN
	K2	(96)	EPNLKPGAALAFHGFNIHYGYIKPSEDDHVFVMVAPKGPCHIVRREYANG
	K7	(95)	APNLEEGNALVFHGFNIHFGQVPPSYVDVFMVAPKGPCHIVRRTYTEG
	S2	(100)	KDNLKPGSALAFHGFNIHFGLEPRKIDVFMIAPKGPCHIVRSEYVRG
30	K25	(92)	APNLSAGDALAFHGFNIHFEVITPPKEVDVFLVAPKGPCHIVRRTFSEG
	K26	(95)	APNLEAGNAVGFHGFNIHFEFIKVPADVDFMCAKGPCHIVRRTFEEG
	엘. 락티스 (L. Lactis)	(95)	KPNLKAGSALGFHGFNIHFGYIKVPEDVDVFMVAPKGPCHIVRRTYTEG
	K9	(150)	KGVPCLVAVEQDATGKALDMALAYALAIIGGARAGVLETT T ETETDLFG
	K1	(146)	RGVPVVVAVEQDPDGKWTWFLCLAYAKALGALRAGAIKTT F TEETDLFG
	K2	(146)	RGVPVVVAVEQDPRGDWDITLAYAKALGALRAGAIKTT F KEEETDLFG
	K7	(145)	AGVPCLIAVHQDATGKAKQYALAYANGIGGARAGVLETT F KDEETDLFG

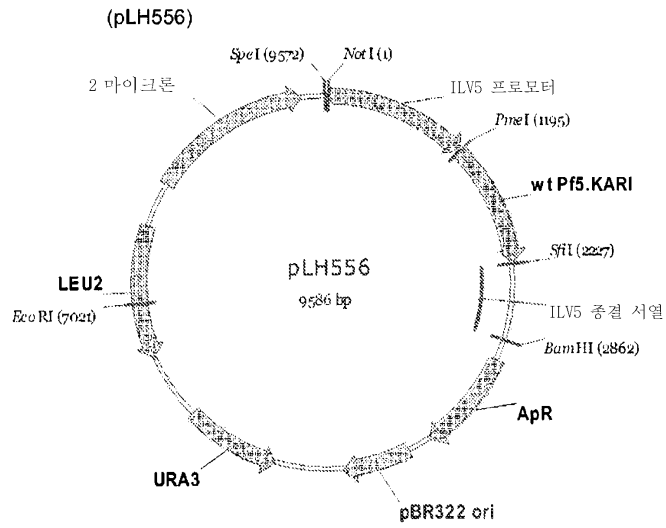
도면3b

5	S2	(150)	GGVPCLVAVDQDASGNAHDIALAYASGIGGGRSGVIETTFREEVETDLFG
	K25	(142)	FAVPALFAVYQDATGNAQETALSAYAKGIGATRVGLVLETT F KEEETDLFG
	K26	(145)	FGVPALYAVYQDATGNAKDIAWDCKGIGARVGLLETTYKEET E EEDLFG
	엘. 락티스 (L. Lactis)	(145)	FGTPALFVSHQNASGHAREIAMDWAKGIGCARVGIETTFKEET E EEDLFG
10	K9	(200)	EQAVLCGGVCAQMAGFETLVEAGYD P RNAYFECIHEMKLIVDLIY Q SGF
	K1	(196)	EQDVLMMGINHLCDMGFDVLTAGYQPEIAYFEVFEHMKMLVDLANEGGL
	K2	(196)	EQNVLMMGINKLVEMGFVLTAGYQPEIAYFEVFEHMKMLVDLMNEGGL
	K7	(195)	EQAVLCGGVSELKAGFETLVEAGYAPENAYFECHEMKLIVDLIYCGGL
	S2	(200)	EQAVLCGGLTALITAGFETLVEAGYAPENAYFECHEMKLIVDLIY E AGI
15	K25	(192)	EQAVLCGGLTSMIETGFETLVEAGYQPELAYFEVCHMKLIVDLIYEGGF
	K26	(195)	EQAVLCGGLTALITGFEVLTAGYAPELAYFEVLHEMKLIVDLIYEGGF
	엘. 락티스 (L. Lactis)	(195)	EQAVLCGGLTALVEAGFETLVEAGYAGELAYFEVLHEMKLIVDLIYEGGF
	K9	(250)	SGMRYISINTAEYGDYITGPKIITEDTKKAMKKIISDIQDGTFAKDFIVD
	K1	(246)	NKARWSCSDTAQYGDYTS--TVITEETKKRMQYQLKRIQDGSFAKEFMDD
20	K2	(246)	NKARWSCSDTAQYGDYTN--TVINEDCRKRMQYVHLGRIQDGSFAKEFI DD
	K7	(245)	SMMRYISISDTAEYGDYQIGRRIITDETKKEMKKVLTEIQDGTFAKNWLE
	S2	(250)	ANMRYISINTAEYGDIVSGPRVINEESKKAMKAILDDIQSGRFVSKFVLD
	K25	(242)	AKMRQISINTAEYGDYVSGSRIITEQTKANMKNVLDIQNGTFAKGFIAD
	K26	(245)	KKMRQISINTAEYGDYVSGPRVITEQVKENMKAVLADIQNGKFANDFVND
25	엘. 락티스 (L. Lactis)	(245)	TKMRQISINTAEYGDYVTGPRIITDEVKKNMKLVLDIQSGKFAQDFVDD
	K9	(300)	MSDAGSQVHFKAMRKLASEHFAEVVGEEIRSLYSWSDDEKLIINN-----
	K1	(294)	QAAG--APKFYKLQEEYSHPHLETVGPKLRAMFSWNNQVDADADMAESFN
	K2	(294)	QDAG--APKFELQEEYGNVRIETVGPKLRAMFSWNNQDDADADMAT-F T
	K7	(295)	NQTN--RPGFNARRRMEAEHPIEKVGEKLRGMMSWIDTAKVD-----
30	S2	(300)	NRAG--QPELKAARKRMAAHPTEQVGARLRKMMFWIASNKLVDKARN-KL
	K25	(292)	NQAG--FPEFKMREQNGNHPIEKVGEELRKMMFVFTKD-----
	K26	(295)	YKAG--RPKLTAYREAAANLEIEKVGAELRKAMPFVGQNDDDAFKIYN--
	엘. 락티스 (L. Lactis)	(295)	FKAG--RPKLIAYREAAKNLEIEKIGAEIRQAMPFTQSGDDDAFKIYQ--

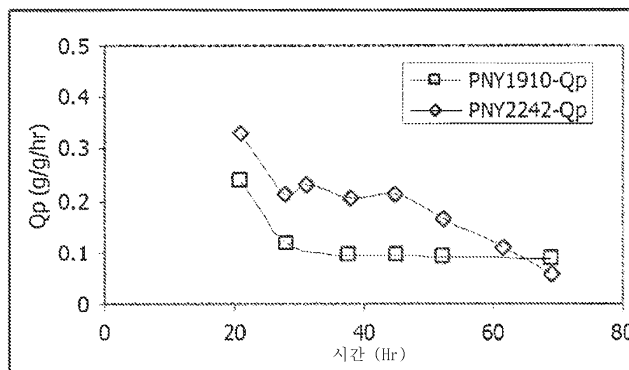
도면3c

5	K9	(344)	-----
	K1	(342)	GKIARTQVQ
	K2	(341)	GKIARSQVQ
	K7	(335)	-----
	S2	(347)	GCF-----
10	K25	(329)	-----
	K26	(341)	-----
	엘. 락티스 (L. Lactis)	(341)	-----

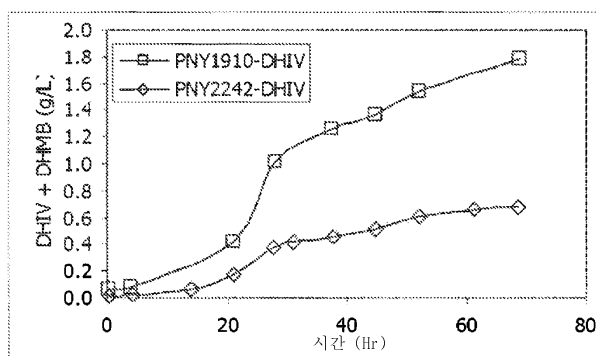
도면4



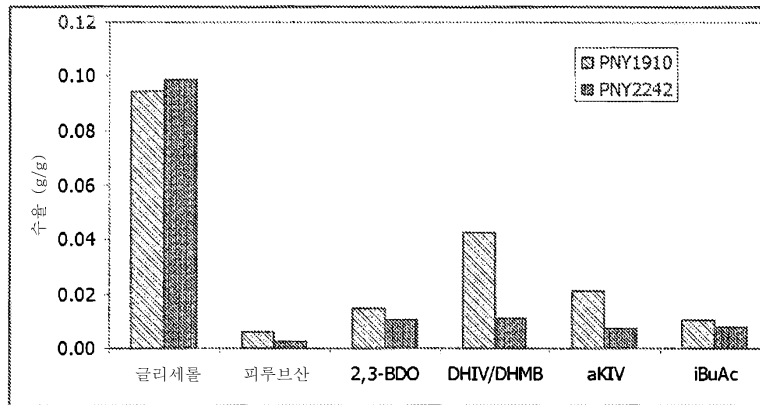
도면5



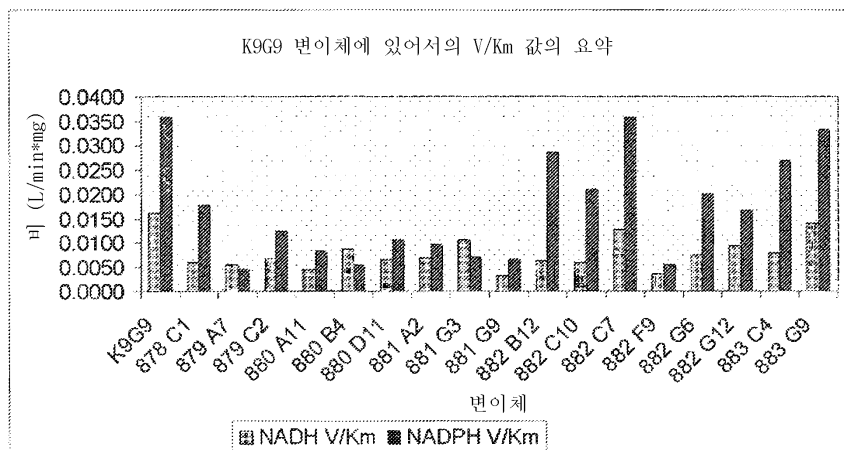
도면6



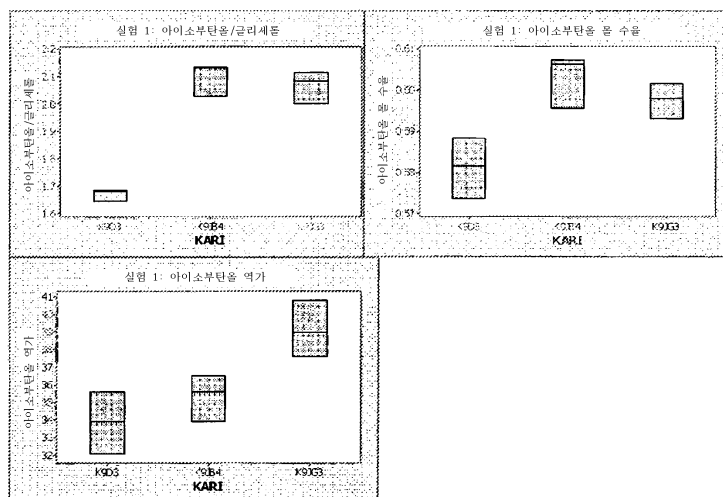
도면7



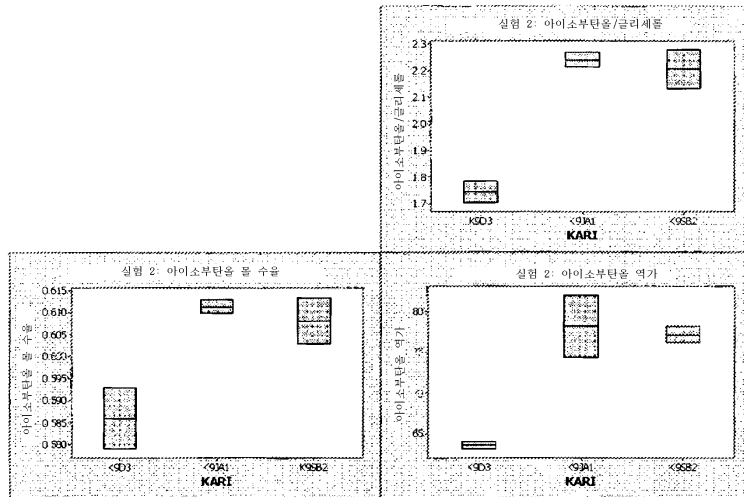
도면8



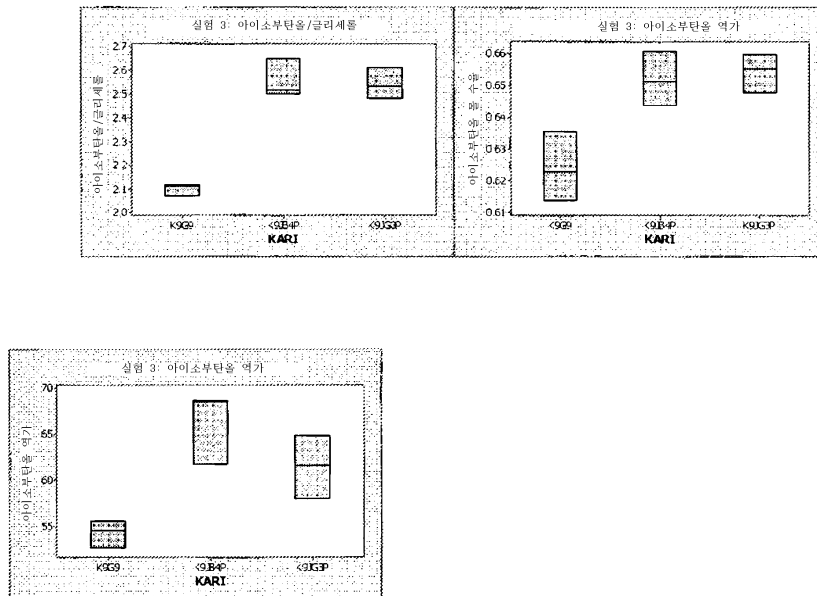
도면9a



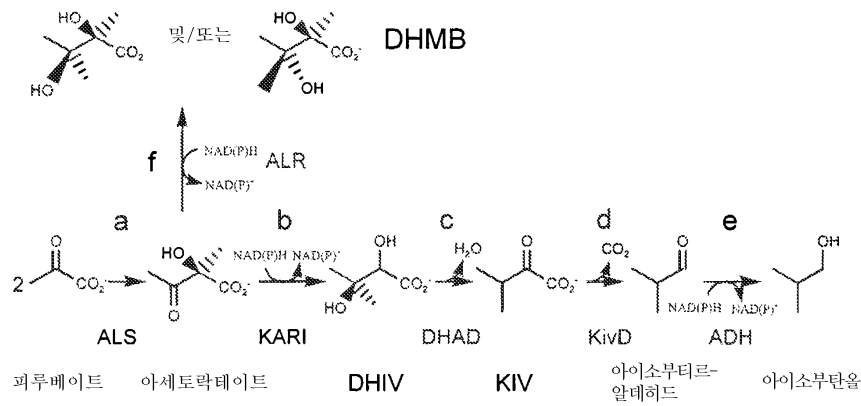
도면9b



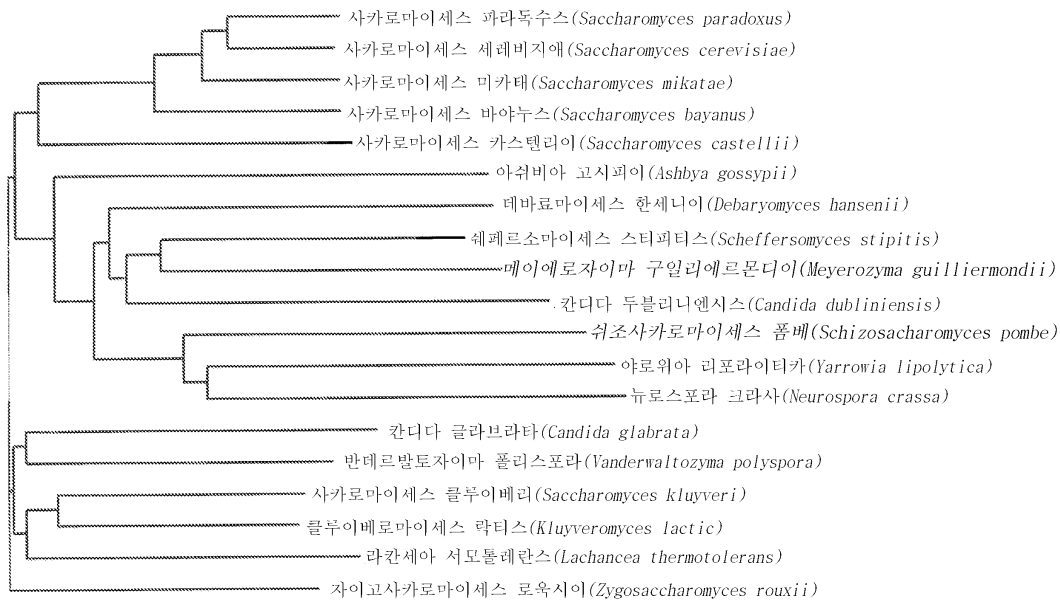
도면9c



도면10



도면11



도면12a

	1				50
AABY01000127					
AACH01000019					
AACA01000631					
AACF01000116					
AE016819					
CR382139					
XM_001387479					
XM_001482184	ATGTGCCTCT	TACCAGCCGG	TAGCACTGTA	TTATGTCATC	ACCCAGTAGT
XM_002419771					
CR380959					
XM_001645671					
SAKLOH04730					
XM_451902					
XM_002553230					
XM_002494574					
YMR226C					
NM_001018495					
XM_501554					
XM_957621					
	51				100
AABY01000127			AT	GTCCCAAGGT	AGAAAAGCTG
AACH01000019			AT	GTCTCAAGGT	AGAAAAGCTG
AACA01000631			AT	GTCCCAAGGT	AGAAAAGCTG
AACF01000116			AT	GTCTCAAGGT	CCTAAAGCTG
AE016819			AT	GTCCCTAGGA	AGAAAAGCAG
CR382139			AT	GTCTACGGA	TCTAAAGCTG
XM_001387479			AT	GTCTTTGGA	AAAAAGCTG
XM_001482184	GAGTGTGGAG	ATTAAATCCT	CAATCTTCAT	GTCTTTGCTG	GCCAAAGCCG
XM_002419771			AT	GTCTTTGCTG	AGAAAAGCTG
CR380959			AT	GTCTCAAGGA	AGAAAAGCTG
XM_001645671			AT	GTCTCAAGGT	AGAAAAGCTT
SAKLOH04730			AT	GTCTCAAGGT	AGAAAAGCTG
XM_451902			AT	GTCTCAAGGT	AGAAAAGCTG
XM_002553230			AT	GTCTCAAGGA	AGAAAAGCAG
XM_002494574			AT	GTCTCAAGGT	GTCAAAGCTG
YMR226C			AT	GTCCCAAGGT	AGAAAAGCTG
NM_001018495					
XM_501554			AT	GTCTTTGCTG	GATAAAGCTG
XM_957621			AT	GTCTCTCAGC	AGTAG
	101				150
AABY01000127	CAGAAA.GAT	TGGCTAACAA	GACCGTGCTC	ATTACGGGTG	CCTCTGCTGG
AACH01000019	CAGAAA.GAT	TGGCTGGCAA	AACCGTTCCTC	ATCAGGGGTG	CCTCTGCTGG
AACA01000631	CAGAAA.GAT	TGGCCAACAA	GACGGTGCTC	ATTACAGGCG	CTTCTGCTGG
AACF01000116	CCGAAA.GAT	TGAATGAGAA	GATTGTGTTT	ATCACTGGTG	CTTCAGCTGG
AE016819	CTGAAA.GAT	TAGCCAACAA	AATTGTGCTT	GTGACTGGTG	CCTCTGCGGG
CR382139	CTGAAC.GTG	TTGCCAATAA	GATTGTCTTA	ATCACTGGTG	CTTCATCTGG
XM_001387479	CTGAAA.GAC	TTGCCAACAA	AATCATCTCT	ATCAGCGGGG	CTTCGTCTGG
XM_001482184	CTGAAC.GCC	TTGCCAACAA	GATCATATTG	ATCACTGGGG	CATCGTCTGG
XM_002419771	CTGAAA.GAT	TAGCCAATAG	ATCCATTCTT	ATCACTGGTG	CTTCATCTGG
CR380959	CTGAGA.GGT	TACAGGGGAA	GATTGCTTTT	ATTACGGGTG	CCTCTGCGGG
XM_001645671	CAGAAA.GGT	TGGCTGGTAA	AATCTGTATTA	ATTACAGGTG	CTTCATCAGG
SAKLOH04730	CAGAAA.GAC	TAGCAAACAA	GACCGTTTTT	ATAACTGGCG	CCTCTGCCGG

도면12b

XM_451902	CTGAAA.GAT	TGCAAAACAA	GACAATTTTC	ATTACCGGTG	CTTCTGCAGG
XM_002553230	CTGAAA.GAC	TGGCAGGAAA	GACTGTCTTC	ATCACAGGCG	CATCAGCCGG
XM_002494574	CTGAAA.GAC	TAGCTGGTAA	GACTGTATTC	ATTACAGGTG	CTTCTGCAGG
YMR226C	CAGAAA.GAT	TGGCTAAGAA	GACTGTCTTC	ATTACAGGTG	CATCTGCTGG
NM_001018495	ATGAGCCGTT	TGGATGGAAA	AACGATTTTA	ATCACTGGTG	CCTCTTCTGG
XM_501554	CTGCTC.GAC	TTGCGGGCAA	GACCGTCTTT	GTTACCGGCG	CCTCGTCCGG
XM_957621	CCAAGC.GCC	TCGCGGGCAA	GACCATTGTC	ATAACCGGAG	CGTCCTCAGG

	151		200
AABY01000127	TATTGGTAAG	GCCACCGCAT	TAGAGTATTT G...GAGGCA TCCAATGGTG
AACH01000019	TATTGGCAAA	GCCACTGCAT	TAGAGTATTT G...GAGGCA TCCAATGGCG
AACA01000631	TATTGGTAAG	GCCACCGCAT	TGGAGTATTT G...GAAGCA TCCAATGGAA
AACF01000116	TATTGGGGCAA	GCCACCGCTT	TGGAATACAT G...GATGCG TCGAACGGTA
AE016819	CATTGGCCGT	GCTACAGCCA	TTAACATATGC A...GACGCG ACGGACGGGG
CR382139	AATTGGTGAA	GCAACTGCCA	AAGAAATTGC A...TCAGCC GCTAATGGCA
XM_001387479	TATTGGTGAA	GCTACAGCTA	GAGAGTTTGC A...TCTGCT GCCAATGGGA
XM_001482184	TATAGGCGAG	GCTACCGCCA	GAGAATTTCGC T...GCTGCT GCCAATGGAA
XM_002419771	GATTGGTGAA	GCATGTGCTA	AAGTTTTCGC T...GAAGCA TCTAATGGTC
CR380959	CATCGGTAAA	GCTACAGCCA	TTGAGTATTT G...GATGCT TCCAATGGTA
XM_001645671	GATTGGGAAA	GCCACTGCAT	TAGAATATCT A...GATGCC TCCAATGGTG
SAKL0H04730	CATTGGCCAA	GCTACTGCTT	TGGAATACTG T...GATGCT TCTAACGGTA
XM_451902	TATTGGTCAA	GCCACAGCAT	TGGAATATCT A...GATGCT GCTAACGGTA
XM_002553230	TATCGGTGAG	GCCACTGCGC	AAGAATACCT G...GAAGCA TCCGAAGGCA
XM_002494574	TATCGGTCAA	GCAACTGCCA	AGGAATATTT G...GATGCA TCCAATGGTC
YMR226C	TATTGGTAAG	GCGACCGCAT	TAGAGTACTT G...GAGGCA TCCAATGGTG
NM_001018495	AATTGGAAAA	AGCACTGCTT	TTGAAATTGC C...AAAGTT GCCAAAG...
XM_501554	CATTGGCCAG	GCCACTGCTC	TCGCTCTAGC C...GAAGCT GCCAAGGGCG
XM_957621	AATCGGGCGT	AGCACAGCCT	TCGAGTTTGC GCGGACGCGC CCCAAGCCACG

	201		250
AABY01000127	ATATGAAACT	GGTCTTAGCT	GCTAGAAGAT TAGAAAAGCT CGAGGAATTA
AACH01000019	ATATGAAATT	AATCTTAGCC	GCTAGAAGAT TAGAAAAGCT CGAGGAATTG
AACA01000631	ACATGAAACT	GATCTTGGCT	GCGAGGAGAT TGGAGAAGCT AGAGGAGCTC
AACF01000116	CTGTGAAATT	GGTTCTAGTT	GCCAGAAGAT TGGAGAAATT ACAACAATTG
AE016819	CAATCAAGTT	GATTTTGGTG	GCAAGACGCG CAGAAAAGCT CACCAAGCTTG
CR382139	ATTTAAAAAT	AGTGTGTGT	GCTAGACGAA AAGAAAAGTT GGATAATTTA
XM_001387479	ATATCAGATT	GATTTTGACA	GCCAGAAGAA AAGAAAAGTT GGCTCAATTG
XM_001482184	AAATTTCTGT	GATTTTGACC	GCTCGGAGAG AAGACAAACT CAAGTCTCTC
XM_002419771	AAGTTAAATT	AGTTTTAGGA	GCAAGAAGAA AAGAACGATT AGTTAAATTA
CR380959	GTGTGAAGCT	AGTTCTTGGT	GCACGTAGAA TGGAGAAATT GGAGGAGTTG
XM_001645671	ATATGAAGTT	AATTTTAGTT	GCAAGAAGAT TAGAAAAATT GCAAGAGTTG
SAKL0H04730	AAATTAAGTT	GGTGTAAAGT	GCCAGAAGGC TGGAAAAATT GCAAGAGTTA
XM_451902	ATGTCAAATT	GATCTTAGCA	GCAAGAAGGT TGGCTAAGTT GGAAGAATTG
XM_002553230	AAATCAAGTT	GATCCTTGCA	GCAAGAAGAC TCGACAAGCT GGAGGAAATC
XM_002494574	AAATTAAGTT	GATCTTGGCT	GCAAGAAGAT TAGAGAAATT ACACGAGTTT
YMR226C	ATATGAAACT	GATCTTGGCT	GCTAGAAGAT TAGAAAAGCT CGAGGAATTG
NM_001018495	...TAAACT	TATTTTGGCT	GCTCGCAGAT TTTCTACCGT TGAGGAAATC
XM_501554	ACCTCAAGTT	TGTGCTTGCT	GCCCCACGAA CCGACCGTCT GGACGAGCTC
XM_957621	GCCTCAAGCT	GATCTTGACG	GCCCCGCCGCG TCGATGCCCT GGAGCAGATT

	251		300
AABY01000127	AAGAAAACTA	TTGATCAGG.	AGTTTCCA.. AACGCCAAG TTCATGTGGC
AACH01000019	AAGAAGACTA	TCGATGAAG.	AGTTTCCA.. AACGCAAGG TCCATGTGAC
AACA01000631	AAGAAGACCA	TCGACGAGG.	AGTTTCCC.. AATGCAAGG TTCACGTTGG
AACF01000116	AAGGAAGTCA	TTGAGGCAA.	AATACCCCT.. AAGAGTAAAG TCTATATTGG
AE016819	AAACAGGAGA	TCGA.AAGCA	AGTATCCC.. AACGCCAAG TCCATGTCGG
CR382139	TCT.AAAGAA	TTGACTGACA	AATATTCA.. TCCATCAAG TTCATGTTGC

도면12c

XM_001387479	TCA.GACTCA	TTGACCAAGG	AATTTCCA..	ACTATCAAAA	TCCATTCTGC
XM_001482184	TCG.CAACAA	TTGAGCCTCA	TTTACCCG..	CAAATTAAAA	TCCATTCTGC
XM_002419771	TCT.GATACT	TTAATTAAAC	AAATACCT..	AATATTAAAA	TTTCATCATGA
CR380959	AAG.AAGGAA	TTGCTGGCTC	AATATCCT..	GATGCAAAGA	TTTCATATAGG
XM_001645671	AAGGAAACAA	TTTGTAAAG.	AATATCCA..	GAATCTAAGG	TTTCATGTTGA
SAKL0H04730	AAGGACAAAA	TCACCAAGG.	AGTATCCT..	GAAGCCAAGG	TTTATATTGG
XM_451902	AAGGAAAAAA	TCAATGCTG.	AATACCCA..	CAAGCTAAAG	TATATATCGG
XM_002553230	AAAGCCAAGG	TTTCTAAAG.	ACTTCCTC..	GAAGCACAGG	TGCATATCGG
XM_002494574	AAAGAACAAA	CTAC.AAAGA	GTTACCCA..	AGCGCTCAAG	TCCACATTGG
YMR226C	AAGAAGACCA	TTGATCAAG.	AGTTTCCA..	AACGCAAAAG	TTTCATGTGGC
NM_001018495	GCAAAGGAGT	TAGA.ATCGA	AATAT....	GAAGTATCGG	TTCTTCCTCT
XM_501554	AAGAAGAAGC	TGGAGACCG.	ACTACAAG..	GGTATCCAGG	TGCTGCCTTT
XM_957621	GCCAAGGAGA	TCCGCCAAGA	GTTTGGCGAG	GGCGTGCAAG	TGCTGCCTGT
301					
AABY01000127	CCAACCTGGAT	ATCACTCAAG	CAGAAAAGAT	CAAGCCCTTT	ATTGAGAATT
AACH01000019	CAAACCTGGAC	ATCACACAGA	CAGAAAAGAT	CAAGCCCTTT	ATTGAAAACCT
AACA01000631	CCAACCTGGAT	ATCACACAGG	CCGAGAAGAT	CAAGCCCTTC	ATTGAAAACCT
AACF01000116	GAAGTTGGAT	GTGACAGAGC	TTGAGACCAT	TCAACCATTC	TTGGATAATC
AE016819	ACAATTGGAT	GTGACCCAAC	TGGACCAGAT	CCGCCCATTT	TTGGAGGGAC
CR382139	TCAACTAGAT	GTATCTAAGC	TCGAGACTAT	CAAGCCATTT	ATCAATGATT
XM_001387479	CAAATTGGAT	GTGACCGAAC	ATGATGGCAT	CAAGCCCTTC	ATTCTCTGGT
XM_001482184	TCGTCCTGAT	GTCTCTGAGT	TTTCGTCAC	TAAGCCGTTT	ATTACTGGGT
XM_002419771	TTTTTTGGAT	GTTACTATTA	AAGATTCAAT	TTCAAATTC	ATTGCTGGAA
CR380959	TAAACTGGAT	GTTACAGACT	TTGAAAACGT	CAAGCAGTTT	TTGGCTGACT
XM_001645671	AGAATTAGAT	ATTTCTGATA	TTAATAGAA	CCCAGAATTT	ATTGCAAAAT
SAKL0H04730	TGTGCTTGAT	GTGACCGAAA	CGGAAAAAAT	CAAACCATTC	TTGGATGGTT
XM_451902	TCAATTGGAC	GTCACCTGAA	CTGAGAAGAT	TCAACCTTTC	ATTGATAACT
XM_002553230	CCAGCTAGAT	GTGACTCAGA	CGGACAAAAT	CCAGCCTTTT	GTGACAAAT
XM_002494574	TAAATTGGAC	GTCACCTGCA	TTGACACCAT	AAAACCATTT	TTGGATAAAT
YMR226C	CCAGCTGGAT	ATCACTCAAG	CAGAAAAAAT	CAAGCCCTTC	ATTGAAAACCT
NM_001018495	TAAATTGGAT	GTTTCTGATT	TGAAGTCTAT	TCCTGGGGTA	ATTGAGTCAT
XM_501554	CAAGCTGGAC	GTGTCCAAGG	TCGAGGAGAC	CGAGAACATT	GTGTCCAAGC
XM_957621	CAAGCTGGAC	GTCAGCCAGC	CGGAGGAGGT	CAGGGGGTTT	GTTGGGAACCT
351					
AABY01000127	TGCCAAAGGA	GTTCAAAGAC	ATTGACATTT	TGGTGAACAA	CGCTGGGAAG
AACH01000019	TGCCAGAGGA	GTTCAAAGAC	ATTGATATTC	TGGTGAACAA	CGCTGGTAAG
AACA01000631	TGCCCGAGGC	ATTCAAGGAT	ATTGACATCC	TGATAAACAA	TGCCGGCAAA
AACF01000116	TTCTCTGAGGA	ATTTAAGGAT	ATTGATATCT	TGATTAATAA	TGCCGGGAAG
AE016819	TACCTGAGGA	GTTCCGAGAC	ATTGATATTT	TAATTAACAA	CGCAGGTAAG
CR382139	TACCGAAGA	ATTTCTCTGAC	GTGGATGTAT	TAGTCAACAA	TGCAGGCTTG
XM_001387479	TACCCAAGGA	TTTCGCCGAC	ATCGATGTGT	TGATCAACAA	TGCTGGAAAA
XM_001482184	TGCCAAAGGA	TTTTGCTAGC	ATCGACGTTT	TGGTGAATAA	TGCCGGGAAA
XM_002419771	TTCCCTCATGA	ATTTG..AAC	CT.GATGTAT	TAATTAATAA	TAGTGGTAAA
CR380959	TGCCAGAAGA	GTTCAAGGAC	ATCGACATCC	TGATCAATAA	CGCTGGTAAA
XM_001645671	TACCTGAAGA	ATTCAAAGAT	ATTGATATAT	TGATTAACAA	TGCAGGTAAA
SAKL0H04730	TACCAGAAGA	ATTTAAAGAT	ATTGACATCT	TGATCAATAA	TGCAGGCAAA
XM_451902	TGCCGAAGA	ATTCAAGGAT	ATCGATATTT	TGATTAACAA	TGCCGGTAAA
XM_002553230	TGCCCAAGA	GTTCAAAGAC	ATCGACATCC	TGATCAACAA	CGCGGGCAAG
XM_002494574	TACCAAGGA	ATTTCAAGAT	ATCGATATTT	TGATCAACAA	TGCCGGTAAG
YMR226C	TGCCACAAGA	GTTCAAGGAT	ATTGACATTC	TGGTGAACAA	TGCCGGAAG
NM_001018495	TGCCAAAGGA	ATTTGCTGAT	ATCGATGCT	TGATTAATAA	TGCTGGACTT
XM_501554	TGCCCAAGGA	GTTTTCCGAG	GTGGACGTGC	TTATCAACAA	CGCCGGCATG
XM_957621	TGCCCGAGGA	GTGGAGGGAT	ATCCATGTGC	TGGTCAATAA	TGCTGGTCTC
401					
450					

도면12d

AABY01000127	GGCCTTGGTA	CCGACCGTGT	GGGGGAGATT	GCAACACAAG	ATATCCAGGA
AACH01000019	GCTCTTGGTA	CGGACCGTGT	TGGGGAGATT	GATACACAGG	ACGTCCAGGA
AACA01000631	GCCCTGGGCT	CCGAACGTGT	CGGGGAAATT	GCCACACAGG	ACATCCAGGA
AACF01000116	GCATTAGGTT	CCGATCGTGT	AGGTGATATT	GATATAAAAG	ATGTGAAGGG
AE016819	CCCCTCGGCA	CTGAGAGGGT	GGGGGAAATC	TCGATGGACC	ATATCCAGGA
CR382139	GCTTTGGGCC	GTGATGAAAT	TGGAACCAAT	GACACAGATG	ATATGTTATC
XM_001387479	GCTCTTGGAA	AAGCATCTGT	TGGTGAAATC	AGTGACAGTG	ATATCCAAGG
XM_001482184	GCATTGGGAA	GAGCCAATGT	TGGTGAAATT	TCCCAAGAGG	AAATCAATGG
XM_002419771	GCCTTGGGGA	AAGAAGAAGT	TGGAGAATTG	AAAGATGAAG	ATATTACGGA
CR380959	GCCTTGGGGT	CTGACAAAAGT	TGGAGACATT	GACCTGAGG	ATATCGCAGG
XM_001645671	GCATTAGGAA	GTGATACTAT	TGGTAATATC	GAGAATGAGG	ATATTAAAGG
SAKLOH04730	CGCTTAGGCT	CTGATCCTGT	TGGTACCATC	AAAACGAAG	ATATTGAAGG
XM_451902	GCCTTGGGAT	CTGATGTTGT	CGGTACCATC	AGTAGCGAGG	ACATCAAAGG
XM_002553230	GCGCTCGGAT	CCGACCCCGT	GGGCACAATC	GACCCCAATG	ATATTCAAGG
XM_002494574	GCATTAGGTA	CTGATAAAGT	TGGTGATATT	GCAGATGAAG	ACGTGGAAGG
YMR226C	GCTCTTGGCA	GTGACCGTGT	GGGCCAGATC	GCAACGGAGG	ATATCCAGGA
NM_001018495	GCTCTAGGTA	CCGATAAAGT	CATTGATCTT	AATATTGATG	ACGCCGTTAC
XM_501554	GTCCACGGCA	CCGAAAAGGT	TGGCTCCATC	AACCAGAACG	ACATTGAGAT
XM_957621	GTAAAAGG..	A..GC	TCCCTCCATC	GCCGAAGAAG	ACATCAACGT

	451		500
AABY01000127	TGTGTTTGAC	ACCAACGTCA	CAGCTTTAAT
AACH01000019	CGTGTTTCGAC	ACCAACGTCT	CGGCTTTGAT
AACA01000631	CGTGTTTCGAC	ACCAACGTCA	CGGCGTTGAT
AACF01000116	AATGATGGAT	ACCAATGTCT	TGGGGTTGAT
AE016819	GGTTTTCAAC	ACTAATGTTA	TCGGCTTGGT
CR382139	GATGTTTCAA	ACTAATGTTT	TAGGGTTAAT
XM_001387479	CATGATTCGAA	ACGAATGTCT	TGGGACTCAT
XM_001482184	CATGTTCCAT	ACCAATGTTT	TTGGGTTGAT
XM_002419771	AATGTTTGAT	ACTAATGTCA	TTGGAGTCAT
CR380959	AATGGTTAAC	ACCAACGTCC	TTGCATTGAT
XM_001645671	TATGTTTGAG	ACTAACGTTT	TTGGATTAAAT
SAKLOH04730	AATGATCAAC	ACCAATGTCT	TAGCTCTTAT
XM_451902	TATGATAGAT	ACTAACGTTG	TTGCCCTTAT
XM_002553230	CATGATCCAG	ACTAACGTTA	TCGGGCTTAT
XM_002494574	TATGTTTCGAC	ACCAATGTCT	TGGGGTTAAT
YMR226C	CGTGTTTGAC	ACCAACGTCA	CGGCTTTAAT
NM_001018495	CATGATTACT	ACCAATGTTT	TTGGTATGAT
XM_501554	CATGTTCCAC	ACCAACGTGC	TCGGACTCAT
XM_957621	CATGTTTGCC	ACCAACGTCA	CGGGCTTGAT

	501		550
AABY01000127	TGCCCATTTT	TCAAGCCAAG	AACTCAG...
AACH01000019	TGCCCATCTT	CCAAGCTAAG	AACTCAG...
AACA01000631	TGCCAATTTT	CCAAGCCAAG	AACTCAG...
AACF01000116	TGCACATTTT	CCAAAAGAAG	AACTCCG...
AE016819	TACCTATTAT	GAAAGCCAAG	AATTCCG...
CR382139	TGCCAATCAT	GAAAAAGAAG	AACAGCG...
XM_001387479	TTCCCATTTT	TAAGGCTAAA	AATTCTG...
XM_001482184	TACCCATCTT	CAAAAAGAAA	AATGCTG...
XM_002419771	TACCTTTACT	TAAAAAAGAA	CCTTATG...
CR380959	TGCCATTATT	CAAGAAGAAG	AACAGTG...
XM_001645671	TTCCAATATT	CAAGGCTAAA	AATGGTG...
SAKLOH04730	TGCCAATCTT	CAAGCCAAG	AATTTCG...
XM_451902	TGCCTATTTT	CAAGCAAAAG	AATTCCG...
XM_002553230	TGCCCATCTT	CAAGGCCAAA	AACTCTG...
XM_002494574	TACCTATCTT	CAAAAAGAAA	AATTCTG...

도면12e

YMR226C	TGCCCATATT	CCAAGCCAAG	AATTCAG...	G	AGATATTGTA
NM_001018495	TTCTATATT	CTACAGCAA	AACAAGG...	G	TGATATTTTG
XM_501554	TCGGCGAGAT	GCGAAAGCGA	AACAAGG...	G	CGACATTGTC
XM_957621	TGCCCATCTT	CAAGGCCCGC	GGCTCCGAGG	GTGGTTCGGG	AGACATTGTG

	551		600
AABY01000127	AACCTGGG.T	TCGGTGGCTG	GCAGGGATGC ATACCCAACG GGTTCATCT
AACH01000019	AACCTGGG.C	TCGGTAGCTG	GCAGAGATGC ATACCCAACG GGTTCATCT
AACA01000631	AACCTGGGGC	TCGGTGGCCG	GCAGAGACGC ATACCCACAC GGTTCATCT
AACF01000116	AACCTAGG.T	TCAGTTGCTG	GAAGAGATGC ATACCCAACA GGTTCATTT
AE016819	AATGTTGG.G	TCGATTGCCG	GCCGCGAAGC CTACCTGGT GGTCTATTT
CR382139	AATATAGG.T	TCAATTGCTG	GAAGAGACTC TTACCTGGA GGTGGAATTT
XM_001387479	AACATCGG.T	TCGATTGCTG	GAAGAGACCC TTACCTGGT GGATCGATCT
XM_001482184	AACATTGG.C	TCAGTGGCCG	GTAGAGAACC TTACCTGGA GGTGCAATAT
XM_002419771	TTCATTGG.A	AGTATTGCTG	GACGTGTTCC TTATAAAAT GGAGGTGGTT
CR380959	AACCTGGG.A	TCGATTGCTG	GTAGAGACGC ATACCCAACG GGTGCTATAT
XM_001645671	AATTTAGG.G	TCAATTGCAG	GCATAGAAGC TTACCCAACA GGATCTATAT
SAKLOH04730	AACCTGGG.G	TCTGTGCTG	GTAGAGATGC TTATCCAACC GGTGCAATCT
XM_451902	AACCTAGG.T	TCTGTTGCCG	GTAGAGATGC ATATCCAACG GGTCTATCT
XM_002553230	AACCTGGG.T	TCTGTGCTG	GCAGAGAAGC TTACCTTACA GGATCTATTT
XM_002494574	AACATTGG.T	TCGGTTGCTG	GTAGAGAGGC ATACCCAGGT GGTTCATTT
YMR226C	AATTTGGG.T	TCAATCGCTG	GCAGAGACGC ATACCCAACA GGTCTATCT
NM_001018495	AACGTTGG.C	AGTATTGCCG	GCAGAGAATC ATACGTAGGC GGCTCCGTTT
XM_501554	AACATTGG.C	TCCATCGCCG	GACGAGAGCC CTACGTTGGA GGAGGAATCT
XM_957621	AACATCGG.T	TCTATTGCCG	GGAGAGAACC ATACGCGGGA GGCTCCATCT

	601		650
AABY01000127	ATTGTGCCTC	CAAGTTTGCC	GTGGGGGCGT TCACTGATAG TTTAAGAAAG
AACH01000019	ATTGTGCATC	TAAGTTTGCC	GTGGGGGCTT TCACTGAGAG TTTGAGAAAG
AACA01000631	ACTGTGCTTC	CAAGTTTGCC	GTGGGTGCGT TCACTGACAG TTTGAGAAAG
AACF01000116	ACTGTGCTTC	TAAATTTGCC	GTGAGGGCCT TTAAGTAAAG TTTGAGAAAG
AE016819	ACTGTGCCAC	GAAACATGCG	GTCAAGGCTT TCACCAAGGC CATGCGGAAG
CR382139	ACTGTCCAAC	TAAGGCAAGT	GTCAAGTCGT TTTGCAAGT TTTAAGAAAG
XM_001387479	ACTGTGCCTC	CAAGGCTGCT	GTTAAGTTCT TCTCGCATTC TTTGAGAAAG
XM_001482184	ACTGTGCTTC	AAAGGCAGCA	GTTAACTACT TTTCTCATTC TTTGAGAAAG
XM_002419771	ATTGTGCATC	TAAAGCTGCT	GTTGCTAGTT TCACCGATAC ATTTAGAAAA
CR380959	ACTGTGCAAC	AAAACATGCT	GTCAGGGCAT TCACACAATC CTTAAGGAAAG
XM_001645671	ATTGTGCAAC	TAAATTTGCA	GTTAAAGCAT TCACTGAAAG TTTAAGAAAG
SAKLOH04730	ACTGTGCTAG	CAAACATGCA	GTCAGAGCCT TCACTCAAAG TTTGAGGAAG
XM_451902	ATTGTGCTTC	GAAGCATGCT	GTCAGAGCGT TCACTCAGTC TTTGAGAAAA
XM_002553230	ACTGCGCTAC	GAAGCACGCG	GTGCGTGCTT TCACCCAGAG CCTGCGCAAG
XM_002494574	ACTGTGCTAC	TAAACACGCT	GTTAAGGCAT TCACTGAAAG TTTGCGTAAG
YMR226C	ATTGTGCCTC	TAAGTTTGCC	GTGGGGGCGT TCACTGATAG TTTGAGAAAG
NM_001018495	ACTGCTCTAC	CAAGTCTGCC	CTTGCTCAAT TCACTTCCGC TTTGCGTAAG
XM_501554	ACTGTGCCAC	CAAGGCCGCC	GTGCGATCTT TCACTGAGAC TCTCCGAAAA
XM_957621	ACTGCGCCAC	CAAGGCTGCT	GTGCGCAGCT TCACCGACGC GCTGCGCAAG

	651		700
AABY01000127	GAGCTTATCA	ACACCAAGAT	CAGAGTCATC CTAATCGCAC CAGGGCTAGT
AACH01000019	GAACCTATAA	ACACTAAGAT	TAGAGTCATT CTAATTGCAC CAGGGTTAGT
AACA01000631	GAACCTGATCA	ACACGAAGAT	CAGAGTTATC TTGATCGCGC CCGGGCTGGT
AACF01000116	GAATTAATTA	ATACCAAGAT	TAGGGTGATA TTGATAGCCC CCGGTATCGT
AE016819	GAGCTCATTA	GCACCAAGAT	CCGGGTCTTC GAAATTGCGC CCGGCTCTGT
CR382139	GAATTGATTA	GCACCAAGAT	TAGAGTTCTT GAGGTTGACC CTGGTAATGT
XM_001387479	GAACCTATTA	ACACCAAGAT	CAGAGTTTGT GAAGTTGATC CAGGTGCTGT
XM_001482184	GAAACTATCA	ATCCCAAAAT	CAGGGTCATG GAGGTGGATC CTGGGCGAGT
XM_002419771	GAAACTATTA	ATACTGGTAT	TAGAGTCATT GAAGTTGATC CAGGTGCAAT

도면12f

```

CR380959    GAATTGATCA ACACCGACAT TAGAGTAATT GAAATTGCTC CTGGTATGGT
XM_001645671 GAATTGATTA ATACAAAGAT CAGAGTTATT GAAATTGCAC CAGGTATGGT
SAKLOH04730 GAATTGGTGA ACACCAATAT CAGAGTGATT GAAATTGCTC CGGGTAATGT
XM_451902    GAATTATCA  ATACTGGTAT TAGGGTCATT GAGATTGCTC CAGGTAATGT
XM_002553230 GAACTGATCA ACACAAACAT CAGGGTTATT GAGGTCGCTC CAGGTAACGT
XM_002494574 GAATTAGTCG ATACAAAAAT CAGAGTCATG AGTATTGATC CTGGTAATGT
YMR226C     GAGCTCATCA ACACTAAAAAT TAGAGTCATT CTAATTGCAC CAGGGCTAGT
NM_001018495 GAGACTATTG ACACTCGCAT TCGTATTATG GAGGTTGATC CTGGCTTGGT
XM_501554   GAGAACATCG ACACTCGAAT CCGAGTCATT GAGGTTGATC CTGGAGCCGT
XM_957621   GAGCTGATCG CCACGCGCAT CCGTGTCATG GAGATTGACC CTGGCCAGGT

```

```

701
AABY01000127 CGAAACTGAA TTTTCTACTGG TTAGATACAG AGGCAACGAG GAGCAAGCCA
AACH01000019 CGAAACTGAG TTTTCCCTGG TTAGATACAG AGGTAACGAA GAACAAGCCA
AACAO1000631 TGAGACCGAG TTCTCACTGG TCAGATACAG AGGTAATGAG GAACAAGCTA
AACFO1000116 CGAAACTGAA TTCTCAGTTG TTAGATACAA GGGTGATAAT GAGCGTGCTA
AE016819      AGAAACGGAA TTCTCCATGG TTCGTATGCG CGGTAACGAA GAGAAATGCCA
CR382139      TGAAACTGAA TTTTCAAATG TCAGATTCAA GGGCGATATG GAAAAGGCAA
XM_001387479 GTTGACCGAG TTCTCTTTGG TTCGTTTCCA CGGTGATCAG GGAGCTGCTG
XM_001482184 AGAGACAGAG TTCTCGTTGG TTCGTTTGG CGGTGATGCC GAGGCTGCGA
XM_002419771 ACTTACTGAA TTTAGTGTGG TTCGTTATAA AGGTGACACT GATGCTGCCG
CR380959      CGAAACCGAG TTTTCTGTGG TCAGGTACAA AGGTGACAAG TCCAAAGCAG
XM_001645671 TAACACTGAA TTTTCTGTAA TTAGATATAA AGGTGACCAA GAAAAGGCAG
SAKLOH04730 TGAAACCGAG TTTTCCCTAG TTAGATATAA AGGTGATACG GACCGTGCTA
XM_451902     CGAAACTGAG TTCTCTCTAG TTAGATACAA GGGTGATGCC GATCGTGCTA
XM_002553230 GGAGACCGAG TTTTCTCTGG TTAGATACAA GGGCGACTCT GAGAAAGCCA
XM_002494574 AGAGACCGAA TTTTCTATGG TTAGATTCCG TGGTGATACA GAAAAGGCAA
YMR226C      CGAGACTGAA TTTTCACTAG TTAGATACAG AGGTAACGAG GAACAAGCCA
NM_001018495 CGAAACCGAA TTCAGCGTTG TGAGATTCCA CGGAGACAAA CAAAAGGCTG
XM_501554    TGAGACCGAG TTCTCCGTGG TGCGATTCCG AGGAGACAAG TCCAAGGCCG
XM_957621    CGAGACCGAG TTCAGCGTGG TGAGGTTCTA TGGTGATAAG AACAAGGCTG

```

```

751
AABY01000127 AGAATGTCTA CAA.GGACAC CACCCCATTA ATGGCTGATG ACGTGGCTGA
AACH01000019 AGAATGTTTA CAA.GGACAC CACTCCGTTG ATGGCCGATG ACGTGGCTGA
AACAO1000631 AAAACGTCTA CAA.GGACAC TACGCCGTTG ATGGCCGACG ACGTGGCTGA
AACFO1000116 AATCTGTCTA CGATGGAGTT CACCCCTTGG AAG.CAGACG ACGTAGCAGA
AE016819      AGAAAGTGTA CCA.GGGATT TGAACCCCTA GATGGTGATG ATATCGCTGA
CR382139      AGCTGGTTTA CGC.GGGTAC TGAACCATTA TTATCCGAAG ACGTAGCTGA
XM_001387479 ATGCTGTTTA TGA.AGGTAC CCAACCTTTG GATGCCTCTG ATATCGCAGA
XM_001482184 AAAAGGTGTA TGA.GGGAAC CGAGCCTTTG GGCCGAGAAG ATATTGCGGA
XM_002419771 ATGCTGTTTA TAC.TGGTAC TGAACCATTA ACACCAGAAG ATGTTGCTGA
CR380959      ACGACGTCTA CAG.AGGTAC AACACCACTA TATGCCGATG ATATCGCGGA
XM_001645671 AATAAGTTTA TGA.AAACAC TACTCCTTTA TATGCAGATG ACATCGCTGA
SAKLOH04730 AAAAGGTGTA TGA.AGGTAC TAACCCATTA TATGCAGATG ACATTGCAGA
XM_451902     AACAGGTTTA CAA.AGGTAC TACTCCTCTA TATGCAGATG ACATTGCTGA
XM_002553230 AGAAGGTTTA CGA.AGGCAC ACAACCCCTT TACGCTGACG ATATCGCAGA
XM_002494574 AGAAGGTTTA CCA.AGACAC TGTCCCATTA TATGCAGATG ACATTGCAGA
YMR226C      AGAATGTTTA CAA.GGATAC TACCCCATTT ATGGCTGATG ACGTGGCTGA
NM_001018495 ATAATGTTTA CAA.AAATAG TGAGCCTTTG ACACCCGAAG ACATTGCTGA
XM_501554    ACGCTGTTTA CGC.TGGAAC CGAGCCTCTG GTCGCTGACG ATATTGCCGA
XM_957621    ATGCGCTCTA TGC.CGGTGT CGATCCCTTG ACGCCCGATG ATATCGCAGA

```

```

801
AABY01000127 TTTGATCGTG TACGCAACTT CCAGGAAACA AAACACTGTA ATTGCAGACA
AACH01000019 TTTGATTGTG TATGCGACTT CAAGGAAGCA GAACACTGTA ATTGCAGACA
AACAO1000631 CTTAATCGTA TATTCACCTT CCAGAAAGCA GAACACCGTG GTTGCCGACA

```

도면12g

```

AACF01000116   TTTAATTGTA TACACCACTT CAAGAAAACA GAACACAGTA ATTGCTGACA
AE016819        TACAATTGTC TATGCCACAT CCAGAAAGATC CAACACCGTA GTTGCGAGAGA
CR382139        GGTGTGCGTA TTCGGACTTA CAAGAAAGCA AAATACCGTT ATTGCTGAGA
XM_001387479    AGTTATCGTG TTTGGTATCA CCAGAAAGCA GAACACCGTC ATAGCCGAAA
XM_001482184    AATCATTGTG TTTGCTGTGT CGAGAAAGAGC CAAAACTGTC ATTGCGGAAA
XM_002419771    AGTGGTTGTT TTTGCATCTT CAAGAAAACA AAATACCGTT ATTGCTGATA
CR380959        TTTGATTGTG TACTCTACCA GCAGAAAGCC AAACATGGTG GTAGCAGATG
XM_001645671    TTTGATAGTT TACACCACTT CTAGAAAATC GAATACCGTT ATCGCTGATG
SAKL0H04730     CCTTATTGTG TATGCTACTT CTAGAAAGCC TAATACTGTC ATTGCGGATG
XM_451902       CTTGATCGTT TATGCCACTT CAAGAAAGCC TAATACTGTC ATCGCTGATG
XM_002553230    CCTAATCGTT TACGCAACCT CGAGAAAGCC AAACACCGTC ATCGCGGACG
XM_002494574    TTTAATCGTC TATGCAACCT CTAGAAAGCA AAACACTGTC ATTGCTGACA
YMR226C         TCTGATCGTC TATGCAACTT CCAGAAAACA AAATACTGTA ATTGCGAGACA
NM_001018495    GGTGATTCTT TTTGCCCTCA CTCGCAGAGA AAACGTCGTT ATTGCCGATA
XM_501554       GTTCATCACC TACACTCTCA CTCGCAGAGA GAATGTCGTC ATTGCCGATA
XM_957621       GATCGTGGTG TTCGTCGTAA CACGACGGGA GAACGTTGTT GTTGCTGATA

```

851

900

```

AABY01000127   CGCTAATCTT TCCAACCAAC CAAGCATCGC CTCACCACAT CTTCCGTGGA
AACH01000019   CACTAATCTT TCCTACCAAC CAAGCGTCAC CTTACCATAT CTTTCGCGGG
AACA01000631   CCTGATCTTT CCCCACCAAC CAAGCCTCGC CCTACCACAT CTTTCGCGGT
AACF01000116   CTTTGATATT CCCAACCTCT CAAGGTTCCG CATTCCACGT CCATCGCGAT
AE016819        TGGTCTGTTT CCCATCCGCG CAAGGTTCTC TGTACGATAC TCACCGCAAC
CR382139        CATTAAGTCTT TTCAACCAAT CAAGCCAGCT CATCTCACTT ATACCGTGAA
XM_001387479    CCTTGGTATT CCCAAGTCAC CAGGCTTCTG CCTCTCATGT TTACAAGGCT
XM_001482184    CTTTGGTGTT TCCTACCCAT CAGGCTGGAG CAGTTCATGT TCATAGAGGG
XM_002419771    CTTTGATTTT CCCAATCAT CAAGCTTCTC CAGATCATGT TTATAGAAAA
CR380959        TCCTGGTCTT CCCAACACAC CAGGCATCGG CTCGCACAT CTACAGGGGC
XM_001645671    TTTTGGTATT CCCAACATGC CAAGCTTCTG CATCCCATAT CTATCGTGGA
SAKL0H04730     TTTTGGTTTT TGCTTCAAAC CAAGCATCCC CTTACCATAT CTATCGCGGT
XM_451902       TTTTGGTATT TGCTTCCAAC CAAGCATCTC CTTACCACAT TTACCGTGGC
XM_002553230    TTTTGGTTTT CGCTTCGAAC CAGGCTTCGC CTTACCACAT TTACCGTGGT
XM_002494574    CTTTGATCTT CTCTTCTAAC CAGGCATCAC CATACCACCT CTACAGAGGC
YMR226C         CTTTAATCTT TCCAACAAAC CAAGCGTCAC CTCATCATAT CTTCCGTGGA
NM_001018495    CACTTGTTTT CCCATCCCAT CAAGGTGGTG CCAATCATGT GTACAGAAAG
XM_501554       CTCTCATTTT CCCCACCCAC CAGGCTTCTC CTACTCACGT CTACCGAAAG
XM_957621       CGTTGGTCTT CCCTAGCCAT CAGGCTGGCG CTGGTATTAT GCACCGCAAG

```

901

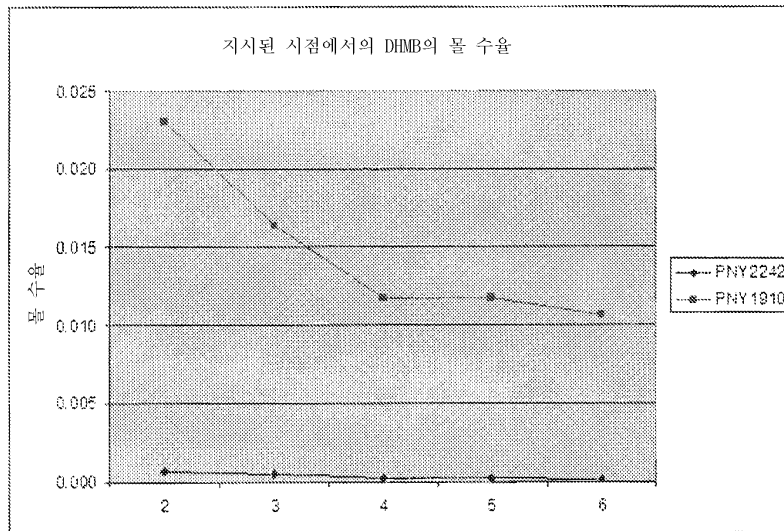
921

```

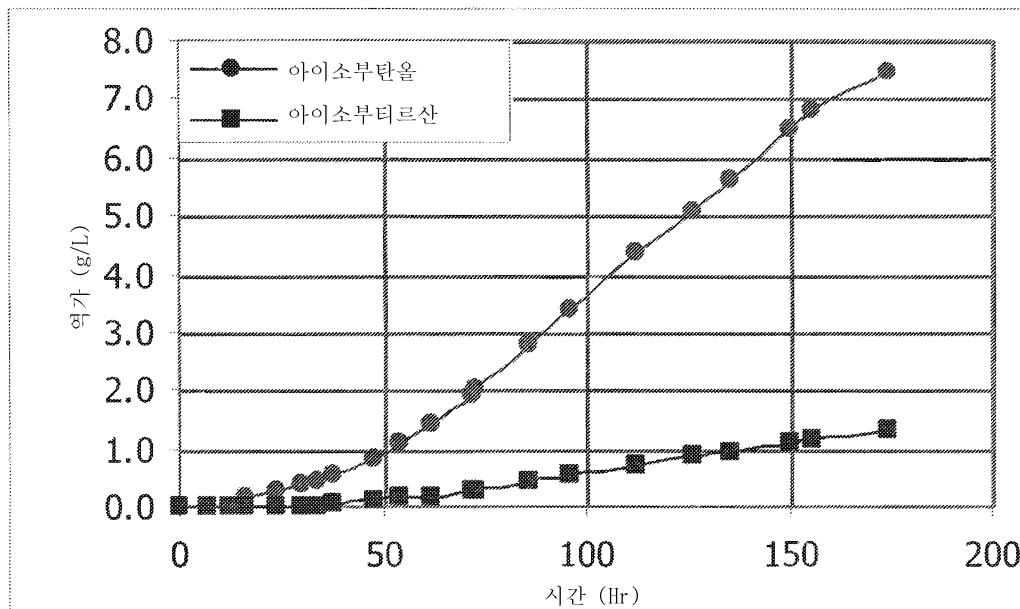
AABY01000127   TGA..... (SEQ ID NO:41)
AACH01000019   TGA..... (SEQ ID NO:47)
AACA01000631   TAA..... (SEQ ID NO:43)
AACF01000116   TAA..... (SEQ ID NO:45)
AE016819        TAA..... (SEQ ID NO:49)
CR382139        AGCGATAAAT AA..... (SEQ ID NO:53)
XM_001387479    CCTAAGTAG. .... (SEQ ID NO:55)
XM_001482184    CCGCTTGAGT GA..... (SEQ ID NO:57)
XM_002419771    CCTAATTAA. .... (SEQ ID NO:61)
CR380959        GACTAA.... (SEQ ID NO:51)
XM_001645671    TAA..... (SEQ ID NO:59)
SAKL0H04730     GACTAA.... (SEQ ID NO:69)
XM_451902       GAATAG.... (SEQ ID NO:67)
XM_002553230    TAG..... (SEQ ID NO:65)
XM_002494574    TCTCAAGACA AAACCAATTG A (SEQ ID NO:63)
YMR226C         TAA..... (SEQ ID NO:24)
NM_001018495    CAAGCGTAG. .... (SEQ ID NO:73)
XM_501554       AACTGA.... (SEQ ID NO:71)
XM_957621       TCGACATGA. .... (SEQ ID NO:75)

```

도면13



도면14



서열 목록

SEQUENCE LISTING

<110> BUTAMAX(TM) Advanced Biofuels, LLC

<120> Host Cells and Methods for Production of Isobutanol

<130> CL5367USNA

<150> 61/472,484

<151> 2011-04-06

<150> 61/467,261

<151> 2011-03-24

<150> 61/472,487

<151> 2011-04-06

<150> 61/476,271

<151> 2011-03-24

<150> 61/570,513

<151> 2011-12-14

<150> 61/467,249

<151> 2011-03-24

<150> 61/472,497

<151> 2011-04-06

<150> 61/472,474

<151> 2011-04-06

<160> 965

<170

> PatentIn version 3.5

<210> 1

<211> 335

<212> PRT

<213> Sulfolobus solfataricus

<400> 1

Met Lys Cys Thr Ser Lys Ile Tyr Thr Asp Asn Asp Ala Asn Leu Asp

1 5 10 15

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

20 25 30

Arg Ala Trp Ala Gln Asn Leu Arg Asp Ser Gly Leu Asn Val Val Val

35 40 45

Gly Leu Glu Arg Glu Gly Lys Ser Trp Glu Leu Ala Lys Ser Asp Gly

50 55 60

Ile Thr Pro Leu His Thr Lys Asp Ala Val Lys Asp Ala Asp Ile Ile

65 70 75 80

Ile Phe Leu Val Pro Asp Met Val Gln Arg Thr Leu Trp Leu Glu Ser

85 90 95

Val Gln Pro Tyr Met Lys Lys Gly Ala Asp Leu Val Phe Ala His Gly

100 105 110
Phe Asn Ile His Tyr Lys Leu Ile Asp Pro Pro Lys Asp Ser Asp Val

115 120 125
Tyr Met Ile Ala Pro Lys Gly Pro Gly Pro Thr Val Arg Glu Tyr Tyr

130 135 140
Lys Ala Gly Gly Gly Val Pro Ala Leu Val Ala Val His Gln Asp Val

145 150 155 160
Ser Gly Thr Ala Leu His Lys Ala Leu Ala Ile Ala Lys Gly Ile Gly

165 170 175
Ala Thr Arg Ala Gly Val Ile Pro Thr Thr Phe Lys Glu Glu Thr Glu

180 185 190
Thr Asp Leu Phe Gly Glu Gln Val Ile Leu Val Gly Gly Ile Met Glu

195 200 205
Leu Met Arg Ala Ala Phe Glu Thr Leu Val Glu Glu Gly Tyr Gln Pro

210 215 220
Glu Val Ala Tyr Phe Glu Thr Ile Asn Glu Leu Lys Met Leu Val Asp

225 230 235 240
Leu Val Tyr Glu Lys Gly Ile Ser Gly Met Leu Lys Ala Val Ser Asp

245 250 255
Thr Ala Lys Tyr Gly Gly Met Thr Val Gly Lys Phe Val Ile Asp Glu

260 265 270
Ser Val Arg Lys Arg Met Lys Glu Ala Leu Gln Arg Ile Lys Ser Gly

275 280 285
Lys Phe Ala Glu Glu Trp Val Glu Glu Tyr Gly Arg Gly Met Pro Thr

290 295 300
Val Val Asn Gly Leu Ser Asn Val Gln Asn Ser Leu Glu Glu Lys Ile

305 310 315 320
Gly Asn Gln Leu Arg Asp Leu Val Gln Lys Gly Lys Pro Lys Ser

325 330 335

<210> 2

<211> 328

<212> PRT

<213> *Pyrobaculum aerophilum*

<400> 2

Met Ala Lys Ile Tyr Thr Asp Arg Glu Ala Ser Leu Glu Pro Leu Lys

1 5 10 15

Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ile Gln Gly Arg Ala Gln

20 25 30

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

35 40 45

Arg Gly Gly Lys Ser Trp Glu Leu Ala Thr Ser Glu Gly Phe Arg Val

50 55 60

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

65 70 75 80

Ile Pro Asp Met Glu Gln Pro Lys Val Trp Gln Glu Gln Ile Ala Pro

85 90 95

Asn Leu Lys Glu Gly Val Val Val Asp Phe Ala His Gly Phe Asn Val

100 105 110

His Phe Gly Leu Ile Lys Pro Pro Lys Asn Ile Asp Val Ile Met Val

115 120 125

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

130 135 140

Arg Gly Val Pro Ala Leu Val Ala Val Tyr Gln Asp Tyr Ser Gly Ser

145 150 155 160

Ala Leu Lys Tyr Ala Leu Ala Leu Ala Lys Gly Ile Gly Ala Thr Arg

165 170 175

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

180 185 190

Ile Gly Glu Gln Ile Val Leu Val Gly Gly Leu Met Glu Leu Ile Lys

195 200 205

Lys Gly Phe Glu Val Leu Val Glu Met Gly Tyr Gln Pro Glu Val Ala

210 215 220

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

Lys Glu Trp Val Glu Glu Tyr Asn Arg Gly Ala Pro Thr Leu Arg Lys
 290 295 300
 Leu Met Glu Glu Ala Arg Thr His Pro Ile Glu Lys Val Gly Glu Glu
 305 310 315 320
 Met Arg Lys Leu Leu Phe Gly Pro
 325

<210> 3

<211> 338

<212> PRT

<213> *Ralstonia solanacearum*

<400> 3

Met Lys Val Phe Tyr Asp Lys Asp Ala Asp Leu Ser Leu Ile Lys Gly
 1 5 10 15

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

Pro Asp Glu Gln Ile Ala Asp Val Tyr Lys Asn Glu Val His Gly Asn
 85 90 95
 Ile Lys Gln Gly Ala Ala Leu Ala Phe Ala His Gly Phe Asn Val His

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

<210> 4

<211> 338

<212> PRT

<213> *Pseudomonas aeruginosa*

<400> 4

Met Arg Val Phe Tyr Asp Lys Asp Cys Asp Leu Ser Ile Ile Gln Gly

1 5 10 15

Lys Lys Val Ala Ile Ile Gly Tyr Gly Ser Gln Gly His Ala His Ala

20 25 30

Cys Asn Leu Lys Asp Ser Gly Val Asp Val Thr Val Gly Leu Arg Ser

35 40 45

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

50 55 60

Asp Val Lys Thr Ala Val Ala Ala Ala Asp Val Val Met Ile Leu Thr

65 70 75 80

Pro Asp Glu Phe Gln Gly Arg Leu Tyr Lys Glu Glu Ile Glu Pro Asn

85 90 95

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

100 105 110

Tyr Asn Gln Val Val Pro Arg Ala Asp Leu Asp Val Ile Met Ile Ala

115 120 125

Pro Lys Ala Pro Gly His Thr Val Arg Ser Glu Phe Val Lys Gly Gly

130 135 140

Gly Ile Pro Asp Leu Ile Ala Ile Tyr Gln Asp Ala Ser Gly Asn Ala

145 150 155 160

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

165 170 175

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

180 185 190

Gly Glu Gln Ala Val Leu Cys Gly Gly Cys Val Glu Leu Val Lys Ala

195 200 205

Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr Ala Pro Glu Met Ala Tyr

210 215 220
 Phe Glu Cys Leu His Glu Leu Lys Leu Ile Val Asp Leu Met Tyr Glu
 225 230 235 240
 Gly Gly Ile Ala Asn Met Asn Tyr Ser Ile Ser Asn Asn Ala Glu Tyr

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

 305 310 315 320
 Arg Ala Met Met Pro Trp Ile Ala Ala Asn Lys Ile Val Asp Lys Ser
 325 330 335
 Lys Asn

 <210> 5
 <211> 338
 <212> PRT
 <213> Pseudomonas fluorescens
 <400> 5
 Met Lys Val Phe Tyr Asp Lys Asp Cys Asp Leu Ser Ile Ile Gln Gly
 1 5 10 15
 Lys Lys Val Ala Ile Ile Gly Tyr Gly Ser Gln Gly His Ala Gln Ala
 20 25 30

 Cys Asn Leu Lys Asp Ser Gly Val Asp Val Thr Val Gly Leu Arg Lys
 35 40 45
 Gly Ser Ala Thr Val Ala Lys Ala Glu Ala His Gly Leu Lys Val Thr
 50 55 60
 Asp Val Ala Ala Ala Val Ala Gly Ala Asp Leu Val Met Ile Leu Thr
 65 70 75 80

Pro Asp Glu Phe Gln Ser Gln Leu Tyr Lys Asn Glu Ile Glu Pro Asn
85 90 95

Ile Lys Lys Gly Ala Thr Leu Ala Phe Ser His Gly Phe Ala Ile His
100 105 110

Tyr Asn Gln Val Val Pro Arg Ala Asp Leu Asp Val Ile Met Ile Ala
115 120 125

Pro Lys Ala Pro Gly His Thr Val Arg Ser Glu Phe Val Lys Gly Gly
130 135 140

Gly Ile Pro Asp Leu Ile Ala Ile Tyr Gln Asp Ala Ser Gly Asn Ala
145 150 155 160

Lys Asn Val Ala Leu Ser Tyr Ala Ala Gly Val Gly Gly Gly Arg Thr
165 170 175

Gly Ile Ile Glu Thr Thr Phe Lys Asp Glu Thr Glu Thr Asp Leu Phe
180 185 190

Gly Glu Gln Ala Val Leu Cys Gly Gly Thr Val Glu Leu Val Lys Ala
195 200 205

Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr Ala Pro Glu Met Ala Tyr
210 215 220

Phe Glu Cys Leu His Glu Leu Lys Leu Ile Val Asp Leu Met Tyr Glu
225 230 235 240

Gly Gly Ile Ala Asn Met Asn Tyr Ser Ile Ser Asn Asn Ala Glu Tyr
245 250 255

Gly Glu Tyr Val Thr Gly Pro Glu Val Ile Asn Ala Glu Ser Arg Gln
260 265 270

Ala Met Arg Asn Ala Leu Lys Arg Ile Gln Asp Gly Glu Tyr Ala Lys
275 280 285

Met Phe Ile Ser Glu Gly Ala Thr Gly Tyr Pro Ser Met Thr Ala Lys
290 295 300

Arg Arg Asn Asn Ala Ala His Gly Ile Glu Ile Ile Gly Glu Gln Leu
305 310 315 320

Arg Ser Met Met Pro Trp Ile Gly Ala Asn Lys Ile Val Asp Lys Ala

325 330 335

Lys Asn

<210> 6

<211> 338

<212> PRT

<213> Spinacia oleracea

<400> 6

Met Arg Val Phe Tyr Asp Lys Asp Cys Asp Leu Ser Ile Ile Gln Gly

1 5 10 15

Lys Lys Val Ala Ile Ile Gly Tyr Gly Ser Gln Gly His Ala His Ala

20 25 30

Cys Asn Leu Lys Asp Ser Gly Val Asp Val Thr Val Gly Leu Arg Ser

35 40 45

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

50 55 60

Asp Val Lys Thr Ala Val Ala Ala Ala Asp Val Val Met Ile Leu Thr

65 70 75 80

Pro Asp Glu Phe Gln Gly Arg Leu Tyr Lys Glu Glu Ile Glu Pro Asn

85 90 95

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

100 105 110

Tyr Asn Gln Val Val Pro Arg Ala Asp Leu Asp Val Ile Met Ile Ala

115 120 125

Pro Lys Ala Pro Gly His Thr Val Arg Ser Glu Phe Val Lys Gly Gly

130 135 140

Gly Ile Pro Asp Leu Ile Ala Ile Tyr Gln Asp Ala Ser Gly Asn Ala

145 150 155 160

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

165 170 175

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

180 185 190

Gly Glu Gln Ala Val Leu Cys Gly Gly Cys Val Glu Leu Val Lys Ala

195 200 205

Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr Ala Pro Glu Met Ala Tyr

210 215 220

Phe Glu Cys Leu His Glu Leu Lys Leu Ile Val Asp Leu Met Tyr Glu

225 230 235 240

Gly Gly Ile Ala Asn Met Asn Tyr Ser Ile Ser Asn Asn Ala Glu Tyr

245 250 255

Gly Glu Tyr Val Thr Gly Pro Glu Val Ile Asn Ala Glu Ser Arg Ala

260 265 270

Ala Met Arg Asn Ala Leu Lys Arg Ile Gln Asp Gly Glu Tyr Ala Lys

275 280 285

Met Phe Ile Thr Glu Gly Ala Ala Asn Tyr Pro Ser Met Thr Ala Tyr

290 295 300

Arg Arg Asn Asn Ala Ala His Pro Ile Glu Gln Ile Gly Glu Lys Leu

305 310 315 320

Arg Ala Met Met Pro Trp Ile Ala Ala Asn Lys Ile Val Asp Lys Ser

325 330 335

Lys Asn

<210> 7

<211> 331

<212> PRT

<213> Natronomonas pharaanis

<400> 7

Met Thr Asp Ala Thr Ile Tyr Tyr Asp Asp Asp Ala Glu Ser Thr Val

1 5 10 15

Leu Asp Asp Lys Thr Val Ala Val Leu Gly Tyr Gly Ser Gln Gly His

20 25 30

Ala His Ala Gln Asn Leu Asp Asp Ser Gly Val Asp Val Val Val Gly

35 40 45

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

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

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

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

290 295 300

Gln Leu Arg Ala Ala Glu Lys Asn His Asp Ile Glu Ala Val Gly Glu

305 310 315 320

Asp Leu Arg Ala Leu Phe Ala Trp Gly Asp Asp

325 330

<210> 8

<211> 342

<212> PRT

<213> Bacillus subtilis

<400> 8

Met Val Lys Val Tyr Tyr Asn Gly Asp Ile Lys Glu Asn Val Leu Ala

1 5 10 15

Gly Lys Thr Val Ala Val Ile Gly Tyr Gly Ser Gln Gly His Ala His

20 25 30

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

35 40 45

Gln Gly Lys Ser Phe Thr Gln Ala Gln Glu Asp Gly His Lys Val Phe

50 55 60

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

65 70 75 80

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

85 90 95

Leu Thr Ala Gly Lys Ser Leu Val Phe Ala His Gly Phe Asn Val His

100 105 110

Phe His Gln Ile Val Pro Pro Ala Asp Val Asp Val Phe Leu Val Ala

115 120 125

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

130 135 140

Gly Val Pro Ala Leu Phe Ala Ile Tyr Gln Asp Val Thr Gly Glu Ala

145 150 155 160

Arg Asp Lys Ala Leu Ala Tyr Ala Lys Gly Ile Gly Gly Ala Arg Ala

165 170 175
 Gly Val Leu Glu Thr Thr Phe Lys Glu Glu Thr Glu Thr Asp Leu Phe
 180 185 190
 Gly Glu Gln Ala Val Leu Cys Gly Gly Leu Ser Ala Leu Val Lys Ala
 195 200 205
 Gly Phe Glu Thr Leu Thr Glu Ala Gly Tyr Gln Pro Glu Leu Ala Tyr
 210 215 220

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

 Glu Trp Ile Val Glu Asn Gln Val Asn Arg Pro Arg Phe Asn Ala Ile
 290 295 300
 Asn Ala Ser Glu Asn Glu His Gln Ile Glu Val Val Gly Arg Lys Leu
 305 310 315 320
 Arg Glu Met Met Pro Phe Val Lys Gln Gly Lys Lys Lys Glu Ala Val
 325 330 335
 Val Ser Val Ala Gln Asn
 340
 <210> 9
 <211> 338
 <212> PRT
 <213> Corynebacterium glutamicum

 <400> 9
 Met Ala Ile Glu Leu Leu Tyr Asp Ala Asp Ala Asp Leu Ser Leu Ile
 1 5 10 15
 Gln Gly Arg Lys Val Ala Ile Val Gly Tyr Gly Ser Gln Gly His Ala
 20 25 30

His Ser Gln Asn Leu Arg Asp Ser Gly Val Glu Val Val Ile Gly Leu
35 40 45

Arg Glu Gly Ser Lys Ser Ala Glu Lys Ala Lys Glu Ala Gly Phe Glu
50 55 60

Val Lys Thr Thr Ala Glu Ala Ala Ala Trp Ala Asp Val Ile Met Leu
65 70 75 80

Leu Ala Pro Asp Thr Ser Gln Ala Glu Ile Phe Thr Asn Asp Ile Glu
85 90 95

Pro Asn Leu Asn Ala Gly Asp Ala Leu Leu Phe Gly His Gly Leu Asn
100 105 110

Ile His Phe Asp Leu Ile Lys Pro Ala Asp Asp Ile Ile Val Gly Met
115 120 125

Val Ala Pro Lys Gly Pro Gly His Leu Val Arg Arg Gln Phe Val Asp
130 135 140

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

Thr Ala Gln Ala Leu Thr Leu Ser Tyr Ala Ala Ala Ile Gly Gly Ala
165 170 175

Arg Ala Gly Val Ile Pro Thr Thr Phe Glu Ala Glu Thr Val Thr Asp
180 185 190

Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Thr Glu Glu Leu Val
195 200 205

Lys Val Gly Phe Glu Val Leu Thr Glu Ala Gly Tyr Glu Pro Glu Met
210 215 220

Ala Tyr Phe Glu Val Leu His Glu Leu Lys Leu Ile Val Asp Leu Met
225 230 235 240

Phe Glu Gly Gly Ile Ser Asn Met Asn Tyr Ser Val Ser Asp Thr Ala
245 250 255

Glu Phe Gly Gly Tyr Leu Ser Gly Pro Arg Val Ile Asp Ala Asp Thr
260 265 270

Lys Ser Arg Met Lys Asp Ile Leu Thr Asp Ile Gln Asp Gly Thr Phe

275 280 285
 Thr Lys Arg Leu Ile Ala Asn Val Glu Asn Gly Asn Thr Glu Leu Glu
 290 295 300
 Gly Leu Arg Ala Ser Tyr Asn Asn His Pro Ile Glu Glu Thr Gly Ala
 305 310 315 320

Lys Leu Arg Asp Leu Met Ser Trp Val Lys Val Asp Ala Arg Ala Glu
 325 330 335
 Thr Ala

<210> 10

<211> 339

<212> PRT

<213> Phaeospririllum molischianum

<220><221> misc_feature

<222> (310)..(310)

<223> Xaa can be any naturally occurring amino acid

<400> 10

Met Arg Val Tyr Tyr Asp Arg Asp Ala Asp Val Asn Leu Ile Lys Ser
 1 5 10 15
 Lys Lys Val Ala Val Ile Gly Tyr Gly Ser Gln Gly His Ala His Val

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

85 90 95
 Asn Leu Lys Pro Gly Ala Ala Leu Val Phe Ala His Gly Leu Ala Ile
 100 105 110
 His Phe Lys Leu Ile Glu Ala Arg Ala Asp Leu Asp Val Phe Met Val

115 120 125
 Ala Pro Lys Gly Pro Gly His Thr Val Arg Gly Glu Tyr Leu Lys Gly
 130 135 140
 Gly Gly Val Pro Cys Leu Val Ala Val Ala Gln Asn Pro Thr Gly Asn

 145 150 155 160
 Ala Leu Glu Leu Ala Leu Ser Tyr Ala Ser Ala Ile Gly Gly Gly Arg
 165 170 175
 Ser Gly Ile Ile Glu Thr Thr Phe Arg Glu Glu Cys Glu Thr Asp Leu
 180 185 190
 Phe Gly Glu Gln Val Val Leu Cys Gly Gly Leu Ser Lys Leu Ile Gln
 195 200 205
 Tyr Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr Ala Pro Glu Met Ala

 210 215 220
 Tyr Phe Glu Cys Leu His Glu Val Lys Leu Ile Val Asp Leu Ile Tyr
 225 230 235 240
 Glu Gly Gly Ile Ala Asn Met Arg Tyr Ser Ile Ser Asn Thr Ala Glu
 245 250 255
 Tyr Gly Asp Tyr Val Thr Gly Ser Arg Ile Ile Thr Glu Ala Thr Lys
 260 265 270
 Ala Glu Met Lys Arg Val Leu Ala Asp Ile Gln Ser Gly Arg Phe Val

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

<210> 11

<211> 339

<212> PRT

<213> *Zymomonas mobilis*

<400> 11

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

1 5 10 15

Lys Lys Ile Ala Ile Leu Gly Tyr Gly Ser Gln Gly His Ala His Ala

20 25 30

Gln Asn Leu Arg Asp Ser Gly Val Ala Glu Val Ala Ile Ala Leu Arg

35 40 45

Pro Asp Ser Ala Ser Val Lys Lys Ala Gln Asp Ala Gly Phe Lys Val

50 55 60

Leu Thr Asn Ala Glu Ala Ala Lys Trp Ala Asp Ile Leu Met Ile Leu

65 70 75 80

Ala Pro Asp Glu His Gln Ala Ala Ile Tyr Ala Glu Asp Leu Lys Asp

85 90 95

Asn Leu Arg Pro Gly Ser Ala Ile Ala Phe Ala His Gly Leu Asn Ile

100 105 110

His Phe Gly Leu Ile Glu Pro Arg Lys Asp Ile Asp Val Phe Met Ile

115 120 125

Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser Glu Tyr Val Arg Gly

130 135 140

Gly Gly Val Pro Cys Leu Val Ala Val Asp Gln Asp Ala Ser Gly Asn

145 150 155 160

Ala His Asp Ile Ala Leu Ala Tyr Ala Ser Gly Ile Gly Gly Gly Arg

165 170 175

Ser Gly Val Ile Glu Thr Thr Phe Arg Glu Glu Val Glu Thr Asp Leu

180 185 190

Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Leu Thr Ala Leu Ile Thr

195 200 205

Ala Gly Phe Glu Thr Leu Thr Glu Ala Gly Tyr Ala Pro Glu Met Ala

210 215 220

Phe Phe Glu Cys Met His Glu Met Lys Leu Ile Val Asp Leu Ile Tyr

225 230 235 240
Glu Ala Gly Ile Ala Asn Met Arg Tyr Ser Ile Ser Asn Thr Ala Glu

245 250 255
Tyr Gly Asp Ile Val Ser Gly Pro Arg Val Ile Asn Glu Glu Ser Lys

260 265 270
Lys Ala Met Lys Ala Ile Leu Asp Asp Ile Gln Ser Gly Arg Phe Val

275 280 285
Ser Lys Phe Val Leu Asp Asn Arg Ala Gly Gln Pro Glu Leu Lys Ala

290 295 300
Ala Arg Lys Arg Met Ala Ala His Pro Ile Glu Gln Val Gly Ala Arg

305 310 315 320
Leu Arg Lys Met Met Pro Trp Ile Ala Ser Asn Lys Leu Val Asp Lys

325 330 335
Ala Arg Asn

<210> 12

<211> 338

<212> PRT

<213> Alkalilimnicola ehrlichei

<400> 12

Met Gln Val Tyr Tyr Asp Lys Asp Ala Asp Leu Ser Ile Ile Gln Gly

1 5 10 15

Lys Lys Val Ala Val Ile Gly Tyr Gly Ser Gln Gly His Ala His Ala

20 25 30

Asn Asn Leu Lys Glu Ser Gly Val Asp Val Val Val Gly Leu Arg Glu

35 40 45

Gly Ser Ser Ser Ala Ala Lys Ala Gln Lys Ala Gly Leu Ala Val Ala

50 55 60

Ser Ile Glu Asp Ala Ala Ala Gln Ala Asp Val Val Met Ile Leu Ala

65 70 75 80

Pro Asp Glu His Gln Ala Val Ile Tyr His Asn Gln Ile Ala Pro Asn

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

Lys Asn

<210> 13

<211> 340

<212> PRT

<213> Campylobacter lari

<400> 13

Met Ala Val Ser Ile Tyr Tyr Asp Lys Asp Cys Asp Ile Asn Leu Ile

1 5 10 15
Lys Ser Lys Lys Val Ala Ile Ile Gly Phe Gly Ser Gln Gly His Ala

20 25 30
His Ala Met Asn Leu Arg Asp Ser Gly Val Glu Val Ile Ile Gly Leu

35 40 45
Lys Glu Gly Gly Gln Ser Trp Ala Lys Ala Gln Lys Ala Asn Phe Ile

50 55 60
Val Lys Ser Val Lys Glu Ala Thr Lys Glu Ala Asp Leu Ile Met Ile

65 70 75 80
Leu Ala Pro Asp Glu Ile Gln Ser Glu Ile Phe Asn Glu Glu Ile Lys

85 90 95
Pro Glu Leu Lys Ala Gly Lys Thr Leu Ala Phe Ala His Gly Phe Asn

100 105 110
Ile His Tyr Gly Gln Ile Val Ala Pro Lys Gly Ile Asp Val Ile Met

115 120 125
Ile Ala Pro Lys Ala Pro Gly His Thr Val Arg His Glu Phe Ser Ile

130 135 140
Gly Gly Gly Thr Pro Cys Leu Ile Ala Ile His Gln Asp Glu Ser Lys

145 150 155 160
Asn Ala Lys Asn Leu Ala Leu Ser Tyr Ala Ser Ala Ile Gly Gly Gly

165 170 175
Arg Thr Gly Ile Ile Glu Thr Thr Phe Lys Ala Glu Thr Glu Thr Asp

180 185 190

Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Leu Ser Ala Leu Ile

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

260 265 270
Lys Glu Ala Met Lys Gly Val Leu Lys Asp Ile Gln Asn Gly Ser Phe
275 280 285
Ala Lys Asp Phe Ile Leu Glu Arg Arg Ala Asn Phe Ala Arg Met His
290 295 300
Ala Glu Arg Lys Leu Met Asn Asp Ser Leu Ile Glu Lys Thr Gly Arg
305 310 315 320
Glu Leu Arg Ala Met Met Pro Trp Ile Ser Ala Lys Lys Leu Val Asp

325 330 335
Lys Asp Lys Asn
340

<210> 14

<211> 338

<212> PRT

<213> Marinobacter aquaeolei

<400> 14

Met Gln Val Tyr Tyr Asp Lys Asp Cys Asp Leu Ser Ile Ile Gln Gly
1 5 10 15
Lys Lys Val Ala Ile Leu Gly Phe Gly Ser Gln Gly His Ala His Ala
20 25 30
Cys Asn Leu Lys Asp Ser Gly Val Asp Val Val Val Gly Leu Arg Ala

35

40

45

Gly Ser Ser Ser Ile Ala Lys Ala Glu Ala Tyr Gly Leu Lys Thr Ser
 50 55 60
 Asp Val Ala Ser Ala Val Ala Ser Ala Asp Val Val Met Val Leu Thr
 65 70 75 80
 Pro Asp Glu Phe Gln Ala Gln Leu Tyr Arg Glu Glu Ile Glu Pro Asn
 85 90 95
 Leu Lys Gln Gly Ala Thr Leu Ala Phe Ala His Gly Phe Ala Ile His
 100 105 110
 Tyr Asn Gln Ile Val Pro Arg Lys Asp Leu Asp Val Ile Met Val Ala
 115 120 125
 Pro Lys Ala Pro Gly His Thr Val Arg Thr Glu Phe Thr Lys Gly Gly
 130 135 140
 Gly Ile Pro Asp Leu Ile Ala Ile Phe Gln Asp Ala Ser Gly Asn Ala
 145 150 155 160
 Lys Asn Val Ala Leu Ser Tyr Ala Ser Gly Ile Gly Gly Gly Arg Thr
 165 170 175
 Gly Ile Ile Glu Thr Thr Phe Lys Asp Glu Thr Glu Thr Asp Leu Phe
 180 185 190
 Gly Glu Gln Ala Val Leu Cys Gly Gly Ala Val Glu Leu Val Lys Ala
 195 200 205
 Gly Phe Glu Thr Leu Thr Glu Ala Gly Tyr Ala Pro Glu Met Ala Tyr
 210 215 220
 Phe Glu Cys Leu His Glu Leu Lys Leu Ile Val Asp Leu Met Tyr Glu
 225 230 235 240
 Gly Gly Ile Ala Asn Met Asn Tyr Ser Ile Ser Asn Asn Ala Glu Tyr
 245 250 255
 Gly Glu Tyr Val Thr Gly Pro Glu Val Ile Asn Glu Gln Ser Arg Glu
 260 265 270
 Ala Met Arg Asn Ala Leu Lys Arg Ile Gln Ser Gly Glu Tyr Ala Lys
 275 280 285
 Met Phe Ile Ser Glu Gly Ala Leu Asn Tyr Pro Ser Met Thr Ala Arg

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

 <210> 15
 <211> 338
 <212> PRT
 <213> Psychrobacter arcticus
 <400> 15
 Met Asn Val Tyr Tyr Asp Lys Asp Cys Asp Leu Ser Ile Val Gln Gly
 1 5 10 15

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

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

 Gly Ile Pro Asp Leu Ile Ala Ile Tyr Gln Asp Ala Ser Gly Gln Ala
 145 150 155 160

Lys Gln Leu Ala Leu Ser Tyr Ala Ala Gly Val Gly Gly Gly Arg Ser
165 170 175
Gly Ile Ile Glu Thr Thr Phe Lys Asp Glu Thr Glu Thr Asp Leu Phe
180 185 190
Gly Glu Gln Ala Val Leu Cys Gly Gly Ala Val Glu Leu Val Lys Met
195 200 205

Gly Phe Glu Thr Leu Thr Glu Ala Gly Tyr Ala Pro Glu Met Ala Tyr
210 215 220
Phe Glu Cys Leu His Glu Leu Lys Leu Ile Val Asp Leu Met Tyr Glu
225 230 235 240
Gly Gly Ile Ala Asp Met Asn Tyr Ser Ile Ser Asn Asn Ala Glu Tyr
245 250 255
Gly Glu Tyr Val Thr Gly Pro Glu Val Ile Asn Glu Gln Ser Arg Glu
260 265 270

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

Lys Asn

<210> 16

<211> 338

<212> PRT

<213> Hahella chejuensis

<400> 16

Met Gln Val Tyr Tyr Asp Lys Asp Cys Asp Leu Ser Ile Ile Gln Gly
1 5 10 15
Lys Lys Val Ala Ile Ile Gly Tyr Gly Ser Gln Gly His Ala His Ala

20 25 30
 Asn Asn Leu Lys Asp Ser Gly Val Asp Val Cys Val Gly Leu Arg Lys
 35 40 45
 Gly Ser Gly Ser Trp Ala Lys Ala Glu Asn Ala Gly Leu Ala Val Lys

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

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

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

 245 250 255
 Gly Glu Tyr Val Thr Gly Pro Glu Val Ile Asn Asp Gln Ser Arg Ala
 260 265 270

Ala Met Arg Asn Ala Leu Lys Arg Ile Gln Asp Gly Glu Tyr Ala Lys
275 280 285

Met Phe Ile Ala Glu Gly Ala His Asn Tyr Pro Ser Met Thr Ala Tyr
290 295 300

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

305 310 315 320
Arg Ser Met Met Pro Trp Ile Ala Ser Asn Lys Ile Val Asp Lys Ser
325 330 335

Lys Asn

<210> 17

<211> 338

<212> PRT

<213> Thiobacillus denitrificans

<400> 17

Met Lys Val Tyr Tyr Asp Lys Asp Ala Asp Leu Ser Leu Ile Lys Gln
1 5 10 15
Arg Lys Val Ala Ile Val Gly Tyr Gly Ser Gln Gly His Ala His Ala
20 25 30

Asn Asn Leu Lys Asp Ser Gly Val Asp Val Thr Val Ala Leu Arg Pro
35 40 45

Gly Ser Ala Ser Ala Lys Lys Ala Glu Asn Ala Gly Leu Thr Val Lys
50 55 60

Ser Val Pro Glu Ala Val Ala Gly Ala Asp Leu Val Met Ile Leu Thr
65 70 75 80

Pro Asp Glu Phe Gln Ser Arg Leu Tyr Arg Asp Glu Ile Glu Pro Asn
85 90 95

Ile Lys Gln Gly Ala Thr Leu Ala Phe Ala His Gly Phe Ser Ile His
100 105 110

Tyr Asn Gln Val Val Pro Arg Ala Asp Leu Asp Val Ile Met Ile Ala
115 120 125

Pro Lys Ala Pro Gly His Thr Val Arg Ser Glu Phe Val Lys Gly Gly

130 135 140
 Gly Ile Pro Asp Leu Ile Ala Ile Tyr Gln Asp Ala Ser Gly Lys Ala
 145 150 155 160

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

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

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

<210> 18

<211> 338

<212> PRT

<213> Azotobacter vinelandii

<400> 18

Met Lys Val Tyr Tyr Asp Lys Asp Cys Asp Leu Ser Ile Ile Gln Ser

1 5 10 15
Lys Lys Val Ala Ile Ile Gly Tyr Gly Ser Gln Gly His Ala His Ala
20 25 30
Cys Asn Leu Lys Asp Ser Gly Val Asp Val Tyr Val Gly Leu Arg Ala
35 40 45
Gly Ser Ala Ser Val Ala Lys Ala Glu Ala His Gly Leu Thr Val Lys
50 55 60
Ser Val Lys Asp Ala Val Ala Ala Ala Asp Val Val Met Ile Leu Thr

65 70 75 80
Pro Asp Glu Phe Gln Gly Arg Leu Tyr Lys Asp Glu Ile Glu Pro Asn
85 90 95
Leu Lys Lys Gly Ala Thr Leu Ala Phe Ala His Gly Phe Ser Ile His
100 105 110
Tyr Asn Gln Val Val Pro Arg Ala Asp Leu Asp Val Ile Met Ile Ala
115 120 125
Pro Lys Ala Pro Gly His Thr Val Arg Ser Glu Phe Val Arg Gly Gly

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

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

245 250 255
Gly Glu Tyr Val Thr Gly Pro Glu Val Ile Asn Glu Gln Ser Arg Gln

260 265 270
Ala Met Arg Asn Ala Leu Lys Arg Ile Gln Asp Gly Glu Tyr Ala Lys

275 280 285
Met Phe Ile Thr Glu Gly Ala Ala Asn Tyr Pro Ser Met Thr Ala Tyr

290 295 300
Arg Arg Asn Asn Ala Ala His Gln Ile Glu Val Val Gly Glu Lys Leu

305 310 315 320
Arg Thr Met Met Pro Trp Ile Ala Ala Asn Lys Ile Val Asp Lys Thr

325 330 335
Lys Asn

<210> 19

<211> 338

<212> PRT

<213> Pseudomonas syringae

<400> 19

Met Lys Val Phe Tyr Asp Lys Asp Cys Asp Leu Ser Ile Ile Gln Gly
1 5 10 15

Lys Lys Val Ala Ile Ile Gly Tyr Gly Ser Gln Gly His Ala Gln Ala
20 25 30

Cys Asn Leu Lys Asp Ser Gly Val Asp Val Thr Val Gly Leu Arg Lys
35 40 45

Gly Ser Ala Thr Val Ala Lys Ala Glu Ala His Gly Leu Lys Val Thr
50 55 60

Asp Val Ala Ser Ala Val Ala Ala Ala Asp Leu Val Met Ile Leu Thr
65 70 75 80

Pro Asp Glu Phe Gln Ser Gln Leu Tyr Lys Asn Glu Val Glu Pro Asn
85 90 95

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

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

<210> 20

<211> 338

<212> PRT

<213> *Pseudomonas syringae*

<400> 20

Met Lys Val Phe Tyr Asp Lys Asp Cys Asp Leu Ser Ile Ile Gln Gly

1 5 10 15

Lys Lys Val Ala Ile Ile Gly Tyr Gly Ser Gln Gly His Ala Gln Ala

20 25 30

Cys Asn Leu Lys Asp Ser Gly Val Asp Val Thr Val Gly Leu Arg Lys

35 40 45

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

50 55 60

Asp Val Ala Ser Ala Val Ala Ala Ala Asp Leu Val Met Ile Leu Thr

65 70 75 80

Pro Asp Glu Phe Gln Ser Gln Leu Tyr Lys Asn Glu Val Glu Pro Asn

85 90 95

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

100 105 110

Tyr Asn Gln Val Val Pro Arg Ala Asp Leu Asp Val Ile Met Ile Ala

115 120 125

Pro Lys Ala Pro Gly His Thr Val Arg Thr Glu Phe Val Lys Gly Gly

130 135 140

Gly Ile Pro Asp Leu Ile Ala Val Tyr Gln Asp Ala Ser Gly Asn Ala

145 150 155 160

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

165 170 175

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

180 185 190

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

195 200 205

Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr Ala Pro Glu Met Ala Tyr

210

215

220

Phe Glu Cys Leu His Glu Leu Lys Leu Ile Val Asp Leu Met Tyr Glu

225

230

235

240

Gly Gly Ile Ala Asn Met Asn Tyr Ser Ile Ser Asn Asn Ala Glu Tyr

245

250

255

Gly Glu Tyr Val Thr Gly Pro Glu Val Ile Asn Ala Glu Ser Arg Gln

260

265

270

Ala Met Arg Asn Ala Leu Lys Arg Ile Gln Asp Gly Glu Tyr Ala Lys

275

280

285

Met Phe Ile Thr Glu Gly Ala Thr Gly Tyr Pro Ser Met Thr Ala Lys

290

295

300

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

305

310

315

320

Arg Ser Met Met Pro Trp Ile Ala Ala Asn Lys Ile Val Asp Lys Asp

325

330

335

Lys Asn

<210> 21

<211> 338

<212>

> PRT

<213> Pseudomonas putida

<400> 21

Met Lys Val Phe Tyr Asp Lys Asp Cys Asp Leu Ser Ile Ile Gln Gly

1

5

10

15

Lys Lys Val Ala Ile Ile Gly Tyr Gly Ser Gln Gly His Ala Gln Ala

20

25

30

Cys Asn Leu Lys Asp Ser Gly Val Asp Val Thr Val Gly Leu Arg Lys

35

40

45

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

50

55

60

Asp Val Ala Thr Ala Val Ala Ala Ala Asp Leu Val Met Ile Leu Thr
 65 70 75 80
 Pro Asp Glu Phe Gln Gly Ala Leu Tyr Lys Asn Glu Ile Glu Pro Asn
 85 90 95
 Ile Lys Lys Gly Ala Thr Leu Ala Phe Ser His Gly Phe Ser Ile His
 100 105 110
 Tyr Asn Gln Val Val Pro Arg Ala Asp Leu Asp Val Ile Met Ile Ala
 115 120 125
 Pro Lys Ala Pro Gly His Thr Val Arg Ser Glu Phe Val Lys Gly Gly
 130 135 140
 Gly Ile Pro Asp Leu Ile Ala Ile Tyr Gln Asp Ala Ser Gly Asn Ala
 145 150 155 160
 Lys Asn Val Ala Leu Ser Tyr Ala Ser Gly Val Gly Gly Gly Arg Thr
 165 170 175
 Gly Ile Ile Glu Thr Thr Phe Lys Asp Glu Thr Glu Thr Asp Leu Phe
 180 185 190
 Gly Glu Gln Ala Val Leu Cys Gly Gly Thr Val Glu Leu Val Lys Ala
 195 200 205
 Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr Ala Pro Glu Met Ala Tyr
 210 215 220
 Phe Glu Cys Leu His Glu Leu Lys Leu Ile Val Asp Leu Met Tyr Glu
 225 230 235 240
 Gly Gly Ile Ala Asn Met Asn Tyr Ser Ile Ser Asn Asn Ala Glu Tyr
 245 250 255
 Gly Glu Tyr Val Thr Gly Pro Glu Val Ile Asn Glu Glu Ser Arg Lys
 260 265 270
 Ala Met Arg Asn Ala Leu Lys Arg Ile Gln Asp Gly Glu Tyr Ala Lys
 275 280 285
 Met Phe Ile Ser Glu Gly Ala Thr Asn Tyr Pro Ser Met Thr Ala Lys
 290 295 300
 Arg Arg Asn Asn Ala Ala His Gly Ile Glu Ile Ile Gly Glu Gln Leu

305 310 315 320
Arg Ser Met Met Pro Trp Ile Ser Ala Asn Lys Ile Val Asp Lys Thr

 325 330 335
Lys Asn

<210> 22

<211> 338

<212> PRT

<213> Pseudomonas entomophila

<400> 22

Met	Lys	Val	Phe	Tyr	Asp	Lys	Asp	Cys	Asp	Leu	Ser	Ile	Ile	Gln	Gly
1				5					10					15	
Lys	Lys	Val	Ala	Ile	Ile	Gly	Tyr	Gly	Ser	Gln	Gly	His	Ala	Gln	Ala
			20					25					30		

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

Ile	Lys	Lys	Gly	Ala	Thr	Leu	Ala	Phe	Ser	His	Gly	Phe	Ala	Ile	His
			100					105					110		
Tyr	Asn	Gln	Val	Val	Pro	Arg	Ala	Asp	Leu	Asp	Val	Ile	Met	Ile	Ala
			115					120					125		
Pro	Lys	Ala	Pro	Gly	His	Thr	Val	Arg	Ser	Glu	Phe	Val	Lys	Gly	Gly
			130					135					140		
Gly	Ile	Pro	Asp	Leu	Ile	Ala	Ile	Tyr	Gln	Asp	Ala	Ser	Gly	Asn	Ala
			145					150					155		160

Lys Asn Val Ala Leu Ser Tyr Ala Ser Gly Val Gly Gly Gly Arg Thr
165 170 175

Gly Ile Ile Glu Thr Thr Phe Lys Asp Glu Thr Glu Thr Asp Leu Phe
 180 185 190
 Gly Glu Gln Ala Val Leu Cys Gly Gly Thr Val Glu Leu Val Lys Ala
 195 200 205
 Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr Ala Pro Glu Met Ala Tyr
 210 215 220

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

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

<210> 23

<211> 338

<212> PRT

<213> Pseudomonas mendocina

<400> 23

Met Lys Val Tyr Tyr Asp Lys Asp Cys Asp Leu Ser Ile Ile Gln Gly

1 5 10 15
 Lys Lys Val Ala Ile Ile Gly Tyr Gly Ser Gln Gly His Ala Gln Ala
 20 25 30
 Cys Asn Leu Lys Asp Ser Gly Val Asp Val Thr Ile Gly Leu Arg Lys

35 40 45
 Gly Ser Ala Thr Val Ala Lys Ala Glu Ala His Gly Leu Lys Val Thr
 50 55 60
 Asp Val Ala Ser Ala Val Ala Ala Asp Leu Val Met Ile Leu Thr

 65 70 75 80
 Pro Asp Glu Phe Gln Gly Gln Leu Tyr Lys Asn Glu Ile Glu Pro Asn
 85 90 95
 Ile Lys Lys Gly Ala Thr Leu Ala Phe Ser His Gly Phe Ala Ile His
 100 105 110
 Tyr Asn Gln Val Val Pro Arg Ala Asp Leu Asp Val Ile Met Ile Ala
 115 120 125
 Pro Lys Ala Pro Gly His Thr Val Arg Thr Glu Phe Val Lys Gly Gly

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

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

 260 265 270
 Ala Met Arg Asn Ala Leu Lys Arg Ile Gln Asp Gly Glu Tyr Ala Lys
 275 280 285

Met Phe Ile Ser Glu Gly Ala Thr Gly Tyr Pro Ser Met Thr Ala Lys
290 295 300

Arg Arg Asn Asn Ala Ala His Gly Ile Glu Val Ile Gly Glu Gln Leu
305 310 315 320

Arg Ala Met Met Pro Trp Ile Ala Ala Asn Lys Ile Val Asp Lys Thr

325 330 335

Lys Asn

<210> 24

<211> 336

<212> PRT

<213> Bacillus cereus

<400> 24

Met Ala Lys Val Tyr Tyr Glu Lys Asp Val Thr Val Asn Val Leu Lys
1 5 10 15

Glu Lys Lys Val Ala Ile Ile Gly Tyr Gly Ser Gln Gly His Ala His
20 25 30

Ala Gln Asn Leu Arg Asp Asn Gly Phe Asp Val Val Val Gly Leu Arg
35 40 45

Lys Gly Lys Ser Trp Asp Lys Ala Lys Glu Asp Gly Phe Ser Val Tyr
50 55 60

Thr Val Ala Glu Ala Ala Lys Gln Ala Asp Val Val Met Ile Leu Leu
65 70 75 80

Pro Asp Glu Leu Gln Pro Glu Val Tyr Glu Ala Glu Ile Ala Pro Asn
85 90 95

Leu Gln Ala Gly Asn Ser Leu Val Phe Ala His Gly Phe Asn Val His
100 105 110

Phe Asp Gln Val Lys Pro Pro Ala Asn Val Asp Val Phe Leu Val Ala
115 120 125

Pro Lys Gly Pro Gly His Leu Val Arg Arg Thr Phe Ser Glu Gly Gly
130 135 140

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

145 150 155 160
 Thr Glu Lys Ala Leu Ser Tyr Ala Asp Gly Ile Gly Ala Thr Arg Ala
 165 170 175

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

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

 Asn Glu Lys Glu Asn Glu His Glu Ile Glu Val Val Gly Arg Lys Leu
 305 310 315 320
 Arg Glu Met Met Pro Phe Val Gln Pro Arg Val Lys Val Gly Met Lys
 325 330 335

 <210> 25
 <211> 335
 <212> PRT
 <213> Bacillus cereus
 <400> 25
 Met Lys Thr Tyr Tyr Glu Lys Asp Ala Asn Val Glu Leu Leu Lys Gly
 1 5 10 15
 Lys Thr Val Ala Val Ile Gly Tyr Gly Ser Gln Gly His Ala Gln Ala

20 25 30
 Gln Asn Leu Arg Asp Ser Gly Val Glu Val Val Val Gly Val Arg Pro
 35 40 45
 Gly Lys Ser Phe Glu Val Ala Lys Thr Asp Gly Phe Glu Val Met Ser
 50 55 60
 Val Ser Glu Ala Val Arg Thr Ala Gln Val Val Gln Met Leu Leu Pro
 65 70 75 80
 Asp Glu Gln Gln Ala His Val Tyr Lys Ala Gly Val Glu Glu Asn Leu
 85 90 95
 Arg Glu Gly Gln Met Leu Leu Phe Ser His Gly Phe Asn Ile His Phe
 100 105 110
 Gly Gln Ile Asn Pro Pro Ser Tyr Val Asp Val Ala Met Val Ala Pro
 115 120 125
 Lys Ser Pro Gly His Leu Val Arg Arg Val Phe Gln Glu Gly Asn Gly
 130 135 140
 Val Pro Ala Leu Val Ala Val His Gln Asp Ala Thr Gly Thr Ala Leu
 145 150 155 160
 His Val Ala Leu Ala Tyr Ala Lys Gly Val Gly Cys Thr Arg Ala Gly
 165 170 175
 Val Ile Glu Thr Thr Phe Gln Glu Glu Thr Glu Thr Asp Leu Phe Gly
 180 185 190
 Glu Gln Thr Val Leu Cys Gly Gly Val Thr Ala Leu Val Lys Ala Gly
 195 200 205
 Phe Glu Thr Leu Thr Glu Gly Gly Tyr Arg Pro Glu Ile Ala Tyr Phe
 210 215 220
 Glu Cys Leu His Glu Leu Lys Leu Ile Val Asp Leu Met Tyr Glu Gly
 225 230 235 240
 Gly Leu Thr Asn Met Arg His Ser Ile Ser Asp Thr Ala Glu Phe Gly
 245 250 255
 Asp Tyr Val Thr Gly Ser Arg Ile Val Thr Asp Glu Thr Lys Lys Glu
 260 265 270

Met Lys Arg Val Leu Thr Glu Ile Gln Gln Gly Glu Phe Ala Lys Lys

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

<210> 26

<211> 1032

<212> DNA

<213> Anaerostipes caccae

<400> 26

atggaagaat gtaagatggc taagatttac taccaagaag actgtaactt gtccttggtg	60
gatggtaaga ctatcgccgt tatcggttac ggttctcaag gtcacgctca tgccctgaat	120
gctaaggaat ccggttgtaa cgttatcatt ggtttatagc aaggttctaa gtcctggaaa	180
agagctgaag aacaaggttt cgaagtctac accgctgctg aagctgctaa gaaggctgac	240
atcattatga tcttgatcaa cgatgaaaag caggctacca tgtacaaaaa cgacatcgaa	300
ccaaacttgg aagccggtaa catgttgatg ttcgctcacg gtttcaacat ccatttcggt	360
tgtattgttc caccaaagga cgttgatgtc actatgatcg ctccaaaggg tccaggtcac	420
accgttagat ccgaatacga agaaggtaaa ggtgtcccat gcttggttgc tgtcgaacaa	480
gacgctactg gcaaggtttt ggatatggct ttggcctacg ctttagccat cgggtggtgct	540
agagccgggt tcttggaac taccttcaga accgaaactg aaaccgactt gttcggtgaa	600
caagctgttt tatgtgggtg tgtctgcgct ttgatgcagg ccggttttga aaccttggtt	660
gaagccggtt acgaccaag aaacgcttac ttcgaatgta tccacgaaat gaagttgatc	720
gttgacttga tctaccaatc tggtttctcc ggtatgcgtt actctatctc caacactgct	780
gaatacgggt actacattac cggtcctaaag atcattactg aagatacaca gaaggctatg	840
aagaagattt tgcttgacat tcaagatggt acctttgcca aggacttctt ggttgacatg	900
tctgatgctg gttcccaggt ccacttcaag gctatgagaa agttggcctc cgaacaccca	960
gctgaagttg tcggtgaaga aattagatcc ttgtactcct ggtccgacga agacaagttg	1020
attaacaact ga	1032

<210> 27

<211> 343

<212> PRT

<213> *Anaerostipes caccae*

<400> 27

Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn

1 5 10 15

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

20 25 30

Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val

35 40 45

Ile Ile Gly Leu Tyr Glu Gly Ser Lys Ser Trp Lys Arg Ala Glu Glu

50 55 60

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

65 70 75 80

Ile Ile Met Ile Leu Ile Asn Asp Glu Lys Gln Ala Thr Met Tyr Lys

85 90 95

Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala

100 105 110

His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val

115 120 125

Asp Val Thr Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser

130 135 140

Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln

145 150 155 160

Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala

165 170 175

Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu

180 185 190

Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val

195 200 205

Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr
 210 215 220
 Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile
 225 230 235 240
 Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile
 245 250 255
 Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile
 260 265 270

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

Glu Asp Lys Leu Ile Asn Asn
 340

<210> 28

<211> 1029

<212> DNA

<213> Artificial sequence

<220><223> Anaerostipes caccae KARI-AB1D1 variant

<400> 28

atggaagaat gtaagatggc taagatttac taccaagaag actgtaactt gtccttgttg	60
gatggtaaga ctatcgccgt tatcggttac ggttctcaag gtcacgtca tgccctgaat	120
gctaaggaat ccggttgtaa cgttatcatt ggtttatacg aagggtcgaa gtcctggaaa	180
agagctgaag aacaaggttt cgaagtctac accgtgctg aagctgctaa gaaggtgac	240
atcattatga tcttgatcaa cgatgaaaag caggctacca tgtacaaaaa cgacatcgaa	300
ccaaacttgg aagccggtaa catgttgatg ttcgtcacg gtttcaacat ccatttcggt	360
tgtattgttc caccaaagga cgttgatgtc actatgatcg ctccaaaggg tccaggtcac	420
accgttagat ccgaatacga agaaggtaaa ggtgtcccat gcttggttgc tgtcgaacaa	480

gacgctactg gcaaggcttt ggatatggct ttggcctacg ctttagccat cggtagtgct 540
 agagccggtg tcttggaac taccttcaga accgaaactg aaaccgactt gttcggtgaa 600
 caagctgttt tatgtggagg tgtctgcgct ttgatgcagg ccggttttga aaccttggtt 660
 gaagccggtt acgacccaag aaacgcttac ttcgaatgta tccacgaaat gaagttgatc 720

gttgacttga tctaccaatc tggtttctcc ggtagcggtt actctatctc caacactgct 780
 gaatacgggtg actacattac cggtcctaaag atcattactg aagataccaa gaaggtatg 840
 aagaagattt tgtctgacat tcaagatggg acctttgcca aggacttctt ggttgacatg 900
 tctgatgctg gttcccaggc ccacttcaag gctatgagaa agttggcctc cgaacaccca 960
 gctgaagttg tcggtgaaga aattagatcc ttgtactcct ggtccgacga agacaagttg 1020
 attaacaac 1029

<210> 29

<211> 343

<212> PRT

<213> Artificial sequence

<220><223> Anaerostipes caccae KARI-AB1D1 variant

<400> 29

Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn

1 5 10 15

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

20 25 30

Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val

35 40 45

Ile Ile Gly Leu Tyr Glu Gly Ala Lys Ser Trp Lys Arg Ala Glu Glu

50 55 60

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

65 70 75 80

Ile Ile Met Ile Leu Ile Asn Asp Glu Lys Gln Ala Thr Met Tyr Lys

85 90 95

Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala

100 105 110

His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val

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

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

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

305 310 315 320
 Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
 325 330 335
 Glu Asp Lys Leu Ile Asn Asn

340

<210> 30

<211> 1029

<212> DNA

<213> Artificial sequence

<220><223> Anaerostipes caccae KARI-495B5 variant

<400> 30

atggaagaat gtaagatggc taagatttac taccaagaag actgtaactt gtccttggtg 60

gatggtaaga ctatcgccgt tatcggttac ggttctcaag gtcacgtca tgcctgaat 120

gctaaggaat cgggttgtaa cggttatcatt gggtttatagc aaggtgctaa gcattggaaa 180

agagctgaag aacaagggtt cgaagtctac accgctgctg aagctgctaa gaaggctgac 240

atcattatga tcttgatcaa cgatgaaaag caggctacca tgtacaaaaa cgacatcgaa 300

ccaaacttgg aagccggtaa catgttgatg ttcgctcacg gtttcaacat ccatttcggt 360

tgtattgttc caccaaagga cggtgatgtc actatgatcg ctccaaaggg tccaggtcac 420

accgttagat ccgaatacga agaaggtaaa ggtgtcccat gcttggttgc tgcgaacaa 480

gacgctactg gcaaggcttt ggatatggct ttggcctacg ctttagccat cggtggtgct 540

agagccggtg tcttggaac taccttcaga accgaaactg aaaccgactt gttcggtgaa 600

caagctgttt tatgtgggtg tgtctgcgtt ttgatgcagg ccggttttga aaccttggtt 660

gaagccggtt acgacccaag aaacgcttac ttcgaatgta tccacgaaat gaagttgatc 720

gttgacttga tctaccaatc tggtttctcc ggtatgcgtt actctatctc caacactgct 780

gaatacgggt actacattac cgggtccaaag atcattactg aagatacaca gaaggtatg 840

aagaagattt tgtctgacat tcaagatggt acctttgcca aggacttctt ggttgacatg 900

tctgatgctg gttcccaggt ccacttcaag gctatgagaa agttggcctc cgaacaccca 960

gctgaagttg tcggtgaaga aattagatcc ttgtactcct ggtccgacga agacaagttg 1020

attaacaac 1029

<210> 31

<211> 343

<212> PRT

<213> Artificial sequence

<220><223> Anaerostipes caccae KARI 495B5 variant

<400> 31

Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn

1 5 10 15

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

20 25 30

Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val
35 40 45

Ile Ile Gly Leu Tyr Glu Gly Ala Lys His Trp Lys Arg Ala Glu Glu
50 55 60

Gln Gly Phe Glu Val Tyr Thr Ala Ala Glu Ala Ala Lys Lys Ala Asp
65 70 75 80

Ile Ile Met Ile Leu Ile Asn Asp Glu Lys Gln Ala Thr Met Tyr Lys
85 90 95

Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala
100 105 110

His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val
115 120 125

Asp Val Thr Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser
130 135 140

Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln
145 150 155 160

Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala
165 170 175

Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu
180 185 190

Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val
195 200 205

Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr
210 215 220

Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile
225 230 235 240

Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile
245 250 255

Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile
260 265 270

Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln

	275		280		285	
Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly						
290		295		300		
Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro						
305		310		315		320
Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp						
	325		330		335	
Glu Asp Lys Leu Ile Asn Asn						
	340					
<210>	32					
<211>	1029					
<212>	DNA					
<213>	Artificial sequence					
<220><223>	Anaerostipes caccae KARI-AB1D3 variant					
<400>	32					
atggaagaat gtaagatggc taagatttac taccaagaag actgtaactt gtccttgttg						60
gatggtaaga ctatcgccgt tatcggttac ggttctcaag gtcacgctca tgccttgaat						120
gctaaggaat ccggttgtaa cgttatcatt ggtttatacg aagggtctaa ggattggaaa						180
agagctgaag aacaaggttt cgaagtctac accgctgctg aagctgctaa gaaggctgac						240
atcattatga tcttgatcaa cgatgaaaag caggctacca tgtacaaaaa cgacatcgaa						300
ccaaacttgg aagccggtaa catgttgatg ttcgctcacg gtttcaacat ccatttcggt						360
tgtattgttc caccaaagga cgttgatgtc actatgatcg ctccaaaggg tccaggtcac						420
accgttagat ccgaatacga agaaggtaaa ggtgtcccat gcttggttgc tgtcgaacaa						480
gacgctactg gcaaggcttt ggatatggct ttggcctacg ctttagccat cgggtggtgct						540
agagccggtg tcttggaaac taccttcaga accgaaactg aaaccgactt gttcggtgaa						600
caagctgttt tatgtggtgg tgtctgcgct ttgatgcagg ccggttttga aaccttggtt						660
gaagccggtt acgaccaag aaacgcttac ttcgaatgta tccacgaaat gaagttgatc						720
gttgacttga tctaccaate tggtttctcc ggtatgcgtt actctatctc caaactgct						780
gaatacggtg actacattac cggtccaaag atcattactg aagataccaa gaaggctatg						840
aagaagattt tgtctgacat tcaagatggg acctttgccca aggacttctt ggttgacatg						900
tctgatgctg gtccccaggt ccacttcaag gctatgagaa agttggcctc cgaacacca						960

gctgaagttg tcggtgaaga aattagatcc ttgtactcct ggtccgacga agacaagttg 1020
 attaacaac 1029
 <210> 33
 <211> 343
 <212> PRT
 <213> Artificial sequence
 <220><223> Anaerostipes caccae KARI-AB1D3 variant
 <400> 33

Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn
 1 5 10 15
 Leu Ser Leu Leu Asp Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser
 20 25 30
 Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val
 35 40 45
 Ile Ile Gly Leu Tyr Glu Gly Ala Lys Asp Trp Lys Arg Ala Glu Glu
 50 55 60

Gln Gly Phe Glu Val Tyr Thr Ala Ala Glu Ala Ala Lys Lys Ala Asp
 65 70 75 80
 Ile Ile Met Ile Leu Ile Asn Asp Glu Lys Gln Ala Thr Met Tyr Lys
 85 90 95
 Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala
 100 105 110
 His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val
 115 120 125

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

180	185	190	
Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val			
195	200	205	
Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr			
210	215	220	
Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile			
225	230	235	240
Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile			
245	250	255	
Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile			
260	265	270	
Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln			
275	280	285	
Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly			
290	295	300	
Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro			
305	310	315	320
Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp			
325	330	335	
Glu Asp Lys Leu Ile Asn Asn			
340			
<210>	34		
<211>	1029		
<212>	DNA		
<213>	Artificial sequence		
<220><223>	Anaerostipes caccae KARI-AB1G9 variant		
<400>	34		
atggaagaat gtaagatggc taagatttac taccaagaag actgtaactt gtccttggtg		60	
gatggtgaaga ctatcgccgt tatcggttac ggttctcaag gtcacgtca tgcctgaat		120	
gctaaggaat ccggttgtaa cgttatcatt ggtttatcag aaggtgcgaa ggagtggaaa		180	
agagctgaag aacaaggttt cgaagtctac accgctgctg aagctgctaa gaaggtgac		240	

atcattatga tcttgatcaa cgatgaaaag caggctacca tgtacaaaaa cgacatcgaa 300
 ccaaacttgg aagccggtaa catgttgatg ttgcgtcacg gtttcaacat ccatttcggt 360
 tgtattgttc caccaaagga cgttgatgtc actatgatcg ctccaaaggg tccaggtcac 420
 accgttagat ccgaatacga agaaggtaaa ggtgtcccat gcttgggtgc tgcgaacaa 480
 gacgctactg gcaaggcttt ggatatggct ttggcctacg ctttagccat cggtggtgct 540
 agagccggtg tcttggaac taccttcaga accgaaactg aaaccgactt gttcggtgaa 600

caagctgttt tatgtggagg tgtctgcgct ttgatgcagg ccggttttga aaccttggtt 660
 gaagccggtt acgacccaag aaacgcttac ttcgaatgta tccacgaaat gaagttgatc 720
 gttgacttga tctaccaatc tggtttctcc ggtatgcgct actctatctc caacactgct 780
 gaatacgggtg actacattac cggtccaaag atcattactg aagatacaca gaaggctatg 840
 aagaagattt tctctgacat tcaagatggt acctttgcca aggacttctt ggttgacatg 900
 tctgatgctg gtccccaggt ccacttcaag gctatgagaa agttggcctc cgaacaccca 960
 gctgaagttg tcggtgaaga aattagatcc ttgtactcct ggtccgacga agacaagttg 1020

attaacaac 1029

<210> 35

<211> 343

<212> PRT

<213> Artificial sequence

<220><223> Anaerostipes caccae KARI AB1G9 variant

<400> 35

Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn

1 5 10 15

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

20 25 30

Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val

35 40 45

Ile Ile Gly Leu Tyr Glu Gly Ala Lys Glu Trp Lys Arg Ala Glu Glu

50 55 60

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

65 70 75 80

Ile Ile Met Ile Leu Ile Asn Asp Glu Lys Gln Ala Thr Met Tyr Lys

85 90 95

Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala

100 105 110

His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val

115 120 125

Asp Val Thr Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser

130 135 140

Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln

145 150 155 160

Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala

165 170 175

Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu

180 185 190

Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val

195 200 205

Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr

210 215 220

Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile

225 230 235 240

Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile

245 250 255

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

260 265 270

Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln

275 280 285

Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly

290 295 300

Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro

305 310 315 320

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

325 330 335

Glu Asp Lys Leu Ile Asn Asn

340

<210> 36

<211> 1029

<212> DNA

<213> Artificial sequence

<220><223> Anaerostipes caccae KARI-variant A07B5

<400> 36

```
atggaagaat gtaagatggc taagatttac taccaagaag actgtaactt gtccttggtg      60
gatggtaaga ctatcgccgt tatcggttac ggttctcaag gtcacgtca tgcctgaat      120
gctaaggaat ccggttgtaa cgttatcatt ggtttatagc aaggctgtaa ggattggaaa      180
agagctgaag aacaagggtt cgaagtctac accgctgctg aagctgctaa gaaggctgac      240
atcattatga tcttgggtcc tgatgaaaag caggctacca tgtacaaaaa cgacatcgaa      300
ccaaacttgg aagccggtaa catgttgatg ttcgctcacg gtttcaacat ccatttcggt      360
tgtattgttc caccaaagga cgttgatgtc actatgatcg ctccaaaggg tccaggctac      420
```

```
accgttagat ccgaatacga agaaggtaaa ggtgtcccat gcttggttgc tgtcgaacaa      480
gacgctactg gcaaggcttt ggatatggct ttggcctacg ctttagccat cgggtgtgct      540
agagccgggtg tcttggaac taccttcaga accgaaactg aaaccgactt gttcggtgaa      600
caagctgttt tatgtgggtg tgtctgcgct ttgatgcagg ccggttttga aaccttggtt      660
gaagccggtt acgaccaag aaacgcttac ttcgaatgta tccacgaaat gaagttgatc      720
gttgacttga tctaccaatc tggtttctcc ggtatgcgtt actctatctc caaactgct      780
gaatacgggtg actacattac cggtccaaag atcattactg aagatacaca gaaggctatg      840
```

```
aagaagattt tgtctgacat tcaagatggt acctttgcca aggacttctt ggttgacatg      900
tctgatgctg gtgccagggt ccacttcaag gctatgagaa agttggcctc cgaacaccca      960
gctgaagttg tcggtgaaga aattagatcc ttgtactcct ggtccgacga agacaagttg      1020
attaacaac                                     1029
```

<210> 37

<211> 343

<212> PRT

<213> Artificial sequence

<220><223> Anaerostipes caccae KARI A07B5 variant

<400> 37

Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn

1 5 10 15

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

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

 Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln
 145 150 155 160
 Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala
 165 170 175
 Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu
 180 185 190
 Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val
 195 200 205

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

Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile
260 265 270

Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln
275 280 285

Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly
290 295 300

Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro
305 310 315 320

Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
325 330 335

Glu Asp Lys Leu Ile Asn Asn
340

<210> 38

<211> 1029

<212> DNA

<213> Artificial sequence

<220><223> Anaerostipes caccae KARI-A07A9 variant

<400> 38

atggaagaat gtaagatggc taagatttac taccaagaag actgtaactt gtccttggtg 60
gatggtaaga ctatcgccgt tatcggttac ggttctcaag gtcacgtca tgccctgaat 120
gctaaggaat ccggttgtaa cgttatcatt ggtttatcag aaggctgtaa ggattggaaa 180
agagctgaag aacaagggtt cgaagtctac accgctgctg aagctgtaa gaagctgac 240
atcattatga tcttgacgcc tgatgaaaag caggctacca tgtacaaaa cgacatcgaa 300

ccaaacttgg aagccggtta catgttgatg ttcgctcag gtttcaacat ccatttcggt 360
tgtattgttc caccaaagga cgttgatgc actatgatcg ctccaaaggg tccaggtcac 420
accgttagat ccgaatacga agaaggtaaa ggtgtcccat gcttggttgc tgtcgaacaa 480
gacgctactg gcaaggcttt ggatatggct ttggcctacg ctttagccat cgggtgtgct 540
agagccggtg tcttggaac taccttcaga accgaaactg aaaccgactt gttcggtgaa 600
caagctgttt tatgtgggtg tgtctgcgct ttgatgcagg ccggttttga aaccttggtt 660
gaagccggtt acgaccaag aaacgcttac ttcgaatgta tccacgaaat gaagttgatc 720

gttgacttga tctaccaatc tggtttctcc ggtatgcgtt actctatctc caacactgct 780

gaatacgggtg actacattac cggtcctaaag atcattactg aagataccaa gaaggctatg 840
aagaagattt tgtctgacat tcaagatggt acctttgccca aggacttctt ggttgacatg 900
tctgatgctg gttcccaggt ccacttcaag gctatgagaa agttggcctc cgaacaccca 960
gctgaagttg tcggtgaaga aattagatcc ttgtactcct ggtccgacga agacaagttg 1020
attaacaac 1029

<210> 39

<211> 343

<212> PRT

<213> Artificial sequence

<220><223> Anaerostipes caccae KARI A07A9 variant

<400> 39

Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn

1 5 10 15

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

20 25 30

Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val

35 40 45

Ile Ile Gly Leu Tyr Glu Gly Ala Lys Asp Trp Lys Arg Ala Glu Glu

50 55 60

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

65 70 75 80

Ile Ile Met Ile Leu Thr Pro Asp Glu Lys Gln Ala Thr Met Tyr Lys

85 90 95

Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala

100 105 110

His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val

115 120 125

Asp Val Thr Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser

130 135 140

Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln

145 150 155 160

Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala

165 170 175
Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu

180 185 190
Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val

195 200 205
Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr

210 215 220
Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile

225 230 235 240
Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile

245 250 255
Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile

260 265 270
Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln

275 280 285
Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly

290 295 300
Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro

305 310 315 320
Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp

325 330 335
Glu Asp Lys Leu Ile Asn Asn

340

<210> 40

<211> 1029

<212> DNA

<213> Artificial sequence

<220><223> Anaerostipes caccae KARI-A07H8 variant

<400> 40

atggaagaat gtaagatggc taagatttac taccaagaag actgtaactt gtccttggtg 60

gatggtaaga ctatcgccgt tatcggttac ggttctcaag gtcacgtca tgccctgaat 120

gctaaggaat ccggttgtaa cgttatcatt ggtttatacg aaggtgctaa ggattggaaa 180
 agagctgaag aacaaggttt cgaagtctac accgctgctg aagctgctaa gaaggctgac 240
 atcattatga tcttgatccc cgatgaaaag caggctacca tgtacaaaaa cgacatcgaa 300
 ccaaacttgg aagccggtaa catgttgatg ttcgctcacg gtttcaacat ccatttcggt 360
 tgtattgttc caccaaagga cgttgatgtc actatgatcg ctccaaaggg tccaggtcac 420
 accgttagat ccgaatacga agaaggtaaa ggtgtcccat gcttggttgc tgcgaacaa 480
 gacgctactg gcaaggcttt ggatatggct ttggcctacg ctttagccat cggtaggtgct 540

agagccggtg tcttggaac taccttcaga accgaaactg aaaccgactt gttcggtgaa 600
 caagctgttt tatgtggagg tgtctgcgct ttgatgcagg ccggttttga aaccttggtt 660
 gaagccggtt acgaccaag aaacgcttac ttcgaatgta tccacgaaat gaagttgatc 720
 gttgacttga tctaccaatc tggtttctcc ggtatgcggt actctatctc caacactgct 780
 gaatacgggt actacattac cggtccaaag atcattactg aagatacaca gaaggctatg 840
 aagaagattt tgtctgacat tcaagatggt acctttgcca aggacttctt ggttgacatg 900
 tctgatgctg gttcccaggc ccacttcaag gctatgagaa agttggcctc cgaacaccca 960

gctgaagttg tcggtgaaga aattagatcc ttgtactcct ggtccgacga agacaagttg 1020
 attaacaac 1029

<210> 41

<211> 343

<212> PRT

<213> Artificial sequence

<220><223> Anaerostipes caccae KARI A07H8 variant

<400> 41

Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn

1 5 10 15

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

20 25 30

Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val

35 40 45

Ile Ile Gly Leu Tyr Glu Gly Ala Lys Asp Trp Lys Arg Ala Glu Glu

50 55 60

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

65 70 75 80

Ile Ile Met Ile Leu Ile Pro Asp Glu Lys Gln Ala Thr Met Tyr Lys
85 90 95

Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala
100 105 110

His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val
115 120 125

Asp Val Thr Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser
130 135 140

Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln
145 150 155 160

Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala
165 170 175

Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu
180 185 190

Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val
195 200 205

Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr
210 215 220

Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile
225 230 235 240

Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile
245 250 255

Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile
260 265 270

Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln
275 280 285

Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly
290 295 300

Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro
305 310 315 320

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

325	330	335
Glu Asp Lys Leu Ile Asn Asn		
340		
<210> 42		
<211> 1029		
<212> DNA		
<213> Artificial sequence		
<220><223> Anaerostipes caccae KARI-A07D8 variant		
<400> 42		
atggaagaat gtaagatggc taagatttac taccaagaag actgtaactt gtccttgttg	60	
gatggtaaga ctatcgccgt tatcggttac ggttctcaag gtcacgctca tgcctgaat	120	
gctaaggaat ccggttgtaa cgttatcatt ggtttatacg aaggtgctaa ggattggaaa	180	
agagctgaag aacaagggtt cgaagtctac accgctgctg aagctgctaa gaaggctgac	240	
atcattatga tcttgatcaa cgatgaaaag caggctacca tgtacaaaaa cgacatcgaa	300	
ccaaacttgg aagccggtaa catgttgatg ttcgctcacg gtttcaacat ccatttcggt	360	
tgtattgttc caccaaagga cgttgatgtc tgtatgatcg ctcaaaggg tccaggtcac	420	
accgttagat ccgaatacga agaaggtaaa ggtgtcccat gcttgggtgc tgtcgaacaa	480	
gacgctactg gcaaggcttt ggatatggct ttggcctacg ctttagccat cgggtgtgct	540	
agagccggtg tcttggaac taccttcaga tcggaaactg aaaccgactt gttcggtgaa	600	
caagctgttt tatgtgggtg tgtctgcgct ttgatgcagg ccggttttga aaccttggtt	660	
gaagccggtt acgacccaag aaacgcttac ttcgaatgta tccacgaaat gaagttgatc	720	
gttgacttga tctaccaatc tggtttctcc ggtatgcgtt actctatctc caacactgct	780	
gaatacgggtg actacattac cggtccaaag atcattactg aagataccaa gaaggctatg	840	
aagaagatgtt tgtctgacat tcaagatggt acctttgcca aggacttctt ggttgacatg	900	
tctgatgctg gttcccaggc ccacttcaag gctatgagaa agttggcctc cgaacaccca	960	
gctgaagttg tcggtgaaga aattagatcc ttgtactcct ggtccgacga agacaagttg	1020	
attaacaac	1029	
<210> 43		
<211> 343		
<212> PRT		
<213> Artificial sequence		
<220><223> Anaerostipes caccae KARI A07D8 variant		

<400> 43

```

Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn
1           5           10          15
Leu Ser Leu Leu Asp Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser
          20          25          30
Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val
          35          40          45
Ile Ile Gly Leu Tyr Glu Gly Ala Lys Asp Trp Lys Arg Ala Glu Glu
          50          55          60

Gln Gly Phe Glu Val Tyr Thr Ala Ala Glu Ala Ala Lys Lys Ala Asp
65           70           75           80
Ile Ile Met Ile Leu Ile Asn Asp Glu Lys Gln Ala Thr Met Tyr Lys
          85          90          95
Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala
          100         105         110
His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val
          115         120         125

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

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

```

Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile
245 250 255

Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile
260 265 270

Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln
275 280 285

Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly
290 295 300

Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro
305 310 315 320

Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
325 330 335

Glu Asp Lys Leu Ile Asn Asn
340

<210> 44

<211> 1029

<212> DNA

<213> Artificial sequence

<220><223> Anaerostipes caccae KARI- A07F7 variant

<400> 44

atggaagaat gtaagatggc taagatttac taccaagaag actgtaactt gtccttgttg 60
gatggtgaaga ctatcgccgt tatcggttac ggttctcaag gtcacgtca tgccctgaat 120
gctaaggaat ccggttgtaa cgttatcatt ggtttatacg aaggtgctaa ggattggaaa 180

agagctgaag aacaaggttt cgaagtctac accgctgctg aagctgctaa gaaggtgac 240
atcattatga tcttgatcaa cgatgaaaag caggctacca tgtacaaaa cgacatcgaa 300
ccaaacttgg aagccggtaa catgttgatg ttcgctcacg gtttcaacat ccatttcggt 360
tgtattgttc caccaaagga cgttgatgtc gttatgatcg ctccaaaggg tccaggtcac 420
accgttagat ccgaatacga agaaggtaaa ggtgtcccat gcttggttgc tgtcgaacaa 480
gacgctactg gcaaggcttt ggatatggct ttggcctacg ctttagccat cgggtgtgct 540
agagccggtg tcttggaaac taccttcaga gctgaaactg aaaccgactt gttcggtgaa 600

caagctgttt tatgtgggtg tgtctgcgct ttgatgcagg ccggttttga aaccttggtt 660

gaagccggtt acgacccaag aaacgcttac ttcgaatgta tccacgaaat gaagttgatc 720
gttgacttga tctaccaatc tggtttctcc ggtatgcgtt actctatctc caacactgct 780
gaatacgggtg actacattac cgggtccaaag atcattactg aagataccaa gaaggctatg 840
aagaagattt tgictgacat tcaagatggt acctttgcca aggacttctt gggtgacatg 900
tctgatgctg gttcccaggt ccacttcaag gctatgagaa agttggcctc cgaacaccca 960
gctgaagttg tcggtgaaga aattagatcc ttgtactcct ggtccgacga agacaagttg 1020

attaacaac 1029

<210> 45

<211> 343

<212> PRT

<213> Artificial sequence

<220><223> Anaerostipes caccae KARI A07F7 variant

<400> 45

Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn

1 5 10 15

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

20 25 30

Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val

35 40 45

Ile Ile Gly Leu Tyr Glu Gly Ala Lys Asp Trp Lys Arg Ala Glu Glu

50 55 60

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

65 70 75 80

Ile Ile Met Ile Leu Ile Asn Asp Glu Lys Gln Ala Thr Met Tyr Lys

85 90 95

Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala

100 105 110

His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val

115 120 125

Asp Val Val Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser

130 135 140

Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln

145 150 155 160
Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala

165 170 175
Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Ala Glu

180 185 190
Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val

195 200 205
Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr

210 215 220
Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile

225 230 235 240
Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile

245 250 255
Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile

260 265 270
Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln

275 280 285
Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly

290 295 300
Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro

305 310 315 320
Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp

325 330 335
Glu Asp Lys Leu Ile Asn Asn

340

<210> 46

<211> 1029

<212> DNA

<213> Artificial sequence

<220><223> Anaerostipes caccae KARI-A07H7 variant

<400> 46

atggaagaat gtaagatggc taagatttac taccaagaag actgtaactt gtccttggtg 60
gatggaaga ctatcgccgt tateggttac ggttctcaag gtcacgtca tgcctgaat 120
gctaaggaat ccggttgtaa cgttatcatt ggtttatagc aaggtgctaa ggattggaaa 180
agagctgaag aacaaggttt cgaagtctac accgctgctg aagctgctaa gaaggctgac 240
atcattatga tcttgatcaa cgatgaaaag caggctacca tgtacaaaaa cgacatcgaa 300
ccaaacttgg aagccggtaa catgttgatg ttcgctcacg gtttcaacat ccatttcggt 360
tgtattgttc caccaaagga cgttgatgtc actatgatcg ctccaaaggg tccaggtcac 420

accgttagat ccgaatacga agaaggtaaa ggtgtcccat gcttggttgc tgcgaacaa 480
gacgctactg gcaaggcttt ggatatggct ttggcctacg ctttagccat cgggtgtgct 540
agagccggtg tcttggaac taccttcaga agtgaaactg aaaccgactt gttcggtgaa 600
caagctgttt tatgtggtgg tgtctgcgct ttgatgcagg ccggttttga aaccttggtt 660
gaagccggtt acgacccaag aaacgcttac ttcgaatgta tccacgaaat gaagttgatc 720
gttgacttga tctaccaatc tggtttctcc ggtatgcgtt actctatctc caacactgct 780
gaatacgggtg actacattac cgggtccaaag atcattactg aagataccaa gaaggctatg 840

aagaagattt tgctgacat tcaagatggt acctttgcca aggacttctt ggttgacatg 900
tctgatgctg gttcccaggt ccacttcaag gctatgagaa agttggcctc cgaacaccca 960
gtgaagttg tcggtgaaga aattagatcc ttgtactcct ggtccgacga agacaagttg 1020
attaacaac 1029

<210> 47

<211> 343

<212> PRT

<213> Artificial sequence

<220><223> Anaerostipes caccae-KARI A07H7 variant

<400> 47

Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn

1 5 10 15

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

20 25 30

Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val

35 40 45

Ile Ile Gly Leu Tyr Glu Gly Ala Lys Asp Trp Lys Arg Ala Glu Glu

50 55 60

Gln Gly Phe Glu Val Tyr Thr Ala Ala Glu Ala Ala Lys Lys Ala Asp
65 70 75 80

Ile Ile Met Ile Leu Ile Asn Asp Glu Lys Gln Ala Thr Met Tyr Lys
85 90 95

Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala
100 105 110

His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val
115 120 125

Asp Val Thr Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser
130 135 140

Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln
145 150 155 160

Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala
165 170 175

Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Ser Glu
180 185 190

Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val
195 200 205

Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr
210 215 220

Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile
225 230 235 240

Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile
245 250 255

Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile
260 265 270

Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln
275 280 285

Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly
290 295 300

Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro

<213> Artificial sequence

<220><223> Anaerostipes caccae KARI-AS6F1 variant

<400> 49

```

Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn
1           5           10           15
Leu Ser Leu Leu Asp Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser
           20           25           30
Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val
           35           40           45
Ile Ile Gly Leu Tyr Glu Gly Ala Lys Asp Trp Lys Arg Ala Glu Glu

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

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

           180          185          190
Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val
           195          200          205
Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr
           210          215          220

```

Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile
 225 230 235 240
 Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile

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

305 310 315 320
 Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
 325 330 335
 Glu Asp Lys Leu Ile Asn Asn
 340

<210> 50

<211> 1029

<212> DNA

<213> Artificial sequence

<220><223> Anaerostipes caccae AWG4 variant

<400> 50

atggaagaat gtaagatggc taagatttac taccaagaag actgtaactt gtccttggtg 60

gatggtaaga ctatcgccgt tategggttac ggttctcaag gtcacgtca tgcctgaat 120

gctaaggaat cgggttgtaa cggttatcatt gggtttatagc aaggtgctaa ggattggaaa 180

agagctgaag aacaagggtt cgaagtctac accgctgctg aagctgctaa gaaggctgac 240

atcattatga tcttggtgcc ggatgaaaag caggctacca tgtacaaaaa cgacatcgaa 300

ccaaacttgg aagccggtaa catgttgatg ttcgctcacg gtttcaacat ccatttcggt 360

tgtattgttc caccaaagga cggttgatgc atgatgatcg ctccaaaggg tccaggtcac 420

accgttagat ccgaatacga agaaggtaaa ggtgtcccat gcttggttgc tgcgaacaa 480

gacgctactg gcaaggcttt ggatatggct ttggcctacg ctttagccat cgggtggtgct 540

agagccggtg tcttggaac taccttcaga accgaaactg aaaccgactt gttcggtgaa 600
 caagctgttt tatgtgggtg tgtctgcgct ttgatgcagg ccggttttga aaccttggtt 660
 gaagccggtt acgaccaag aaacgcttac ttcgaatgta tccacgaaat gaagttgatc 720
 gttgacttga tctaccaatc tggtttctcc ggtatgcggt actctatctc caacactgct 780
 gaatacgggt actacattac cggtccaaag atcattactg aagatacaa gaaggctatg 840
 aagaagattt tgtctgacat tcaagatggt acctttgcca aggacttctt ggttgacatg 900
 tctgatgctg gttcccaggt ccacttcaag gctatgagaa agttggcctc cgaacaccca 960

gctgaagttg tcggtgaaga aattagatcc ttgtactcct ggtccgacga agacaagttg 1020
 attaacaac 1029

<210> 51

<211> 343

<212> PRT

<213> Artificial sequence

<220><223> Anaerostipes caccae KARI AW4G variant

<400> 51

Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn

1 5 10 15

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

20 25 30

Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val

35 40 45

Ile Ile Gly Leu Tyr Glu Gly Ala Lys Asp Trp Lys Arg Ala Glu Glu

50 55 60

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

65 70 75 80

Ile Ile Met Ile Leu Val Pro Asp Glu Lys Gln Ala Thr Met Tyr Lys

85 90 95

Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala

100 105 110

His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val

115 120 125

Asp Val Met Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser

130 135 140
 Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln
 145 150 155 160

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

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

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

 <210> 52
 <211> 1029
 <212> DNA
 <213> Artificial sequence

<220><223> Anaerostipes caccae AWB9 variant

<400> 52

atggaagaat gtaagatggc taagatttac taccaagaag actgtaactt gtccttggtg 60
gatggaaga ctatcgccgt tatcggttac ggttctcaag gtcacgtca tgccctgaat 120
gctaaggaat cgggttgtaa cggttatcatt ggtttatagc aagggtgtaa ggattggaaa 180
agagctgaag aacaagggtt cgaagtctac accgtgctg aagctgtaa gaaggtgac 240
atcattatga tcttggtgcc ggatgaaaag caggctacca tgtacaaaaa cgacatcgaa 300
ccaaacttgg aagccggtaa catgttgatg ttcgctcacg gtttcaacat ccatttcggt 360

tgtattgttc caccaaagga cggtgatgc gcaatgatcg ctccaaaggg tccaggtcac 420
accgttagat ccgaatacga agaaggtaaa ggtgtcccat gcttggttgc tgtcgaacaa 480
gacgctactg gcaaggcttt ggatatggct ttggcctacg ctttagccat cgggtgtgct 540
agagccggtg tcttggaac taccttcaga accgaaactg aaaccgactt gttcggtgaa 600
caagctgttt tatgtggtgg tgtctgcgct ttgatgcagg ccggttttga aaccttggtt 660
gaagccggtt acgaccaag aaacgcttac ttcgaatgta tccacgaaat gaagttgatc 720
gttgacttga tctaccaatc tggtttctcc ggtatgcgtt actctatctc caacactgct 780

gaatacgggtg actacattac cgggtccaaag atcattactg aagatacaca gaaggctatg 840
aagaagattt tgctgacat tcaagatggt accttgcca aggacttctt ggttgacatg 900
tctgatgctg gttcccaggt ccacttcaag gctatgagaa agttggcctc cgaacaccca 960
gctgaagttg tcggtgaaga aattagatcc ttgtactcct ggtccgacga agacaagttg 1020
attaacaac 1029

<210> 53

<211> 343

<212> PRT

<213> Artificial sequence

<220><223> Anaerostipes caccae KARI AWB9 variant

<400> 53

Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn
1 5 10 15
Leu Ser Leu Leu Asp Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser
20 25 30
Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val
35 40 45

Ile Ile Gly Leu Tyr Glu Gly Ala Lys Asp Trp Lys Arg Ala Glu Glu
50 55 60

Gln Gly Phe Glu Val Tyr Thr Ala Ala Glu Ala Ala Lys Lys Ala Asp
65 70 75 80

Ile Ile Met Ile Leu Val Pro Asp Glu Lys Gln Ala Thr Met Tyr Lys
85 90 95

Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala
100 105 110

His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val
115 120 125

Asp Val Ala Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser
130 135 140

Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln
145 150 155 160

Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala
165 170 175

Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu
180 185 190

Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val
195 200 205

Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr
210 215 220

Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile
225 230 235 240

Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile
245 250 255

Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile
260 265 270

Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln
275 280 285

Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly

290 295 300
 Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro
 305 310 315 320
 Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
 325 330 335
 Glu Asp Lys Leu Ile Asn Asn
 340
 <210> 54
 <211> 1029
 <212> DNA
 <213> Artificial sequence
 <220><223> Anaerostipes caccae KARI AWC1 variant
 <400> 54
 atggaagaat gtaagatggc taagatttac taccaagaag actgtaactt gtccttggtg 60
 gatggtaaga ctatcgccgt tatcggttac ggttctcaag gtcacgtca tgccctgaat 120
 gctaaggaat ccggttgtaa cgttatcatt ggtttatacg aagggtgctaa ggattggaaa 180
 agagctgaag aacaaggttt cgaagtctac accgtgctg aagctgctaa gaaggctgac 240
 atcattatga tcttggtgcc ggatgaaaag caggctacca tgtacaaaa cgacatcgaa 300
 ccaaacttg aagccggtaa catgttgatg ttcgtcacg gtttcaacat ccatttcggt 360
 tgtattgttc caccaaagga cgttgatgtc gttatgatcg ctccaaaggg tccaggtcac 420
 accgttagat ccgaatacga agaaggtaaa ggtgtcccat gcttggttgc tgcgaacaa 480
 gacgtactg gcaaggcttt ggatatggct ttggcctacg ctttagccat cgggtggtgct 540
 agagccggtg tcttggaac taccttcaga accgaaactg aaaccgactt gttcggtgaa 600
 caagctgttt tatgtggtgg tgtctgcgct ttgatgcagg ccggttttga aaccttggtt 660
 gaagccggtt acgaccaag aaacgcttac ttcgaatgta tccacgaaat gaagttgatc 720
 gttgacttga tctaccaatc tggtttctcc ggtatgcgtt actctatctc caacactgct 780
 gaatacgggt actacattac cggtcctaaag atcattactg aagataccta gaaggctatg 840
 aagaagattt tgctgacat tcaagatggt acctttgcca aggacttctt ggttgacatg 900
 tctgatgctg gttcccaggt ccacttcaag gctatgagaa agttggcctc cgaacacca 960
 gctgaagttg tcggtgaaga aattagatcc ttgtactcct ggtccgacga agacaagttg 1020
 attaacaac 1029

<210> 55

<211> 343

<212> PRT

<213> Artificial sequence

<220><223> Anaerostipes caccae KARI AWC1 variant

<400> 55

Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn

1 5 10 15

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

20 25 30

Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val

35 40 45

Ile Ile Gly Leu Tyr Glu Gly Ala Lys Asp Trp Lys Arg Ala Glu Glu

50 55 60

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

65 70 75 80

Ile Ile Met Ile Leu Val Pro Asp Glu Lys Gln Ala Thr Met Tyr Lys

85 90 95

Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala

100 105 110

His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val

115 120 125

Asp Val Val Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser

130 135 140

Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln

145 150 155 160

Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala

165 170 175

Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu

180 185 190

Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val

195 200 205

Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr

210 215 220

Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile

225 230 235 240

Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile

245 250 255

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

260 265 270

Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln

275 280 285

Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly

290 295 300

Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro

305 310 315 320

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

325 330 335

Glu Asp Lys Leu Ile Asn Asn

340

<210> 56

<211> 1029

<212> DNA

<213> Artificial sequence

<220><223> Anaerostipes caccae KARI AWF4 variant

<400> 56

atggaagaat gtaagatggc taagatttac taccaagaag actgtaactt gtccttggtg	60
gatggtaaga ctatcgccgt tatcggttac ggttctcaag gtcacgtca tgccctgaat	120
gctaaggaat cgggttgtaa cggttatcatt ggtttatagc aaggtgctaa ggattggaaa	180
agagctgaag aacaagggtt cgaagtctac accgctgctg aagctgctaa gaaggtgac	240
atcattatga tcttggtgcc ggatgaaaag caggctacca tgtacaaaaa cgacatcgaa	300
ccaaacttgg aagccggtag catgttgatg ttcgctcagc gtttcaacat ccatttcggt	360
tgtattgttc caccaaagga cgttgatgac gttatgatcg ctccaaaggg tccaggtcac	420

accgttagat ccgaatacga agaaggtaaa ggtgtcccat gcttggttgc tgtcgaacaa 480
gacgctactg gcaaggcttt ggatatggct ttggcctacg ctttagccat cgggtgtgct 540
agagccggtg tcttggaac taccttcaga accgaaactg aaaccgactt gttcggtgaa 600
caagctgttt taigtgttgg tgtctgcgct ttgatgcagg ccggttttga aaccttggtt 660
gaagccggtt acgaccaag aaacgcttac ttcgaatgta tccacgaaat gaagttgatc 720
gttgacttga tctaccaatc tggtttctcc ggtatgcgtt actctatctc caaactgct 780
gaatacgggtg actacattac cggtcctaaag atcattactg aagataccaa gaaggctatg 840

aagaagattt tgtctgacat tcaagatggt acctttgcca aggacttctt ggttgacatg 900
tctgatgctg gttcccaggc ccacttcaag gctatgagaa agttggcctc cgaacaccca 960
gctgaagttg tcggtgaaga aattagatcc ttgtactcct ggtccgacga agacaagttg 1020
attaacaac 1029

<210> 57

<211> 343

<212> PRT

<213> Artificial sequence

<220><223> Anaerostipes caccae KARI AWF4 variant

<400> 57

Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn

1 5 10 15

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

20 25 30

Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val

35 40 45

Ile Ile Gly Leu Tyr Glu Gly Ala Lys Asp Trp Lys Arg Ala Glu Glu

50 55 60

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

65 70 75 80

Ile Ile Met Ile Leu Val Pro Asp Glu Lys Gln Ala Thr Met Tyr Lys

85 90 95

Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Ser Met Leu Met Phe Ala

100 105 110

His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val

115 120 125
 Asp Val Val Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser
 130 135 140

Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln
 145 150 155 160
 Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala
 165 170 175
 Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu
 180 185 190
 Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val
 195 200 205

Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr
 210 215 220
 Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile
 225 230 235 240
 Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile
 245 250 255
 Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile
 260 265 270

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

Glu Asp Lys Leu Ile Asn Asn
 340

<210> 58

<211> 1029

<212> DNA

<213> Artificial sequence

<220><223> Anaerostipes caccae KARI AWF6 variant

<400> 58

atggaagaat gtaagatggc taagatttac taccaagaag actgtaactt gtccttggtg	60
gatggtaaga ctatgccgt tatcggttac ggttctcaag gtcacgtca tgcctgaat	120
gctaaggaat ccggttgtaa cgttatcatt ggtttatagc aaggtgctaa ggattggaaa	180
agagctgaag aacaagggtt cgaagtctac accgctgctg aagctgctaa gaaggtgac	240
atcattatga tcttggtgcc ggatgaaaag caggctacca tgtacaaaaa cgacatcgaa	300
ccaaacttgg aagccggtaa catgttgatg ttcgctcacg gtttcaacat ccatttcggt	360
tgtattgttc caccaaagga cgttgatgtc gttatgatcg ctccaaaggg tccaggtcac	420
accgttagat ccgaatacga agaaggtaaa ggtgtcccat gcttggttgc tgcgaacaa	480
gacgctactg gcaaggcttt ggatatggct ttggcctacg ctttagccat cgggtgtgct	540
agagccggtg tcttggaaac taccttcaga gacgaaactg aaaccgactt gttcggtgaa	600
caagctgttt taigtggagg tgtctgcgct ttgatgcagg ccggttttga aaccttggtt	660
gaagccggtt acgaccaag aaacgcttac ttcgaatgta tccacgaaat gaagttgatc	720
gttgacttga tctaccaatc tggtttctcc ggtagcgtt actctatctc caacactgct	780
gaatacgggtg actacattac cggtccaaag atcattactg aagatacaca gaaggctatg	840
aagaagattt tgtctgacat tcaagatggt acctttgcca aggacttctt ggttgacatg	900
tctgatgctg gttcccaggc ccacttcaag gctatgagaa agttggcctc cgaacaccca	960
gctgaagttg tcggtgaaga aattagatcc ttgtactcct ggtccgacga agacaagttg	1020
attaacaac	1029

<210> 59

<211> 343

<212> PRT

<213> Artificial sequence

<220><223> Anaerostipes caccae KARI AWF6 variant

<400> 59

Met	Glu	Glu	Cys	Lys	Met	Ala	Lys	Ile	Tyr	Tyr	Gln	Glu	Asp	Cys	Asn
1					5						10				15
Leu	Ser	Leu	Leu	Asp	Gly	Lys	Thr	Ile	Ala	Val	Ile	Gly	Tyr	Gly	Ser
				20				25							30

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

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

305 310 315 320
 Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
 325 330 335
 Glu Asp Lys Leu Ile Asn Asn

340

<210> 60

<211> 1029

<212> DNA

<213> Artificial sequence

<220><223> Anaerostipes caccae KARI AWH3 variant

<400> 60

atggaagaat gtaagatggc taagatttac taccaagaag actgtaactt gtccttgttg 60

gatggtaaga ctatcgccgt tatcggttac ggttctcaag gtcacgtca tgccctgaat 120

gctaaggaat ccggttgtaa cgttatcatt ggtttatcag aaggtagtaa ggattggaaa 180

agagctgaag aacaagggtt cgaagtctac accgctgctg aagctgctaa gaaggctgac 240

atcattatga tcttggtgcc ggatgaaaag caggctacca tgtacaaaa cgacatcgaa 300

ccaaacttgg aagccggtaa catgttgatg ttgctcacg gtttcaacat ccatttcggt 360

tgtattgttc caccaaagga cgttgatgtc gttatgatcg ctccaaaggg tccaggtcac 420

accgttagat ccgaatacga agaaggtaaa ggtgtcccat gcttgggttc tgcgaacaa 480

gacgctactg gcaaggcttt ggatatggct ttggcctacg ctttagccat cggtaggtgct 540

agagccggtg tcttggaac taccttcaga ggcgaaactg aaaccgactt gttcgtgaa 600

caagctgttt tatgtgggtg tgtctgcgct ttgatgcagg ccggttttga aaccttggtt 660

gaagccggtt acgaccaag aaacgcttac ttgaatgta tccacgaaat gaagttgatc 720

gttgacttga tctaccaatc tggtttctcc ggtatgcgtt actctatctc caaacctgct 780

gaatacgggtg actacattac cgggtccaaag atcattactg aagatacga gaaggctatg 840

aagaagattt tgtctgacat tcaagatggt accttgcca aggacttctt ggttgacatg 900

tctgatgctg gttcccagggt ccacttcaag gctatgagaa agttggcctc cgaacacca 960

gctgaagttg tcggtgaaga aattagatcc ttgtactcct ggtccgacga agacaagttg 1020
attaacaac 1029

<210> 61
<211> 343
<212> PRT
<213> Artificial sequence
<220><223> Anaerostipes caccae KARI AWH3 variant
<400> 61

Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn
1 5 10 15
Leu Ser Leu Leu Asp Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser
20 25 30

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

Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala
100 105 110
His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val
115 120 125
Asp Val Val Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser
130 135 140
Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln
145 150 155 160

Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala
165 170 175
Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Gly Glu
180 185 190

Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val
195 200 205

Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr
210 215 220

Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile
225 230 235 240

Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile
245 250 255

Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile
260 265 270

Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln
275 280 285

Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly
290 295 300

Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro
305 310 315 320

Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
325 330 335

Glu Asp Lys Leu Ile Asn Asn
340

<210> 62

<211> 1029

<212> DNA

<213> Artificial sequence

<220><223> Anaerostipes caccae KARI AWD6 variant

<400> 62

atggaagaat gtaagatggc taagatttac taccaagaag actgtaactt gtccttggtg	60
gatggaaga ctatcgccgt tatcggttac ggttctcaag gtcacgtca tgccctgaat	120
gctaaggaat cgggttgtaa cgttatcatt ggtttatagc aaggctgtaa ggattggaaa	180
agagctgaag aacaaggttt cgaagtctac accgctgctg aagctgctaa gaaggctgac	240
atcattatga tcttggtgcc ggatgaaaag caggctacca tgtacaaaaa cgacatcgaa	300

ccaaacttgg aagccggtaa catgttgatg ttcgtcacg gtttcaacat ccatttcggt 360

tgtattgttc caccaaagga cgttgatgtc gttatgatcg ctcaaaggg tccaggtcac 420

accgttagat ccgaatacga agaaggtaaa ggtgtcccat gcttggttgc tgcgaacaa 480

gacgctactg gcaaggcttt ggatatggct ttggcctacg ctttagccat cgggtgtgct 540

agagccggtg tcttggaac taccttcaga tcggaaactg aaaccgactt gttcgtgaa 600

caagctgttt tatgtggagg tgtctgcgct ttgatgcagg ccggttttga aaccttggtt 660

gaagccggtt acgaccaag aaacgcttac ttcgaatga tccacgaaat gaagttgatc 720

gttgacttga tctaccaatc tggtttctcc ggtatgcgtt actctatctc caacactgct 780

gaatacgggtg actacattac cggtcctaaag atcattactg aagatacctg gaaggctatg 840

aagaagattt tgtctgacat tcaagatggt acctttgcca aggacttctt ggttgacatg 900

tctgatgctg gttcccaggc ccacttcaag gctatgagaa agttggcctc cgaacaccca 960

gctgaagtgg tcggtgaaga aattagatcc ttgtactcct ggtccgacga agacaagttg 1020

attaacaac 1029

<210> 63

<211> 343

<212> PRT

<213> Artificial sequence

<220><223> Anaerostipes caccae KARI AWD6 variant

<400> 63

Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn

1 5 10 15

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

20 25 30

Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val

35 40 45

Ile Ile Gly Leu Tyr Glu Gly Ala Lys Asp Trp Lys Arg Ala Glu Glu

50 55 60

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

65 70 75 80

Ile Ile Met Ile Leu Val Pro Asp Glu Lys Gln Ala Thr Met Tyr Lys

85 90 95

Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala
 100 105 110
 His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val
 115 120 125

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

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

 Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile
 260 265 270
 Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln
 275 280 285
 Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly
 290 295 300
 Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro
 305 310 315 320

 Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
 325 330 335
 Glu Asp Lys Leu Ile Asn Asn

340

<210> 64

<211> 1029

<212> DNA

<213> Artificial sequence

<220><223> Anaerostipes caccae KARI AWD10 variant

<400> 64

atggaagaat gtaagatggc taagatttac taccaagaag actgtaactt gtccttggtg 60

gatggtaaga ctatgccgt tatcggttac ggttctcaag gtcacgtca tgccctgaat 120

gctaaggaat ccggttgtaa cgttatcatt ggtttatagc aagggtgtaa ggattggaaa 180

agagctgaag aacaagggtt cgaagtctac accgtgctg aagctgctaa gaagctgac 240

atcattatga tcttggtgcc ggatgaaaag caggctacca tgtacaaaaa cgacatcgaa 300

ccaaacttgg aagccggtaa catgttgatg ttcgctcacg gtttcaacat ccatttcggt 360

tgtattgttc caccaaagga cgttgatgtc gcaatgatcg ctccaaaggg tccaggtcac 420

accgttagat ccgaatacga agaaggtaaa ggtgtcccat gcttggttgc tgcgaacaa 480

gacgctactg gcaaggcttt ggatatggct ttggcctacg ctttagccat cgggtgtgct 540

agagccggtg tcttggaaac taccttcaga tgcgaaactg aaaccgactt gtccggtgaa 600

caagctgttt tatgtgggtg tgtctgcgct ttgatgcagg ccggttttga aaccttggtt 660

gaagccggtt acgaccaag aaacgcttac ttcgaatgta tccacgaaat gaagttgatc 720

gttgacttga tctaccaatc tggtttctcc ggtatgcgtt actctatctc caaactgct 780

gaatacgggt actacattac cgggtccaaag atcattactg aagatacaca gaaggctatg 840

aagaagattt tgtctgacat tcaagatggt acctttgcca aggacttctt ggttgacatg 900

tctgatgctg gtccccaggt ccacttcaag gctatgagaa agttggcctc cgaacaccca 960

gctgaagttg tcggtgaaga aattagatcc ttgtactcct ggtccgacga agacaagttg 1020

attaacaac 1029

<210> 65

<211> 343

<212> PRT

<213> Artificial sequence

<220><223> Anaerostipes caccae KARI AWD10 variant

<400> 65

Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn

1 5 10 15
 Leu Ser Leu Leu Asp Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser
 20 25 30
 Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val

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

 100 105 110
 His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val
 115 120 125
 Asp Val Ala Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser
 130 135 140
 Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln
 145 150 155 160
 Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala

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

 225 230 235 240
 Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile
 245 250 255

Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile
260 265 270

Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln
275 280 285

Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly

290 295 300

Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro

305 310 315 320

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

325 330 335

Glu Asp Lys Leu Ile Asn Asn

340

<210> 66

<211> 1002

<212> DNA

<213> Staphylococcus capitis

<400> 66

atgacaaaag ttiattatga tcaatcagtt gaaaaagatg ctttacaagg taaaaaaatt 60

gcagtaattg gttacgggtc acaaggacat gcacacgctc aaaattttaa agacaatggt 120

tatgatgtag tcgttggat tcgacctggt cattctttcg atagagctaa agaagatgga 180

tttgatgttt atccagttaa tgaggctact aaacaagcag atgtagtaat gattttatta 240

cctgatgaaa tccaaggtaa tgtatataaa aatgaaattg aacctaactt agaagcaggc 300

aatgcattag cttttgcgca cggatttaac attcactttg gagttattca acccctgct 360

gatgtagatg tattcttagt agcacctaaa ggtcctggtc atttagtaag acgtaccttt 420

gtagaaggta ctgctgtacc agcattattc ggtgtacaac aagatgattc aggtcatgca 480

cgtgatatcg cactaagcta tgccaaaggt attggtgcaa cacgtgctgg tgttattgaa 540

actacattca aagaagaac tgttactgat ttattcgggtg aacaagctgt actttgcggg 600

ggatatccata aattaattca aagtggtttt gaaacttttag tagaagctgg gtatcaaaaa 660

gagttagcat acttcgaagt attacatgaa atgaaattaa tcgttgattt aatgtacgaa 720

ggcggtatgg aaaacgttcg ctactctatt tcaaatagag ctgaatttgg tgattatgta 780

tctggaccac gtgttatcac tccagaagtt aaaaacaaca tgaaaactgt attagaagat 840

attcaaaacg gtaactttgc taatcgtttt gttaaagata atgaaaatgg ctttaaagaa 900

ttttaccaat tacgtgaaca acaacatggt catgaaatcg aagcggttgg tcgtgaatta 960

agaaaaatga tgccatttat taaagctaaa agtattcaaa aa 1002

<210> 67

<211> 334

<212> PRT

<213> Staphylococcus capitis

<400> 67

Met Thr Lys Val Tyr Tyr Asp Gln Ser Val Glu Lys Asp Ala Leu Gln

1 5 10 15

Gly Lys Lys Ile Ala Val Ile Gly Tyr Gly Ser Gln Gly His Ala His

20 25 30

Ala Gln Asn Leu Lys Asp Asn Gly Tyr Asp Val Val Val Gly Ile Arg

35 40 45

Pro Gly His Ser Phe Asp Arg Ala Lys Glu Asp Gly Phe Asp Val Tyr

50 55 60

Pro Val Asn Glu Ala Thr Lys Gln Ala Asp Val Val Met Ile Leu Leu

65 70 75 80

Pro Asp Glu Ile Gln Gly Asn Val Tyr Lys Asn Glu Ile Glu Pro Asn

85 90 95

Leu Glu Ala Gly Asn Ala Leu Ala Phe Ala His Gly Phe Asn Ile His

100 105 110

Phe Gly Val Ile Gln Pro Pro Ala Asp Val Asp Val Phe Leu Val Ala

115 120 125

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

130 135 140

Ala Val Pro Ala Leu Phe Gly Val Gln Gln Asp Asp Ser Gly His Ala

145 150 155 160

Arg Asp Ile Ala Leu Ser Tyr Ala Lys Gly Ile Gly Ala Thr Arg Ala

165 170 175

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

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

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

290 295 300
 Arg Glu Gln Gln His Gly His Glu Ile Glu Ala Val Gly Arg Glu Leu
 305 310 315 320
 Arg Lys Met Met Pro Phe Ile Lys Ala Lys Ser Ile Gln Lys
 325 330

<210> 68

<211> 1002

<212> DNA

<213> Staphylococcus epidermidis

<400> 68

atgacaaaag tttattatga tggatctgta actaaagacg cattacaagg taaaaaaatt 60

gcagtaatcg gatatggctc acaaggacat gccacgcgc aaaacttaaa agatagtggc 120

tatgatgtca tcatcggtat ccgctctgga cactcatacg ataaagctaa aaaagatggt 180

ttcgatgtgt atccagtaag tgaagctaca aaacaagcag atgtcggtat ggttttatta 240

ccagatgaaa tccaaggtaa cgtatacaaa aatgaaatcg cacctaactt agaagctggt 300

aatgccctag ctttcgcccc cggttttaac attcactttg gtgttattca acctcctgct 360

gatgtagatg tatttttagt agctcctaaa ggaccaggtc accttgtaag acgtacattc 420

gttgaaggca gtgcagtacc ttcattattt ggtgttcaac aagatgctac tggaaatgca 480

cgatgatatcg ctttaagcta tgccaaaggt attggtgcta cacgtgctgg tgttattgaa 540

acaacattta aagaagaaac tgaaactgat ttattcgggtg agcaagctgt actttgtggt 600

ggatatccata aattaattca aagtgggtttt gaaacttttag tagaagctgg ttatcaaaaa 660

gaattagctt acttgaagt attacatgaa atgaaattaa tcgttgattt aatgtatgaa 720

ggcggtatgg aaaacgttcg ctattctatt tcaaatactg ccgaatttgg tgattatgta 780

tctggaccac gtgtcatcac tccagaagtt aaaaacaaca tgaaaactgt attagaagat 840

atccaaaacg gtaactttgc taatcgtttt gttaaagata atgaaaatgg ctttaagaa 900

ttctatcaat tacgtgaaca acaacatggt catgaaattg aagcggtagg tcgtgagctt 960

cgtgaaatga tgccatttat taaatctaaa agcattgaaa aa 1002

<210> 69

<211> 334

<212> PRT

<213> Staphylococcus epidermidis

<400> 69

Met Thr Lys Val Tyr Tyr Asp Gly Ser Val Thr Lys Asp Ala Leu Gln

1 5 10 15

Gly Lys Lys Ile Ala Val Ile Gly Tyr Gly Ser Gln Gly His Ala His

20 25 30

Ala Gln Asn Leu Lys Asp Ser Gly Tyr Asp Val Ile Ile Gly Ile Arg

35 40 45

Pro Gly His Ser Tyr Asp Lys Ala Lys Lys Asp Gly Phe Asp Val Tyr

50 55 60

Pro Val Ser Glu Ala Thr Lys Gln Ala Asp Val Val Met Val Leu Leu

65 70 75 80

Pro Asp Glu Ile Gln Gly Asn Val Tyr Lys Asn Glu Ile Ala Pro Asn

85 90 95

Leu Glu Ala Gly Asn Ala Leu Ala Phe Ala His Gly Phe Asn Ile His

100 105 110

Phe Gly Val Ile Gln Pro Pro Ala Asp Val Asp Val Phe Leu Val Ala

115 120 125

Pro Lys Gly Pro Gly His Leu Val Arg Arg Thr Phe Val Glu Gly Ser

130 135 140
Ala Val Pro Ser Leu Phe Gly Val Gln Gln Asp Ala Thr Gly Asn Ala
145 150 155 160
Arg Asp Ile Ala Leu Ser Tyr Ala Lys Gly Ile Gly Ala Thr Arg Ala
165 170 175

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

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

Arg Glu Gln Gln His Gly His Glu Ile Glu Ala Val Gly Arg Glu Leu
305 310 315 320
Arg Glu Met Met Pro Phe Ile Lys Ser Lys Ser Ile Glu Lys
325 330

<210> 70

<211> 1002

<212> DNA

<213> Staphylococcus aureus

<400> 70

atgacaacag tttattatga tcaagacgta aaaacggacg ctttacaagg caaaaaaatt 60

gcagtagtag gttatggatc acaaggccac gcgcatgcac aaaacttaaa agacaatgga 120

tatgatgtag tcacgggat tcgcccaggt cgttcttttg acaaggctaa agaagatgga 180

tttgatgtgt tccctgttgc agaagcagtt aagcaagctg atgtaattat ggtgctatta 240

cctgatgaaa ttcaagggtga tgtatacaaa aacgaaattg aaccaaattt agaaaaacat 300

aatgtgcttg catttgctca tggctttaac attcattttg gtgttattca accaccagct 360

gatgttgatg tatttttagt agctcctaaa ggaccgggtc atttagttag acgtacattt 420

gttgaagggt ctgctgtacc atcactattt ggtattcaac aagacgcttc aggtcaagca 480

cgtaaatattg cttaagtta tgcaaaaggt attggtgcaa ctcgtgcagg tgttattgaa 540

acaacattta aagaagaaac tgagacagat ttatttggtg aacaagcagt actttgcggt 600

ggtgtatcga aattaattca aagtggcttt gaaacattag tagaagcggg ttatcaacca 660

gaattagctt attttgaagt attacatgaa atgaaattaa tcgttgattt gatgtatgaa 720

ggcggtatgg aaaaagtacg ttactcaatt tcaaatactg ctgaatttgg tgactatgtt 780

tcaggaccac gtgttatcac accagatgtt aaagaaaata tgaaagctgt attaactgat 840

atccaaaatg gtaacttcag taatcgcttt atcgaagaca ataaaaatgg attcaaagaa 900

ttttataaat tacgcgaaga acaacatggt catcaaattg aaaaagttgg tcgtgaatta 960

cgcgaaatga tgccttttat taaatctaaa agcattgaaa aa 1002

<210> 71

<211> 334

<212> PRT

<213> Staphylococcus aureus

<400> 71

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

1 5 10 15

Gly Lys Lys Ile Ala Val Val Gly Tyr Gly Ser Gln Gly His Ala His

20 25 30

Ala Gln Asn Leu Lys Asp Asn Gly Tyr Asp Val Val Ile Gly Ile Arg

35 40 45

Pro Gly Arg Ser Phe Asp Lys Ala Lys Glu Asp Gly Phe Asp Val Phe

50 55 60

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

65 70 75 80

Pro Asp Glu Ile Gln Gly Asp Val Tyr Lys Asn Glu Ile Glu Pro Asn

85 90 95
 Leu Glu Lys His Asn Val Leu Ala Phe Ala His Gly Phe Asn Ile His
 100 105 110
 Phe Gly Val Ile Gln Pro Pro Ala Asp Val Asp Val Phe Leu Val Ala

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

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

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

 305 310 315 320
 Arg Glu Met Met Pro Phe Ile Lys Ser Lys Ser Ile Glu Lys
 325 330

<210> 72

<211> 1002

<212> DNA

<213> Staphylococcus warneri

<400> 72

```

atgacaacag tttattatga tgaatcagta aaaacagacg ctttagaagg taaaaaaatt      60
gctgtgattg gatatgggtc acaaggccat gctcacgcac aaaacttaaa agataacgga      120
tatgatgttg ttatcggtat ccgtccagga cgctcttttg ataaagctaa ggaaaatggt      180
ttcgatgtat atcctgtggc agaagcaacg aaacaagctg acgttatcat ggttttaatt      240

ccagatgaaa tccaagggtga tgtatataaa aatgaaattg aacctaattt agaagctggt      300
aacgcacttg catttgctca tggtttcaat attcactttg gtgttattca accccctgca      360
gatgtcgatg ttttcttagt tgcacaaaaa ggaccaggtc acttagttag acgtacattt      420
gaagaaggta gtgcagtgcc atctttattt ggtgttcaac aagatgctac aggtcatgca      480
agagatattg cattaagtta tgcaaaaggt atcggatcta caagagctgg tgtaattgaa      540
acaacattta aagaagaaac agaaacggat ttatttgggtg aacaagctgt actatgtggc      600
ggtatccata aattaattca aagtggattt gaaacattag ttgaagctgg ttaccaacca      660

gaattagcct attttgaagt attacatgaa atgaaattaa tcgttgattt aatgtatgaa      720
ggcggtatgg ataatgtccg ttattcaatt tctaatactg cagagtatgg tgactatgta      780
tctggacgtc gtatcatcac tccagatgtc aaagacaata tgaaagctgt attaacagat      840
attcaaaacg gtaaatttgc taatagtttt gttgaagata ataaaaatgg atttaaagaa      900
ttctataaaa tgcgtgaaga acaagctggt catccaattg aaaaagtggg ccgtaaatta      960
cgtgatatga tgccatttat taaatcaaaa agcattgaaa ag                                1002

```

<210> 73

<211> 334

<212> PRT

<213> Staphylococcus warneri

<400> 73

```

Met Thr Thr Val Tyr Tyr Asp Glu Ser Val Lys Thr Asp Ala Leu Glu
1           5           10           15
Gly Lys Lys Ile Ala Val Ile Gly Tyr Gly Ser Gln Gly His Ala His
           20           25           30
Ala Gln Asn Leu Lys Asp Asn Gly Tyr Asp Val Val Ile Gly Ile Arg

```

35	40	45
Pro Gly Arg Ser Phe Asp Lys Ala Lys Glu Asn Gly Phe Asp Val Tyr		
50	55	60
Pro Val Ala Glu Ala Thr Lys Gln Ala Asp Val Ile Met Val Leu Ile		
65	70	75
80		
Pro Asp Glu Ile Gln Gly Asp Val Tyr Lys Asn Glu Ile Glu Pro Asn		
85	90	95
Leu Glu Ala Gly Asn Ala Leu Ala Phe Ala His Gly Phe Asn Ile His		
100	105	110
Phe Gly Val Ile Gln Pro Pro Ala Asp Val Asp Val Phe Leu Val Ala		
115	120	125
Pro Lys Gly Pro Gly His Leu Val Arg Arg Thr Phe Glu Glu Gly Ser		
130	135	140
Ala Val Pro Ser Leu Phe Gly Val Gln Gln Asp Ala Thr Gly His Ala		
145	150	155
160		
Arg Asp Ile Ala Leu Ser Tyr Ala Lys Gly Ile Gly Ser Thr Arg Ala		
165	170	175
Gly Val Ile Glu Thr Thr Phe Lys Glu Glu Thr Glu Thr Asp Leu Phe		
180	185	190
Gly Glu Gln Ala Val Leu Cys Gly Gly Ile His Lys Leu Ile Gln Ser		
195	200	205
Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr Gln Pro Glu Leu Ala Tyr		
210	215	220
Phe Glu Val Leu His Glu Met Lys Leu Ile Val Asp Leu Met Tyr Glu		
225	230	235
240		
Gly Gly Met Asp Asn Val Arg Tyr Ser Ile Ser Asn Thr Ala Glu Tyr		
245	250	255
Gly Asp Tyr Val Ser Gly Arg Arg Ile Ile Thr Pro Asp Val Lys Asp		
260	265	270
Asn Met Lys Ala Val Leu Thr Asp Ile Gln Asn Gly Lys Phe Ala Asn		
275	280	285

Ser Phe Val Glu Asp Asn Lys Asn Gly Phe Lys Glu Phe Tyr Lys Met
290 295 300

Arg Glu Glu Gln Ala Gly His Pro Ile Glu Lys Val Gly Arg Lys Leu
305 310 315 320

Arg Asp Met Met Pro Phe Ile Lys Ser Lys Ser Ile Glu Lys
325 330

<210> 74

<211> 1002

<212> DNA

<213> Staphylococcus epidermidis

<400> 74

atgacaaaag tatattacga tgaacagta actcaagatg cattacaagg taaaaaatt 60
gctgtcattg gttatggctc acaaggacat gcacatgcac aaaatttaaa ggacaatggt 120
tatgatgtag tcattggctc acgtccagga cgatcattta ataaagctaa agaagatgga 180
tttgatgttt atacggtaag tgatgctaca caacaagctg atgtagtgtat ggtactattg 240
cctgatgaaa ttcaaggtaa agtatataac gaggaaatta aaccacattt agaaaaagga 300

aatgccttag cattcgaca cggttttaat atccacttca gtgttatcga accacctagt 360
gatgtcgatg tctttttagt agcacctaaa ggaccaggtc atttagttag acgtacattt 420
gttgaaggaa gtgccgtacc agcattattt ggtgttcaac aagatgctac aggccaagca 480
agaaacattg ctttaagcta cgcaaaaggc attggtgcta ctctgtgccgg ggtcattgaa 540
acgacattta aagaagaac tgaacagat ttattcgggtg aacaagctgt tctttgtgga 600
ggagtttcca aattaattca gagtggtttc gaaacacttg tagaagcagg ttaccaacct 660
gaattagctt attttgaagt cttacacgaa atgaaattaa ttgttgattt aatgtatgaa 720

ggcggaatgg aaaacgttcg ttattctatc tctaacactg ctgaatttgg cgactatgtt 780
tctggaccaa gagtaattac acctgatgtt aaagaaaata tgaaaaaagt acttgaagat 840
attcaaatg gtaacttttag ccgtagattt gttgaagata acaaaaatgg ctttaaagaa 900
ttctatcaat tacgtgaaga tcaacatggt catcaaattg aacaagttgg acgtgaatta 960
agagaaatga tgccattcat taaatctaaa agtattgaaa aa 1002

<210> 75

<211> 334

<212> PRT

<213> Staphylococcus epidermidis

<400> 75

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

1 5 10 15

Gly Lys Lys Ile Ala Val Ile Gly Tyr Gly Ser Gln Gly His Ala His

20 25 30

Ala Gln Asn Leu Lys Asp Asn Gly Tyr Asp Val Val Ile Gly Leu Arg

35 40 45

Pro Gly Arg Ser Phe Asn Lys Ala Lys Glu Asp Gly Phe Asp Val Tyr

50 55 60

Thr Val Ser Asp Ala Thr Gln Gln Ala Asp Val Val Met Val Leu Leu

65 70 75 80

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

85 90 95

Leu Glu Lys Gly Asn Ala Leu Ala Phe Ala His Gly Phe Asn Ile His

100 105 110

Phe Ser Val Ile Glu Pro Pro Ser Asp Val Asp Val Phe Leu Val Ala

115 120 125

Pro Lys Gly Pro Gly His Leu Val Arg Arg Thr Phe Val Glu Gly Ser

130 135 140

Ala Val Pro Ala Leu Phe Gly Val Gln Gln Asp Ala Thr Gly Gln Ala

145 150 155 160

Arg Asn Ile Ala Leu Ser Tyr Ala Lys Gly Ile Gly Ala Thr Arg Ala

165 170 175

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

180 185 190

Gly Glu Gln Ala Val Leu Cys Gly Gly Val Ser Lys Leu Ile Gln Ser

195 200 205

Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr Gln Pro Glu Leu Ala Tyr

210 215 220

Phe Glu Val Leu His Glu Met Lys Leu Ile Val Asp Leu Met Tyr Glu

225 230 235 240

Gly Gly Met Glu Asn Val Arg Tyr Ser Ile Ser Asn Thr Ala Glu Phe

245 250 255

Gly Asp Tyr Val Ser Gly Pro Arg Val Ile Thr Pro Asp Val Lys Glu

260 265 270

Asn Met Lys Lys Val Leu Glu Asp Ile Gln Asn Gly Asn Phe Ser Arg

275 280 285

Arg Phe Val Glu Asp Asn Lys Asn Gly Phe Lys Glu Phe Tyr Gln Leu

290 295 300

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

305 310 315 320

Arg Glu Met Met Pro Phe Ile Lys Ser Lys Ser Ile Glu Lys

325 330

<210> 76

<211> 1002

<212> DNA

<213> Staphylococcus saprophyticus

<400> 76

atgacaacag ttattatga ccaatcagtt acgaaagacg ctttacaagg taaaaaagta 60
gcaatcatag gctatggttc tcaaggacac gcacatgctc aaaattttaa agacaacggt 120
tatgatgtca ttatcggaat cagacctgga tgttcttttg acaaagctaa agatgatggc 180
ttcgaagtat atcctgtaga tgaagcagca aaacaagctg atgtcatcat ggtgttatta 240
ccagatgaaa ttcaaggcca agtatataaa gaagaaattg aaccaaactt agaagcaaac 300

aatgcattag tattcgcgca tggttttaat attcatttcg gtgttattca accaccagaa 360
aacgtagatg tattcttagt agcgccctaaa ggacctggac atttagtacg tcgtacattt 420
gtgaaggaa gcgcagtcctc tgccctattc gcagttgaac aagatccgag tggatgaagct 480
agagatttag cattaagcta tgctaaaggt atcgggtgcaa cacgtgcagg tgtattagaa 540
acatcattta aagaagaac agaaacagat ttattcgggtg aacaagcagt gctttgtggt 600
ggtacgacta aattagtaca atctggtttc gaaacgtag tagaagcagg ttaccaacca 660
gaaattgcat actttgaagt attgcatgaa atgaaattga ttgttgattt aatgtatgaa 720

ggcggtatgg aaaatatgcg ctattcaatt tcaaatagc ctgaatttgg tgactatgtt 780
tctggaccac gtattatcac accggatgtt aaagataata tgaaagctgt attagatgat 840

attcaaaaag gaaacttcag tgatcgattc attaaagata atcaaaataa ttttgaagaa 900
 ttccataaat taagagaaga acaacatggt catcaaatcg aagcggttgg tagagaactc 960
 cgtgatatga tgccattcat caaatctaag agcattgaaa aa 1002

<210> 77

<211> 334

<212> PRT

<213> Staphylococcus saprophyticus

<400> 77

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

1 5 10 15
 Gly Lys Lys Val Ala Ile Ile Gly Tyr Gly Ser Gln Gly His Ala His

20 25 30
 Ala Gln Asn Leu Lys Asp Asn Gly Tyr Asp Val Ile Ile Gly Ile Arg

35 40 45
 Pro Gly Cys Ser Phe Asp Lys Ala Lys Asp Asp Gly Phe Glu Val Tyr

50 55 60
 Pro Val Asp Glu Ala Ala Lys Gln Ala Asp Val Ile Met Val Leu Leu

65 70 75 80
 Pro Asp Glu Ile Gln Gly Gln Val Tyr Lys Glu Glu Ile Glu Pro Asn

85 90 95
 Leu Glu Ala Asn Asn Ala Leu Val Phe Ala His Gly Phe Asn Ile His

100 105 110
 Phe Gly Val Ile Gln Pro Pro Glu Asn Val Asp Val Phe Leu Val Ala

115 120 125
 Pro Lys Gly Pro Gly His Leu Val Arg Arg Thr Phe Ala Glu Gly Ser

130 135 140
 Ala Val Pro Ala Leu Phe Ala Val Glu Gln Asp Pro Ser Gly Glu Ala

145 150 155 160
 Arg Asp Leu Ala Leu Ser Tyr Ala Lys Gly Ile Gly Ala Thr Arg Ala

165 170 175
 Gly Val Leu Glu Thr Ser Phe Lys Glu Glu Thr Glu Thr Asp Leu Phe

180 185 190

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

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

260 265 270
Asn Met Lys Ala Val Leu Asp Asp Ile Gln Lys Gly Asn Phe Ser Asp
275 280 285
Arg Phe Ile Lys Asp Asn Gln Asn Asn Phe Glu Glu Phe His Lys Leu
290 295 300
Arg Glu Glu Gln His Gly His Gln Ile Glu Ala Val Gly Arg Glu Leu
305 310 315 320
Arg Asp Met Met Pro Phe Ile Lys Ser Lys Ser Ile Glu Lys

325 330
<210> 78
<211> 1002
<212> DNA
<213> Staphylococcus carnosus
<400> 78
atgacaacag tttattatga cgaaacagta aaagaagatg ctttaaaagg taaaaaaatt 60
gcagtagttg gttatgggtc tcaaggccat gctcatgcac aaaacttaaa agacaatggt 120
tatgacgttg taatcgggtat ccgcccaggt aattctttcg ataaagctaa aaaagacggc 180
tttgatgtat atccagtga tgaagcagtg aaacaagctg atgtagtaat ggtattatta 240
cctgatgaaa tccaaggtaa agtatatgaa gaagaaatcg caccaaactt agaagctggc 300

aacgctttag cttttgcaca tggtttcaac atccactttg atgttatcca accacctgaa 360
gatgttgatg tattcttagt tgccccctaaa ggaccaggac atttagtaag acgtacattt 420
actgaaggaa gtcagtagc tgccctattc ggtgtagaac aagatgcttc tggcaatgct 480

ttcaacttag ctttaagtta tgcaaaaggt atcggagcaa caaaagctgg ggttatcgaa 540
 actacattta aagaagaaac tgaaactgat ttatttggtg aacaagctgt actttgcggc 600
 ggtgtaacaa gattaatcca aagtggattt gaaacattag ttgaagctgg ttaccaacca 660
 gaattagctt attttgaagt attacatgaa atgaaattaa tcgttgattt aatgtatgaa 720

ggcggaatgg aaacaatgcg ttattctatc tcaaatacag cagaatacgg cgattatggt 780
 tctgggccac gtgtcatcac tgatgatggt aaacaaaaca tgaaagctgt cttagaagat 840
 atccaaaacg gtaaattcgc aaacagtttc attgaagata ataaaaatgg tttcaagag 900
 ttctacaaat tgcgtgaaga agcacaagat catcaaattc aagaagttgg taaagaactt 960
 cgtgaaatga tgccttttgt tgaaaaataa actattgaaa aa 1002

<210> 79

<211> 334

<212> PRT

<213> Staphylococcus carnosus

<400> 79

Met Thr Thr Val Tyr Tyr Asp Glu Thr Val Lys Glu Asp Ala Leu Lys

1 5 10 15

Gly Lys Lys Ile Ala Val Val Gly Tyr Gly Ser Gln Gly His Ala His

20 25 30

Ala Gln Asn Leu Lys Asp Asn Gly Tyr Asp Val Val Ile Gly Ile Arg

35 40 45

Pro Gly Asn Ser Phe Asp Lys Ala Lys Lys Asp Gly Phe Asp Val Tyr

50 55 60

Pro Val Asp Glu Ala Val Lys Gln Ala Asp Val Val Met Val Leu Leu

65 70 75 80

Pro Asp Glu Ile Gln Gly Lys Val Tyr Glu Glu Glu Ile Ala Pro Asn

85 90 95

Leu Glu Ala Gly Asn Ala Leu Ala Phe Ala His Gly Phe Asn Ile His

100 105 110

Phe Asp Val Ile Gln Pro Pro Glu Asp Val Asp Val Phe Leu Val Ala

115 120 125

Pro Lys Gly Pro Gly His Leu Val Arg Arg Thr Phe Thr Glu Gly Ser

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

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

 260 265 270
 Asn Met Lys Ala Val Leu Glu Asp Ile Gln Asn Gly Lys Phe Ala Asn
 275 280 285
 Ser Phe Ile Glu Asp Asn Lys Asn Gly Phe Lys Glu Phe Tyr Lys Leu
 290 295 300
 Arg Glu Glu Ala Gln Asp His Gln Ile Gln Glu Val Gly Lys Glu Leu
 305 310 315 320
 Arg Glu Met Met Pro Phe Val Glu Asn Lys Thr Ile Glu Lys

 325 330

<210> 80

<211> 993

<212> DNA

<213> *Listeria monocytogenes*

<400> 80

atgacaaaag tttattatga agatgcagta aaaaacaacg cactagaagg taaaacagta 60

gcagtaatcg ggtacgggtc gcaaggtcac gcacattctc aaaatctacg tgacaatggc 120

aataacgtta ttatcgcat tcgcgaagga aaatctgccg aatctgctag aaacgatggc 180
 ttgatgttt attctgttag cgaagccgct gataaagctg atgtcatcat gattcttttg 240
 ccagatgaaa cacaaggcga aacatacgaa aatgaaatta aacctaacct aaaagctggt 300

aattcacttg ttttcgctca cggttttaac attcattttg acgtaattaa tctccaagc 360
 gatgtagatg tttttctagt agctccaaaa ggtcctggtc acttagttcg ccgcacattt 420
 gtgaagggtg gcgtgttcc ttccctattc gctatctatc aagatgccac tggaaacgca 480
 cgcgacacag ccttttccta tgcaaaaggt attggcgcaa ctcgtgctgg cgttatcgaa 540
 accattttca aagaagaaac cgaaaccgat ctatttggcg aacaagcagt tctttgtggg 600
 ggtgcgactc accttatcca agctgggttc gaaacacttg tagaggctgg ctaccaacca 660
 gaacttgctt atttgaagt attacatgaa atgaaactaa ttgttgattt gatgtatgaa 720

ggtggtatgg aaaaaatgcg ccactcgatc tccaatacag cagaatatgg tgattatgtt 780
 tccggtcctc gaggttgtac agccgatacg aaaaaagcaa tgaaagaagt acttaccgac 840
 attcaaatg gtaactttgc taaatctttc atcaatgaca acaaaaatgg ttttaaagaa 900
 ttccatagaa tgcgtaaaga acaacaaggt catcaaatcg aaaaagttgg tgcagaactt 960
 cgcgaaatga tgccatttgt caaaccacaa cat 993

<210> 81

<211> 331

<212> PRT

<213> *Listeria monocytogenes*

<400> 81

Met Thr Lys Val Tyr Tyr Glu Asp Ala Val Lys Asn Asn Ala Leu Glu

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

65	70	75	80
Pro	Asp	Glu	Thr
Gln	Gly	Glu	Thr
Tyr	Glu	Asn	Glu
Ile	Lys	Pro	Asn

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

<210> 82

<211> 993

<212> DNA

<213> *Listeria grayi*

<400> 82

```

atgacaaaag tatattatga agatgctgta aaagaaaatg cactaacggg gaagaaagtt      60
gccgttatcg gttatggatc acaaggacat gcgcattcgc aaaatctgcg tgacaatggc      120
aatgacgtta tcatcggtat ccgcgaagga cgttctaaag aaagtgcgaa aaacgacggc      180
ttcgatgttt actccgttgc cgaagcgagt aaacaagcag atgtcatcat gatccttcta      240
ccagacgaaa tccaaggaga tacatagcaa aaagaaatca gcccaaactt ggaagcagga      300
aatgccattg ttttcgcca tggtttcaac atccattttg atgtgatcaa acctcctaca      360

gatgtagatg ttttcctagt tgcacaaaaa ggccccggac atctagtaag acgtaccttt      420
gtagaaggcg gtgccgtccc ttccctcttc gccgtccacc aagatgccac tggcaaagcg      480
caggacctcg ccttttctta tgccaaagga atcggagcta cacgtgccgg cgttatcgaa      540
acaactttca aagaagaaac agaaacggat ctattcgggtg aacaagctgt cctatgcggc      600
ggtgccactc accttatcca agcaggcttt gaaacacttg tagaagcagg ctatcaaccg      660
gaactcgctt acttcgaagt actgcatgag atgaaactga tcgtagacct gatgtacgaa      720
ggcggaatgg aaaaaatgcg ctattcgatc tccaatacag cagaatatgg cgactatgta      780

tccggtccgc gcatcgtgac cagcgatacg aagaaagcga tgaaagaagt attgaccgat      840
atccaaaatg gtaattttgc caaatcggtt atcgatgaca ataaaaacgg cttcaaagaa      900
tttcaccgaa tgcgcaaaga acagcaaggt catcaaatcg aagcagtcgg caaaaatctc      960
cgcgaaatga tgccgttcgt aaatccaaga aaa                                993

```

<210> 83

<211> 331

<212> PRT

<213> *Listeria grayi*

<400> 83

```

Met Thr Lys Val Tyr Tyr Glu Asp Ala Val Lys Glu Asn Ala Leu Thr
1           5           10           15

Gly Lys Lys Val Ala Val Ile Gly Tyr Gly Ser Gln Gly His Ala His
           20           25           30

Ser Gln Asn Leu Arg Asp Asn Gly Asn Asp Val Ile Ile Gly Ile Arg

```

35	40	45	
Glu Gly Arg Ser Lys Glu Ser Ala Lys Asn Asp Gly Phe Asp Val Tyr			
50	55	60	
Ser Val Ala Glu Ala Ser Lys Gln Ala Asp Val Ile Met Ile Leu Leu			
65	70	75	80
Pro Asp Glu Ile Gln Gly Asp Thr Tyr Glu Lys Glu Ile Ser Pro Asn			
85	90	95	
Leu Glu Ala Gly Asn Ala Ile Val Phe Ala His Gly Phe Asn Ile His			
100	105	110	
Phe Asp Val Ile Lys Pro Pro Thr Asp Val Asp Val Phe Leu Val Ala			
115	120	125	
Pro Lys Gly Pro Gly His Leu Val Arg Arg Thr Phe Val Glu Gly Gly			
130	135	140	
Ala Val Pro Ser Leu Phe Ala Val His Gln Asp Ala Thr Gly Lys Ala			
145	150	155	160
Gln Asp Leu Ala Leu Ser Tyr Ala Lys Gly Ile Gly Ala Thr Arg Ala			
165	170	175	
Gly Val Ile Glu Thr Thr Phe Lys Glu Glu Thr Glu Thr Asp Leu Phe			
180	185	190	
Gly Glu Gln Ala Val Leu Cys Gly Gly Ala Thr His Leu Ile Gln Ala			
195	200	205	
Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr Gln Pro Glu Leu Ala Tyr			
210	215	220	
Phe Glu Val Leu His Glu Met Lys Leu Ile Val Asp Leu Met Tyr Glu			
225	230	235	240
Gly Gly Met Glu Lys Met Arg Tyr Ser Ile Ser Asn Thr Ala Glu Tyr			
245	250	255	
Gly Asp Tyr Val Ser Gly Pro Arg Ile Val Thr Ser Asp Thr Lys Lys			
260	265	270	
Ala Met Lys Glu Val Leu Thr Asp Ile Gln Asn Gly Asn Phe Ala Lys			
275	280	285	

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

<210> 84

<211> 987

<212> DNA

<213> Enterococcus casseliflavus

<400> 84

atggtaaaag tatattacaa cgaggaagtt acagataatt taggagacaa gaagattgcg 60
 atcgtcggct acggctctca agggcacgct catgcgcaaa atttacggga cacaggtcat 120
 gaagtcttta tcggcattcg tgaaggaaaa tcagcgggaag ccgcaagaaa cgatggattt 180
 gaggttttcc cagtagctga agccgttcaa aaagcggatg tagtcatgat cttagcaccc 240
 gatgagatcc aaggatcggt gtataaaaat gagatcgaac cgaatctttc tgcgggcaat 300
 gcattggcat ttgctcatgg ttccaatatt cattttgatg tgatcgcgcc accaaaagag 360

 gtggatgttt ttctagtgc gccaaaagga ccaggtcatt tggcgcgtcg aacctttacc 420
 gaagggtttg cgttaccagc cttgtttgca gtttatcaag atgcgactgg taacgcacaa 480
 gcaacggctt tatcctatgc aaaaggcatt ggtgccacac gggtcggcgt tttggaaacg 540
 acctttaag aagagaccga aaccgatctt ttcggtgagc aagccgtact ttgtggcgggt 600
 ttgaccagta tgatcgaagc gggctttgaa acattggtag aagcggggta ccaaccggaa 660
 ttggcttatt ttgaagtttg tcatgaattg aaactgatcg ttgatttgat ttatgaagga 720
 ggcttcaaaa aaatgcgcca ttgatctcg aatacggctg agtatggtga ctacgtttct 780

 ggtgctcggg tgatcactgc acaagccaaa gaaaacatga aagccgtatt gacagacatc 840
 caaacggcc aatttgcgaa aggattcatt gaagacaatc aaaacggctt ccctgaattc 900
 cataaaatgc gtgaggaaaa cggcaatcat caaattgaaa aagtcggcag cgagttacgg 960
 aaaatgatgc cgtttgttgc gaaagat 987

<210> 85

<211> 329

<212> PRT

<213> Enterococcus casseliflavus

<400> 85

Met Val Lys Val Tyr Tyr Asn Glu Glu Val Thr Asp Asn Leu Gly Asp
1 5 10 15

Lys Lys Ile Ala Ile Val Gly Tyr Gly Ser Gln Gly His Ala His Ala
20 25 30

Gln Asn Leu Arg Asp Thr Gly His Glu Val Phe Ile Gly Ile Arg Glu
35 40 45

Gly Lys Ser Ala Glu Ala Ala Arg Asn Asp Gly Phe Glu Val Phe Pro
50 55 60

Val Ala Glu Ala Val Gln Lys Ala Asp Val Val Met Ile Leu Ala Pro
65 70 75 80

Asp Glu Ile Gln Gly Ser Leu Tyr Lys Asn Glu Ile Glu Pro Asn Leu
85 90 95

Ser Ala Gly Asn Ala Leu Ala Phe Ala His Gly Phe Asn Ile His Phe
100 105 110

Asp Val Ile Ala Pro Pro Lys Glu Val Asp Val Phe Leu Val Ala Pro
115 120 125

Lys Gly Pro Gly His Leu Val Arg Arg Thr Phe Thr Glu Gly Phe Ala
130 135 140

Val Pro Ala Leu Phe Ala Val Tyr Gln Asp Ala Thr Gly Asn Ala Gln
145 150 155 160

Ala Thr Ala Leu Ser Tyr Ala Lys Gly Ile Gly Ala Thr Arg Val Gly
165 170 175

Val Leu Glu Thr Thr Phe Lys Glu Glu Thr Glu Thr Asp Leu Phe Gly
180 185 190

Glu Gln Ala Val Leu Cys Gly Gly Leu Thr Ser Met Ile Glu Ala Gly
195 200 205

Phe Glu Thr Leu Val Glu Ala Gly Tyr Gln Pro Glu Leu Ala Tyr Phe
210 215 220

Glu Val Cys His Glu Leu Lys Leu Ile Val Asp Leu Ile Tyr Glu Gly
225 230 235 240

Gly Phe Lys Lys Met Arg His Ser Ile Ser Asn Thr Ala Glu Tyr Gly
245 250 255
Asp Tyr Val Ser Gly Ala Arg Val Ile Thr Ala Gln Ala Lys Glu Asn
260 265 270
Met Lys Ala Val Leu Thr Asp Ile Gln Asn Gly Gln Phe Ala Lys Gly
275 280 285
Phe Ile Glu Asp Asn Gln Asn Gly Phe Pro Glu Phe His Lys Met Arg
290 295 300
Glu Glu Asn Gly Asn His Gln Ile Glu Lys Val Gly Ser Glu Leu Arg
305 310 315 320
Lys Met Met Pro Phe Val Ala Lys Asp
325

<210> 86

<211> 984

<212> DNA

<213> Enterococcus gallinarum

<400> 86

atggtaaaag tatattataa cgaagtgaca gacaatttag gagataagac tattgcaatc	60
attggttacg gttctcaagg acatgcccac gcacagaatt tacgggatac aggacacaac	120
gtcgtgatcg ggattcgtgc aggaataatca gctgagaacg caaagaatga tggatttacc	180
gtgcttcag tttctgaagc tgtacaacaa gcaaatgttg tgatgatctt agctcctgat	240
gagatccaag gagaacttta cgagaaggaa attgcgccta atctttctgc aggggatgca	300
ttagcctttg ctcatggctt taatattcat tttagagtca tcacgcctcc aaaagaggtg	360
gatgtcttcc tgggtggcacc taaaggacca ggccatttgg tgcgccggac ctttagcgaa	420
ggatttgctg tccttgcgtt gtttctgtt tatcaagatg ctactggcaa tgcacaagag	480
acagccttgt cctatgccaag agaatcggg gctacaagag tcggtgtact agaaacaacg	540
ttcaaggaag agacggaac agacttgttt gggaacaag ccgtattatg tggcggactg	600
acaagtatga tcgagactgg gtttgaaacg ctggtagaag ccggtatca gccagaacta	660
gcctattttg aagtttgta cgaatgaaa ctgattgttg acttgattta tgaaggagga	720
ttcgccaaga tgcgtgattc aatctccaat acagcagaat atggggatta tgtctctggt	780
tcacgaatca ttaccgaaca aacaaaagcc aatatgaaga atgtcttgaa agatattcaa	840

aacggcacgt ttgccaaagg ttttattgcg gataatcaag ctggttttcc ggaattcaaa 900
 aaaatgcggg aacaaaacgg gaatcatcca atcgaaaaag ttgggtcaga attaagaaaa 960
 atgatgcctt ttgttacgaa agat 984

<210> 87

<211> 328

<212> PRT

<213> Enterococcus gallinarum

<400> 87

Met Val Lys Val Tyr Tyr Asn Glu Val Thr Asp Asn Leu Gly Asp Lys

1 5 10 15

Thr Ile Ala Ile Ile Gly Tyr Gly Ser Gln Gly His Ala His Ala Gln

20 25 30

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

35 40 45

Lys Ser Ala Glu Asn Ala Lys Asn Asp Gly Phe Thr Val Leu Pro Val

50 55 60

Ser Glu Ala Val Gln Gln Ala Asn Val Val Met Ile Leu Ala Pro Asp

65 70 75 80

Glu Ile Gln Gly Glu Leu Tyr Glu Lys Glu Ile Ala Pro Asn Leu Ser

85 90 95

Ala Gly Asp Ala Leu Ala Phe Ala His Gly Phe Asn Ile His Phe Glu

100 105 110

Val Ile Thr Pro Pro Lys Glu Val Asp Val Phe Leu Val Ala Pro Lys

115 120 125

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

130 135 140

Pro Ala Leu Phe Ala Val Tyr Gln Asp Ala Thr Gly Asn Ala Gln Glu

145 150 155 160

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

165 170 175

Leu Glu Thr Thr Phe Lys Glu Glu Thr Glu Thr Asp Leu Phe Gly Glu

180 185 190

Gln Ala Val Leu Cys Gly Gly Leu Thr Ser Met Ile Glu Thr Gly Phe

195 200 205

Glu Thr Leu Val Glu Ala Gly Tyr Gln Pro Glu Leu Ala Tyr Phe Glu

210 215 220

Val Cys His Glu Met Lys Leu Ile Val Asp Leu Ile Tyr Glu Gly Gly

225 230 235 240

Phe Ala Lys Met Arg Asp Ser Ile Ser Asn Thr Ala Glu Tyr Gly Asp

245 250 255

Tyr Val Ser Gly Ser Arg Ile Ile Thr Glu Gln Thr Lys Ala Asn Met

260 265 270

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

275 280 285

Ile Ala Asp Asn Gln Ala Gly Phe Pro Glu Phe Lys Lys Met Arg Glu

290 295 300

Gln Asn Gly Asn His Pro Ile Glu Lys Val Gly Ser Glu Leu Arg Lys

305 310 315 320

Met Met Pro Phe Val Thr Lys Asp

325

<210> 88

<211> 990

<212> DNA

<213> *Macrococcus caseolyticus*

<400> 88

atggctaaag tatattatga taacgatgta aataaggaat acttaaaaaa taagaaaatt 60

gcagtattag ggtacggctc tcaaggacac gcgcatgcat tgaatcttaa ggacagtggc 120

tatgacgtag tcatcggagt acgtgaaggg cagtccagaa caaaggcaga agaagatgga 180

tttcaagcat atgacgtcag tgaagctgtg gaacgtgcgg atgttaccgt cgtccttatg 240

ccggacgaag tacagcagcg tgtgttcaac gaggaaatcg tacctcatct caaaccacaa 300

tcggctctag ttttcgctca tggatttaat gtgcatttcg gtgcaataaa agtccttgac 360

gatgttgatg tcttcttagt cgcacacaaa ggaccaggac atcttggtcg acgcgaatat 420

gttaagggca gtgctgtccc tgcgttattt gcagtcgac aggatgtaac tggatgatgca 480

aaggcccttg cactgaacta tgcacacggc atcggtgcaa cgcgcgccgg cgtcatcgaa 540
 acgaccttca aggaagaaac agagacagac ctttttggtg agcaagcggc actctgcggc 600
 ggtatcacia agctcatcca ctacggtttt gaaacactga cggaagcagg ttatcagcca 660
 gaactcgcct atttcgaagt actgcatgaa atgaagctga ttgtggacct tatgtacgag 720
 ggcggtatgg agaagatgag acattccatc tcaaataccg ctgagtttgg agattatgtc 780
 tctggtgccg gcgttattac ccccgaagtt aaggaaaata tgaaacaggt attagctgat 840
 atccagtctg gcacattcag caggcagttt atcgaagaca ataataatgg atttgagtct 900

tttaaactca tgagaagtac ccaggcagga catcctatag agaaagtcgg tgcagaccta 960
 agaatgatga tgccgtttat caataataac 990

<210> 89

<211> 330

<212> PRT

<213> *Macrococcus caseolyticus*

<400> 89

Met Ala Lys Val Tyr Tyr Asp Asn Asp Val Asn Lys Glu Tyr Leu Lys

1 5 10 15

Asn Lys Lys Ile Ala Val Leu Gly Tyr Gly Ser Gln Gly His Ala His

20 25 30

Ala Leu Asn Leu Lys Asp Ser Gly Tyr Asp Val Val Ile Gly Val Arg

35 40 45

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

50 55 60

Asp Val Ser Glu Ala Val Glu Arg Ala Asp Val Thr Val Val Leu Met

65 70 75 80

Pro Asp Glu Val Gln Gln Arg Val Phe Asn Glu Glu Ile Val Pro His

85 90 95

Leu Lys Pro Gln Ser Ala Leu Val Phe Ala His Gly Phe Asn Val His

100 105 110

Phe Gly Ala Ile Lys Ala Pro Asp Asp Val Asp Val Phe Leu Val Ala

115 120 125

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

130 135 140

Ala Val Pro Ala Leu Phe Ala Val Asp Gln Asp Val Thr Gly Asp Ala
 145 150 155 160
 Lys Ala Leu Ala Leu Asn Tyr Ala His Gly Ile Gly Ala Thr Arg Ala

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

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

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

<210> 90

<211> 1020

<212> DNA

<213> Streptococcus vestibularis

<400> 90

atggcagttc aaatggaata tgaaaaagac gtaaaagtag cagcacttga cggtaaaaaa 60

atcgccgtta tcggttacgg ttcacaaggt catgcacatg cacaaaactt gcgtgattca 120

ggtcacgatg tgatcatcgg tgttcgtcca ggtaaatacat ttgataaagc taaagaagat 180
 ggttttgata cttacacagt agcagaagca actaaattgg ctgatatcat tatggttttg 240
 gcaccagatg aaattcaaaa agatatctat aaagatgaaa ttgagccaaa cttgagtgc 300
 ggcaaagctc ttggttttgc acacggattt aacatccact tcgaattcat caaggttcct 360
 aaagatgttg atgtctttat ggttgcctcct aaaggaccag gtcacttggc acgccgtaca 420
 tatacagaag gctttgggtg tccagcactt tacgcagtat accaagatgc tacaggcaat 480
 gctaaagaca tcgctatgga ctgggctaaa ggtattgggt ctgcacgtgt tggattactt 540

gtaacaacat ttaaagaaga aactgaagaa gatttgtttg gtgagcaagc ggttcttatg 600
 ggtgggttga cacatcttat tgaagctggt tttgaagtct tgactgaagc aggctatgca 660
 ccacaattgg ctactttga agtgcttcac gaaatgaaat tgatcgttga cttgatctac 720
 gaagtggtat tcaagaaaa gcgtcaatca tgttcaaata ctgcagaatt tggtgacttt 780
 gtaactggtc cacgtgtcat cggccagaa gtcaaagaaa acatgaaagc tgcccttgcg 840
 gatatccaat caggtaaatt tgcacgtgaa ttcgttgaag accatgatgc tggcttccca 900
 cgtttgaaag cttaccgtaa agaagctgaa gagcttgaaa ttgaaaagat tgggtgctgaa 960

ttgcgtaaag caatgccatt cgttgggtcaa aatgacgatg acgcattcaa aatctataac 1020

<210> 91

<211> 340

<212> PRT

<213> Streptococcus vestibularis

<400> 91

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

1 5 10 15

Asp Gly Lys Lys Ile Ala Val Ile Gly Tyr Gly Ser Gln Gly His Ala

20 25 30

His Ala Gln Asn Leu Arg Asp Ser Gly His Asp Val Ile Ile Gly Val

35 40 45

Arg Pro Gly Lys Ser Phe Asp Lys Ala Lys Glu Asp Gly Phe Asp Thr

50 55 60

Tyr Thr Val Ala Glu Ala Thr Lys Leu Ala Asp Ile Ile Met Val Leu

65 70 75 80

Ala Pro Asp Glu Ile Gln Lys Asp Ile Tyr Lys Asp Glu Ile Glu Pro

85 90 95

Asn Leu Ser Ala Gly Lys Ala Leu Gly Phe Ala His Gly Phe Asn Ile
100 105 110

His Phe Glu Phe Ile Lys Val Pro Lys Asp Val Asp Val Phe Met Val
115 120 125

Ala Pro Lys Gly Pro Gly His Leu Val Arg Arg Thr Tyr Thr Glu Gly
130 135 140

Phe Gly Val Pro Ala Leu Tyr Ala Val Tyr Gln Asp Ala Thr Gly Asn
145 150 155 160

Ala Lys Asp Ile Ala Met Asp Trp Ala Lys Gly Ile Gly Ser Ala Arg
165 170 175

Val Gly Leu Leu Val Thr Thr Phe Lys Glu Glu Thr Glu Glu Asp Leu
180 185 190

Phe Gly Glu Gln Ala Val Leu Met Gly Gly Leu Thr His Leu Ile Glu
195 200 205

Ala Gly Phe Glu Val Leu Thr Glu Ala Gly Tyr Ala Pro Gln Leu Ala
210 215 220

Tyr Phe Glu Val Leu His Glu Met Lys Leu Ile Val Asp Leu Ile Tyr
225 230 235 240

Glu Gly Gly Phe Lys Lys Met Arg Gln Ser Cys Ser Asn Thr Ala Glu
245 250 255

Phe Gly Asp Phe Val Thr Gly Pro Arg Val Ile Gly Pro Glu Val Lys
260 265 270

Glu Asn Met Lys Ala Ala Leu Ala Asp Ile Gln Ser Gly Lys Phe Ala
275 280 285

Arg Glu Phe Val Glu Asp His Asp Ala Gly Phe Pro Arg Leu Lys Ala
290 295 300

Tyr Arg Lys Glu Ala Glu Glu Leu Glu Ile Glu Lys Ile Gly Ala Glu
305 310 315 320

Leu Arg Lys Ala Met Pro Phe Val Gly Gln Asn Asp Asp Asp Ala Phe
325 330 335

Lys Ile Tyr Asn

340

<210> 92

<211> 1020

<212> DNA

<213> Streptococcus mutans

<400> 92

atggcagtag aaatgttata tgaagcagat gtaaaagtag cagcacttga tggtaagaag 60
attgctgtta ttggttatgg atcacaagga catgctcatg cccaaaactt gagagattca 120

ggacacgatg ttatcatcgg tgtgcgtcat gggaaatcgt ttgacaaggc taaagaagat 180
ggttttgata cctatgaagt gggatgaagct actaaattgg ctgatattat catggtttta 240
gcaccgatg aaattcaaaa agatatctac aaagatgaaa ttgctcctaa ttttaagtgt 300
ggtaaggctc ttggttttgc tcatggtttt aatatccatt tcggttatat taaggcacca 360
gaagatgtag atgttttcat ggttgccct aaaggaccgg gccatcttgt acgtcgtact 420
tacacagaag gatttggagt accttcttta tatgctgttt atcaaaatcc aactggtaat 480
gctgaaaata ttgctttaga ttgggctaag ggtattggtt ctgcgcgtgt aggtctcctt 540

gtaaccactt tcaaagaaga aacggaagaa gacctttttg gtgaacaagc ggttcttatg 600
ggtaggttta ctcaccttat tgaagcaggt ttgaagtgc ttactgaagc aggtatgct 660
ccacaattgg cttattttga agtgcttcat gaaatgaaat tgattgtaga cttatctat 720
gaaggtggtt tcaagaaaat gcgtcaatca tgttcaaata ccgctgaatt tggtagcttt 780
gtgacaggcc cacgtgttat cggtccagaa gtaaaagaaa acatgaaggc tgctcttgct 840
gatattcaat caggtaaatt tgcccgtgaa ttcgtggaag atcatgatgc tggcttccca 900
cgtttgaaag ctttccgtaa agaagctgaa ggactcgaaa tcgaaaaaat tggtagcagaa 960

cttcgtaaag caatgccatt tgttaatcaa aacgatgacg atgctttcaa aatttataac 1020

<210> 93

<211> 340

<212> PRT

<213> Streptococcus mutans

<400> 93

Met Ala Val Glu Met Leu Tyr Glu Ala Asp Val Lys Val Ala Ala Leu

1 5 10 15

Asp Gly Lys Lys Ile Ala Val Ile Gly Tyr Gly Ser Gln Gly His Ala

20 25 30

His Ala Gln Asn Leu Arg Asp Ser Gly His Asp Val Ile Ile Gly Val
35 40 45

Arg His Gly Lys Ser Phe Asp Lys Ala Lys Glu Asp Gly Phe Asp Thr
50 55 60

Tyr Glu Val Gly Glu Ala Thr Lys Leu Ala Asp Ile Ile Met Val Leu
65 70 75 80

Ala Pro Asp Glu Ile Gln Lys Asp Ile Tyr Lys Asp Glu Ile Ala Pro
85 90 95

Asn Leu Ser Ala Gly Lys Ala Leu Gly Phe Ala His Gly Phe Asn Ile
100 105 110

His Phe Gly Tyr Ile Lys Ala Pro Glu Asp Val Asp Val Phe Met Val
115 120 125

Ala Pro Lys Gly Pro Gly His Leu Val Arg Arg Thr Tyr Thr Glu Gly
130 135 140

Phe Gly Val Pro Ser Leu Tyr Ala Val Tyr Gln Asn Pro Thr Gly Asn
145 150 155 160

Ala Glu Asn Ile Ala Leu Asp Trp Ala Lys Gly Ile Gly Ser Ala Arg
165 170 175

Val Gly Leu Leu Val Thr Thr Phe Lys Glu Glu Thr Glu Glu Asp Leu
180 185 190

Phe Gly Glu Gln Ala Val Leu Met Gly Gly Leu Thr His Leu Ile Glu
195 200 205

Ala Gly Phe Glu Val Leu Thr Glu Ala Gly Tyr Ala Pro Gln Leu Ala
210 215 220

Tyr Phe Glu Val Leu His Glu Met Lys Leu Ile Val Asp Leu Ile Tyr
225 230 235 240

Glu Gly Gly Phe Lys Lys Met Arg Gln Ser Cys Ser Asn Thr Ala Glu
245 250 255

Phe Gly Asp Phe Val Thr Gly Pro Arg Val Ile Gly Pro Glu Val Lys
260 265 270

Glu Asn Met Lys Ala Ala Leu Ala Asp Ile Gln Ser Gly Lys Phe Ala

275	280	285	
Arg Glu Phe Val Glu Asp His Asp Ala Gly Phe Pro Arg Leu Lys Ala			
290	295	300	
Phe Arg Lys Glu Ala Glu Gly Leu Glu Ile Glu Lys Ile Gly Ala Glu			
305	310	315	320
Leu Arg Lys Ala Met Pro Phe Val Asn Gln Asn Asp Asp Asp Ala Phe			
325	330	335	
Lys Ile Tyr Asn			
340			
<210>	94		
<211>	1020		
<212>	DNA		
<213>	Streptococcus gordonii		
<400>	94		
atggcagttc aaatggaata cgaaaaagat gtaaaagtag cagcgcttga cggtaagaaa		60	
attgccgtaa tcggctatgg ttcacaaggt catgctcatg cacaaaatct gcgtgataca		120	
ggccatgatg tgattatcgg tgttcgccct ggtaagtcac tcgataaggc taaagaagat		180	
ggctttgata cttacacagt agcagaggca gctaagttgg ctgatgtaat catgattttg		240	
gcgccagatg aaattcaaca agaactctac gaagcagaaa tcgctccaaa cttggaagct		300	
ggtaatgcag ttggatttgc ccatggtttc aatatccatt ttgaatttat caaagttcct		360	
gctgatgtag atgtctttat gtgtgccctt aaaggtcctg gtcacttggt tcgtcgtact		420	
tttgaagaag gatttgggtg tccggtcttt tatgcagtct accaagatgc aacaggaaat		480	
gcgaaaaaca ttgctatgga ctgggtgtaa ggtgttggtg cagctcgtgt tggtttgctt		540	
gaaacaactt acaaagaaga aacagaagaa gatctctttg gtgaacaagc tgtcctttgt		600	
ggtaggtttga ctgcccttat cgaagcaggt tttgaagtct tgacagaagc aggttacgct		660	
cctgaattgg cttactttga agttcttcat gaaatgaaat tgatcgttga cttgatctac		720	
gagggtggat tcaagaaaa gcgccaatca atttcaaata ctgctgaata tggtgattac		780	
glatcaggcc cacgtgtaat tactgagcaa gtcaaagaaa acatgaaggc tgttttggca		840	
gacattcaaa atggtaaatt tgcaaatgac tttgtaaatg actacaaggc tggccgtcca		900	
aaactcactg cttaccgtga acaagcagcc aacctgaaa tcgaaaaagt aggtgctgaa		960	
ttgcgtaaag caatgccatt tgttggtaaa aatgacgacg acgcatttaa aatctacaac		1020	

<210> 95

<211> 340

<212> PRT

<213> Streptococcus gordonii

<400> 95

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

1 5 10 15

Asp Gly Lys Lys Ile Ala Val Ile Gly Tyr Gly Ser Gln Gly His Ala

20 25 30

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

35 40 45

Arg Pro Gly Lys Ser Phe Asp Lys Ala Lys Glu Asp Gly Phe Asp Thr

50 55 60

Tyr Thr Val Ala Glu Ala Ala Lys Leu Ala Asp Val Ile Met Ile Leu

65 70 75 80

Ala Pro Asp Glu Ile Gln Gln Glu Leu Tyr Glu Ala Glu Ile Ala Pro

85 90 95

Asn Leu Glu Ala Gly Asn Ala Val Gly Phe Ala His Gly Phe Asn Ile

100 105 110

His Phe Glu Phe Ile Lys Val Pro Ala Asp Val Asp Val Phe Met Cys

115 120 125

Ala Pro Lys Gly Pro Gly His Leu Val Arg Arg Thr Phe Glu Glu Gly

130 135 140

Phe Gly Val Pro Ala Leu Tyr Ala Val Tyr Gln Asp Ala Thr Gly Asn

145 150 155 160

Ala Lys Asn Ile Ala Met Asp Trp Cys Lys Gly Val Gly Ala Ala Arg

165 170 175

Val Gly Leu Leu Glu Thr Thr Tyr Lys Glu Glu Thr Glu Glu Asp Leu

180 185 190

Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Leu Thr Ala Leu Ile Glu

195 200 205

Ala Gly Phe Glu Val Leu Thr Glu Ala Gly Tyr Ala Pro Glu Leu Ala

210 215 220
Tyr Phe Glu Val Leu His Glu Met Lys Leu Ile Val Asp Leu Ile Tyr
225 230 235 240

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

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

<210> 96

<211> 1020

<212> DNA

<213> Streptococcus suis

<400> 96

atgacagtaa caatgcaata tgaaaaagat gtaacagtgg cagcacttga cggtaaactg 60

atcgccgtta tcggttatgg ttcacaaggt catgcccatg ctcaaaactt gcgtgatata 120

ggacacgatg tcatcatcgg tgtgcgtgca ggtaagtcac ttgacaaggc aaaagaagac 180

ggttttgaaa cttttgaagt agcagaggca gcaaaacaag ctgatgtcat tatgatattg 240

gctccagacg aaattcaagc agatttgtat gagaacgaag ttgctccaaa cttggaagct 300

ggaaatgcgc ttggttttgc tcatggtttc aacgttcact ttgaatttat caaggtacca 360

gcagatgtgg atgtcttcat gtgtgcacca aaaggtccag gtcacttggt ccgccgtacc 420

tttgaagaag gttttggagt accagccctc tatgcggtct accaagatgc aactggcaat 480

gcaaaacata ttgcgatgga ctgggcaaaa ggtgttggtt cagcccgtgt tgggtctttg 540

gaaacaactt tcaaagaaga gacagaagaa gatttgtttg gtgagcaagc tgttctctgc 600
 ggtggtttga cagcccttat gcaggcaggt tttgaagtct tgacagaagc tggttatgca 660
 ccagaattgg cttatttcga agttctccac gagatgaaac tcatcgttga ccttgtctac 720
 gaagggtggct ttaagaaaaat gcgccaatct atttcaaaca ctgctgaatt tggtgactat 780
 giatcaggtc cagcgtgat tacagaccaa gtcaaagaaa acatgaaggc agtccttgca 840
 gatattcaat ctggtaaatt tgcaaatgac tttgtaaatg actataaagc aggccgtcca 900
 cgtatggaag cttatcgtaa agaagcagag aatcttgaaa tcgaaaaagt aggtgcagaa 960

ctccgcaaag caatgccatt tgtcggacgc aatgatgacg acgcattcaa aatctataat 1020

<210> 97

<211> 340

<212> PRT

<213> Streptococcus suis

<400> 97

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

1 5 10 15

Asp Gly Lys Arg Ile Ala Val Ile Gly Tyr Gly Ser Gln Gly His Ala

20 25 30

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

35 40 45

Arg Ala Gly Lys Ser Phe Asp Lys Ala Lys Glu Asp Gly Phe Glu Thr

50 55 60

Phe Glu Val Ala Glu Ala Ala Lys Gln Ala Asp Val Ile Met Ile Leu

65 70 75 80

Ala Pro Asp Glu Ile Gln Ala Asp Leu Tyr Glu Asn Glu Val Ala Pro

85 90 95

Asn Leu Glu Ala Gly Asn Ala Leu Gly Phe Ala His Gly Phe Asn Val

100 105 110

His Phe Glu Phe Ile Lys Val Pro Ala Asp Val Asp Val Phe Met Cys

115 120 125

Ala Pro Lys Gly Pro Gly His Leu Val Arg Arg Thr Phe Glu Glu Gly

130 135 140

Phe Gly Val Pro Ala Leu Tyr Ala Val Tyr Gln Asp Ala Thr Gly Asn

145 150 155 160
Ala Lys His Ile Ala Met Asp Trp Ala Lys Gly Val Gly Ser Ala Arg
 165 170 175

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

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

Tyr Arg Lys Glu Ala Glu Asn Leu Glu Ile Glu Lys Val Gly Ala Glu
305 310 315 320
Leu Arg Lys Ala Met Pro Phe Val Gly Arg Asn Asp Asp Asp Ala Phe
 325 330 335
Lys Ile Tyr Asn
 340

<210> 98

<211> 1020

<212> DNA

<213> Streptococcus infantarius

<400> 98

atggcagtaa caatggaata cgaaaaagac gtaaaagtag cagctcttga tggtaaaaaa 60

attgccgtta ttggttatgg atcacaaggt catgctcatg ctcaaaactt gcgtgactca 120

ggtcacgatg ttatcattgg ggttcgccat ggtaaatcat tcgataaagc aaaagaagat 180

ggatttgata cttatgaagt agcagaagca acaaaacttg ctgatgttat catgggtattg 240

gctcctgatg aaatccaagc taaactttat gctgaagaaa tcgctccaaa ctttgaagct 300

ggtaacgctc ttggatttgc acatggtttt aatatccgtt ttggatacat taaagctcca 360

gaaacagtag atgtcttcat gtgtgctcct aaaggaccag gtcaccttgt tcgtcgtact 420

tacacagaag gatttgggtg accagcactt tacgtgttt accaagatgc tactggtaat 480

gctaaagaca tcgcaatgga ttggctctaaa ggtatcggtg ctgcacgtgt tggacttctt 540

gaaacaacat ttaaagaaga aactgaagaa gacctctttg gtgaacaagc agtactttgt 600

ggtgggttaa ctgctcttat cgaagctggt ttggaagttc ttactgaagc tggctatgct 660

ccagaattgg cttactttga agttcttcat gaaatgaaac ttatcggtga ctttatctac 720

gaaggtggat tcaagaaaat gcgtcaatca atttcaaata cagctgaatt tggtgactac 780

gtatctggac cactgtttat cactaaagat gttaaagaaa atatgaaagc tgttcttgct 840

gatatccaat caggtaaatt cgctgaagat ttgtttaacg actaccaagc aggtcgtcca 900

aaacttgaag cataccgtaa agaagctgca gctcttgaaa ttgaaaaagt gggtgctgaa 960

cttcgtaaag caatgccttt tgtaaccaa aacgatgacg atgcattcaa aatttataac 1020

<210> 99

<211> 340

<212> PRT

<213> Streptococcus infantarius

<400> 99

Met Ala Val Thr Met Glu Tyr Glu Lys Asp Val Lys Val Ala Ala Leu

1 5 10 15

Asp Gly Lys Lys Ile Ala Val Ile Gly Tyr Gly Ser Gln Gly His Ala

20 25 30

His Ala Gln Asn Leu Arg Asp Ser Gly His Asp Val Ile Ile Gly Val

35 40 45

Arg His Gly Lys Ser Phe Asp Lys Ala Lys Glu Asp Gly Phe Asp Thr

50 55 60

Tyr Glu Val Ala Glu Ala Thr Lys Leu Ala Asp Val Ile Met Val Leu

65 70 75 80

Ala Pro Asp Glu Ile Gln Ala Lys Leu Tyr Ala Glu Glu Ile Ala Pro
85 90 95

Asn Leu Glu Ala Gly Asn Ala Leu Gly Phe Ala His Gly Phe Asn Ile
100 105 110

Arg Phe Gly Tyr Ile Lys Ala Pro Glu Thr Val Asp Val Phe Met Cys
115 120 125

Ala Pro Lys Gly Pro Gly His Leu Val Arg Arg Thr Tyr Thr Glu Gly
130 135 140

Phe Gly Val Pro Ala Leu Tyr Ala Val Tyr Gln Asp Ala Thr Gly Asn
145 150 155 160

Ala Lys Asp Ile Ala Met Asp Trp Ser Lys Gly Ile Gly Ala Ala Arg
165 170 175

Val Gly Leu Leu Glu Thr Thr Phe Lys Glu Glu Thr Glu Glu Asp Leu
180 185 190

Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Leu Thr Ala Leu Ile Glu
195 200 205

Ala Gly Phe Glu Val Leu Thr Glu Ala Gly Tyr Ala Pro Glu Leu Ala
210 215 220

Tyr Phe Glu Val Leu His Glu Met Lys Leu Ile Val Asp Leu Ile Tyr
225 230 235 240

Glu Gly Gly Phe Lys Lys Met Arg Gln Ser Ile Ser Asn Thr Ala Glu
245 250 255

Phe Gly Asp Tyr Val Ser Gly Pro Arg Val Ile Thr Lys Asp Val Lys
260 265 270

Glu Asn Met Lys Ala Val Leu Ala Asp Ile Gln Ser Gly Lys Phe Ala
275 280 285

Glu Asp Phe Val Asn Asp Tyr Gln Ala Gly Arg Pro Lys Leu Glu Ala
290 295 300

Tyr Arg Lys Glu Ala Ala Ala Leu Glu Ile Glu Lys Val Gly Ala Glu
305 310 315 320

Leu Arg Lys Ala Met Pro Phe Val Asn Gln Asn Asp Asp Ala Phe

	325	330	335
Lys Ile Tyr Asn			
	340		
<210>	100		
<211>	1020		
<212>	DNA		
<213>	Lactococcus lactis		
<400>	100		
atggcagtta caatgtatta tgaagaagat gtagaagtag cgcactcgc gggttaagaaa			60
atcgcagtga ttggatatgg ctcaacaagga cacgctcatg cacaaaactt gcgtgattct			120
ggctcatgatg tgattattgg tgtccgtcag gggaaatctt ttgataaagc aaaagaagat			180
ggttttgaaa catttgaagt aggagaagca gtagctaaag ctgacgtcat tatggttctg			240
gcacctgatg aacttcaaca atctatttat gaagaggaca taaaaccaa tttgaaagca			300
ggttcagcac ttggttttgc ccatggtttc aatattcatt ttggctatat tgaagttcca			360
gaagatgttg atgtcttcat ggttgcgcca aaagcgccgg gacatctcgt tcggcggact			420
tttaccgaag gtttcggaac gccagctttg ttcgtttcgc atcaaaatgc cactggtcac			480
gcgcgtgaaa ttgccatgga ctgggccaac ggaattggct gtgcccgtgt cggtatcatt			540
gaaacaactt tcaaagaaga aacagaagaa gatttgtttg gcgaacaggc cgtgctttgt			600
ggcggtttga cagcacttgt tgaagctggt ttgaaacac tgacagaagc tggatatgct			660
ggcgaattgg ctactttga agtgctgcat gaaatgaaat tgattgttga cttatgtac			720
gaagtggtt tcaactaaat gcgtcagtca atctcaaca ctgccgaatt tggtgattat			780
gtgactggac cagcattat tactgacgaa gttaaaaaga atatgaaact cgtgttggct			840
gacattcaat caggaaaatt tgcgcaagat ttcgttgatg atttcaaagc tggacgtcca			900
aaattaactg cttatcgtga agcagctaaa aatctggaaa ttgaaaaaat cgggtgcagaa			960
ctacgtaaag caatgccatt tacacaatct ggtgatgacg acgccttta aatttatcaa			1020
<210>	101		
<211>	340		
<212>	PRT		
<213>	Lactococcus lactis		
<400>	101		
Met Ala Val Thr Met Tyr Tyr Glu Glu Asp Val Glu Val Ala Ala Leu			
1	5	10	15

Ala Gly Lys Lys Ile Ala Val Ile Gly Tyr Gly Ser Gln Gly His Ala
20 25 30

His Ala Gln Asn Leu Arg Asp Ser Gly His Asp Val Ile Ile Gly Val
35 40 45

Arg Gln Gly Lys Ser Phe Asp Lys Ala Lys Glu Asp Gly Phe Glu Thr
50 55 60

Phe Glu Val Gly Glu Ala Val Ala Lys Ala Asp Val Ile Met Val Leu
65 70 75 80

Ala Pro Asp Glu Leu Gln Gln Ser Ile Tyr Glu Glu Asp Ile Lys Pro
85 90 95

Asn Leu Lys Ala Gly Ser Ala Leu Gly Phe Ala His Gly Phe Asn Ile
100 105 110

His Phe Gly Tyr Ile Glu Val Pro Glu Asp Val Asp Val Phe Met Val
115 120 125

Ala Pro Lys Ala Pro Gly His Leu Val Arg Arg Thr Phe Thr Glu Gly
130 135 140

Phe Gly Thr Pro Ala Leu Phe Val Ser His Gln Asn Ala Thr Gly His
145 150 155 160

Ala Arg Glu Ile Ala Met Asp Trp Ala Lys Gly Ile Gly Cys Ala Arg
165 170 175

Val Gly Ile Ile Glu Thr Thr Phe Lys Glu Glu Thr Glu Glu Asp Leu
180 185 190

Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Leu Thr Ala Leu Val Glu
195 200 205

Ala Gly Phe Glu Thr Leu Thr Glu Ala Gly Tyr Ala Gly Glu Leu Ala
210 215 220

Tyr Phe Glu Val Leu His Glu Met Lys Leu Ile Val Asp Leu Met Tyr
225 230 235 240

Glu Gly Gly Phe Thr Lys Met Arg Gln Ser Ile Ser Asn Thr Ala Glu
245 250 255

Phe Gly Asp Tyr Val Thr Gly Pro Arg Ile Ile Thr Asp Glu Val Lys

260	265	270	
Lys Asn Met Lys Leu Val Leu Ala Asp Ile Gln Ser Gly Lys Phe Ala			
275	280	285	
Gln Asp Phe Val Asp Asp Phe Lys Ala Gly Arg Pro Lys Leu Thr Ala			
290	295	300	
Tyr Arg Glu Ala Ala Lys Asn Leu Glu Ile Glu Lys Ile Gly Ala Glu			
305	310	315	320
Leu Arg Lys Ala Met Pro Phe Thr Gln Ser Gly Asp Asp Asp Ala Phe			
325	330	335	
Lys Ile Tyr Gln			
340			
<210>	102		
<211>	1032		
<212>	DNA		
<213>	Lactococcus lactis		
<400>	102		
atggcagtta caatgtatta tgaagatgat gtagaagtat cagcacttgc tggaaagcaa	60		
attgcagtaa tcggttatgg ttcacaagga catgctcacg cacagaattt gcgtgattct	120		
ggtcacaacg ttatcattgg tgtgcgccac ggaaaatctt ttgataaagc aaaagaagat	180		
ggctttgaaa catttgaagt aggagaagca gtagctaaag ctgatgttat tatggttttg	240		
gcaccagatg aacttcaaca atccatttat gaagaggaca tcaaaccaaa ctgaaagca	300		
ggttcagcac ttggttttgc tcacggattt aatatccatt ttggctatat taaagtacca	360		
gaagacgttg acgtctttat ggttgcgcct aaggtccag gtcaccttgt ccgtcggact	420		
tatactgaag gttttggtac accagctttg ttgtttcac accaaaatgc aagtggatcat	480		
gcgcgtgaaa tcgcaatgga ttgggccaaa ggaattgggt gtgctcgagt gggaattatt	540		
gaaacaactt ttaaagaaga aacagaagaa gatttgtttg gagaacaagc tgttctatgt	600		
ggaggtttga cagcacttgt tgaagccggt ttgaaacac tgacagaagc tggatacgct	660		
ggcgaattgg cttactttga agttttgcac gaaatgaaat tgattgttga cctcatgtat	720		
gaaggtgggt ttactaaaat gcgtcaatcc atctcaata ctgctgagtt tggcgattat	780		
gtgactggtc cacggattat tactgacgaa gttaaaaaga atatgaagct tgttttggct	840		
gatattcaat ctggaaaatt tgctcaagat ttcgttgatg acttcaaagc ggggcgtcca	900		

aaattaatag cctatcgcg agctgcaaaa aatcttgaaa ttgaaaaaat tggggcagag 960

cacgtcaagc aatgccattc acacaatctg gtgatgacga tgcctttaaa atctatcagt 1020

aattttctctt at 1032

<210> 103

<211> 344

<212> PRT

<213> Lactococcus lactis

<400> 103

Met Ala Val Thr Met Tyr Tyr Glu Asp Asp Val Glu Val Ser Ala Leu

1 5 10 15

Ala Gly Lys Gln Ile Ala Val Ile Gly Tyr Gly Ser Gln Gly His Ala

20 25 30

His Ala Gln Asn Leu Arg Asp Ser Gly His Asn Val Ile Ile Gly Val

35 40 45

Arg His Gly Lys Ser Phe Asp Lys Ala Lys Glu Asp Gly Phe Glu Thr

50 55 60

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

65 70 75 80

Ala Pro Asp Glu Leu Gln Gln Ser Ile Tyr Glu Glu Asp Ile Lys Pro

85 90 95

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

100 105 110

His Phe Gly Tyr Ile Lys Val Pro Glu Asp Val Asp Val Phe Met Val

115 120 125

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

130 135 140

Phe Gly Thr Pro Ala Leu Phe Val Ser His Gln Asn Ala Ser Gly His

145 150 155 160

Ala Arg Glu Ile Ala Met Asp Trp Ala Lys Gly Ile Gly Cys Ala Arg

165 170 175

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

180 185 190
Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Leu Thr Ala Leu Val Glu
195 200 205
Ala Gly Phe Glu Thr Leu Thr Glu Ala Gly Tyr Ala Gly Glu Leu Ala
210 215 220
Tyr Phe Glu Val Leu His Glu Met Lys Leu Ile Val Asp Leu Met Tyr

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

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

<210> 104

<211> 1041

<212> DNA

<213> Leuconcstoc mesenteroides

<400> 104

atgactacaa aaatgtttta tgataaagat attgatacaa caccattgga aaacaaaaag 60

attgcagtta ttggctatgg tgcacaaggg catgctcaag cgaataatct acgcgactca 120

ggattttgatg tcatcatggg attgcgccca ggaaaatctt ttgatagtgc taaaaaagat 180

gggttcgaag ttactcagc tgctgaagca acggcacaag cggatgttgt tatgatggaa 240

accctgacg aattacaagc ggcagtttgg gagaaagaag ttgagcctaa ccttaaggca 300

ggttcttacc ttggattttc tcatgggttc aacattgttt atggtttgat taagcccaat 360

gctgatatta atgcatgat cattgcgcca aagggtccag gaaacattga acgtcgccaa 420
 ttcgttgaag ggggtgggat tccttctttg tatggtgttc atcaagatcc tacgggtgat 480

 acagctgaag tggccaaggc ctatgcgaaa ggtattgggt caggtcgcg aggtatcttg 540
 gaaacgactt ttgaagaaga aacaacagaa gacttgtttg gtgaacaagc tgtactttgt 600
 ggtggcttaa cacaattgat tgaggcagga tttaacacat tgggtggaagc aggttactca 660
 cctgaattgg cttatttcga aacatctcat gaaatgaaga tgattgttga ttgatcttt 720
 gaaggtgggt tcgagaagat gagacacgat tgctcaaata cttgtgaata tggtgaaatg 780
 ttgaacggac cacgtatcat cactgaagaa tcaaagcaag gaatgcgtga tgttttgaag 840
 gatatccaag atggtactta tgctaagaag tggttggcag aatacaattc tggtttgaag 900

 gatttggaaa agatgcggac agagtataag agcggtttgt acgagcaaac aggtaagaag 960
 gttcgtgcta tgatgcatg gatttcagat gcagataagt actcgacagc agcagatact 1020
 gagcaattct cagcagcaaa g 1041

<210> 105

<211> 347

<212> PRT

<213> Leuconostoc mesenteroides

<400> 105

Met Thr Thr Lys Met Phe Tyr Asp Lys Asp Ile Asp Thr Thr Pro Leu

1 5 10 15

Glu Asn Lys Lys Ile Ala Val Ile Gly Tyr Gly Ala Gln Gly His Ala

20 25 30

Gln Ala Asn Asn Leu Arg Asp Ser Gly Phe Asp Val Ile Met Gly Leu

35 40 45

Arg Pro Gly Lys Ser Phe Asp Ser Ala Lys Lys Asp Gly Phe Glu Val

50 55 60

Tyr Ser Ala Ala Glu Ala Thr Ala Gln Ala Asp Val Val Met Met Glu

65 70 75 80

Thr Pro Asp Glu Leu Gln Ala Ala Val Trp Glu Lys Glu Val Glu Pro

85 90 95

Asn Leu Lys Ala Gly Ser Tyr Leu Gly Phe Ser His Gly Phe Asn Ile

100 105 110

Val Tyr Gly Leu Ile Lys Pro Asn Ala Asp Ile Asn Val Met Ile Ile
115 120 125
Ala Pro Lys Gly Pro Gly Asn Ile Glu Arg Arg Gln Phe Val Glu Gly
130 135 140
Gly Gly Ile Pro Ser Leu Tyr Gly Val His Gln Asp Pro Thr Gly Asp
145 150 155 160
Thr Ala Glu Val Ala Lys Ala Tyr Ala Lys Gly Ile Gly Ser Gly Arg
165 170 175
Ala Gly Ile Leu Glu Thr Thr Phe Glu Glu Glu Thr Thr Glu Asp Leu
180 185 190
Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Leu Thr Gln Leu Ile Glu
195 200 205
Ala Gly Phe Asn Thr Leu Val Glu Ala Gly Tyr Ser Pro Glu Leu Ala
210 215 220
Tyr Phe Glu Thr Ser His Glu Met Lys Met Ile Val Asp Leu Ile Phe
225 230 235 240
Glu Gly Gly Phe Glu Lys Met Arg His Asp Cys Ser Asn Thr Cys Glu
245 250 255
Tyr Gly Glu Met Leu Asn Gly Pro Arg Ile Ile Thr Glu Glu Ser Lys
260 265 270
Gln Gly Met Arg Asp Val Leu Lys Asp Ile Gln Asp Gly Thr Tyr Ala
275 280 285
Lys Lys Trp Leu Ala Glu Tyr Asn Ser Gly Leu Lys Asp Leu Glu Lys
290 295 300
Met Arg Thr Glu Tyr Lys Ser Gly Leu Tyr Glu Gln Thr Gly Lys Lys
305 310 315 320
Val Arg Ala Met Met Pro Trp Ile Ser Asp Ala Asp Lys Tyr Ser Thr
325 330 335
Ala Ala Asp Thr Glu Gln Phe Ser Ala Ala Lys
340 345

<210> 106

<211> 1044

<212> DNA

<213> *Lactobacillus buchneri*

<400> 106

atgagtgtag aaatgttgta tgacaaggat gttaccacca attatcttca aggaaagaaa	60
attgctttta tcggttatgg atcacaaggt catgcccaag ctaacaactt aagagattca	120
ggttatgacg ttatcgttgg tgttcgccca ggacaatcat ttgaaaatgc caaagttgat	180
ggatttgatg tatacacacc agccgaagca gctcgtcgtg cagactggat tcaaagtta	240
acgcctgatg aagttatggg cgacgtttat aaaaatgaaa ttgcacctaa cctagaggaa	300
ggcaatgtat taggcttttc acatggcttc aacattcatt ataaagaaat cgtaccacca	360
gcaaacgttg atgttggtat gatggcacct aaaggcccag gtaatctttg tcgtcgtaca	420
tacgttgaag gctccggtgt tccagcatta tatggctatt tccaagacta ttcaggtcat	480
gccgaagatt tatccaagga attcgccaaa ggaaatgggtg cagcacgtgc cggactattg	540
aagacaacct tcaaagaaga aactgaagaa gatttgtttg gtgagcaaaa cgtccttatg	600
ggtaggtgta ccgcgcttat cgaaaccggt tatgaagtct taaccgaagc aggatactca	660
ccacagttgg catacttcga agtggatcac gaaatgaaat tgatctgtga ccttatctac	720
gaaggtgggt tcaacaagat gtatgtgac tgcctgaaca cttctgaata tggttcatat	780
glagtgggtc caaaggtcgt tggttaaggaa tccaagcaag ctatgaaaga cgctttgaag	840
cggattcaag acggttcatt tgctaaagaa ttcattggatg actaccgtaa cggtttcaag	900
aagctttaca agatgcgtga acgttcagct aactcacttc tttcacgtgt tgggtgctgat	960
cttcgtgatc acatgtcatt tggttggtgaa gccgacaagt acagtacacc tactgaagaa	1020
aaagcagaag cagaagctga aaag	1044

<210> 107

<211> 348

<212> PRT

<213> *Lactobacillus buchneri*

<400> 107

Met Ser Val Glu Met Leu Tyr Asp Lys Asp Val Thr Thr Asn Tyr Leu

1	5	10	15
Gln Gly Lys Lys Ile Ala Phe Ile Gly Tyr Gly Ser Gln Gly His Ala			
	20	25	30
Gln Ala Asn Asn Leu Arg Asp Ser Gly Tyr Asp Val Ile Val Gly Val			

35 40 45
 Arg Pro Gly Gln Ser Phe Glu Asn Ala Lys Val Asp Gly Phe Asp Val
 50 55 60
 Tyr Thr Pro Ala Glu Ala Ala Arg Arg Ala Asp Trp Ile Gln Met Leu

 65 70 75 80
 Thr Pro Asp Glu Val Met Gly Asp Val Tyr Lys Asn Glu Ile Ala Pro
 85 90 95
 Asn Leu Glu Glu Gly Asn Val Leu Gly Phe Ser His Gly Phe Asn Ile
 100 105 110
 His Tyr Lys Glu Ile Val Pro Pro Ala Asn Val Asp Val Val Met Met
 115 120 125
 Ala Pro Lys Gly Pro Gly Asn Leu Cys Arg Arg Thr Tyr Val Glu Gly

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

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

 260 265 270
 Gln Ala Met Lys Asp Ala Leu Lys Arg Ile Gln Asp Gly Ser Phe Ala
 275 280 285

Lys Glu Phe Met Asp Asp Tyr Arg Asn Gly Phe Lys Lys Leu Tyr Lys
 290 295 300
 Met Arg Glu Arg Ser Ala Asn Ser Leu Leu Ser Arg Val Gly Ala Asp
 305 310 315 320
 Leu Arg Asp His Met Ser Phe Val Gly Glu Ala Asp Lys Tyr Ser Thr

325 330 335
 Pro Thr Glu Glu Lys Ala Glu Ala Glu Ala Glu Lys
 340 345

<210> 108

<211> 1002

<212> DNA

<213> Staphylococcus haemolyticus

<400> 108

atgactaaag tttattacga tcaatcagtt gagaaagatg cattacaagg taaaaaaatt	60
gcaattatcg gttacggttc acaaggacac gcgcatgcac aaaaccttaa agacaacggt	120
tatgacgtta ttgttggtat tcgccctggt cattcttttg ataaagctaa agaagatggc	180
ttcgatgttt atccagttgc agaagctggt aaacaagcag atgtcatcat ggtgttattg	240
ccagatgaaa tccaaggaaa tgtctataaa aatgaaattg aacctaactt agaagctgga	300
aatgccctag catttgcgca tggttttaat atccactttg aagttatcaa accacctaaa	360
gatgtggatg tcttcttagt agctcctaaa ggacctggtc atttagtaag acgtaccttt	420
gltgaaggaa ctgccgtacc agcactatit ggtgtacaac aagatgctac tggtaagca	480
cgcgatatit cattaagtta tgcgaaaggt atcggagcta cgcgtgcagg tgtaattgaa	540
acaacattta aagaagaac tgaaacagat ttattcgggtg agcaagcagt actttgtggt	600
ggtattcata aattaatcca aagtggcttt gaaactttag ttgaagcggg ctatcaaaaa	660
gaattagcat acttcaagt attacatgaa atgaagttaa ttgttgattt aatgtatgaa	720
ggcggtatgg agaacgttcg ttactcaatt tctaacacag cagagtatgg tgactatgtt	780
tcaggaccta gagtcacac tctgatgtt aaagacaata tgaaagcagt attaaaagat	840
attcaaaatg gtaattttgc aaatagcttt gttaaagaca atgaaaatgg ctttaaagaa	900
ttctataaat tacgtgaaca acaacacggt catgaaatcg aagcagttgg tcgcgaatta	960
agaaaaatga tgccattcat taaatctaaa agtattcaaa aa	1002

<210> 109

<211> 334

<212> PRT

<213> *Staphylococcus haemolyticus*

<400> 109

Met Thr Lys Val Tyr Tyr Asp Gln Ser Val Glu Lys Asp Ala Leu Gln

1 5 10 15

Gly Lys Lys Ile Ala Ile Ile Gly Tyr Gly Ser Gln Gly His Ala His

20 25 30

Ala Gln Asn Leu Lys Asp Asn Gly Tyr Asp Val Ile Val Gly Ile Arg

35 40 45

Pro Gly His Ser Phe Asp Lys Ala Lys Glu Asp Gly Phe Asp Val Tyr

50 55 60

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

65 70 75 80

Pro Asp Glu Ile Gln Gly Asn Val Tyr Lys Asn Glu Ile Glu Pro Asn

85 90 95

Leu Glu Ala Gly Asn Ala Leu Ala Phe Ala His Gly Phe Asn Ile His

100 105 110

Phe Glu Val Ile Lys Pro Pro Lys Asp Val Asp Val Phe Leu Val Ala

115 120 125

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

130 135 140

Ala Val Pro Ala Leu Phe Gly Val Gln Gln Asp Ala Thr Gly Gln Ala

145 150 155 160

Arg Asp Ile Ser Leu Ser Tyr Ala Lys Gly Ile Gly Ala Thr Arg Ala

165 170 175

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

180 185 190

Gly Glu Gln Ala Val Leu Cys Gly Gly Ile His Lys Leu Ile Gln Ser

195 200 205

Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr Gln Lys Glu Leu Ala Tyr

210 215 220

Phe Glu Val Leu His Glu Met Lys Leu Ile Val Asp Leu Met Tyr Glu
 225 230 235 240
 Gly Gly Met Glu Asn Val Arg Tyr Ser Ile Ser Asn Thr Ala Glu Tyr
 245 250 255

Gly Asp Tyr Val Ser Gly Pro Arg Val Ile Thr Pro Asp Val Lys Asp
 260 265 270
 Asn Met Lys Ala Val Leu Lys Asp Ile Gln Asn Gly Asn Phe Ala Asn
 275 280 285
 Ser Phe Val Lys Asp Asn Glu Asn Gly Phe Lys Glu Phe Tyr Lys Leu
 290 295 300
 Arg Glu Gln Gln His Gly His Glu Ile Glu Ala Val Gly Arg Glu Leu
 305 310 315 320

Arg Lys Met Met Pro Phe Ile Lys Ser Lys Ser Ile Gln Lys
 325 330

<210> 110

<211> 1002

<212> DNA

<213> Staphylococcus epidermidis

<400> 110

atgacaaaag tatattacga cgaaacagta actcaggatg cattacaagg taaaaaaatt	60
gctgtcattg gttatggctc acaaggacat gcacatgcac aaaattttaa ggacaatggt	120
tatgatgtag tcattggcct gcgtccagga cgatcattta ataaagctaa agaagatgga	180
tttgatgttt atacggtaag tgaagcaaca caacaagcag atgtagtgat ggtactattg	240
cctgatgaaa ttcaaggtga agtatataac aaggaaatta aaccatattt agaaaaagga	300
aatgcttttag cattcgcaca cggttttaat atccatttca gtgttatcga accacctagt	360
gatgtcgatg tcitttttagt agcacctaaa ggacctggtc atttagttag acgtacattt	420
gttgaaggaa gtgccgtacc agcattattt ggtgttcaac aagatgctac aggccaagct	480
agaaacattg cttaaagcta cgcaaaaggc attgggtgcta ctctgtgccg ggtcattgaa	540
acgacattta aagaagaac tgaaacagat ttattcgggtg aacaagctgt actttgtgga	600
ggagtttcca aattaattca gagtggattc gaaacacttg tggaagcagg ttaccaacct	660
gaattagctt attttgaagt cttacacgaa atgaaattaa ttgttgattt aatgtatgaa	720

ggcggaatgg aaaatgtccg ttattctatc tctaactctg ctgaatttgg cgactatgtt 780
tctggaccaa gagtaattac acctaattgt aaagaaaata tgaaaaaagt acttgaagat 840
attcaaatg gtaacttttag ccgtagattt gttgaagata acaaaaatgg ctttaaagaa 900
ttctatcaat tacgtgaaga tcaacatggt catcaaattg aacaagttgg acgtgaatta 960
agagaaatga tgccattcat taaatctaaa agtattgaaa aa 1002

<210> 111

<211> 334

<212> PRT

<213> Staphylococcus epidermidis

<400> 111

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

1 5 10 15

Gly Lys Lys Ile Ala Val Ile Gly Tyr Gly Ser Gln Gly His Ala His

20 25 30

Ala Gln Asn Leu Lys Asp Asn Gly Tyr Asp Val Val Ile Gly Leu Arg

35 40 45

Pro Gly Arg Ser Phe Asn Lys Ala Lys Glu Asp Gly Phe Asp Val Tyr

50 55 60

Thr Val Ser Glu Ala Thr Gln Gln Ala Asp Val Val Met Val Leu Leu

65 70 75 80

Pro Asp Glu Ile Gln Gly Glu Val Tyr Asn Lys Glu Ile Lys Pro Tyr

85 90 95

Leu Glu Lys Gly Asn Ala Leu Ala Phe Ala His Gly Phe Asn Ile His

100 105 110

Phe Ser Val Ile Glu Pro Pro Ser Asp Val Asp Val Phe Leu Val Ala

115 120 125

Pro Lys Gly Pro Gly His Leu Val Arg Arg Thr Phe Val Glu Gly Ser

130 135 140

Ala Val Pro Ala Leu Phe Gly Val Gln Gln Asp Ala Thr Gly Gln Ala

145 150 155 160

Arg Asn Ile Ala Leu Ser Tyr Ala Lys Gly Ile Gly Ala Thr Arg Ala

165 170 175

Gly Val Ile Glu Thr Thr Phe Lys Glu Glu Thr Glu Thr Asp Leu Phe
180 185 190

Gly Glu Gln Ala Val Leu Cys Gly Gly Val Ser Lys Leu Ile Gln Ser
195 200 205

Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr Gln Pro Glu Leu Ala Tyr
210 215 220

Phe Glu Val Leu His Glu Met Lys Leu Ile Val Asp Leu Met Tyr Glu
225 230 235 240

Gly Gly Met Glu Asn Val Arg Tyr Ser Ile Ser Asn Thr Ala Glu Phe
245 250 255

Gly Asp Tyr Val Ser Gly Pro Arg Val Ile Thr Pro Asn Val Lys Glu
260 265 270

Asn Met Lys Lys Val Leu Glu Asp Ile Gln Asn Gly Asn Phe Ser Arg
275 280 285

Arg Phe Val Glu Asp Asn Lys Asn Gly Phe Lys Glu Phe Tyr Gln Leu
290 295 300

Arg Glu Asp Gln His Gly His Gln Ile Glu Gln Val Gly Arg Glu Leu
305 310 315 320

Arg Glu Met Met Pro Phe Ile Lys Ser Lys Ser Ile Glu Lys
325 330

<210> 112

<211> 1021

<212> DNA

<213> Streptococcus pneumoniae

<400> 112

atgacagttc aaatggaata tgaaaaagat gttaaagtag cagcacttga cggtaaaaaa 60

atcgccgtta tcggttatgg ttcacaaggg catgcgcatg ctcaaaactt gcgtgattca 120

ggtcgtgacg ttattatcgg tgtacgtcca ggtaaacttt ttgataaagc aaaagaagat 180

ggatttgata cttacacagt aacagaagct actaagttgg ctgatgttat catgatcttg 240

gcgccagacg aaattcaaca agaattgtac gaagcagaaa tcgctccaaa cttggaagct 300

ggaaacgcag ttggatttgc ccatggtttc aacatccact ttgaatttat caaagttcct 360

gcggatgtag atgtcttcatt gtgtgctcct aaaggaccag gacacttggt acgtcgtact 420
 tacgaagaag gatttgggtg tccagctctt tatgcagtat accaagatgc aacaggaaat 480
 gctaaaaaca ttgctatgga ctggtgtaaa ggtgttgag cggctcgtgt aggtcttctt 540
 gaaacaactt acaaagaaga aactgaagaa gatttgtttg gtgaacaagc tgtactttgt 600
 ggtggtttga ctgcccttat cgaagcaggt ttcgaagtct tgacagaagc aggttacgct 660
 ccagaattgg cttactttga agttcttcac gaaatgaaat tgatcgttga cttgatctac 720

 gaagggtggat tcaagaaaat gcgtcaatct atttcaaaca ctgctgaata cggtgactat 780
 gtatcaggtc cactgttaat cactgaacaa gttaaagaaa atatgaagc tgtcttggca 840
 gacatccaaa atggtaaatt tgcaaatgac ttgttaaatg actataaagc tggacgtcca 900
 aaattgactg cttaccgtga acaagcagct aaccttgaaa ttgaaaaagt tggcgcagaa 960
 ttgcgtaaag caatgccatt cgttggtaaa aacgatgatg atgcattcaa aatctataac 1020
 t 1021

<210> 113

<211> 340

<212> PRT

<213> Streptococcus pneumoniae

<400> 113

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

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

115	120	125	
Ala Pro Lys Gly Pro Gly His Leu Val Arg Arg Thr Tyr Glu Glu Gly			
130	135	140	
Phe Gly Val Pro Ala Leu Tyr Ala Val Tyr Gln Asp Ala Thr Gly Asn			
145	150	155	160
Ala Lys Asn Ile Ala Met Asp Trp Cys Lys Gly Val Gly Ala Ala Arg			
	165	170	175
Val Gly Leu Leu Glu Thr Thr Tyr Lys Glu Glu Thr Glu Glu Asp Leu			
180	185	190	
Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Leu Thr Ala Leu Ile Glu			
195	200	205	
Ala Gly Phe Glu Val Leu Thr Glu Ala Gly Tyr Ala Pro Glu Leu Ala			
210	215	220	
Tyr Phe Glu Val Leu His Glu Met Lys Leu Ile Val Asp Leu Ile Tyr			
225	230	235	240
Glu Gly Gly Phe Lys Lys Met Arg Gln Ser Ile Ser Asn Thr Ala Glu			
245	250	255	
Tyr Gly Asp Tyr Val Ser Gly Pro Arg Val Ile Thr Glu Gln Val Lys			
260	265	270	
Glu Asn Met Lys Ala Val Leu Ala Asp Ile Gln Asn Gly Lys Phe Ala			
275	280	285	
Asn Asp Phe Val Asn Asp Tyr Lys Ala Gly Arg Pro Lys Leu Thr Ala			
290	295	300	
Tyr Arg Glu Gln Ala Ala Asn Leu Glu Ile Glu Lys Val Gly Ala Glu			
305	310	315	320
Leu Arg Lys Ala Met Pro Phe Val Gly Lys Asn Asp Asp Ala Phe			
	325	330	335
Lys Ile Tyr Asn			
	340		
<210> 114			
<211> 1008			

<212> DNA

<213> Streptococcus pneumoniae

<400> 114

atggaatatg aaaaagatgt taaagtagca gcacttgacg gtaaaaaaat cgccgttatac 60
ggttatgggt cacaaggga tgcgcacgtc caaaacttgc gtgattcagg tcgtgacgtt 120
attatcgggt tacgtccagg taaatctttt gataaagcaa aagaagatgg atttgatact 180
tacacagtag cagaagctac taagttggct gatgttatca tgatcttggc gccagacgaa 240

attcaacaag aattgtacga agcagaaatc gtcctaaact tggaagctgg aaacgcagtt 300
ggatttgccc atggtttcaa catccacttt gaatttatca aagttcctgc ggatgtagat 360
gtcttcatgt gtgctcctaa aggaccagga cacttggtag gtcgtactta cgaagaagga 420
tttgggtgtc cagctcttta tgcagtatac caagatgcaa caggaaatgc taaaaacatt 480
gctatggact gggttaaagg tgttggagcg gtcgtgtag gtcttcttga aacaacttac 540
aaagaagaaa ctgaagaaga tttgtttgtt gaacaagctg tactttgtgg tggtttgact 600
gcccttatcg aagcaggttt cgaagtcttg acagaagcag gttacgtcc agaattggct 660

tactttgaag ttcttcacga aatgaaattg atcgttgact tgatctacga aggtggattc 720
aagaaaatgc gtcaatctat ttcaaacact gctgaatacg gtgactatgt atcaggcca 780
cgtgtaatac ctgaacaagt taaagaaaat atgaaggctg tcttggcaga catccaaaat 840
ggtaaatgtt caaatgactt tgtaaatgac tataaagctg gacgtccaaa attgactgct 900
taccgtgaac aagcagctaa ccttgaaatt gaaaaagttg gtgcagaatt gcgtaaagca 960
atgccattcg ttggtaaaaa cgacgatgat gcattcaaaa tctataac 1008

<210> 115

<211> 336

<212> PRT

<213> Streptococcus pneumoniae

<400> 115

Met Glu Tyr Glu Lys Asp Val Lys Val Ala Ala Leu Asp Gly Lys Lys
1 5 10 15
Ile Ala Val Ile Gly Tyr Gly Ser Gln Gly His Ala His Ala Gln Asn
20 25 30
Leu Arg Asp Ser Gly Arg Asp Val Ile Ile Gly Val Arg Pro Gly Lys
35 40 45
Ser Phe Asp Lys Ala Lys Glu Asp Gly Phe Asp Thr Tyr Thr Val Ala

50 55 60
 Glu Ala Thr Lys Leu Ala Asp Val Ile Met Ile Leu Ala Pro Asp Glu
 65 70 75 80
 Ile Gln Gln Glu Leu Tyr Glu Ala Glu Ile Ala Pro Asn Leu Glu Ala
 85 90 95
 Gly Asn Ala Val Gly Phe Ala His Gly Phe Asn Ile His Phe Glu Phe
 100 105 110
 Ile Lys Val Pro Ala Asp Val Asp Val Phe Met Cys Ala Pro Lys Gly
 115 120 125

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

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

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

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

305 310 315 320

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

325 330 335

<210> 116

<211> 1020

<212> DNA

<213> Streptococcus sanguinis

<400> 116

atggcagtaa caatggaata cgaaaaagat gttaaagtag cagcgcttga cggtaaaaaa 60
atcgccgtta tcggttatgg ttcaacaagga catgcgcatg cacaaaactt gcgtgataca 120
ggccacgatg tgattatcgg tggttcgtcct ggtaagtcac ttgacaaggc taaagaagat 180
ggctttgata cttatacagt agcagaagca gctaaattag ctgatgtcat catgattttg 240

gtccagatg aaatccaaca agatctctat gaagcagaaa tcgctccaaa cttggaagct 300
ggaaacgcag ttgggtttgc tcatggtttc aacatccatt ttgagtttat caaagttcct 360
gccgatgtag atgtctttat gtgtgctcct aaaggccctg gtcacttggt tcgccgtact 420
ttcgaagaag gttttgggtg accagctctg tatgcagtct accaagacgc tactggaaat 480
gccaaagaca tcgcaatgga ctggtgtaaa ggtgttggtt cagctcgtgt tggacttctt 540
gaaacaactt ataaagaaga aacagaagaa gatctctttg gtgaacaagc tgtgctttgt 600
ggtggcttga ctgccctgat tgaagcagga tttgaagttc tgactgaagc aggttatgcg 660

ccagaattgg cttactttga agtgcctgcac gaaatgaaac tgattgtaga cttgatctat 720
gaaggtggct tcaagaagat gcgccaatct atctcaaata cagctgaata tgggtgactat 780
glatctggct cgcgcgtcat tactgagcaa gtcaaagaaa acatgaaggc agtcttggct 840
gatatccaaa acggtaaatt tgcaaacgac ttcgttgatg attacaaggc tggtcgtcca 900
aaactcactg cttaccgcga gcaagctgct aatctggaga ttgaaaaggt cgggtgcagaa 960
ttgcgtaaag caatgccatt cgttggtaaa aacgacgacg atgctttcaa gatttataat 1020

<210> 117

<211> 340

<212> PRT

<213> Streptococcus sanguinis

<400> 117

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

245										250										255														
Tyr	Gly	Asp	Tyr	Val	Ser	Gly	Pro	Arg	Val	Ile	Thr	Glu	Gln	Val	Lys																			
260					265					270																								
Glu	Asn	Met	Lys	Ala	Val	Leu	Ala	Asp	Ile	Gln	Asn	Gly	Lys	Phe	Ala																			
275					280					285																								
Asn	Asp	Phe	Val	Asp	Asp	Tyr	Lys	Ala	Gly	Arg	Pro	Lys	Leu	Thr	Ala																			
290					295					300																								
Tyr	Arg	Glu	Gln	Ala	Ala	Asn	Leu	Glu	Ile	Glu	Lys	Val	Gly	Ala	Glu																			
305					310					315					320																			
Leu	Arg	Lys	Ala	Met	Pro	Phe	Val	Gly	Lys	Asn	Asp	Asp	Asp	Ala	Phe																			
325					330					335																								
Lys	Ile	Tyr	Asn																															
340																																		
<210> 118																																		
<211> 1020																																		
<212> DNA																																		
<213> Streptococcus salivarius																																		
<400> 118																																		
atggcagttc aaatggaata tgaaaaagac gtaaaagtag cagcacttga cggtaaaaaa															60																			
atcgccgtta tcggttacgg ttcacaaggt catgcacatg cacaaaactt gcgtgataca															120																			
ggtcacgatg tgattatcgg tgttcgtcca ggtaaatcat ttgataaagc taaagaagat															180																			
ggttttgaca cttacacagt agcagaagca actaaattgg ctgatgttat catgattttg															240																			
gctccagacg aaatccaaca agaactttac gaagcagaaa ttgctccaaa cttgaagct															300																			
ggtaatgctg ttggttttgc acacggattt aacatccact tcgaattcat caaagttcct															360																			
aaagatgttg atgtcttcat gtgtgctcct aaaggaccag gtcacttggc acgtcgtaca															420																			
tttgaagaag gatttgggtg tctgcactt tatgcagtat accaagatgc tacaggtaac															480																			
gctaaagata tcgctatgga ctggtgtaaa ggtgttggcg cagcacgtgt tggctcttctt															540																			
gaaacaacat acaaagaaga aactgaagaa gatttgttcg gtgaacaagc agttctttgt															600																			
ggtggtttga ctgcccttat cgaagctggt tttgaagtct tgactgaagc tggatatgct															660																			
ccagaattgg cttactttga agtgcttcac gaaatgaaat tgatcgttga cttgatttac															720																			
gaaggtggat tcaagaaaaa gcgtcaatca atttctaaca ctgctgaatt cggtgactac															780																			

glatcaggtc cactgtttat cactgaacaa gtgaaagaaa acatgaaagc agttcttgct 840
gacatccaaa acggtaaatt cgctaacgac ttcgttaacg actacaaagc tggacgtcca 900
aaacttactg cttaccgtga agaagctgct aaccttgaaa tcgaaaaagt tgggtgctgaa 960
ttgcgtaaag caatgccatt cgttgggtcaa aacgacgatg acgcattcaa aatctataac 1020

<210> 119

<211> 340

<212> PRT

<213> Streptococcus salivarius

<400> 119

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

1 5 10 15

Asp Gly Lys Lys Ile Ala Val Ile Gly Tyr Gly Ser Gln Gly His Ala

20 25 30

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

35 40 45

Arg Pro Gly Lys Ser Phe Asp Lys Ala Lys Glu Asp Gly Phe Asp Thr

50 55 60

Tyr Thr Val Ala Glu Ala Thr Lys Leu Ala Asp Val Ile Met Ile Leu

65 70 75 80

Ala Pro Asp Glu Ile Gln Gln Glu Leu Tyr Glu Ala Glu Ile Ala Pro

85 90 95

Asn Leu Glu Ala Gly Asn Ala Val Gly Phe Ala His Gly Phe Asn Ile

100 105 110

His Phe Glu Phe Ile Lys Val Pro Lys Asp Val Asp Val Phe Met Cys

115 120 125

Ala Pro Lys Gly Pro Gly His Leu Val Arg Arg Thr Phe Glu Glu Gly

130 135 140

Phe Gly Val Pro Ala Leu Tyr Ala Val Tyr Gln Asp Ala Thr Gly Asn

145 150 155 160

Ala Lys Asp Ile Ala Met Asp Trp Cys Lys Gly Val Gly Ala Ala Arg

165 170 175

Val Gly Leu Leu Glu Thr Thr Tyr Lys Glu Glu Thr Glu Glu Asp Leu

180	185	190	
Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Leu Thr Ala Leu Ile Glu			
195	200	205	
Ala Gly Phe Glu Val Leu Thr Glu Ala Gly Tyr Ala Pro Glu Leu Ala			
210	215	220	
Tyr Phe Glu Val Leu His Glu Met Lys Leu Ile Val Asp Leu Ile Tyr			
225	230	235	240
Glu Gly Gly Phe Lys Lys Met Arg Gln Ser Ile Ser Asn Thr Ala Glu			
245	250	255	
Phe Gly Asp Tyr Val Ser Gly Pro Arg Val Ile Thr Glu Gln Val Lys			
260	265	270	
Glu Asn Met Lys Ala Val Leu Ala Asp Ile Gln Asn Gly Lys Phe Ala			
275	280	285	
Asn Asp Phe Val Asn Asp Tyr Lys Ala Gly Arg Pro Lys Leu Thr Ala			
290	295	300	
Tyr Arg Glu Glu Ala Ala Asn Leu Glu Ile Glu Lys Val Gly Ala Glu			
305	310	315	320
Leu Arg Lys Ala Met Pro Phe Val Gly Gln Asn Asp Asp Asp Ala Phe			
325	330	335	
Lys Ile Tyr Asn			
340			
<210>	120		
<211>	1020		
<212>	DNA		
<213>	Streptococcus thermophilus		
<400>	120		
atggcagttc aaatggaata tgaaaaagac gtaaaagtac cagcacttga tggtaaaaaa		60	
atcgccgtta tcggttacgg ttcacaaggt cacgctcact cacaaaactt gcgtgataca		120	
ggtcacgatg ttatcatcgg tgttcgtcca ggtaaactcat ttgataaagc taaagaagat		180	
ggttttgata cttacacagt agcagaagca actaaattgg ctgatgttat catgattttg		240	
gctccagacg aaatccaaca agaactttac gaagcagaaa tcgctccaaa tcttgaagct		300	

ggtaacgctg ttggttttgc tcacggtttc aacatccact ttgaatttat taaagtacca 360
gcagatgttg atgtatttat gtgtgcacct aaaggaccag gtcacttggt acgtcgtaca 420
tttgaagaag gttttggtgt acctgccctt tacgcagtat accaagatgc cacaggtaac 480
gctaaagaca tcgctatgga ctggigtataa ggtattggtg cagcacgtgt aggtcttctt 540
gaaacaacat acaaagaaga aactgaagaa gatttggttg gtgaacaagc agttctttgt 600
ggtggtttga ctgcccttat cgaaactggt tttgaagtct tgactgaagc tggatatgca 660

ccagaattgg cttactttga agtacttcac gaaatgaaat taatcgttga cttgatttac 720
gaaggtggat tcaagaaaat gcgtcaatca atttctaaca ctgctgaatt cggtgactat 780
glatcaggtc cacgtgttat cactgaacaa gttaaagaaa acatgaaagc agttcttgct 840
gatatccaaa atggtaaatt cgctaagac ttcgttaacg actacaaagc tggacgtcca 900
aaacttactg cttaccgtga agaagctgct aaccttgaaa tcgaaaaagt tggctgctgaa 960
ttgcgtaaag caatgccatt cgttggtcaa aacgacgatg atgcattcaa aatctataac 1020

<210> 121

<211> 340

<212> PRT

<213> Streptococcus thermophilus

<400> 121

Met Ala Val Gln Met Glu Tyr Glu Lys Asp Val Lys Val Pro Ala Leu

1 5 10 15

Asp Gly Lys Lys Ile Ala Val Ile Gly Tyr Gly Ser Gln Gly His Ala

20 25 30

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

35 40 45

Arg Pro Gly Lys Ser Phe Asp Lys Ala Lys Glu Asp Gly Phe Asp Thr

50 55 60

Tyr Thr Val Ala Glu Ala Thr Lys Leu Ala Asp Val Ile Met Ile Leu

65 70 75 80

Ala Pro Asp Glu Ile Gln Gln Glu Leu Tyr Glu Ala Glu Ile Ala Pro

85 90 95

Asn Leu Glu Ala Gly Asn Ala Val Gly Phe Ala His Gly Phe Asn Ile

100 105 110

His Phe Glu Phe Ile Lys Val Pro Ala Asp Val Asp Val Phe Met Cys

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

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

 Phe Gly Asp Tyr Val Ser Gly Pro Arg Val Ile Thr Glu Gln Val Lys
 260 265 270
 Glu Asn Met Lys Ala Val Leu Ala Asp Ile Gln Asn Gly Lys Phe Ala
 275 280 285
 Asn Asp Phe Val Asn Asp Tyr Lys Ala Gly Arg Pro Lys Leu Thr Ala
 290 295 300
 Tyr Arg Glu Glu Ala Ala Asn Leu Glu Ile Glu Lys Val Gly Ala Glu
 305 310 315 320

 Leu Arg Lys Ala Met Pro Phe Val Gly Gln Asn Asp Asp Asp Ala Phe
 325 330 335
 Lys Ile Tyr Asn
 340
 <210> 122
 <211> 975

<212> DNA

<213> Streptococcus pneumoniae

<400> 122

ttgacggtaa aaaaatacgc cgttatcggg tatggttcac aaggcatgc gcatgctcaa 60
aacttgcgtg attcaggctg tgacgttatt ataggtgtac gtccaggtaa atctttgat 120
aaagcaaaag aagatggatt tgatacttac acagtagcag aagctactaa gtggctgat 180
gttatcatga tcttggcgcc agacgaaatt caacaagaat tgtacgaagc agaaatcgct 240

ccaaacttgg aagctggaaa cgcagttgga ttgcccattg gtttcaacat ccactttgaa 300
tttatcaaag ttcttgcgga tgtagatgtc ttcattgttg ctctaaagg accaggacac 360
ttggtacgtc gtacttacga agaaggattt ggtgttccag ctctttatgc agtatacca 420
gatgcaacag gaaatgctaa aaacattgct atggactggg gtaaagggtg tggagcggct 480
cgtgtaggtc ttcttgaac aacttacaaa gaagaaactg aagaagattt gtttggtgaa 540
caagctgtac tttgtggagg ttgactgcc cttatcgaag caggtttcga agtcttgaca 600
gaagcaggtt acgtccaga attggcttac ttgaagtgc ttcacgaaat gaaattgac 660

gttgacttga tctacgaagg tggattcaag aaaatgcgtc aatctatttc aaacactgct 720
gaatacgggt actatgtatc aggtccacgt gtaatcactg aacaagttaa agaaaatatg 780
aaggctgtct tggcagacat ccaaaatggg aaatttgcaa atgactttgt aaatgactat 840
aaagctggac gtccaaaatt gactgcttac cgtgaacaag cagctaacct tgaaattgaa 900
aaagtgggtg cagaattgcg taaagcaatg ccattcgttg gtaaaaacga cgatgatgca 960
ttcaaaatct ataac 975

<210> 123

<211> 325

<212> PRT

<213> Streptococcus pneumoniae

<400> 123

Met Thr Val Lys Lys Tyr Ala Val Ile Gly Tyr Gly Ser Gln Gly His
1 5 10 15
Ala His Ala Gln Asn Leu Arg Asp Ser Gly Arg Asp Val Ile Ile Gly
20 25 30
Val Arg Pro Gly Lys Ser Phe Asp Lys Ala Lys Glu Asp Gly Phe Asp
35 40 45
Thr Tyr Thr Val Ala Glu Ala Thr Lys Leu Ala Asp Val Ile Met Ile

50 55 60
 Leu Ala Pro Asp Glu Ile Gln Gln Glu Leu Tyr Glu Ala Glu Ile Ala
 65 70 75 80
 Pro Asn Leu Glu Ala Gly Asn Ala Val Gly Phe Ala His Gly Phe Asn
 85 90 95
 Ile His Phe Glu Phe Ile Lys Val Pro Ala Asp Val Asp Val Phe Met
 100 105 110
 Cys Ala Pro Lys Gly Pro Gly His Leu Val Arg Arg Thr Tyr Glu Glu
 115 120 125

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

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

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

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

305 310 315 320

Phe Lys Ile Tyr Asn

325

<210> 124

<211> 1020

<212> DNA

<213> Lactococcus lactis

<400> 124

atggcagtta caatgtatta tgaagatgat gtagaagtat cagcacttgc tggaaagcaa 60
attgcagtaa tcggttatgg ttcacaagga catgctcacg cacagaattt gcgtgattct 120
ggtcacaacg ttatcattgg tgtgcgccac ggaaaatctt ttgataaagc aaaagaagat 180
ggctttgaaa catttgaagt aggagaagcg gtagctaaag ctgatgttat tatggttttg 240
gcgccagatg aacttcaaca atccatttat gaagaggaca tcaaaccaaa cttgaaagca 300
ggttcagcac ttggttttgc tcacggattt aatatccatt ttggctatat taaagtacca 360

gaagacgttg acgtctttat ggttgcacct aaggctccag gtcaccttgt ccgtcggact 420
tatactgaag gttttggtac accagctttg ttgtttcac accaaaatgc aagtggatcat 480
gcgcgtgaaa tcgcaatgga ttgggccaaa ggaattgggt gtgctcgagt gggaattatt 540
gaaacaacct ttaaagaaga aacagaagaa gatattgttg gagaacaagc tgttctatgt 600
ggaggtttga cagcacttgt tgaagccggt ttgaaacac tgacagaagc tggatacgct 660
ggcgaattgg ctactttga agttttgcac gaaatgaaat tgattgttga cctcatgtat 720
gaaggtgggt ttactaaaat gcgtcaatcc atctcaaata ctgctgagtt tggcgattat 780

gtgactggtc caaggattat tactgacgca gttaaaaaga atatgaagct tgttttggct 840
gatattcaat ctggaaaatt tgctcaagat ttcgttgatg acttcaaagc ggggcgtcca 900
aaattaacag cctatcgca agctgctaaa aatcttgaaa ttgaaaaaat tggggcagaa 960
ttacgtaaag caatgccatt cacacaatct ggtgatgacg atgcctttaa aatctatcag 1020

<210> 125

<211> 340

<212> PRT

<213> Lactococcus lactis

<400> 125

Met Ala Val Thr Met Tyr Tyr Glu Asp Asp Val Glu Val Ser Ala Leu

1 5 10 15

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

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

 Phe Gly Thr Pro Ala Leu Phe Val Ser His Gln Asn Ala Ser Gly His
 145 150 155 160
 Ala Arg Glu Ile Ala Met Asp Trp Ala Lys Gly Ile Gly Cys Ala Arg
 165 170 175
 Val Gly Ile Ile Glu Thr Thr Phe Lys Glu Glu Thr Glu Glu Asp Leu
 180 185 190
 Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Leu Thr Ala Leu Val Glu
 195 200 205

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

Phe Gly Asp Tyr Val Thr Gly Pro Arg Ile Ile Thr Asp Ala Val Lys
260 265 270

Lys Asn Met Lys Leu Val Leu Ala Asp Ile Gln Ser Gly Lys Phe Ala
275 280 285

Gln Asp Phe Val Asp Asp Phe Lys Ala Gly Arg Pro Lys Leu Thr Ala
290 295 300

Tyr Arg Glu Ala Ala Lys Asn Leu Glu Ile Glu Lys Ile Gly Ala Glu
305 310 315 320

Leu Arg Lys Ala Met Pro Phe Thr Gln Ser Gly Asp Asp Asp Ala Phe
325 330 335

Lys Ile Tyr Gln
340

<210> 126

<211> 1041

<212> DNA

<213> Leuconostoc mesenteroides

<400> 126

atgactacaa aaatgtttta tgacaaagat attgatacaa aaccattgga aaacaaaaag	60
attgcagtta ttggctatgg tgcacaaggg catgctcaag cgaataatct tcgcgactca	120
ggatttgatg tcatcatggg attgcgcccc ggaaaatatt ttgatagtgc taaaaaagat	180
gggttcgaag ttactcagc tgctgaagca acggcacaag cggattttgt tatgatggaa	240
accctgacg aattacaagc ggcagtttgg gagaaagaag ttgagcctaa ccttaaggca	300
ggttcttaca ttggattttc tcatgggttc aacattgttt atggtttgat taagccaat	360
gctgatatta atgcatgat catcgcgcca aagggtccag gaaacattga acgtcgccaa	420
ttcgttgaag gggttgggat tccttctttg tatggtgttc atcaagatcc tacgggtgat	480
acagctgaag tggccaaggc ctatgcgaaa ggtattggtt caggtcgcg aggtatcttg	540
gaaacgactt ttgaagaaga aacaacagaa aacttgtttg gtgaacaggc tgtactttgt	600
ggtggcttaa cacaattgat tgaggcagga tttaacacat tgggtggaagc aggttactca	660
cctgaattgg cttatttcga aacatctcat gaaatgaaga tgattgttga ttgatcttt	720
gaaggtggtt tcgagaagat gagacatgat tgcacaaata cttgtgaata tgggtgaaatg	780
ttgaacggac cacgtatcat cactgaagaa tcaaagcaag gaatgcgtga tgttttgaag	840

gatatccaag atggtactta tgctaagaag tggttggcag aatacaattc tggtttgaag 900
gatttggaag agatgcggac agagtataag agcggtttgt acgagcaaac aggtagaag 960
gttcgtgcta tgatgcatg gatttcagat gcagataagt actcgacagc agcagatact 1020
gagcaattct cagcagcaaa g 1041

<210> 127

<211> 347

<212> PRT

<213> Leuconostoc mesenteroides

<400> 127

Met Thr Thr Lys Met Phe Tyr Asp Lys Asp Ile Asp Thr Lys Pro Leu

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

65 70 75 80
Thr Pro Asp Glu Leu Gln Ala Ala Val Trp Glu Lys Glu Val Glu Pro
85 90 95
Asn Leu Lys Ala Gly Ser Tyr Ile Gly Phe Ser His Gly Phe Asn Ile
100 105 110
Val Tyr Gly Leu Ile Lys Pro Asn Ala Asp Ile Asn Val Met Ile Ile
115 120 125
Ala Pro Lys Gly Pro Gly Asn Ile Glu Arg Arg Gln Phe Val Glu Gly

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

180	185	190	
Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Leu Thr Gln Leu Ile Glu			
195	200	205	
Ala Gly Phe Asn Thr Leu Val Glu Ala Gly Tyr Ser Pro Glu Leu Ala			
210	215	220	
Tyr Phe Glu Thr Ser His Glu Met Lys Met Ile Val Asp Leu Ile Phe			
225	230	235	240
Glu Gly Gly Phe Glu Lys Met Arg His Asp Cys Ser Asn Thr Cys Glu			
245	250	255	
Tyr Gly Glu Met Leu Asn Gly Pro Arg Ile Ile Thr Glu Glu Ser Lys			
260	265	270	
Gln Gly Met Arg Asp Val Leu Lys Asp Ile Gln Asp Gly Thr Tyr Ala			
275	280	285	
Lys Lys Trp Leu Ala Glu Tyr Asn Ser Gly Leu Lys Asp Leu Glu Lys			
290	295	300	
Met Arg Thr Glu Tyr Lys Ser Gly Leu Tyr Glu Gln Thr Gly Lys Lys			
305	310	315	320
Val Arg Ala Met Met Pro Trp Ile Ser Asp Ala Asp Lys Tyr Ser Thr			
325	330	335	
Ala Ala Asp Thr Glu Gln Phe Ser Ala Ala Lys			
340	345		
<210> 128			
<211> 1041			
<212> DNA			
<213> Leuconostoc mesenteroides			
<400> 128			
atgactacaa aaatgtttta tgacaaagat attgatacaa aaccattgga aaacaaaaag	60		
attgcagtta ttggctatgg tgcacaaggg catgctcaag cgaataatct tcgcgtctca	120		
ggattttgatg tcatcatggg attgcgccca ggaaaatctt ttgatagtgt taaaaaagat	180		
gggttcgaag ttactcagc tgctgaagca acggcacaag cggatgttgt tatgatggaa	240		
accctgacg aattacaagc ggcagtttgg gagaaagaag ttgagcctaa ccttaaggca	300		

ggttctgacc ttggattttc tcatgggttc aacattgttt atggtttgat taagcccaat 360
gctgatatta atgcatgat cattgcgcca aagggtccag gaaacattga acgtcgccaa 420
ttcgttgaag ggggtgggat tccttctttg tatggtgttc atcaagatcc tacgggtgat 480
acagctgaag tggccaaggc ctatgcgaaa ggtattggtt caggttgccg aggtatcttg 540
gaaacgactt ttgaagaaga aacaacagaa gacttgtttg gtgaacaagc tgtactttgt 600
ggtggcttaa cacaattgat tgaggcagga tttaacacat tgggtgaagc aggttactca 660

cctgaattgg cttatttcga aacatctcat gaaatgaaga tgattgttga tttgatcttt 720
gaagtggtt tcgagaagat gagacatgat tgcctaaata cttgtgaata tggtgaaatg 780
ttgaacggac cacgtatcat cactgaagaa tcaaagcaag gaatgcgtga tgttttgaag 840
gatatccaag atggtactta tgctaagaag tggttggcag aatacaattc tggtttgaag 900
gatttggaag agatcggag agagtataag agcggtttgt acgagcaaac aggtgaagaag 960
gttcgtgcta tgatgcatg gatttcagat gcagataagt actcgacagc agcagatact 1020
gagcaattct cagcagcaaa g 1041

<210> 129

<211> 347

<212> PRT

<213> *Leuconostoc mesenteroides*

<400> 129

Met Thr Thr Lys Met Phe Tyr Asp Lys Asp Ile Asp Thr Lys Pro Leu
1 5 10 15
Glu Asn Lys Lys Ile Ala Val Ile Gly Tyr Gly Ala Gln Gly His Ala
20 25 30
Gln Ala Asn Asn Leu Arg Val Ser Gly Phe Asp Val Ile Met Gly Leu
35 40 45
Arg Pro Gly Lys Ser Phe Asp Ser Val Lys Lys Asp Gly Phe Glu Val
50 55 60
Tyr Ser Ala Ala Glu Ala Thr Ala Gln Ala Asp Val Val Met Met Glu
65 70 75 80
Thr Pro Asp Glu Leu Gln Ala Ala Val Trp Glu Lys Glu Val Glu Pro
85 90 95
Asn Leu Lys Ala Gly Ser Asp Leu Gly Phe Ser His Gly Phe Asn Ile
100 105 110

Val Tyr Gly Leu Ile Lys Pro Asn Ala Asp Ile Asn Val Met Ile Ile

115 120 125

Ala Pro Lys Gly Pro Gly Asn Ile Glu Arg Arg Gln Phe Val Glu Gly

130 135 140

Gly Gly Ile Pro Ser Leu Tyr Gly Val His Gln Asp Pro Thr Gly Asp

145 150 155 160

Thr Ala Glu Val Ala Lys Ala Tyr Ala Lys Gly Ile Gly Ser Gly Cys

165 170 175

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

180 185 190

Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Leu Thr Gln Leu Ile Glu

195 200 205

Ala Gly Phe Asn Thr Leu Val Glu Ala Gly Tyr Ser Pro Glu Leu Ala

210 215 220

Tyr Phe Glu Thr Ser His Glu Met Lys Met Ile Val Asp Leu Ile Phe

225 230 235 240

Glu Gly Gly Phe Glu Lys Met Arg His Asp Cys Ser Asn Thr Cys Glu

245 250 255

Tyr Gly Glu Met Leu Asn Gly Pro Arg Ile Ile Thr Glu Glu Ser Lys

260 265 270

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

275 280 285

Lys Lys Trp Leu Ala Glu Tyr Asn Ser Gly Leu Lys Asp Leu Glu Lys

290 295 300

Met Arg Thr Glu Tyr Lys Ser Gly Leu Tyr Glu Gln Thr Gly Lys Lys

305 310 315 320

Val Arg Ala Met Met Pro Trp Ile Ser Asp Ala Asp Lys Tyr Ser Thr

325 330 335

Ala Ala Asp Thr Glu Gln Phe Ser Ala Ala Lys

340 345

<210> 130

<211> 1038

<212> DNA

<213> *Lactococcus brevis*

<400> 130

atgagtgtag aaatgttgta tgacaaggat gttaccacca attatcttca aggaaagaaa	60
attgctttta tcggttatgg atctcaaggc catgcccaag ctaacaactt aagagattca	120
ggttatgacg ttatcgttgg tgttcgccca ggacaatcat ttgaaaatgc caaaatcgat	180
ggattcgtat tttacacacc agccgaagca gctcgtcgtg cagactggat tcaaagtta	240
acgcctgatg aagttaatggg tgatgtttat aaaaatgaaa ttgcacctaa ccttgaggaa	300
ggcaatgtat taggcttttc acatggcttc aacattcatt ataaagaaat cgtaccacca	360
gcaaactgtt atgttggtat gatggcacct aaaggcccag gtaatctttg tcgtcgtaca	420
tacgttgaag gctccggtgt tccagcatta tatggttatt tccaagacta ttcaggccat	480
gccgaagatt tatccaagga attcgccaaa ggaaatggtg cagcacgtgc cggactattg	540
aagacaacct tcaagaaga aactgaagaa gatttgtttg gtgagcaaaa cgtccttatg	600
ggtggtgtta ccgcgtcat tgaaaccggt tatgaagtct taaccgaagc aggatactca	660
ccacagttgg catactttga agtagatcat gaaatgaaat tgatctgtga ctttatctac	720
gaaggtgggt tcaacaagat gtatgtgac tgctcgaaca cttctgaata tggttcatat	780
glagttggtc caaaggtcgt tggtaaggaa tccaagcaag ctatgaaaga cgctttgaag	840
cggattcaag acggttcatt tgctaaagaa ttcatggatg actaccgtaa cggttcaag	900
aagctttaca agatgcgtga acgttcagct aactcattc tttcacgtgt tgggtctgat	960
cttcgtgac acatgtcatt tgttggtgaa gccgacaagt acagtacacc tactgaagaa	1020
aaagcagaag ctgaaaag	1038

<210> 131

<211> 346

<212> PRT

<213> *Lactococcus brevis*

<400> 131

Met	Ser	Val	Glu	Met	Leu	Tyr	Asp	Lys	Asp	Val	Thr	Thr	Asn	Tyr	Leu
1				5				10						15	
Gln	Gly	Lys	Lys	Ile	Ala	Phe	Ile	Gly	Tyr	Gly	Ser	Gln	Gly	His	Ala
				20				25						30	

Gln Ala Asn Asn Leu Arg Asp Ser Gly Tyr Asp Val Ile Val Gly Val

35 40 45

Arg Pro Gly Gln Ser Phe Glu Asn Ala Lys Ile Asp Gly Phe Asp Val

50 55 60

Tyr Thr Pro Ala Glu Ala Ala Arg Arg Ala Asp Trp Ile Gln Met Leu

65 70 75 80

Thr Pro Asp Glu Val Met Gly Asp Val Tyr Lys Asn Glu Ile Ala Pro

85 90 95

Asn Leu Glu Glu Gly Asn Val Leu Gly Phe Ser His Gly Phe Asn Ile

100 105 110

His Tyr Lys Glu Ile Val Pro Pro Ala Asn Val Asp Val Val Met Met

115 120 125

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

130 135 140

Ser Gly Val Pro Ala Leu Tyr Gly Tyr Phe Gln Asp Tyr Ser Gly His

145 150 155 160

Ala Glu Asp Leu Ser Lys Glu Phe Ala Lys Gly Asn Gly Ala Ala Arg

165 170 175

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

180 185 190

Phe Gly Glu Gln Asn Val Leu Met Gly Gly Val Thr Ala Leu Ile Glu

195 200 205

Thr Gly Tyr Glu Val Leu Thr Glu Ala Gly Tyr Ser Pro Gln Leu Ala

210 215 220

Tyr Phe Glu Val Asp His Glu Met Lys Leu Ile Cys Asp Leu Ile Tyr

225 230 235 240

Glu Gly Gly Phe Asn Lys Met Tyr Ala Asp Cys Ser Asn Thr Ser Glu

245 250 255

Tyr Gly Ser Tyr Val Val Gly Pro Lys Val Val Gly Lys Glu Ser Lys

260 265 270

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

275 280 285
Lys Glu Phe Met Asp Asp Tyr Arg Asn Gly Phe Lys Lys Leu Tyr Lys

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

<210> 132
<211> 1020
<212> DNA
<213> Lactococcus lactis
<400> 132

atggcagtta caatgtatta tgaagatgat gtagaagtat cagcacttgc tggaaagcaa 60

attgcagtaa tcggttatgg ttcacaagga catgctcacg cacagaattt gcgtgattct 120
ggtcacaacg ttatcattgg tgtgcgccac ggaaaatctt ttgataaagc aaaagaagat 180
ggctttgaaa catttgaagt aggagaagca gtagctaaag ctgatgttat tatggttttg 240
gcaccagatg aacttcaaca atccatttat gaagaggaca tcaaaccaaa cttgaaagca 300
ggttcagcac ttggttttgc tcacggattt aatatccatt ttggctatat taaagtacca 360
gaagacgttg acgtctttat ggttgcgcct aaggctccag gtcaccttgt ccgtcggact 420
tatactgaag gttttgttac accagctttg tttgtttcac accaaaatgc aagtggtcac 480

gcgcgtgaaa tcgcaatgga ttgggccaaa ggaattggtt gtgctcgagt gggaattatt 540
gaaacaactt ttaaagaaga aacagaagaa gattttgttg gagaacaagc tgttctatgt 600
ggaggtttga cagcacttgt tgaagccggt tttgaaacac tgacagaagc tggatacgct 660
ggcgaattgg ctactttga agttttgcac gaaatgaaat tgattgttga cctcatgtat 720
gaagtggtt ttactaaaa gcgtcaatcc atctcaaata ctgctgagtt tggcgattat 780
gtgactggtc cacggattat tactgacgaa gttaaaaaga atatgaagct tgttttggct 840
gatattcaat ctggaaaatt tgctcaagat ttcgttgatg acttcaaagc ggggcgtcca 900

aaattaatag cctatcgca agctgcaaaa aatcttgaaa ttgaaaaaat tggggcagag 960
ctacgtcaag caatgccatt cacacaatct ggtgatgacg atgcctttaa aatctatcag 1020

<210> 133

<211> 340

<212> PRT

<213> *Lactococcus lactis*

<400> 133

Met Ala Val Thr Met Tyr Tyr Glu Asp Asp Val Glu Val Ser Ala Leu

1 5 10 15

Ala Gly Lys Gln Ile Ala Val Ile Gly Tyr Gly Ser Gln Gly His Ala

20 25 30

His Ala Gln Asn Leu Arg Asp Ser Gly His Asn Val Ile Ile Gly Val

35 40 45

Arg His Gly Lys Ser Phe Asp Lys Ala Lys Glu Asp Gly Phe Glu Thr

50 55 60

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

65 70 75 80

Ala Pro Asp Glu Leu Gln Gln Ser Ile Tyr Glu Glu Asp Ile Lys Pro

85 90 95

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

100 105 110

His Phe Gly Tyr Ile Lys Val Pro Glu Asp Val Asp Val Phe Met Val

115 120 125

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

130 135 140

Phe Gly Thr Pro Ala Leu Phe Val Ser His Gln Asn Ala Ser Gly His

145 150 155 160

Ala Arg Glu Ile Ala Met Asp Trp Ala Lys Gly Ile Gly Cys Ala Arg

165 170 175

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

180 185 190

Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Leu Thr Ala Leu Val Glu

195 200 205

Ala Gly Phe Glu Thr Leu Thr Glu Ala Gly Tyr Ala Gly Glu Leu Ala

210 215 220
Tyr Phe Glu Val Leu His Glu Met Lys Leu Ile Val Asp Leu Met Tyr

225 230 235 240
Glu Gly Gly Phe Thr Lys Met Arg Gln Ser Ile Ser Asn Thr Ala Glu

245 250 255
Phe Gly Asp Tyr Val Thr Gly Pro Arg Ile Ile Thr Asp Glu Val Lys

260 265 270
Lys Asn Met Lys Leu Val Leu Ala Asp Ile Gln Ser Gly Lys Phe Ala

275 280 285
Gln Asp Phe Val Asp Asp Phe Lys Ala Gly Arg Pro Lys Leu Ile Ala

290 295 300
Tyr Arg Glu Ala Ala Lys Asn Leu Glu Ile Glu Lys Ile Gly Ala Glu

305 310 315 320
Leu Arg Gln Ala Met Pro Phe Thr Gln Ser Gly Asp Asp Asp Ala Phe

325 330 335
Lys Ile Tyr Gln

340

<210> 134

<211> 1002

<212> DNA

<213> Staphylococcus hominis

<400> 134

atgacaaaag tttattatga tcaatcagta gaaaaagatg cactacaagg taaaaaaatc 60

gctattatcg gttacgggtc tcaaggccat gcacatgctc aaaaccttaa agataacggt 120

tacgacgtaa gagtgggtat tcgtccaggt gattcattta ataaagctaa agaagatgga 180

ttagatgttt atccagtggc tgaagcagtt aaacaagcag acgttattat ggtattactt 240

cctgatgaaa ttcaaggtaa tgtatataaa aatgaaattg ctccaaattt agaagctggt 300

aatgcattag cttttgcaca tggatttaat attcatttta acgtcatcga accacctaaa 360

gatgtagatg tcttcttagt agcacctaaa ggtcctggtc atctagtaag acgtacattt 420

gttgaaggtt cagctgtgcc agcattattt ggtgttcaac aagatgctac aggtcatgca 480

cgtgacattg cttaaagtta tgcaaaaggt attggtgcta cacgtgctgg tgttattgaa 540

acaacattta aagaggaaac tgaactgac ttattcgggtg aacaagctgt actttgtggg 600
 ggaattcata aactgattca aagtggtttc gagacattag tggaagctgg atatcaaaaa 660
 gaattagcat attttgaagt actacatgaa atgaaactta tcgttgattt aatgtatgaa 720
 ggcggtatgg aaaatgtacg ttattcaatt tcaaatactg ctgaatttgg tgactacgta 780
 tctgggtccac gtgtcattac tccagatgtt aaagacaata tgaaagaagt acttaaagat 840
 attcaaatg gtaatttcgc taatcgcttt gttaaagata acgaaaatgg atttaaagaa 900

ttttatgaat tgcgtgaaca acaacacggg catgaaattg aagcagtggg tcgtgaacta 960
 agaaaaatga tgccattcat taaatctaaa agcattcaaa aa 1002

<210> 135

<211> 334

<212> PRT

<213> Staphylococcus hominis

<400> 135

Met Thr Lys Val Tyr Tyr Asp Gln Ser Val Glu Lys Asp Ala Leu Gln

1 5 10 15

Gly Lys Lys Ile Ala Ile Ile Gly Tyr Gly Ser Gln Gly His Ala His

20 25 30

Ala Gln Asn Leu Lys Asp Asn Gly Tyr Asp Val Arg Val Gly Ile Arg

35 40 45

Pro Gly Asp Ser Phe Asn Lys Ala Lys Glu Asp Gly Leu Asp Val Tyr

50 55 60

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

65 70 75 80

Pro Asp Glu Ile Gln Gly Asn Val Tyr Lys Asn Glu Ile Ala Pro Asn

85 90 95

Leu Glu Ala Gly Asn Ala Leu Ala Phe Ala His Gly Phe Asn Ile His

100 105 110

Phe Asn Val Ile Glu Pro Pro Lys Asp Val Asp Val Phe Leu Val Ala

115 120 125

Pro Lys Gly Pro Gly His Leu Val Arg Arg Thr Phe Val Glu Gly Ser

130 135 140

Ala Val Pro Ala Leu Phe Gly Val Gln Gln Asp Ala Thr Gly His Ala

145 150 155 160
Arg Asp Ile Ala Leu Ser Tyr Ala Lys Gly Ile Gly Ala Thr Arg Ala

 165 170 175
Gly Val Ile Glu Thr Thr Phe Lys Glu Glu Thr Glu Thr Asp Leu Phe

 180 185 190
Gly Glu Gln Ala Val Leu Cys Gly Gly Ile His Lys Leu Ile Gln Ser

 195 200 205
Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr Gln Lys Glu Leu Ala Tyr

 210 215 220
Phe Glu Val Leu His Glu Met Lys Leu Ile Val Asp Leu Met Tyr Glu

225 230 235 240
Gly Gly Met Glu Asn Val Arg Tyr Ser Ile Ser Asn Thr Ala Glu Phe

 245 250 255
Gly Asp Tyr Val Ser Gly Pro Arg Val Ile Thr Pro Asp Val Lys Asp

 260 265 270
Asn Met Lys Glu Val Leu Lys Asp Ile Gln Asn Gly Asn Phe Ala Asn

 275 280 285
Arg Phe Val Lys Asp Asn Glu Asn Gly Phe Lys Glu Phe Tyr Glu Leu

 290 295 300
Arg Glu Gln Gln His Gly His Glu Ile Glu Ala Val Gly Arg Glu Leu

305 310 315 320
Arg Lys Met Met Pro Phe Ile Lys Ser Lys Ser Ile Gln Lys

 325 330

<210> 136

<211> 11

<212> DNA

<213> Artificial sequence

<220><223> PmeI restriction site and additional 3 bp (AGT)

<400> 136

gtttaaacag t

11

<210> 137

<211> 13

<212> DNA

<213> Artificial sequence

<220><223> SfiI restriction site

<400> 137

ggccctgcag gcc

13

<210> 138

<211> 9586

<212> DNA

<213> Artificial sequence

<220><223> pLH556

<400> 138

ggccgcacct ggtaaacct ctagtggagt agtagatgta atcaatgaag cggaagccaa 60

aagaccagag tagaggccta tagaagaaac tgcgatacct tttgtgatgg ctaaacaac 120

agacatcttt ttatatgttt ttacttctgt atatcgtgaa gtagtaagtg ataagcgaat 180

ttggctaaga acgttgtaag tgaacaaggg acctcttttg cctttcaaaa aaggattaaa 240

tggagttaat cattgagatt tagttttcgt tagattctgt atccctaaat aactccctta 300

cccgaaggga aggcacaaaa gacttgaata atagcaaacg gccagtagcc aagaccaaat 360

aatactagag ttaactgatg gtcttaaaca ggcattacgt ggtgaactcc aagaccaata 420

tacaaaaat cgataagtta ttcttgccca ccaatttaag gagcctacat caggacagta 480

gtaccattcc tcagagaaga ggtatacata acaagaaaat cgcgtaaca cttatataa 540

cttagcccggt tattgagcta aaaaaccttg caaaatttcc tatgaataag aatacttcag 600

acgtgataaa aatttacttt ctaactcttc tcacgtgcc cctatctgtt ctccgctct 660

accgtgagaa ataaagcatc gagtacggca gtctcgtgtc actgaactaa aacaataagg 720

ctagtgcgaa tgatgaactt gcttgctgtc aaacttctga gttgccgctg atgtgacact 780

gtgacaataa attcaaaccg gttatagcgg tctcctccgg taccggttct gccacctcca 840

atagagctca gtaggagtc gaacctctgc ggtggctgtc agtgactcat ccgcgtttcg 900

taagttgtgc gcgtgcacat ttgcgccgtt ccgcgtcatc ttgcagcagg cggaaatatt 960

catcacgctg taggacgcaa aaaaaaata attaatcgta caagaatctt ggaaaaaaa 1020

ttgaaaaatt ttgtataaaa gggatgacct aacttgactc aatggctttt acaccagta 1080

ttttcccttt ccttgtttgt tacaattata gaagcaagac aaaaacatat agacaaccta 1140

ttcctaggag ttatatTTTT ttaccctacc agcaatataa gtaaaaaact gttaaacag 1200

tatgaaagtt ttctacgata aagactgcga cctgtcgatc atccaaggta agaaagttgc	1260
catcatcggc tacggttccc agggccacgc tcaagcatgc aacctgaagg attccggcgt	1320
agacgtgact gttggcctgc gtaaaggctc ggctaccgtt gccaaaggctg aagcccacgg	1380
cttgaaagtg accgacgttg ctgcagccgt tgcgggtgcc gacttgggtca tgatcctgac	1440
cccggacgag ttccagtccc agctgtacaa gaacgaaatc gagccgaaca tcaagaaggg	1500
cgccactctg gccttctccc acggtctcgc gatccactac aaccaggttg tgcctcgtgc	1560
cgacctcgac gtgatcatga tcgcgccgaa ggctccaggc cacaccgtac gttccgagtt	1620
cgtcaagggc ggtgggtatt ctgacctgat cgcgatctac caggacgctt ccggcaacgc	1680
caagaacgtt gccctgtcct acgccgcagg cgtgggcggc ggccgtaccg gcatcatcga	1740
aaccaccttc aaggacgaga ctgaaaccga cctgttcggt gagcaggctg ttctgtgtgg	1800
cggtagcgtc gagctgggtca aagccggttt cgaaaccctg gttgaagctg gctacgctcc	1860
agaaatggcc tacttcgagt gcctgcacga actgaagctg atcgttgacc tcatgtacga	1920
aggcggatc gccaacatga actactgat ctccaacaac gctgaatacg gcgagtacgt	1980
gactgggtcca gaagtcatca acgccgaatc ccgtcaggcc atgcgcaatg ctctgaagcg	2040
catccaggac gggaatacgc cgaagatgtt catcagcgaa ggcgctaccg gctacccatc	2100
gatgaccgcc aagcgtcgta acaacgtgc tcacggatc gaaatcatcg gcgagcaact	2160
gcgtcgatg atgccttggc tcgggtgcaa caaatcgtc gacaaagcca agaactaagg	2220
ccctgcagcg ctatcaagtg ctggaaactt tttctcttgg aatttttgca acatcaagtc	2280
atagtcaatt gaattgacct aatttcacat ttaagatttt tttttttca tccgacatac	2340
atctgtacac taggaagccc tgtttttctg aagcagcttc aaatataat attttttaca	2400
tatttattat gattcaatga acaatctaataaatcgaaa acaagaaccg aaacgcgaat	2460
aaataattta tttagatggt gacaagtgtg taagtcttca tcgggacagc tacgatttct	2520
ctttcggttt tggctgagct actggttgct gtgacgcagc ggcattagcg cggcgttatg	2580
agctaccctc gtggcctgaa agatggcggg aataaagcgg aactaaaaat tactgactga	2640
gccatattga ggtcaatttg tcaactcgtc aagtcacgtt tggtaggacgg cccctttcca	2700
acgaatcgta tatactaaca tgcgcgcgct tcctatatac acatatacat atatatatat	2760
atatatatgt gtgcgtgat gtgtacacct gtatttaatt tcttactcg cgggtttttc	2820
ttttttctca attcttggct tctcttttct cgagcggacc ggatcctccg cggtagcggc	2880
agatctattt aaatggcgcg ccgacgtcag gtggcacttt tcggggaaat gtgcgcggaa	2940
cccctatttg tttatttttc taaatacatt caaatatgta tccgtcatg agacaataac	3000
cctgataaat gcttcaataa tattgaaaaa ggaagagtat gagtattcaa catttccgtg	3060

tcgcccttat tccctttttt gcggcatttt gccttctgt ttttgctcac ccagaaacgc	3120
tggtgaaagt aaaagatgct gaagatcagt tgggtgcacg agtgggttac atcgaactgg	3180
atctcaacag cggtaaagatc cttgagagtt ttcgccccga agaactgttt ccaatgatga	3240
gcacttttaa agttctgcta tgtggcgcggt tattatcccg tattgacgcc gggcaagagc	3300
aactcggtcg ccgcatacac tattctcaga atgacttggg tgagtactca ccagtcacag	3360
aaaagcatct tacggatggc atgacagtaa gagaattatg cagtgtgcc ataaccatga	3420
gtgataacac tgcggccaac ttacttctga caacgatcgg aggaccgaag gagctaaccg	3480
cttttttgca caacatgggg gatcatgtaa ctgccttga tcgttgggaa ccggagctga	3540
atgaagccat accaaacgac gagcgtgaca ccacgatgcc tgtagcaatg gcaacaacgt	3600
tgcgcaaact attaaactggc gaactactta ctctagcttc ccggcaacaa ttaatagact	3660
ggatggaggc ggataaagt gtaggaccac ttctgcgctc ggcccttcg gctggctggg	3720
ttattgctga taaatctgga gccggtgagc gtgggtctcg cggatcatt gcagcactgg	3780
ggccagatgg taagccctcc cgtatcgtag ttatctacac gacggggagt caggcaacta	3840
tggatgaacg aaatagacag atcgtgaga taggtgcctc actgattaag cattggtaac	3900
gtcagacca agtttactca tataacttt agattgattt aaaacttcat ttttaattta	3960
aaaggatcta ggtgaagatc ctttttgata atctcatgac caaaatccct taacgtgagt	4020
tttcgttcca ctgagcgtca gaccccgtag aaaagatcaa aggatcttct tgagatcctt	4080
tttttctgcg cgtaatctgc tgccttgcaaa caaaaaaacc accgtacca gcggtggttt	4140
gtttgccgga tcaagagcta ccaactcttt ttccgaaggc aactggcttc agcagagcgc	4200
agataccaaa tactgttctt ctagtgtagc cgtagttagg ccaccacttc aagaactctg	4260
tagcaccgcc tacatacctc gctctgctaa tctgttacc agtggctgct gccagtggcg	4320
ataagtcgtg tcttaccggg ttggactcaa gacgatagtt accggataag gcgcagcggg	4380
cgggctgaac ggggggttcg tgcacacagc ccagcttga gcgaacgacc tacaccgaac	4440
tgagatacct acagcgtgag ctatgagaaa gcgccacgt tcccgaaggg agaaaggcgg	4500
acaggtatcc ggtaagcggc agggctcgga caggagagcg cacgaggag cttccagggg	4560
gaaacgcctg gtatctttat agtcctgtcg ggtttcgcca cctctgactt gagcgtcgat	4620
ttttgtgatg ctctcaggg gggcggagcc tatggaaaaa cgccagcaac gcggcctttt	4680
tacggttcct ggccttttgc tggccttttg ctacatgtt ctttctgcg ttatccctg	4740
attctgtgga taaccgtatt accgcctttg agtgagctga taccgtctgc cgcagccgaa	4800

cgaccgagcg cagcgagtica gtgagcgagg aagcggaaga gcgccaata cgcaaaccgc	4860
ctctccccgc gcgttggccg attcattaat gcagctggca cgacaggttt cccgactgga	4920
aagcgggcag tgagcgcaac gcaattaatg tgagtttagct cactcattag gcaccccagg	4980
ctttacactt taigtctccg gctcgtatgt tgtgtggaat tgtgagcgga taacaatttc	5040
acacaggaag cagctatgac catgattacg ccaagctttt tctttccaat tttttttttt	5100
tcgtcattat aaaaatcatt acgaccgaga ttcccgggta ataactgata taattaaatt	5160
gaagctctaa ttigtgagtt tagtatacat gcattttactt ataatacagt ttttttagttt	5220
tgctggccgc atctttctca atattgttcc cagcctgctt ttctgtaacg ttcaccctct	5280
accttagcat cctttccctt tgcaaatagt ccttttccaa caataataat gtcagatcct	5340
glagagacca catcatccac ggttctatac tgttgacca atgcgtctcc cttgtcatct	5400
aaaccacac cgggtgtcat aatcaacaa tcgtaacctt catctcttcc acctatgtct	5460
ctttgagcaa taaagccgat aacaaaatct ttgtcgtctc tcgcaatgtc aacagtaccc	5520
ttagtatatt ctccagtaga tagggagccc ttgcatgaca attctgctaa catcaaaagg	5580
cctctagggt cctttgttac ttcttctgcc gcctgcttca aaccgtaac aatacctggg	5640
cccaccacac cgtgtgcatt cgtaatgtct gccattctg ctattctgta tacaccgca	5700
gagtactgca atttgactgt attaccaatg tcagcaaatt ttctgtcttc gaagagtaaa	5760
aaattgtact tggcggataa tgcctttagc ggcttaactg tgcctccat ggaaaaatca	5820
gtcaagatat ccacatgtgt ttttagtaaa caaatTTTgg gacctaatgc ttcaactaac	5880
tccagtaatt ccttggttgt acgaacatcc aatgaagcac acaagtttgt ttgcttttcg	5940
tgcagtatat taaatagctt ggcagcaaca ggactaggat gagtagcagc acgttcctta	6000
tatgtagctt tcgacatgat ttatcttcgt ttcttgcagg tttttgttct gtgcagttgg	6060
gttaagaata ctgggcaatt tcatgtttct tcaacactac atatgcgtat atataccaat	6120
ctaagtctgt gctccttctt tegtcttcc ttctgttcgg agattaccga atcaaaaaaa	6180
tttcaaggaa accgaaatca aaaaaaagaa taaaaaaaaa atgatgaatt gaaaagcttg	6240
catgcctgca ggtcgactct agtatactcc gtctactgta cgatacactt ccgctcaggt	6300
ccttgcctt taacgaggcc ttaccactct ttgttactc tattgatcca gtcagcaaa	6360
ggcagtgta tctaagattc tatcttcgcg atgtagtaaa actagctaga ccgagaaaga	6420
gactagaaat gcaaaaggca cttctacaat ggctgccatc attattatcc gatgtgacgc	6480
tgcatTTTTT tttttttttt tttttttttt tttttttttt tttttttttt tttttgtac	6540
aaatatcata aaaaaagaga atctttttta gcaaggattt tcttaacttc ttcggcgaca	6600
gcatcaccga cttcggttgt actgttgga ccacctaaat caccagttct gatacctgca	6660

tccaaaacct ttttaactgc atcttcaatg gctttacctt cttcaggcaa gttcaatgac	6720
aatttcaaca tcattgcagc agacaagata gtggcgatag ggttgacctt attctttggc	6780
aaatctggag cggaaccatg gcatggttcg tacaaccaa atgcggtgtt cttgtctggc	6840
aaagaggcca aggacgcaga tggcaacaaa cccaaggagc ctgggataac ggaggcttca	6900
tcggagatga tatcaccaa catgttctg gtgattataa taccatttag gtgggttggg	6960
ttcttaacta ggatcatggc ggcagaatca atcaattgat gttgaacttt caatgtaggg	7020
aattcgttct tgatggtttc ctccacagt tttctccata atcttgaaga ggccaaaaca	7080
ttagctttat ccaaggacca aataggcaat ggtggctcat gttgtagggc catgaaagcg	7140
gccattcttg tgattctttg cacttctgga acggtgtatt gttcactatc ccaagcgaca	7200
ccatcaccat cgtcttctt tctcttacca aagtaaatc ctcccactaa tttcttaaca	7260
acaacgaagt cagtacctt agcaaatgt ggcttgattg gagataagtc taaaagagag	7320
tcggatgcaa agttacatgg tcttaagttg gcgtacaatt gaagttcttt acggattttt	7380
agtaaacctt gttcaggctt aacactaccg gtacccatt taggaccacc cacagcacct	7440
aacaaaacgg catcagcctt cttggaggct tccagcgct catctggaag tggaacacct	7500
glagcatcga tagcagcacc accaattaaa tgattttcga aatcgaactt gacattggaa	7560
cgaacatcag aaatagcttt aagaacctta atggcttcgg ctgtgatttc ttgaccaacg	7620
tggtcacctg gcaaacgac gatcttctta ggggcagaca ttacaatggt atatccttga	7680
aatatatata aaaaaaaaa aaaaaaaaaa aaaaaaaaaat gcagcttctc aatgatattc	7740
gaatacgctt tgaggagata cagcctaata tccgacaaac tgttttacag atttacgac	7800
gtacttgta cccatcattg aattttgaac atccgaacct gggagttttc cctgaaacag	7860
atagtatatt tgaacctgta taataatata tagtctagcg ctttacggaa gacaatgtat	7920
gtatttcggt tcttgagaa actattgcat ctattgcata ggtaatcttg cacgtcgcat	7980
ccccggttca tttctgcgt ttccatcttg cacttcaata gcatatcttt gttacgaag	8040
catctgtgct tcattttgta gaacaaaaat gcaacgcgag agcgctaatt tttcaacaa	8100
agaatctgag ctgcattttt acagaacaga aatgcaacgc gaaagcgcta tttaccaac	8160
gaagaatctg tgcttcattt ttgtaaaaca aaaatgcaac gcgagagcgc taatttttca	8220
aacaaagaat ctgagctgca tttttacaga acagaaatgc aacgcgagag cgctatttta	8280
ccaacaaaga atctatactt cttttttgtt ctacaaaaat gcatcccgag agcgctattt	8340
ttctaacaaa gcatcttaga ttactttttt tctctttgt gcgctctata atgcagtcctc	8400
ttgataactt tttgcactgt aggtccgtta aggttagaag aaggetactt tgggtgtctat	8460
tttctcttc ataaaaaag cctgactcca ctcccgctt ttactgatta ctagcgaagc	8520

tgcgggtgca ttttttcaag ataaaggcat ccccgattat attctatacc gatgtggatt 8580

 gcgcatactt tgtgaacaga aagtgatagc gttgatgatt cttcattggt cagaaaatta 8640
 tgaacggttt cttctatttt gtctctatat actacgtata ggaaatgttt acattttcgt 8700
 attgttttcg attcactcta tgaatagttc ttactacaat ttttttgtct aaagagtaat 8760
 actagagata aacataaaaa atgtagaggt cgagtttaga tgcaagttca aggagcgaaa 8820
 ggtggatggg taggttatat agggatatag cacagagata tatagcaaag agatactttt 8880
 gagcaatgtt tgtggaagcg gtattcgcaa tattttagta gctcgttaca gtccggtgcg 8940
 tttttggttt ttgaaagtg cgtcttcaga gcgcttttgg ttttcaaag cgctctgaag 9000

 ttccataact ttctagagaa taggaacttc ggaataggaa cttcaaagcg tttccgaaaa 9060
 cgagcgcttc cgaanaatgca acgcgagctg cgcacatata gctcactgtt cacgtcgcac 9120
 ctatatctgc gtgttgcttg tatatatata tacatgagaa gaacggcata gtgcgtgttt 9180
 atgcttaaat gcgtacttat atgcgtctat ttatgtagga tgaaaggtag tctagtacct 9240
 cctgtgatat tatcccattc catgcggggt atcgtatgct tccttcagca ctacccttta 9300
 gctgttctat atctgcccac tcctcaattg gattagtctc atccttcaat gctatcattt 9360
 cctttgatat tggatcatat gcatagtacc gagaaactag aggatctccc attaccgaca 9420

 tttggcgct atacgtgcat atgttcatgt atgtatctgt atttaaaaca cttttgtatt 9480
 atttttctc atatatgtgt ataggtttat acggatgatt taattattac ttcaccaccc 9540
 tttatttcag gctgatatct tagccttggt actagtcacc ggtggc 9586

 <210> 139
 <211> 15539
 <212> DNA
 <213> Artificial sequence
 <220><223> Plasmid
 <400> 139

 tcgcgcgttt cggtgatgac ggtgaaaacc tctgacacat gcagctcccg gagacggtca 60
 cagcttgtct gtaagcggt gcccggagca gacaagcccg tcagggcgcg tcagcgggtg 120
 ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180

 accataaatt cccgttttaa gagcttggig agcgctagga gtcactgcca ggtatcgttt 240
 gaacacggca ttagtcaggg aagtcataac acagtccttt cccgcaattt tctttttcta 300
 ttactcttgg cctcctctag tacactctat atttttttat gcctcggtaa tgattttcat 360
 tttttttttt ccacctagcg gatgactctt tttttttctt agcgattggc attatcacat 420

aatgaattat acattatata aagtaatgtg atttcttcga agaataact aaaaaatgag 480
caggcaagat aaacgaaggc aaagatgaca gagcagaaag ccctagtaa gcgtattaca 540
aatgaaacca agattcagat tgcgatctct ttaaagggtg gtcccctagc gatagagcac 600

tcgatcttcc cagaaaaaga ggcagaagca gtagcagaac aggccacaca atcgcaagtg 660
attaacgtcc acacaggat agggtttctg gaccatatga tacatgctct ggccaagcat 720
tccggctggt cgtaatcgt tgagtgcatt ggtgacttac acatagacga ccatcacacc 780
actgaagact gcgggattgc tctcggtaa gcttttaaag aggccctagg ggccgtgcgt 840
ggagtaaaaa ggtttggatc aggttttgcg cttttggatg aggcacttcc cagagcgggtg 900
gtagatcttt cgaacaggcc gtacgcagtt gtcgaacttg gtttgcaaag ggagaaagta 960
ggagatctct ctgacgagat gatcccgcat tttcttgaaa gctttgcaga ggctagcaga 1020

attaccctcc acgttgattg tctgcgaggc aagaatgac atcacgtag tgagagtgcg 1080
ttcaaggctc ttgcggttgc cataagagaa gccacctcgc ccaatggtac caacgatgtt 1140
ccctccacca aagggtgtct tatgtagtga caccgattat ttaaagctgc agcatacgtat 1200
atatatacat gtgtatatat gtatacctat gaatgtcagt aagtatgtat acgaacagta 1260
tgatactgaa gatgacaagg taatgcatca ttctatacgt gtcattctga acgaggcgcg 1320
ctttcctttt ttctttttgc tttttctttt tttttctctt gaactcgacg gatctatgcg 1380
gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcaggaaat tgtaagcgtt 1440

aatattttgt taaaattcgc gttaaatttt tgtaaataca gtcattttt taaccaatag 1500
gccgaaatcg gcaaaatccc ttataaatca aaagaataga ccgagatagg gttgagtgtt 1560
gttccagttt ggaacaagag tccactattt aagaacgtgg actccaacgt caaaggcgca 1620
aaaaccgtct atcaggcgca tggcccaacta cgtgaaccat caccctaac aagttttttg 1680
gggtcgaggt gccgtaaagc actaaatcgg aaccctaaag ggagccccg atttagagct 1740
tgacggggaa agccggcgaa cgtggcgaga aaggaaggga agaaagcgaa aggagcgggc 1800
gctagggcgc tggcaagtgt agcggtcacg ctgcgcgtaa ccaccacacc cgccgcgtt 1860

aatgcgccgc tacaggcgcc gtccattcgc cattcaggct gcgcaactgt tgggaagggc 1920
gcggtgcggg cctcttcgct attacgccag ctggcgaaag ggggatgtgc tgcaaggcga 1980
ttaagtggg taacgccagg gttttcccag tcacgacgtt gtaaaacgac ggccagttag 2040
cgcgcgtaat acgactcact ataggcgaa ttgggtaccg ggccccctt cgaggtcgac 2100
ggcgcgccac tggtagagag cgactttgta tgcccaatt gcgaaaccg cgatatcctt 2160
ctcgattctt tagtaccga ccaggacaag gaaaaggagg tcgaaacgtt tttgaagaaa 2220

caagaggaac tacacggaag ctctaaagat ggcaaccagc cagaaactaa gaaaatgaag	2280
ttgatggatc caactggcac cgctggcttg aacaacaata ccagccttcc aacttctgta	2340
aataacggcg gtacgccagt gccaccagta cgtttacctt tcggtatacc tcctttcccc	2400
atgtttccaa tgccttcat gcctccaacg gctactatca caaatcctca tcaagctgac	2460
gcaagcccta agaaatgaat aacaatactg acagtactaa ataattgcct acttggcttc	2520
acatacgttg catacgtcga tatagataat aatgataatg acagcaggat tatcgttaata	2580
cgtaatagct gaaaatctca aaaatgtgtg ggtcattacg taaataatga taggaatggg	2640
attcttctat ttttctttt tccattctag cagccgtcgg gaaaacgtgg catectctct	2700
ttcgggctca atiggagtca cgctgccgtg agcatcctct ctttccatat ctaacaactg	2760
agcacgtaac caatgaaaa gcatgagctt agcgttgctc caaaaaagta ttggatggtt	2820
aataccattt gtctgttctc ttctgacttt gactcctcaa aaaaaaaaaa ctacaatcaa	2880
cagatcgctt caattacgcc ctacaaaaa ctttttctt tcttcttcgc ccacgttaaa	2940
ttttatccct catgttgtct aacggatttc tgcacttgat ttattataaa aagacaaaga	3000
cataatactt ctctatcaat ttcagttatt gttcttctt gcgttattct tctgttcttc	3060
ttttctttt gtcatatata accataacca agtaatacat attcaacta gtatgactga	3120
caaaaaact cttaaagact taagaaatcg tagttctgtt tacgattcaa tggttaaatc	3180
acctaactgt gctatgttgc gtgcaactgg tatgcaagat gaagactttg aaaaacctat	3240
cgtcggtgtc atttcaactt gggctgaaaa cacacctgt aatatccact tacatgactt	3300
tggtaacta gccaaagtcg gtgttaagga agctgggtgt tggccagttc agttcggaac	3360
aatcacggtt tctgatggaa tcgccatggg aaccaagga atgcgtttct ccttgacatc	3420
tcgtgatatt attgcagatt ctattgaagc agccatggga ggtcataatg cggatgcttt	3480
tgtagccatt ggcggttgtg ataaaaacat gcccggttct gttatcgcta tggctaacat	3540
ggatatccca gccatttttg cttacggcgg aacaattgca cctggtaatt tagacggcaa	3600
agatatcgat ttagtctctg tctttgaagg tgcggccat tggaaaccag gcgatatgac	3660
caaagaagaa gttaaagctt tggaatgtaa tgcttgtccc ggtcctggag gctgcggtgg	3720
tatgtatact gtaacacaa tggcgacagc tattgaagtt ttgggactta gccttccggg	3780
ttcatcttct caccggctg aatccgcaga aaagaaagca gatattgaag aagctggtcg	3840
cgctgttgtc aaaatgctcg aaatgggctt aaaaccttct gacattttaa cgcgtgaagc	3900
ttttgaagat gctattactg taactatggc tctgggaggt tcaaccaact caaccttca	3960

cctcttagct attgccatg ctgctaagt ggaattgaca cttgatgatt tcaatacttt 4020
ccaagaaaa gticctcatt tggctgattt gaaaccttct ggtcaatatg tattccaaga 4080
cctttacaag gtcggagggg taccagcagt tatgaaatat ctcttaaaa atggcttctt 4140
tcatggtgac cgtatcactt gtactggcaa aacagtcgct gaaaatttga aggcctttga 4200
tgatttaaca cctgggtcaaa aggttattat gccgcttgaa aatcctaaac gtgaagatgg 4260
tccgctcatt attctccatg gtaacttggc tccagacggg gccgttgcca aagtttctgg 4320
tgtaaaagtg cgtcgtcatg tcggctctgc taaggctctt aattctgaag aagaagccat 4380

tgaagctgtc ttgaatgatg atattgttga tggatgatgt gttgtcgtac gtttttagg 4440
accaaagggc ggtcctggta tgcctgaaat gctttccctt tcatcaatga ttgttggtaa 4500
agggaagggt gaaaaagttg cccttctgac agatggccgc ttctcaggtg gtacttatgg 4560
tcttgtcgtg ggtcatatcg ctctgaagc acaagatggc ggtccaatcg cctacctgca 4620
aacaggagac atagtacta ttgaccaaga cactaaggaa ttacactttg atatctccga 4680
tgaagagtta aaacatcgtc aagagacat tgaattgcca ccgctctatt cacgcggtat 4740
ccttggtaaa tatgctcaca tcgtttcgtc tgccttctagg ggagccgtaa cagacttttg 4800

gaagcctgaa gaaactggca aaaaatgttg tcctgggtgc tgtggttaag cggccgcgtt 4860
aattcaaatt aattgatata gttttttaat gattattgaa tctgtttaga aataatggaa 4920
tattattttt atttatttat ttatattatt ggtcggctct tttcttctga aggtcaatga 4980
caaatgata tgaaggaaat aatgatttct aaaattttac aacgtaagat atttttacaa 5040
aagcctagct catcttttgt catgcactat ttactcacg cttgaaatta acggccagtc 5100
cactgcggag tcatttcaaa gtcaccta taatctatc gtttttgata gctcattttg 5160
gagttcgcga ttgtcttctg ttattcacia ctgttttaat ttttatttca ttctggaact 5220

cttcgagttc ttgttaaagt ctttcatagt agcttacttt atctccaac atatttaact 5280
tcatgtcaat ttcggctctt aaattttcca catcatcaag ttcaacatca tcttttaact 5340
tgaatttatt ctctagctct tccaaccaag cctcattgct ccttgattta ctggtgaaaa 5400
gtgatacact ttgcgcgcaa tccaggtaaa aactttcctg caaagaattc accaatttct 5460
cgacatcata gtacaatttg tttgtttctc ccatcacaat ttaataatcc tgatggattc 5520
ttatgaagcg ctgggtaatg gacgtgtcac tctacttcgc ctttttcctt actcctttta 5580
gtacggaaga caatgcta ataaagagg gtaataataa tattattaat cggcaaaaaa 5640

gattaaacgc caagcgttta attatcagaa agcaaagtc gtaccaatcc ttgaatgctt 5700
cccaatttga tattaagagt catcacagca acatattctt gttattaaat taattattat 5760
tgatttttga tattgtataa aaaaacaaa tatgtataaa aaaagtgaat aaaaaatacc 5820

aagtatggag aaatatatta gaagtcctata cgtaaacca cccgggcccc ccctcgaggt 5880
cgacggtatc gataagcttg atatcgaatt cctgcagccc gggggatcca ctagtcttag 5940
agcggccgct ctagaactag taccacaggt gttgtcctct gaggacataa aatacacacc 6000
gagattcatc aactcattgc tggagttagc atatctacaa ttgggtgaaa tggggagcga 6060

tttgcaggca tttgctcggc atgccggtag aggtgtggtc aataagagcg acctcatgct 6120
atactgaga aagcaacctg acctacagga aagagttact caagaataag aattttcgtt 6180
ttaaaccta agagtcactt taaaatttgi atacacttat tttttttata acttatttaa 6240
taataaaaat cataaatcat aagaaattcg ctactctta attaatcaaa aagttaaaat 6300
tgtacgaata gattcaccac ttcttaacaa atcaaaccct tcattgattt tctcgaatgg 6360
caatacatgt gtaattaaag gatcaagagc aaacttcttc gccataaagt cggcaacaag 6420
ttttggaaca ctatccttgc tcttaaaacc gccaaatata gctcccttcc atgtacgacc 6480

gcttagcaac agcataggat tcatcgacaa attttgtgaa tcaggaggaa cacctacgat 6540
cacactgact ccatatgcct cttagacagca ggacaacgca gttacatag tatcaagacg 6600
gcctataact tcaaaagaga aatcaactcc accgtttgac atttcagtaa ggacttcttg 6660
tattggtttc ttataatctt gagggttaac acattcagta gccccgacct ccttagcttt 6720
tgcaaatitg tcttatttga tgtctacacc tataatcttc gctgcgctg cagctttaca 6780
ccccataata acgcttagtc ctactcctcc taaaccgaat actgcacaag tcgaaccctg 6840
tgtaaccttt gcaactttta ctgcggaacc gtaaccggtg gaaaatccgc accctatcaa 6900

gcaaactttt tccagtgggtg aagctgcacg gattttagcg acagatatct cgtccaccac 6960
tgtgtatttg gaaaatgtag aagtaaccaag gaaatggtgt ataggtttcc ctctgcatgt 7020
aaatctgctt gtaccatcct gcatagtacc tctaggcata gacaaatcat ttttaaggca 7080
gaaattaccc tcaggatgtt tgcagactct acacttacca cattgaggag tgaacagtgg 7140
gatcacttta tcaccaggac gaacagtggg aacaccttca cctatggatt caacgattcc 7200
ggcagcctcg tgtcccgcga ttactggcaa aggagtaact agagtccac tcaccacatg 7260
gtcgtcggat ctacagattc cggtaggcaac catcttgatt ctaacctcgt gtgcttttgg 7320

tggcgtact tctacttctt ctatgctaaa cggttttttc tcttcccaca aaactgccgc 7380
tttacactta ataactttac cggtgttga catctcagc tagctattgt aatatgtgtg 7440
tttgtttga ttattaagaa gaataattac aaaaaaatt acaaggaag gtaattacaa 7500
cagaattaag aaaggacaag aaggaggaag agaatacgtt cattatttct tctttgttat 7560
ataacaaacc caagtagcga tttagccata cattaaaagt tgagaaccac ctccttggc 7620
aacagccaca actcgttacc attgttcac acgatcatga aactcgtgt cagctgaaat 7680

ttcacctcag tggatctctc tttttattct tcatcgttcc actaaccttt tccatcagc	7740
tggcagggaa cggaagtgg aatcccatth agcgagcttc ctcttttctt caagaaaaga	7800
cgaagcttgt gtgtgggtgc gcgcgctagt atctttccac attaagaaat ataccataaa	7860
ggttacttag acatcactat ggctatata atatatatat atatatgtaa cttagcacca	7920
tcgcgcgtgc atcactgcat gtgttaaccg aaaagtttgg cgaacacttc accgacacgg	7980
tcatttagat ctgtcgtctg cattgcacgt cccttagcct taaatcctag gcgggagcat	8040
tctcgtgtaa ttgtgcagcc tgcgtagcaa ctcaacatag cgtagtctac ccagtttttc	8100
aagggtttat cgttagaaga ttctcccttt tcttctgct cacaatctt aaagtcatac	8160
attgcacgac taaatgcaag catgcggatc ccccggtctg caggaattcg atatcaagct	8220
tatcgatacc gtcgactggc cattaatctt tcccatatta gatttcgcca agccatgaaa	8280
gttcaagaaa ggtcttttaga cgaattacc ttcatcttc aaactggcgt caagggatcc	8340
tggtaggtt ttatcgtttt atttctggtt cttatagcat cgttttgac ttctctgttc	8400
ccattaggcg gttcaggagc cagcgcagaa tcattctttg aaggatactt atcctttcca	8460
atittgattg tctgttacgt tggacataaa ctgtatacta gaaattggac ttgatgggtg	8520
aaactagaag atatggatct tgataccggc agaaaacaag tagatttgac tcttcgtagg	8580
gaagaaatga ggattgagcg agaaacatta gcaaaaagat ccttcgtaac aagattttta	8640
catttctggt gtigaaggga aagatatgag ctatacagcg gaatttccat atcactcaga	8700
ttttgttate taattttttc ctccccacgt ccgcgggaat ctgtgtatat tactgcatct	8760
agatatatgt tatcttatct tggcgcgtac atttaatttt caacgtattc tataagaaat	8820
tgcgggagtt tttttcatgt agatgatact gactgcacgc aaatatagc atgatttata	8880
ggcatgattt gatggctgta ccgataggaa cgctaagagt aacttcagaa tcgttatcct	8940
ggcggaaaaa attcatttgt aaactttaaa aaaaaagcc aatatccca aaattattaa	9000
gagcgctcc attattaact aaaatttcac tcagcatcca caatgtatca ggtatctact	9060
acagatatta catgtggcga aaaagacaag aacaatgcaa tagcgcatca agaaaaaaca	9120
caaagctttc aatcaatgaa tcgaaaatgt cattaaaata gtatataaat tgaaactaag	9180
tcataaagct ataaaaagaa aatttattta aatgcaagat ttaaagtaaa ttcacggccc	9240
tgcaggcctc agctcttgtt ttgtttcgca aataacttac ccattttttt caaaacttta	9300
ggtgcacctt cctttgctag aataagttct atccaatata tcctatttgg atctgcttga	9360
gcttctttca tcacggatac gaattcattt tctgttttca caattttgga cacaactctg	9420

tcttccgttg ccccgaaact ttctggcagt tttagtaat tccacatagg aatgtcatta	9480
taactctggt tcggaccatg aatttccctc tcaaccgtgt aaccatcggt attaatgata	9540
aagcagattg ggtttatctt ctctctaagt gctagtccta attcttggac agtcagttgc	9600
aatgatccat ctccgataaa caataaatgt ctagattctt tatctgcaat ttggctgcct	9660
agagctgcgg ggaaagtgtg tcctatagat cccacaagg gttgaccaat aaaatgtgat	9720
ttcgatttca gaaatataga tgaggcaccg aagaaagaag tgccttggtc agccacgac	9780
gtctcattac ttgggtcaa attttcgaca gcttgccaca gtctatcttg tgacaacagc	9840
gcgttagaag gtacaaaatc ttcttgcttt ttatctatgt acttgctttt atattcaatt	9900
tcggacaagt caagaagaga tgatacagg gattcgaagt cgaaattttg gattctttcg	9960
ttgaaaattt taccttcacg gatattcaag gaaatcattt tattttcatt aagatgggtga	10020
gtaaatgcac ccgtactaga atcggttaagc ttacaccca acataagaat aaaatcagca	10080
gattccacaa attccttcaa gtttggtctt gacagagtac cgttgtaaat ccccaaaaat	10140
gagggcaatg ctcatcaac agatgattta ccaaagtta aagtagtaat aggttaactta	10200
gtctttgaaa taaactgagt aacagtcttc tctaggccga acgatataat ttcattggcct	10260
gtgattacaa ttggtttctt ggcatctctc agactttcct gtattttgtt cagaatctct	10320
tgatcagatg tattcgacgt ggaattttcc ttcttaagag gcaaggatgg tttttcagcc	10380
ttagcggcag ctacatctac aggtaaattg atgtaaaccg gctttctttc ctttagtaag	10440
gcagacaaca ctctatcaat ttcaacagt gcatctctcg ctgtcaataa agtcctggca	10500
gcagtaaccg gttcgtgcat cttcataaag tgcctgaaat caccatcagc caacgtatgg	10560
tgaacaaact taccttcgtt ctgcactttc gaggtaggag atcccacgat ctcaacaaca	10620
ggcaggttct cagcatagga gcccgttaag ccattaactg cggataattc gccaacacca	10680
aatgtagtca agaatgccgc agcctttttc gttcttgctg acccgtcggc catataggag	10740
gcatttaact cattagcatt tccacccat ttcatatctt tgtgtgaaat aatttgatct	10800
agaaattgca aattgtagtc acctggtact ccgaatatct cttctatacc taattcgtgt	10860
aatctgtcca acagatagtc acctactgta tacattttgt ttactagttt atgtgtgttt	10920
attcgaaact aagtcttgg tgttttaaaa ctaaaaaaaa gactaactat aaaagtagaa	10980
tttaagaagt ttaagaaata gatttacaga attacaatca atacctaccg tctttatata	11040
cttattagtc aagtagggga ataatttcag ggaactggtt tcaacctttt ttttcagctt	11100
tttccaaatc agagagagca gaagtaata gaaggtgtaa gaaaatgaga tagatacatg	11160
cgtgggtcaa ttgccttggt tcattcattta ctccaggcag gttgcatcac tccattgagg	11220
ttgtgcccgt tttttgcctg ttgtgcccc tgttctctgt agttgcgcta agagaatgga	11280

cctatgaact gatggttggg gaagaaaaca atatttttggg gctgggattc tttttttttc	11340
tggatgccag cttaaaaaagc gggctccatt atatttagtg gatgccagga ataaactgtt	11400
caccagaca cctacgatgt tataatattct gtgtaaccgc cccctatatt tgggcatgta	11460
cgggttacag cagaattaaa aggctaattt ttgactaaa taaagttagg aaaatcacta	11520
ctattaatta tttacgtatt ctttgaaatg gcagtattga taatgataaa ctggaactga	11580
aaaagcgtgt tttttattca aaatgattct aactccctta cgtaatcaag gaatcttttt	11640
gccttggcct ccgcgtcatt aaacttcttg ttgttgacgc taacattcaa cgctagtata	11700
tattcgtttt tttcaggtaa gttcttttca acgggtctta ctgatgaggc agtcgcgtct	11760
gaacctgtta agaggtcaaa tatgtcttct tgaccgtacg tgtcttgcat gttattagct	11820
ttgggaattt gcatcaagtc ataggaaaat ttaaactctg gctctcttgg gctcaagggtg	11880
acaaggtcct cgaaaatagg gcgcgccccca ccgcggtgga gctccagctt ttgttccctt	11940
tagtgagggt taattgcgcg cttggcgtaa tcatggtcac agctgtttcc tgtgtgaaat	12000
tgttatccgc tcacaattcc acacaacata cgagccggaa gcataaagt taaagcctgg	12060
ggtgcctaata gagtgagcta actcacatta attgcgttgc gctcactgcc cgctttccag	12120
tcgggaaacc tgcgtgccca gctgcattaa tgaatcggcc aacgcgcggg gagaggcgggt	12180
ttgcgtattg ggcgtcttc cgcttctctg ctactgact cgctgcgtc ggtcgttcgg	12240
ctgcgcgag cggtatcagc tcaactcaaag gcggtatataa gggtatccac agaatacagg	12300
gataacgcag gaaagaacat gtgagcaaaa ggccagcaaa aggccaggaa ccgtaaaaag	12360
gccgcgttgc tggcgttttt ccataggctc cgccccctg acgagcatca caaaaatcga	12420
cgctcaagtc agagggtggc aaaccgcaca ggactataaa gataccaggc gtttccccct	12480
ggaagctccc tcgtgcgtc tctgttccg accctgccgc ttaccggata cctgtccgc	12540
tttctccctt cgggaagcgt ggcgtttct catagctcac gctgtaggta tctcagttcg	12600
gtgtaggtcg ttcgtccaa gctgggctgt gtgcacgaac ccccggttca gcccaccgc	12660
tgcgccttat ccgtaacta tcgtcttgag tccaaccgg taagacacga cttatcgcca	12720
ctggcagcag ccaactggtaa caggattagc agagcgaggt atgtaggcgg tgctacagag	12780
ttcttgaagt ggtggcctaa ctacggctac actagaagaa cagtatttgg tatctgcgt	12840
ctgctgaagc cagttacctt cggaataaga gttggtagct cttgatccgg caaacaacc	12900
accgctggta gcggtggttt tttgtttgc aagcagcaga ttacgcgcag aaaaaagga	12960
tctcaagaag atcctttgat cttttctacg gggctctgacg ctgagtgga cgaaaactca	13020
cgtaagga ttttggcat gagattatca aaaaggatct tcacctagat ctttttaaat	13080
taaaaatgaa gttttaaatc aatctaaagt atatatgagt aaacttggtc tgacagttac	13140

caatgcttaa tcagtgaggc acctatctca gcgatctgtc tatttcgttc atccatagtt	13200
gcctgactcc ccgtcgtgta gataactacg atacgggagg gcttaccatc tggccccagt	13260
gttgcaatga taccgcgaga cccacgctca ccggctccag atttatcagc aataaaccag	13320
ccagccggaa gggccgagcg cagaagtggc cctgcaactt tatccgcctc catccagtct	13380
attaattgtt gccgggaagc tagagtaagt agttcgccag ttaatagttt gcgcaacgtt	13440
gttgccattg ctacaggcat cgtggtgtca cgctcgtcgt ttggtatggc ttcattcagc	13500
tccggttccc aacgatcaag gcgagttaca tgatcccca tgttgtgcaa aaaagcgggt	13560
agctccttcg gtctccgat cgttgtcaga agtaagtgg ccgcagtgtt atcactcatg	13620
gttatggcag cactgcataa ttctcttact gtcatgccat ccgtaagatg cttttctgtg	13680
actggtgagt actcaaccaa gtcattctga gaatagtga tgcggcgacc gagttgctct	13740
tgccccggt caatacggga taataccgcg ccacatagca gaactttaaa agtgcctc	13800
attggaaaac gtctctcggg gcgaaaactc tcaaggatct taccgctgtt gagatccagt	13860
tcgatgtaac ccaactgtgc acccaactga tcttcagcat cttttacttt caccagcgtt	13920
tctgggtgag caaaaacagg aaggcaaaat gccgcaaaaa agggaataag ggcgacacgg	13980
aaatgttgaa tactcatact ctctcttttt caatattatt gaagcattta tcagggttat	14040
tgtctcatga gcggatacat atttgaatgt atttagaaaa ataaacaaat aggggttccg	14100
cgcacatttc ccgaaaagt gccacctgaa cgaagcatct gtgcttcatt ttgtagaaca	14160
aaaatgcaac gcgagagcgc taatttttca aacaaagaat ctgagctgca tttttacaga	14220
acagaaatgc aacgcgaaag cgctatttta ccaacgaaga atctgtgctt catTTTTgtA	14280
aaacaaaaat gcaacgcgag agcgctaatt ttccaacaa agaactgtag ctgcattttt	14340
acagaacaga aatgcaacgc gagagcgcta tttaccaac aaagaatcta tacttctttt	14400
ttgttctaca aaaatgcac cgcgagcgc tatTTTTcta acaaagcatc ttagattact	14460
ttttttctcc ttgtgcgct ctataatgca gtctcttgat aactttttgc actgtaggtc	14520
cgtaaaggtt agaagaagge tactttggtg tctattttct cttccataaa aaaagcctga	14580
ctccacttcc cgcgtttact gattactagc gaagctgcgg gtgcattttt tcaagataaa	14640
ggcatccccg atttatctct ataccgatgt ggattgcgca tactttgtga acagaaagtg	14700
atagcgttga tgattcttca ttgggtcagaa aattatgaac ggtttcttct attttgtctc	14760
tataactac gtataggaaa tgtttacatt ttcgtattgt tttcgattca ctctatgaat	14820
agttcttact acaatTTTT tgtctaaaga gtaatactag agataaacat aaaaaatgta	14880

gaggtcgagt ttagatgcaa gttcaaggag cgaaaggtagg atgggtaggt tatatagga 14940
tatagcacag agatataatag caaagagata cttttgagca atgtttgtgg aagcggtatt 15000
cgcaatattt tagtagctcg ttacagtccg gtgcgttttt ggTTTTTTga aagtgcgtct 15060
tcagagcgct ttTggTTTTc aaaagcgctc tgaagttcct atactttcta gagaatagga 15120
acttcggaat aggaacttca aagcgtttcc gaaaacgagc gcttccgaaa atgcaacgcg 15180
agctgcgcac atacagctca ctgttcacgt cgcacctata tctgcgtgtt gcctgtatat 15240
atatatacat gagaagaacg gcatagtgcg tgtttatgct taaatgcgta cttatatgcg 15300

tctatttatg taggatgaaa ggtagtctag tacctcctgt gatattatcc cattccatgc 15360
ggggTatcgt atgtttcctt cagcactacc ctttagctgt tctatatgct gccactcctc 15420
aattggatta gtctcatcct tcaatgctat catttccttt gatattggat catactaaga 15480
aaccattatt atcatgacat taacctataa aaataggcgt atcacgaggc ctttcgtc 15539

<210> 140
<211> 1053
<212> DNA
<213> Bifidobacterium angulatum
<400> 140

atggccgcta ctatctggta cgaaaaggat gctgatttgt ccgtcttcga aggtagaag 60
gtcgccgtta tcggttacgg ttcccaaggt cacgctcatg ccttaaactt gagagactct 120

ggtgttgatg ttgtcgttgg ttTgcgtcca acctctaagt ccgtcgaata cgctaaagaa 180
caaggtttgg aagttcagtc tgtcgtgac gccactgctg aagccgatgt cgttatgac 240
ttgctgccag accaatacca ggctgccgtt tacaagtcg aaattgaacc aaacttgaaa 300
ccaggtgctg cttTggcttt cgcccacggt ttcaacatcc attacggtta cattaagcca 360
tctgaagatc acccagtttt catggctcgt ccaaagggtc caggtcacat cgtcagacgt 420
gaatacacca acggtagagg tgttccagtc gttgtcgtg tcgaacaaga cccagatggt 480
aaaacttggc cactgtgttt agcttacgcc aaggctttgg gtgctttgag agctggtgct 540

atcaagacca ctttcaccga agaaaccgaa accgacttgt ttggtgaaca agatgtttta 600
atgggtggta tcaaccactt gtgcgacatg ggtttcgatg tcttgactga agccggttac 660
caaccagaaa ttgcttactt cgaagttttt cacgaattga agatgttggT tgacttggcc 720
aacgaaggTg gtttgaacaa ggctagatgg tcttgttccg acaccgtca atacggtgac 780
tacacctcta ctgttatcac cgaagaaacc aagaagagaa tgcaatacca gttgaagaga 840
atccaagatg gttccttcgc taaagaattc atggacgacT aagctgccgg tgctccaaag 900

ttcaagaagt tgcaggaaga atactctcac ccacacttgg aaaccgttgg tccaaagttg 960

agagctatgt tctcctggaa caatcaagtc gacgccgacg ccgacatggc tgaatccttc 1020

aacggtaaga tcgctagaac ccaagtccaa taa 1053

<210> 141

<211> 350

<212> PRT

<213> Bifidobacterium angulatum

<400> 141

Met Ala Ala Thr Ile Trp Tyr Glu Lys Asp Ala Asp Leu Ser Val Phe

1 5 10 15

Glu Gly Lys Lys Val Ala Val Ile Gly Tyr Gly Ser Gln Gly His Ala

20 25 30

His Ala Leu Asn Leu Arg Asp Ser Gly Val Asp Val Val Val Gly Leu

35 40 45

Arg Pro Thr Ser Lys Ser Val Glu Tyr Ala Lys Glu Gln Gly Leu Glu

50 55 60

Val Gln Ser Val Ala Asp Ala Thr Ala Glu Ala Asp Val Val Met Ile

65 70 75 80

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

85 90 95

Pro Asn Leu Lys Pro Gly Ala Ala Leu Ala Phe Ala His Gly Phe Asn

100 105 110

Ile His Tyr Gly Tyr Ile Lys Pro Ser Glu Asp His Pro Val Phe Met

115 120 125

Val Ala Pro Lys Gly Pro Gly His Ile Val Arg Arg Glu Tyr Thr Asn

130 135 140

Gly Arg Gly Val Pro Val Val Val Ala Val Glu Gln Asp Pro Asp Gly

145 150 155 160

Lys Thr Trp Pro Leu Cys Leu Ala Tyr Ala Lys Ala Leu Gly Ala Leu

165 170 175

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

180 185 190
 Leu Phe Gly Glu Gln Asp Val Leu Met Gly Gly Ile Asn His Leu Cys
 195 200 205
 Asp Met Gly Phe Asp Val Leu Thr Glu Ala Gly Tyr Gln Pro Glu Ile
 210 215 220
 Ala Tyr Phe Glu Val Phe His Glu Leu Lys Met Leu Val Asp Leu Ala

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

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

<210> 142

<211> 1050

<212> DNA

<213> Bifidobacterium dentium

<400> 142

atggccgctc aaatctggta cgaaaacgat ggtgacttgt ccgtcttgga tgtaagaag	60
gtcgctatca ttggttacgg ttctcaaggt cagcccatg ctttgaactt gagagactct	120
ggtgttgatg ttgtcgttgg tctgcgtcca acctctaagt ccgtcgaaca agctaaagaa	180
cagggtttgg aagtcaagtc cgtccagaa gctgccgctg aagctgacat cattatgatc	240
ttggccccag atcaatacca gagaactatt tgggctaacg acatcgaacc aaatattaag	300
ccaggtgccg ctatgcctt cgctcacggt ttcaacatcc attacggtta cattaagcca	360

tctgaagatc acccagtttt catggtcgct ccaaaaggtc caggtcacat tgtagacgt 420

gaatacgcta acggttagagg tgttccagtc gttgtcgctg tcgaacaaga cccaagaggt 480

gacgcttggg atatcacctt ggcctacgct aaggctttgg gtgctttgag agctggtgcc 540

attaagacca ccttcaagga agaaaccgaa accgacttgt ttggtgaaca aaacgtttta 600

atgggtgggtg ttaacaagtt ggtcgaaatg ggtttcgaag tcttgaccga cgctggttac 660

caaccagaaa tcgcctactt cgaagtttgt cacgaattga agatgttagt cgatttgatg 720

aacgaagggtg gtttgaacaa ggctagatgg tcttgttccg acactgccca atacggtgac 780

taccaaca cgttatcaa cgaagactgc agaaagcgta tgcaatacca cttgggtaga 840

atccaagacg gtcttttcgc taaggaattc atcgacgacg aagatgctgg tgccccaag 900

ttcaaggaat tgcaggaaga atacggtaac gttagaatcg aaactgtcgg tccaaagttg 960

cgtgctatgt tctcctggaa caatggtaaa gacgacgacg ccgacatggc taccttcact 1020

ggtaagatcg ctagatccca agtccaataa 1050

<210> 143

<211> 349

<212> PRT

<213> Bifidobacterium dentium

<400> 143

Met Ala Ala Gln Ile Trp Tyr Glu Asn Asp Gly Asp Leu Ser Val Leu

1 5 10 15

Asp Gly Lys Lys Val Ala Ile Ile Gly Tyr Gly Ser Gln Gly His Ala

20 25 30

His Ala Leu Asn Leu Arg Asp Ser Gly Val Asp Val Val Val Gly Leu

35 40 45

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

50 55 60

Val Lys Ser Val Pro Glu Ala Ala Ala Glu Ala Asp Ile Ile Met Ile

65 70 75 80

Leu Ala Pro Asp Gln Tyr Gln Arg Thr Ile Trp Ala Asn Asp Ile Glu

85 90 95

Pro Asn Ile Lys Pro Gly Ala Ala Ile Ala Phe Ala His Gly Phe Asn

100 105 110

Ile His Tyr Gly Tyr Ile Lys Pro Ser Glu Asp His Pro Val Phe Met
115 120 125

Val Ala Pro Lys Gly Pro Gly His Ile Val Arg Arg Glu Tyr Ala Asn
130 135 140

Gly Arg Gly Val Pro Val Val Val Ala Val Glu Gln Asp Pro Arg Gly
145 150 155 160

Asp Ala Trp Asp Ile Thr Leu Ala Tyr Ala Lys Ala Leu Gly Ala Leu
165 170 175

Arg Ala Gly Ala Ile Lys Thr Thr Phe Lys Glu Glu Thr Glu Thr Asp
180 185 190

Leu Phe Gly Glu Gln Asn Val Leu Met Gly Gly Val Asn Lys Leu Val
195 200 205

Glu Met Gly Phe Glu Val Leu Thr Asp Ala Gly Tyr Gln Pro Glu Ile
210 215 220

Ala Tyr Phe Glu Val Cys His Glu Leu Lys Met Leu Val Asp Leu Met
225 230 235 240

Asn Glu Gly Gly Leu Asn Lys Ala Arg Trp Ser Cys Ser Asp Thr Ala
245 250 255

Gln Tyr Gly Asp Tyr Thr Asn Thr Val Ile Asn Glu Asp Cys Arg Lys
260 265 270

Arg Met Gln Tyr His Leu Gly Arg Ile Gln Asp Gly Ser Phe Ala Lys
275 280 285

Glu Phe Ile Asp Asp Gln Asp Ala Gly Ala Pro Lys Phe Lys Glu Leu
290 295 300

Gln Glu Glu Tyr Gly Asn Val Arg Ile Glu Thr Val Gly Pro Lys Leu
305 310 315 320

Arg Ala Met Phe Ser Trp Asn Asn Gly Gln Asp Asp Asp Ala Asp Met
325 330 335

Ala Thr Phe Thr Gly Lys Ile Ala Arg Ser Gln Val Gln
340 345

<210> 144

<211> 26
 <212> DNA
 <213> Artificial sequence
 <220><223> Primer K9_56_58_060210f
 <220><221> misc_feature
 <222> (7)..(8)
 <223> n is a, c, g, or t
 <220><221> misc_feature
 <222> (13)..(14)
 <223> n is a, c, g, or t
 <400> 144

gaaggannka aannktggaa gagagc

26

<210> 145
 <211> 26
 <212> DNA
 <213> Artificial sequence

<220><223> Primer K9_56_58_060210r
 <220><221> misc_feature
 <222> (13)..(14)
 <223> n is a, c, g, or t
 <220><221> misc_feature
 <222> (19)..(20)
 <223> n is a, c, g, or t
 <400> 145

gctctcttcc amnnntttmnn tccttc

26

<210> 146
 <211> 37
 <212> DNA
 <213> Artificial sequence
 <220><223> K9_30_101110f primer
 <220><221> misc_feature
 <222> (20)..(21)
 <223> n is a, c, g, or t
 <400> 146

gactatcgcc gttatcggtg nkggttctca aggtcac

37

<210> 147

<211> 37

<212> DNA

<213> Artificial sequence

<220><223> Primer K9_30_101110r

<220><221> misc_feature

<222> (17)..(18)

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

<400> 147

gtgaccttga gaacmnnac cgataacggc gatagtc

37

<210> 148

<211> 37

<212> DNA

<213> Artificial sequence

<220><223> Primer K9_67_101110f

<220><221> misc_feature

<222> (18)..(19)

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

<400> 148

gagctgaaga acaaggttnk gaagtctaca ccgctgc

37

<210> 149

<211> 37

<212> DNA

<213> Artificial sequence

<220><223> Primer K9_67_101110r

<220><221> misc_feature

<222> (19)..(20)

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

<400> 149

gcagcgggtg agacttcmmn acctgtttct tcagctc

37

<210> 150

<211> 35

<212> DNA
 <213> Artificial sequence
 <220><223> Primer K9_131_101110f
 <220><221> misc_feature
 <222> (18)..(19)
 <223> n is a, c, g, or t
 <400> 150
 caaaggacgt tgatgtcnnk atgacgcctc caaag 35

<210> 151
 <211> 34
 <212> DNA
 <213> Artificial sequence
 <220><223> Primer K9_131_101110r
 <220><221> misc_feature
 <222> (17)..(18)
 <223> n is a, c, g, or t
 <400> 151
 ctttgagcgc atcatmnnga catcaacgtc cttt 34

<210> 152
 <211> 33
 <212> DNA
 <213> Artificial sequence
 <220><223> Primer K9_162_101110f
 <220><221> misc_feature
 <222> (16)..(17)
 <223> n is a, c, g, or t
 <400> 152
 gctgtcgaac aagacnnkac tggcaaggct ttg 33

<210> 153
 <211> 33
 <212> DNA
 <213> Artificial sequence
 <220><223> Primer K9_162_101110r

<220><221> misc_feature
 <222> (17)..(18)
 <223> n is a, c, g, or t
 <400> 153
 caaagccttg ccagtmnngt cttgttcgac agc 33
 <210> 154
 <211> 38
 <212> DNA
 <213> Artificial sequence
 <220><223> Primer K9_176_101110f
 <220><221> misc_feature
 <222> (19)..(20)
 <223> n is a, c, g, or t
 <400> 154
 gctttggcct acgctttann katcggtggt gctagagc 38

 <210> 155
 <211> 38
 <212> DNA
 <213> Artificial sequence
 <220><223> Primer K9_176_101110r
 <220><221> misc_feature
 <222> (19)..(20)
 <223> n is a, c, g, or t
 <400> 155
 gctctagcac caccgatmn taaagcgtag gccaaagc 38
 <210> 156
 <211> 33
 <212> DNA
 <213> Artificial sequence
 <220><223> Primer K9_191_101110f
 <220><221> misc_feature
 <222> (16)..(17)
 <223> n is a, c, g, or t
 <400> 156

gaaactacct tcagannkga aactgaaacc gac 33

<210> 157

<211> 33

<212> DNA

<213> Artificial sequence

<220><223> K9_191_101110r

<220><221> misc_feature

<222> (17)..(18)

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

<400> 157

gtcggtttca gtttcmnntc tgaagtagt ttc 33

<210> 158

<211> 35

<212> DNA

<213> Artificial sequence

<220><223> Primer K9_227_101110f

<220><221> misc_feature

<222> (16)..(17)

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

<400> 158

gccggttacg acccannkaa cgcttacttc gaatg 35

<210> 159

<211> 35

<212> DNA

<213> Artificial sequence

<220><223> Primer K9_227_101110r

<220><221> misc_feature

<222> (19)..(20)

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

<400> 159

cattcgaagt aagcggtmnn tgggtcgtaa ccggc 35

<210> 160

<211> 37

<212> DNA
 <213> Artificial sequence
 <220><223> Primer K9_246_101110f
 <220><221> misc_feature
 <222> (16)..(17)
 <223> n is a, c, g, or t
 <400> 160
 gttgacttga tctacnnktc tggtttctcc ggtatgc

37

<210> 161
 <211> 37
 <212> DNA
 <213> Artificial sequence
 <220><223> Primer K9_246_101110r
 <220><221> misc_feature
 <222> (21)..(22)
 <223> n is a, c, g, or t
 <400> 161
 gcataccgga gaaaccagam nngtagatca agtcaac

37

<210> 162
 <211> 41
 <212> DNA
 <213> Artificial sequence
 <220><223> Primer K9_37&39_101110f
 <220><221> misc_feature
 <222> (18)..(19)
 <223> n is a, c, g, or t
 <220><221> misc_feature
 <222> (24)..(25)
 <223> n is a, c, g, or t
 <400> 162
 gtctcaagg tcacgctnnk gccnnkaatg ctaaggaate c

41

<210> 163
 <211> 41

<212> DNA

<213> Artificial sequence

<220><223> Primer K9_37&39_101110r

<220><221> misc_feature

<222> (17)..(18)

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

<220><221> misc_feature

<222> (23)..(24)

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

<400> 163

ggattcctta gcattmnngg cmnagcgtg accttgagaa c

41

<210> 164

<211> 38

<212> DNA

<213> Artificial sequence

<220><223> Primer

K9_86&87_101110f

<220><221> misc_feature

<222> (19)..(20)

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

<220><221> misc_feature

<222> (22)..(23)

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

<400> 164

gacatcatta tgatcttgnn knnkgatgaa aagcaggc

38

<210> 165

<211> 38

<212> DNA

<213> Artificial sequence

<220><223> Primer K9_86&87_101110r

<220><221> misc_feature

<222> (16)..(17)

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

<220><221> misc_feature

<222> (19)..(20)

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

<400> 165

gcctgctttt catcmnnmnn caagatcata atgatgtc 38

<210> 166

<211> 27

<212> DNA

<213> Artificial sequence

<220><223> Primer K9_131L_112210f

<400> 166

ggacgttgat gtcttgatga tcgctcc 27

<210> 167

<211> 27

<212> DNA

<213> Artificial sequence

<220><223> Primer K9_131A_112210f

<400> 167

ggacgttgat gtcgcaatga tcgctcc 27

<210> 168

<211> 27

<212> DNA

<213> Artificial sequence

<220><223> Primer K9_131V_112210f

<400> 168

ggacgttgat gtcgttatga tcgctcc 27

<210> 169

<211> 27

<212> DNA

<213> Artificial sequence

<220><223> Primer K9_131M_112210f

<400> 169

ggacgttgat gtcatgatga tcgctcc 27

<210> 170

<211> 27

<212> DNA
 <213> Artificial sequence
 <220><223> Primer K9_131C_112210f
 <400> 170
 ggacgttgat gtctgaatga tcgctcc 27

<210> 171
 <211> 30
 <212> DNA
 <213> Artificial sequence
 <220><223> Primer K9_191D_112210r
 <400> 171
 ggtttcagtt tcgtctctga aggtagtttc 30

<210> 172
 <211> 30
 <212> DNA
 <213> Artificial sequence
 <220><223> Primer K9_191C_112210r
 <400> 172
 ggtttcagtt tcgcatctga aggtagtttc 30

<210> 173
 <211> 30
 <212> DNA
 <213> Artificial sequence
 <220><223> Primer K9_191S_112210r
 <400> 173
 ggtttcagtt tccgactctga aggtagtttc 30

<210> 174
 <211> 30
 <212> DNA
 <213> Artificial sequence
 <220><223> Primer K9_191G_112210r
 <400> 174
 ggtttcagtt tcgcctctga aggtagtttc 30

<210> 175

<211> 11864

<212> DNA

<213> Artificial sequence

<220><223> pHR81-PCUP1-AIsS-PILV5-Pf5.KARI (pLH550) plasmid

<400> 175

catcatcaca gtttaaacag tatgaaagt ttctacgata aagactgcga cctgtcgatc	60
atccaaggta agaaagtgc catcatcggc tacggttccc agggccacgc tcaagcatgc	120
aacctgaagg attccggcgt agacgtgact gttggcctgc gtaaaggctc ggctaccgtt	180
gccaaaggctg aagcccacgg cttgaaagt accgacgttg ctgcagccgt tgccggtgcc	240
gacttggtca tgatcctgac cccggacgag ttccagtcac agctgtacaa gaacgaaatc	300
gagccgaaca tcaagaagg cgccactctg gccttctccc acggcttcgc gatccactac	360
aaccaggttg tgcctcgtgc cgacctgac gtgatcatga tcgcgccgaa ggctccaggc	420
cacacgtac gttccgagtt cgtcaagggc ggtgggtatt ctgacctgat cgcgatctac	480
caggacgctt ccggcaacgc caagaacgtt gccctgtcct acgccgcagg cgtgggcggc	540
ggccgtaccg gcatcatcga aaccaccttc aaggacgaga ctgaaaccga cctgttcggt	600
gagcaggctg ttctgtgtgg cgggtaccgtc gagctggtca aagccggttt cgaaaccctg	660
gttgaagctg gctacgtccc agaaatggcc tacttcgagt gccctgcacga actgaagctg	720
atcgttgacc tcatgtacga aggcgggtatc gccaacatga actactcgat ctccaacaac	780
gtcgaatacg gcgagtacgt gactggtcca gaagtcatca acgccgaatc ccgtcaggcc	840
atgcgcaatg ctctgaagcg catccaggac ggccaatacg cgaagatgtt catcagcgaa	900
ggcgtaccg gctaccatc gatgaccgcc aagcgtcgta acaacgtgc tcacggtatc	960
gaaatcatcg gcgagcaact gcgctcgatg atgccttgga tcggtgcca caaatcgtc	1020
gacaaagcca agaactaagg ccctgcaggc cagaggaaaa taatatcaag tgctggaaac	1080
tttttctctt ggaatttttg caacatcaag tcatagtcaa ttgaattgac ccaatttcac	1140
atttaagatt ttttttttt catccgacat acatctgtac actaggaagc cctgtttttc	1200
tgaagcagct tcaaatatat atatttttta catatttatt atgattcaat gaacaatcta	1260
attaaatcga aaacaagaac cgaaacgcga ataaataatt tatttagatg gtgacaagtg	1320
tataagtcct catcgggaca gctacattt ctctttcggt ttggctgag ctactggttg	1380
ctgtgacgca gcggcattag gcggcggtta tgagctaccc tcgtggcctg aaagatggcg	1440
ggaataaagc ggaactaaaa attactgact gagccatatt gaggtcaatt tgtcaactcg	1500

tcaagtcacg ttigtggac ggcccccttc caacgaatcg tataactaa catgcgcgcg	1560
cttcctatat acacatatat atatatatat atatatatat gtgtgcgtgt atgtgtacac	1620
ctgtatttaa tttccttact cgcgggtttt tcttttttct caattcttgg cttectcttt	1680
ctcgagcgga ccggtacctc cgcggtgccg gcagatctat ttaaatggcg gcccgacgtc	1740
aggtggcact tttcggggaa atgtgcgcgg aaccctatt tgtttatatt tctaaataca	1800
ttcaaatatg tatccgctca tgagacaata accctgataa atgcttcaat aatattgaaa	1860
aaggaagagt atgagtattc aacatttccg tgcgccttt attccctttt ttgcggcatt	1920
ttgccttctt gtttttgctc acccagaaac gctggtgaaa gtaaaagatg ctgaagatca	1980
gttgggtgca cgagtgggtt acatcgaact ggatctcaac agcggtaaga tccttgagag	2040
ttttcgcctc gaagaacgtt ttccaatgat gaggcatttt aaagtctctg tatgtggcgc	2100
ggtattatcc cgtattgacg ccgggcaaga gcaactcggc gcgcgcatc actattctca	2160
gaatgacttg gttgagtact caccagtcac agaaaagcat cttacggatg gcatgacagt	2220
aagagaatta tgcagtgtg ccataacat gagtgataa actgcggcca acttacttct	2280
gacaacgac ggaggaccga aggagctaac cgtttttttg cacaacatgg gggatcatgt	2340
aactgcctt gatcgttggg aaccggagct gaatgaagcc ataccaaacg acgagcgtga	2400
caccacgatg cctgtagcaa tggcaacaac gttgcgcaaa ctattactg gcgaactact	2460
tactctagct tccgggcaac aattaataga ctggatggag gcggataaag ttgcaggacc	2520
acttctgcgc tcggcccttc cggctggctg gtttattgct gataaatctg gagccgtga	2580
gcgtgggtct cgcggtatca ttgcagcact ggggccagat ggtaagccct cccgtatcgt	2640
agttatctac acgacgggga gtcaggcaac tatggatgaa cgaaatagac agatcgtga	2700
gatagtgcc tcatgatta agcattggta actgtcagac caagtttact catatatact	2760
ttagattgat ttaaaactc atttttaatt taaaaggatc taggtgaaga tcctttttga	2820
taatctcatg accaaaatcc cttaacgtga gttttcgtt cactgagcgt cagacccgt	2880
agaaaagatc aaaggatctt cttgagatcc tttttttctg cgcgtaact gctgcttgca	2940
aacaaaaaaa ccaccgtac cagcgggtgt ttgtttgccg gatcaagagc taccaactct	3000
ttttccgaag gtaactggct tcagcagagc gcagatacca aatactgttc ttctagtga	3060
gccgtagtta ggccaccact tcaagaactc ttagcaccg cctacatacc tcgctctgct	3120
aatcctgtta ccagtggctg ctgccagtgg cgataagtcg tgtcttacg gtttgactc	3180
aagacgatag ttaccggata aggcgcagcg gtcgggctga acggggggtt cgtgcacaca	3240
gccagcttg gagcgaacga cctacaccga actgagatac ctacagcgtg agctatgaga	3300
aagcgcacg ctccccgaag ggagaaaggc ggacaggtat ccggtlaagc gcagggtcgg	3360

aacaggagag cgcacgaggg agcttccagg gggaacgcc tggatcttt atagtcctgt	3420
cgggtttcgc cacctctgac ttgagcgtcg atttttgtga tgctcgtcag gggggcggag	3480
cctatggaaa aacgccagca acgcggcctt ttacggttc ctggcctttt gctggccttt	3540
tgctcacatg ttctttcctg cgttatcccc tgattctgtg gataaccgta ttaccgcctt	3600
tgagttagct gataccgctc gccgcagccg aacgaccgag cgcagcaggt cagttagcga	3660
ggaagcggaa gagcgcccaa tacgcaaacc gcctctcccc gcgcgttggc cgattcatta	3720
atgcagctgg cagcacaggt ttcccgactg gaaagcgggc agtgagcgca acgcaattaa	3780
tgtgagttag ctactcatt aggcacccca ggctttacac tttatgcttc cggtcgtat	3840
gttgtgtgga attgtgagcg gataacaatt tcacacagga aacagctatg accatgatta	3900
cgccaagctt ttctttcca atttttttt ttctgtcatt ataaaaatca ttacgaccga	3960
gattccccgg taataactga tataattaaa ttgaagctct aatttgtgag tttagtatac	4020
atgcatttac ttataataga gttttttagt ttgctggcc gcattcttc aaatatgctt	4080
cccagcctgc tttctgtaa cgttcacct ctaccttagc atcccttccc ttgcaaata	4140
gtctcttcc aacaataata atgtcagatc ctgtagagac cacatcatcc acggttctat	4200
actgttgacc caatgcgtct ccttgtcat ctaaaccac accgggtgtc ataatacaacc	4260
aatcgtaacc ttcatctctt ccacctatgt ctctttgagc aataaagccg ataacaaaat	4320
ctttgtcgt ctctgcaatg tcaacagtac ccttagtata ttctccagta gatagggagc	4380
ccttgcatga caattctgct aacatcaaaa ggctcttagg ttctttgtt acttcttctg	4440
ccgctgctt caaacgcta acaatactg ggcccaccac accgtgtgca ttcgtaatgt	4500
ctgccattc tgctattctg tatacacccg cagagtactg caatttgact gtattaccaa	4560
tgtcagcaaa tttctgtct tcgaagagta aaaaattgta cttggcggat aatgccttta	4620
gcggcttaac tgtgccctcc atggaaaaat cagtcaagat atccacatgt gtttttagta	4680
aacaaatttt gggaccctaat gcttcaacta actccagtaa ttcttgggtg gtacgaacat	4740
ccaatgaagc acacaagttt gtttgccttt cgtgcatgat attaaatagc ttggcagcaa	4800
caggactagg atgagtagca gcacgttcct tatatgtagc tttcgacatg atttatcttc	4860
gtttcctgca ggtttttgtt ctgtgcagtt gggtaagaa tactgggcaa tttcatgttt	4920
cttcaacact acatatgct atatatacca atctaagtct gtgtctcttc ctctgttctt	4980
ccttctgttc ggagattacc gaatcaaaaa aatttcaagg aaaccgaaat caaaaaaag	5040
aataaaaaaa aaatgatgaa ttgaaaagct tgcattgctg caggtcgact ctagtatact	5100
ccgtctactg tacgatacac ttccgtcag gtccttgtcc tttacgagg cttaccact	5160

cttttgttac tctattgac cagctcagca aaggcagtgt gatctaagat tctatcttcg 5220

cgatgtagta aaactagcta gaccgagaaa gagactagaa atgcaaaagg cacttctaca 5280

atggctgcc a tcattattat ccgatgtgac gctgcatttt tttttttttt tttttttttt 5340

ttttttttt tttttttttt tttttttgt acaaatatca taaaaaaga gaatcttttt 5400

aagcaaggat tttcttaact tcttcggcga cagcatcacc gacttcggtg gtactgttgg 5460

aaccaccta atcaccagtt ctgatactg catccaaaac ctttttaact gcacttcaa 5520

tggctttacc tcttcaggc aagtccaatg acaatttcaa catcattgca gcagacaaga 5580

tagtggcgat agggttgacc ttattctttg gcaaatctgg agcggaacca tggcatggtt 5640

cgtaaaaacc aaatgcggtg ttcttgtctg gcaaagaggc caaggacgca gatggcaaca 5700

aaccaagga gcctgggata acggaggctt catcgagat gatacacca aacatgttgc 5760

tgggtattat aatacattt aggtgggttg ggttcttaac taggatcatg gcgcgagaat 5820

caatcaattg atgttgaact ttcaatgtag ggaattcgtt cttgatggtt tcctccacag 5880

tttttctcca taatcttgaa gaggccaaaa cattagcttt atccaaggac caaataggca 5940

atggtggctc atgtttagg gccatgaaag cgccattct tgtgattctt tgcacttctg 6000

gaacggtgta ttgttacta tccaagcga caccatcacc atcgtcttcc tttctcttac 6060

caaagtaaat acctccact aattctctaa caacaacgaa gtcagtacct ttagcaaatt 6120

gtggcttgat tggagataag tctaaaagag agtcggatgc aaagttacat ggtcttaagt 6180

tggcgtacaa ttgaagtctt ttacggattt ttagtaaac ttgttcaggt ctaacactac 6240

cggtaaccca tttaggacca cccacagcac ctaacaaaac ggcatcagcc ttcttggagg 6300

cttccagcgc ctcatctgga agtgaacac ctgtagcatc gatagcagca ccaccaatta 6360

aatgattttc gaaatcgaac ttgacattgg aacgaacatc agaaatagct ttaagaacct 6420

taatggcttc ggctgtgatt tcttgaccaa cgtggtcacc tggcaaaacg acgatcttct 6480

taggggcaga cattacaatg gtatatcctt gaaatatata taaaaaaaaa aaaaaaaaaa 6540

aaaaaaaaa atgcagcttc tcaatgatat tcgaatacgc tttgaggaga tacagcctaa 6600

tatccgacaa acigtcttac agatttacga tcgtacttgt taccatcat tgaatttga 6660

acatccgaac ctgggagttt tccctgaaac agatagtata tttgaacctg tataataata 6720

tatagtctag cgctttacgg aagacaatgt atgtatttcg gttcctggag aaactattgc 6780

atctattgca taggtaatct tgcacgtcgc atccccggtt cattttctgc gtttccatct 6840

tgcacttcaa tagcatactt ttgttaacga agcatctgtg cttcattttg tagaacaaaa 6900

atgcaacgcg agagcgctaa ttttcaaac aaagaatctg agctgcattt ttacagaaca	6960
gaaatgcaac gcgaaagcgc tattttacca acgaagaatc tgtgcttcat ttttgtaaaa	7020
caaaaatgca acgcgagagc gctaattttt caaacaaga atctgagctg catttttaca	7080
gaacagaaat gcaacgcgag agcgctatit taccaacaaa gaatctatac ttttttttg	7140
ttctacaaaa atgcatcccg agagcgctat ttttctaaca aagcatctta gattactttt	7200
tttctccttt gtgcgctcta taatgcagtc tcttgataac ttttgcact gtaggtccgt	7260
taaggttaga agaaggctac tttggtgtct attttctctt ccataaaaaa agcctgactc	7320
cacttccgc gtttactgat tactagcgaa gctgcgggtg cattttttca agataaaggc	7380
atccccgatt atattctata ccgatgtgga ttgcgcatac tttgtgaaca gaaagtata	7440
gcgttgatga ttcttcattg gtcagaaaat tatgaacggt ttttctctat ttgtctctat	7500
atactacgta taggaaatgt ttacattttc gtattgtttt cgattcactc tatgaatagt	7560
tcttactaca atttttttgt ctaaagagta atactagaga taaacataaa aaatgtagag	7620
gtcgagtta gatgcaagtt caaggagcga aaggtggatg ggtaggttat ataggatat	7680
agcacagaga tatatagcaa agagatactt ttgagcaatg tttgtggaag cggtatcgc	7740
aatattttag tagctcgta cagtcggtg cgtttttggg ttttgaaag tgcgtcttca	7800
gagcgctttt ggttttcaaa agcgctctga agttcctata ctttctagag aataggaact	7860
tcggaatagg aacttcaaa cgtttccgaa aacgagcgct tccgaaaatg caacgcgagc	7920
tgcgcacata cagctcactg ttcacgtcgc acctatatct gcgtgttgcc tgtatatata	7980
tatacatgag aagaacggca tagtgcgtgt ttatgcttaa atgcgtactt atatgcgtct	8040
atttatgtag gatgaaagg agtctagtag ctccgtgat attatccat tccatgcggg	8100
glatcgtag ctcccttcag cactaccctt tagctgttct atatgctgcc actcctcaat	8160
tggattagtc tcatccttca atgctatcat ttcctttgat attggatcat atgcatagta	8220
ccgagaaact agaggatctc ccattaccga catttgggcg ctatacgtgc atatgttcat	8280
gtatgtatct gtatttataa cacttttgta ttatttttcc tcatatatgt gtataggttt	8340
atacggatga ttttaattatt acttcaccac cctttatttc aggtgatat cttagccttg	8400
ttactagtta gaaaaagaca tttttgctgt cagtcactgt caagagattc ttttgctggc	8460
atttcttcta gaagcaaaaa gagcgatgcg tcttttccgc tgaaccgttc cagcaaaaaa	8520
gactaccaac gcaatatgga ttgtcagaat catataaaag agaagcaaat aactccttgt	8580
cttgatcaa ttgcattata atatcttctt gttagtgcaa tatcatatag aagtcacga	8640
aatagatatt aagaaaaaca aactgtacaa tcaatcaatc aatcatcgct gaggatgttg	8700
acaaaagcaa caaagaaca aaaatccctt gtgaaaaaca gaggggcgga gcttgttgtt	8760

gattgcttag tggagcaagg tgtcacacat gtatttggca ttccaggtgc aaaaattgat	8820
gcggtatttg acgctttaca agataaagga cctgaaatta tcgttgcccg gcacgaacaa	8880
aacgcagcat tcatggccca agcagtcggc cgtttaactg gaaaaccggg agtcgtgtta	8940
gtcacatcag gaccgggtgc ctctaacttg gcaacaggcc tgctgacagc gaacactgaa	9000
ggagaccctg tcgttgcgt tgctggaaac gtgatccgtg cagatcgttt aaacggaca	9060
catcaatctt tggataatgc ggcgtattc cagccgatta caaaatacag ttagaagtt	9120
caagatgtaa aaaatatacc ggaagctgtt acaaatgcat ttaggatagc gtcagcaggg	9180
caggctgggg ccgcttttgt gagctttccg caagatgttg tgaatgaagt cacaaatacg	9240
aaaaacgtgc gtgctgttgc agcgccaaaa ctcggtcctg cagcagatga tgcaatcagt	9300
gcggccatag caaaaatcca aacagcaaaa ctctctgtcg ttttggtcgg catgaaaggc	9360
ggaagaccgg aagcaattaa agcgtttcgc aagcttttga aaaaggttca gcttccattt	9420
gttgaacat atcaagctgc cggtaacctt tctagagatt tagaggatca atattttggc	9480
cgtatcggtt tgttccgcaa ccagcctggc gatttactgc tagagcaggc agatgttgtt	9540
ctgacgatcg gctatgacc gattgaatat gatccgaaat tctggaatat caatggagac	9600
cggacaatta tccattttaga cgagattatc gctgacattg atcatgtta ccagcctgat	9660
cttgaattga tcggtgacat tccgtccacg atcaatcata tcgaacacga tgctgtgaaa	9720
gtggaatttg cagagcgtga gcagaaaaac ctttctgatt taaaacaata tatgcatgaa	9780
ggtgagcagg tgctgcaga ttggaaatca gacagagcgc accctcttga aatcgtaaa	9840
gagttgcgta atgcagtcga tgatcatgtt acagtaactt gcgatatcgg ttcgcacgcc	9900
atttgatgt cagcttattt ccgcagctac gagccgttaa cattaatgat cagtaacggt	9960
atgcaaacac tcggcgttgc gcttcttgg gcaatcggcg cttcatttgt gaaaccggga	10020
gaaaaagtgg tttctgtctc tgggtgacggc ggtttcttat tctcagcaat ggaattagag	10080
acagcagttc gactaaaagc accaattgta cacattgtat ggaacgacag cacatatgac	10140
atggttgcat tccagcaatt gaaaaaatat aaccgtacat ctgcggtcga tttcggaat	10200
atcgatatcg tgaatatgc ggaaagcttc ggagcaactg gcttgcgctt agaataacca	10260
gaccagctgg cagatgttct gcgtcaaggc atgaacgtg aaggctctgt catcatcgat	10320
gtcccgttg actacagtga taacattaat ttagcaagtg acaagcttcc gaaagaattc	10380
ggggaactca tgaaaacgaa agctctctag ttaattaatc atgtaattag ttatgtcacg	10440
cttacattca cgcctcccc ccacatccg tctaaccgaa aaggaaggag ttagacaacc	10500
tgaagtctag gtccctatit atttttttat agttatgtta gtattaagaa cgttatttat	10560
atttcaaatt tttctttttt ttctgtacag acgcgtgtac gcatgtaaca ttatactgaa	10620

aaccttgctt gagaagggtt tgggacgctc gaaggcttta atttgcgggc ggccgcacct 10680

ggtaaacct ctagtggagt agtagatgta atcaatgaag cggaagccaa aagaccagag 10740

tagaggccta tagaagaaac tgcgatacct tttgtgatgg ctaaacaaac agacatcttt 10800

ttatatgttt ttacttctgt atatcgtgaa gtagtaagtg ataagcgaat ttggctaaga 10860

acgtttgaag tgaacaaggg acctcttttg cttttcaaaa aaggattaaa tggagttaat 10920

cattgagatt tagttttcgt tagattctgt atccctaaat aactccctta cccgacggga 10980

aggcacaaaa gacttgaata atagcaaagc gccagtagcc aagaccaaat aatactagag 11040

ttaactgatg gtcttaaaca ggcattacgt ggtgaactcc aagaccaata tacaaaatat 11100

cgataagtta ttcttgccca ccaatttaag gagcctacat caggacagta gtaccattcc 11160

tcagagaaga ggtatacata acaagaaaat cgcgtgaaca cttatataa cttagcccgt 11220

tattgagcta aaaaaccttg caaaatttcc tatgaataag aatacttcag acgtgataaa 11280

aatttacttt ctaactcttc tcacgtgcc cctatctgtt cttccgctct accgtgagaa 11340

ataaagcadc gactacggca gtgcgtgtc actgaactaa aacaataagg ctagttcgaa 11400

tgatgaactt gcttgcgtgc aaacttctga gtgcccgtg atgtgacact gtgacaataa 11460

attcaaacgg gttatagcgg tctcctcggg taccggttct gccacctcca atagagctca 11520

gtaggagtea gaacctctgc ggtggctgtc agtgactcat ccgcgtttcg taagttgtgc 11580

gcgtgcacat ttgcccgtt cccgctcatc ttgcagcagg cggaaatctt catcacgctg 11640

taggacgcaa aaaaaaata attaatcgta caagaatctt ggaaaaaaa ttgaaaaatt 11700

ttgtataaaa gggatgacct aacttgactc aatggctttt acaccagta ttttcctttt 11760

ccttgtttgt tacaattata gaagcaagac aaaaacatat agacaaccta ttcctaggag 11820

ttatatTTTT ttacctacc agcaatataa gtaaaaaact gttt 11864

<210> 176

<211> 11871

<212> DNA

<213> Artificial sequence

<220><223> pYZ058 (pHR81-PCUP1-AlsS-PILV5-yeast KARI plasmid

<400> 176

tccattacc gacatttggg cgctatacgt gcatatgttc atgtatgtat ctgtatttaa 60

aacacttttg tattattttt cctcatatat gtgtataggt ttatacggat gatttaatta 120

ttacttcacc accttttatt tcaggctgat atcttagcct tgttactagt tagaaaaaga 180

catttttgct gtcagtcact gtcaagagat tcttttgctg gcatttcttc tagaagcaaa 240

aagagcgatg cgctttttcc gctgaaccgt tccagcaaaa aagactacca acgcaatatg 300
gattgtcaga atcatataaa agagaagcaa ataactcctt gtcttgtatc aattgcatta 360

taatatcttc ttgttagtgc aatatcatat agaagtcata gaaatagata ttaagaaaaa 420
caaactgtac aatcaatcaa tcaatcatcg ctgaggatgt tgacaaaagc aacaaaagaa 480
caaaaatccc ttgtgaaaaa cagaggggcg gagcttgttg ttgattgctt agtggagcaa 540
ggtgtcacac atgtatttgg cattccaggt gcaaaaattg atgcggtatt tgacgcttta 600
caagataaag gacctgaaat tatcgttgcc cggcacgaac aaaacgcagc attcatggcc 660
caagcagtcg gccgtttaac tggaaaaccg ggagtcgtgt tagtcacatc aggaccgggt 720
gcctctaact tggcaacagg cctgctgaca gcgaacactg aaggagaccc tgtcgttgcg 780

cttgcctgaa acgigatccg tgcagatcgt ttaaaacgga cacatcaatc tttggataat 840
gcggcgctat tccagccgat taaaaatac agtgtagaag ttcaagatgt aaaaaatata 900
ccggaagctg ttacaaatgc atttaggata gcgtcagcag ggcaggctgg ggccgctttt 960
gtgagctttc cgcaagatgt tgtgaatgaa gtcacaaata cgaaaaacgt gcgtgctgtt 1020
gcagcgccaa aactcggctc tgcagcagat gatgcaatca gtgcggccat agcaaaaaatc 1080
caaacagcaa aacttcctgt cgttttggtc ggcatgaaag gcggaagacc ggaagcaatt 1140
aaagcggttc gcaagctttt gaaaaagggt cagcttccat ttgttgaaac atatcaagct 1200

gccggatccc ttcttagaga tttagaggat caatatattg gccgatcgg tttgttccgc 1260
aaccagcctg gcgatttact gctagagcag gcagatgttg ttctgacgat cggctatgac 1320
ccgattgaat atgatccgaa attctggaat atcaatggag accggacaat tatccattta 1380
gacgagatta tcgctgacat tgatcatgct taccagcctg atcttgaatt gatcggtgac 1440
attccgtcca cgatcaatca tatcgaacac gatgctgtga aagtggaatt tgcagagcgt 1500
gagcagaaaa tcctttctga tttaaaacaa tatatgcatg aaggtgagca ggtgcctgca 1560
gattggaaat cagacagagc gcacctctt gaaatcgta aagagttgcg taatgcagtc 1620

gatgatcatg ttacagtaac ttgcgatac ggttcgcacg ccatttggat gtcacgttat 1680
ttccgcagct acgagccgtt aacattaatg atcagtaacg gtatgcaaac actcggcgtt 1740
gcgcttcctt gggcaatcgg cgcttcattg gtgaaaccgg gaaaaaagt ggtttctgtc 1800
tctggtgacg gcggtttctt attctcagca atggaattag agacagcagt tcgactaaaa 1860
gcaccaattg tacacattgt atggaacgac agcacatatg acatggttgc attccagcaa 1920
ttgaaaaaat ataaccgtac atctgcggtc gatttcggaa atatcgatat cgtgaaatat 1980

gcggaaagct tcggagcaac tggcttgcgc gtagaatcac cagaccagct ggcagatgtt	2040
ctgcgtcaag gcatgaacgc tgaaggctct gtcacatcgc atgtcccgtg tgactacagt	2100
gataacatta atttagcaag tgacaagctt ccgaaagaat tcggggaact catgaaaacg	2160
aaagctctct agttaattaa tcatgtaatt agttatgtca cgcttacatt cagccctcc	2220
ccccacatcc gctctaaccg aaaaggaagg agttagacaa cctgaagtct aggtccctat	2280
ttattttttt atagttaagt tagtattaag aacgtttatt atatttcaa tttttctttt	2340
ttttctgtac agacgcgtgt acgcatgtaa cattatactg aaaccttgc ttgagaaggt	2400
tttgggacgc tcgaaggctt taatttgcgg gcggccgcac ctggtaaaac ctctagtga	2460
gtagtagatg taatcaatga agcgaagcc aaaagaccag agtagaggcc tatagaagaa	2520
actgcgatac cttttgtgat ggctaaacaa acagacatct ttttatatgt ttttacttct	2580
gtatatcgtg aagtagtaag tgataagcga atttggctaa gaacgttgta agtgaacaag	2640
ggacctcttt tgcctttcaa aaaaggattt aatggagtta atcattgaga tttagttttc	2700
gttagattct gtatccctaa ataactcct taccgcacgg gaaggcaca aagacttgaa	2760
taatagcaaa cgccagtag ccaagaccaa ataatactag agttaactga tggctttaa	2820
caggcattac gtggtgaact ccaagaccaa tatacaaat atcgataagt tattcttgcc	2880
caccaattta aggagcctac atcaggacag tagtaccatt cctcagagaa gaggtataca	2940
taacaagaaa atcgcgtaga cacctatat aacttagccc gttattgagc taaaaaacct	3000
tgcaaaattt cctatgaata agaatacttc agacgtgata aaaatttact ttctaactct	3060
tctcacgtg ccctatctg ttcttccgt ctaccgtgag aaataaagca tcgagtacgg	3120
cagttcgctg tcaactgaact aaaacaataa ggctagttcg aatgatgaac ttgcttgctg	3180
tcaaacttct gaggtagccg tgatgtgaca ctgtgacaat aaattcaaac cggttatagc	3240
ggtctcctcc ggtaccggtt ctgccacctc caatagagct cagtaggagt cagaacctct	3300
gcggtggctg tcagtgactc atccgcgttt cgtaagttgt gcgcgtgcac atttcgcccg	3360
ttcccgtca tcttgacga ggcggaattt ttcacacgc tgtaggacgc aaaaaaaaa	3420
taattaatcg tacaagaatc ttggaaaaaa aattgaaaaa tttgtataa aagggatgac	3480
ctaacttgac tcaatggctt ttacaccag tattttccct ttccttggtt gttacaatta	3540
tagaagcaag acaaaaacat atagacaacc ttttctagg agttatattt ttttacccta	3600
ccagcaatat aagtaaaaaa ctgtttaaac agtatgtga agcaaatcaa cttcgtggt	3660
actgttgaaa ccgtctacga aagagctgac tggccaagag aaaagttgtt ggactacttc	3720

aagaacgaca cttttgcttt gatcggttac gggtcccaag gttacggtca aggtttgaac 3780
ttgagagaca acggtttgaa cgttatcatt ggtgtccgta aagatggtgc ttcttggaag 3840
gctgccatcg aagacggttg gggtccaggc aagaacttgt tcaactgtga agatgctatc 3900
aagagaggta gtiacgttat gaacttgttg tccgatgccg ctcaatcaga aacctggcct 3960
gctatcaagc cattgttgac caagggttaag actttgtact tctcccacgg ttctcccca 4020
gtcttcaagg acttgactca cgttgaacca ccaaaggact tagatgttat ctgggttgc 4080
ccaaagggtt ccggtagaac tgtcagatct ttgttcaagg aaggtcgtgg tattaactct 4140

tcttacgccg tctggaacga tgtcaccggt aaggtcacg aaaaggccca agctttggcc 4200
gttgccattg gttccggtta cgtttacc aaactttcg aaagagaagt caactctgac 4260
ttgtacgggtg aaagaggttg tttaatgggt ggtatccacg gtatgttctt ggctcaatac 4320
gacgtcttga gagaaaaacg tcaactccca tctgaagctt tcaacgaaac cgtcgaagaa 4380
gtacaccaat ctctataccc attgatcgggt aagtacggta tggattacat gtacgatgct 4440
tgttccacca ccgccagaag aggtgctttg gactggtacc caatcttcaa gaatgctttg 4500
aagcctgttt tccaagactt gtacgaatct accaagaacg gtaccgaaac caagagatct 4560

ttggaattca actctcaacc tgactacaga gaaaagctag aaaaggaatt agacaccatc 4620
agaaacatgg aaatctggaa ggttggttaag gaagtcagaa agttgagacc agaaaaccaa 4680
taaggccctg caggccctatc aagtgtgga aactttttct cttggaattt ttgcaacatc 4740
aagtcatagt caattgaatt gacccaattt cacatttaag atttttttt ttcatccga 4800
catacatctg tacactagga agcctgttt ttctgaagca gttcaaata tatatatatt 4860
ttacatattt attatgattc aatgaacaat ctaattaaat cgaaaacaag aaccgaaacg 4920
cgaataaata atttatattg atggtgacaa gtgtataagt cctcatcggg acagctacga 4980

tttctctttc ggttttggct gagctactgg ttgctgtgac gcagcggcat tagcgcggcg 5040
ttatgagcta cctcgtggc ctgaaagatg gcgggaataa agcgggaacta aaaattactg 5100
actgagccat attgaggtca atttgtcaac tcgtcaagtc acgtttggtg gacggccct 5160
ttccaacgaa tcgtatatac taacatgcgc gcgcttccta tatacacata tacatatata 5220
tatatatata tatgtgtgcg tgtatgtgta cacctgtatt taatttcctt actcgcgggt 5280
ttttctttt tctcaattct tggcttctc tttctcgagc ggaccggatc ctccgcggtg 5340
ccggcagatc tatttaaatg gcgcgccgac gtcagggtgc acttttcggg gaaatgtgcg 5400

cggaaccctt atttgtttat ttttctaaat acattcaa atgtatccgc tcatgagaca 5460
ataaccctga taaatgcttc aataatattg aaaaaggaag agtatgagta ttcaacattt 5520
ccgtgtcgcc ctattccct tttttcggc attttgctt cctgtttttg ctacaccaga 5580

aacgctgggtg aaagtaaaag atgctgaaga tcagttgggt gcacgagtgg gttacatcga 5640
 actggatctc aacagcggta agatccttga gagttttcgc cccgaagaac gttttccaat 5700
 gatgagcact tttaaagtgc tgctatgtgg cgcggtatta tcccgattg acgccgggca 5760
 agagcaactc ggicgccgca tacactattc tcagaatgac ttggttgagt actcaccagt 5820

 cacagaaaag catcttacgg atggcatgac agtaagagaa ttatgcagtg ctgccataac 5880
 catgagtgat aacactgcgg ccaacttact tctgacaacg atcggaggac cgaaggagct 5940
 aaccgctttt ttgcacaaca tgggggatca tgtaactcgc cttgatcgtt gggaaccgga 6000
 gctgaatgaa gccataccaa acgacgagcg tgacaccacg atgcctgtag caatggcaac 6060
 aacgttgccg aaactattaa ctggcgaact acttactcta gcttcccggc aacaattaat 6120
 agactggatg gaggcggata aagttgcagg accacttctg cgctcggccc ttccggctgg 6180
 ctggtttatt gctgataaat ctggagccgg tgagcgtggg tctcgcggtc tcattgcagc 6240

 actggggcca gatggaagc cctcccgtat cgtagtatc tacacgacgg ggagtcaggc 6300
 aactatggat gaacgaaata gacagatcgc tgagataggt gcctcactga ttaagcattg 6360
 gtaactgtca gaccaagttt actcatatat actttagatt gatttaaaac ttcattttta 6420
 atttaaaagg atctaggatg agatcctttt tgataatctc atgacaaaaa tcccttaacg 6480
 tgagttttcg ttccactgag cgtcagacc cgtagaaaag atcaaaggat cttcttgaga 6540
 tccttttttt ctgcgcgtaa tctgtctgtt gcaaacaaaa aaaccaccgc taccagcggc 6600
 ggtttgtttg ccggatcaag agctaccaac tctttttccg aaggtaactg gcttcagcag 6660

 agcgcagata ccaataactg ttcttctagt gtagccgtag ttaggccacc acttcaagaa 6720
 ctctgtagca ccgcctacat acctcgtctt gctaactctg ttaccagtgg ctgctgccag 6780
 tggcgataag tcgtgtctta ccgggttgga ctcaagacga tagttaccgg ataaggcgca 6840
 gcggtcgggc tgaacggggg gttcgtgcac acagcccagc ttggagcgaa cgacctacac 6900
 cgaactgaga tacctacagc gtgagctatg agaaagcgcc acgcttcccg aaggagagaa 6960
 ggccgacagg tatccggtaa gcggcagggt cggaacagga gagcgcacga gggagcttcc 7020
 agggggaaac gcctggatc tttatagtcc tgcgggttt cgccacctct gacttgagcg 7080

 tcgatttttg tgatgtcgt caggggggcg gagcctatgg aaaaacgcca gcaacgcggc 7140
 ctttttacgg ttccctggcct ttgtctggcc tttgtctcac atgttcttcc ctgcgttacc 7200
 ccctgattct gtggataacc gtattaccgc ctttgagtga gctgataccg ctgcgccgag 7260
 ccgaacgacc gagcgacgag agtcagttag cgaggaagcg gaagagcgcc caatacgcaa 7320
 accgcctctc cccgcgcgtt ggccgattca ttaatgcagc tggcacgaca ggtttcccga 7380
 ctggaaagcg ggcagtgagc gcaacgcaat taatgtgagt tagctcactc attaggcacc 7440

ccaggcttta cactttatgc ttccggctcg tatgttgtgt ggaattgtga gcggataaca 7500

atttcacaca ggaacacagct atgaccatga ttacgccaag ctttttcttt ccaatttttt 7560

ttttttcgte attataaaaa tcattacgac cgagattccc gggtaataac tgatataatt 7620

aaattgaagc tctaatttgt gagtttagta tacatgcatt tacttataat acagtttttt 7680

agttttgctg gccgcatctt ctcaaatatg ctteccagcc tgcttttctg taacgttcac 7740

cctctacett agcatecctt ccctttgcaa atagtctctt tccaacaata ataatgtcag 7800

atcctgtaga gaccacatca tccacgggtc tatactgttg acccaatgcg tctcccttgt 7860

catctaaacc cacaccgggt gtcataatca accaatcgta accttcatct cttecaccca 7920

tgctcttttg agcaataaag ccgataacaa aatctttgtc gctcttcgca atgtcaacag 7980

tacccttagt atatttccca gtagataggg agcccttgca tgacaattct gtaacatca 8040

aaaggcctct aggttccctt gttacttctt ctgccgctg cttcaaaccg ctaacaatac 8100

ctgggcccac cacaccgtgt gcattcgtaa tgcctgcca ttctgctatt ctgtatacac 8160

ccgcagagta ctgcaatttg actgtattac caatgtcagc aaattttctg tcttcgaaga 8220

gtaaaaaatt gtacttggcg gataatgcct ttagcggtt aactgtgcc tccatggaaa 8280

aatcagtcaa gatatccaca tgtgttttta gtaaacaaat tttgggacct aatgcttcaa 8340

ctaactccag taattccttg gtggtagcaa catccaatga agcacacaag tttgtttgct 8400

tttcgtgcat gatattaaat agcttggcag caacaggact aggatgagta gcagcacgtt 8460

ccttatatgt agctttcgac atgatttatc ttcttttctt gcaggttttt gttctgtgca 8520

gttgggttaa gaatactggg caatttcatg tttcttcaac actacatatg cgtatatata 8580

ccaatctaag tcigtgtccc ttcttctgtt ctctcttctg ttcggagatt accgaatcaa 8640

aaaaatttca aggaaaccga aatcaaaaaa aagaataaaa aaaaaatgat gaattgaaaa 8700

gcttgcatgc ctgcaggctg actctagtat actccgtcta ctgtacgata cacttccgct 8760

caggtccttg tcctttaacg aggccttacc actcttttgt tactctattg atccagctca 8820

gcaaaggcag tgtgatctaa gattctatct tcgcgatgta gtaaaactag ctagaccgag 8880

aaagagacta gaaatgcaaa aggcacttct acaatggctg ccatcattat tatccgatgt 8940

gacgctgcat tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt 9000

gtacaaaata tcataaaaaa agagaatctt ttaagcaag gattttctta acttcttcgg 9060

cgacagcatc accgacttcg gtggactgtg tggaaccacc taaatcacca gttctgatac 9120

ctgcatccaa aaccttttta actgcatctt caatggcttt accttcttca ggcaagttca 9180

atgacaattt caacatcatt gcagcagaca agatagtggc gatagggttg accttattct	9240
ttggcaaadc tggagcggaa ccatggcatg gtctgtacaa accaaatgcg gtgttcttgt	9300
ctggcaaaga ggccaaggac gcagatggca acaaacccaa ggagcctggg ataacggagg	9360
cttcatcgga gatgataca ccaaacatgt tgcctgtgat tataatacca tttaggtggg	9420
ttgggttctt aactaggatc atggcggcag aatcaatcaa ttgatgttga actttcaatg	9480
tagggaattc gttcttgatg gtttcctcca cagtttttct ccataatctt gaagaggcca	9540
aaacattagc tttatccaag gaccaaatag gcaatggtgg ctcatgttgt agggccatga	9600
aagcgcccat tcttgtgatt ctttgactt ctggaacggt gtattgttca ctatccaag	9660
cgacaccatc accatcgtct tcctttctct taccaaagta aatacctccc actaatctc	9720
taacaacaac gaagtcagta cctttagcaa attgtggctt gattggagat aagtctaaaa	9780
gagagtcgga tgcaaagta catggtctta agttggcgta caattgaagt tctttacgga	9840
tttttagtaa accttgttca ggtctaacac taccggtacc ccatttagga ccaccacag	9900
cacctaaca aacggcatca gccttcttgg aggcctccag cgcctcatct ggaagtggaa	9960
cacctgtagc atcgatagca gcaccacaa ttaaatgatt ttcgaaatcg aacttgacat	10020
tggaacgaac atcagaaata gctttaagaa ccttaatggc ttcggctgtg atttcttgac	10080
caacgtggtc acctggcaaa acgacgatct tcttaggggc agacattaca atggtatatac	10140
cttgaaatat atataaaaaa aaaaaaaaaa aaaaaaaaaa aaaatgcagc ttctcaatga	10200
tattcgaata cgctttgagg agatacagcc taatatccga caaactgttt tacagattta	10260
cgatcgtact tgttaccat cattgaattt tgaacatccg aacctgggag ttttcctga	10320
aacagatagt atatttgaa ctgtataata atatatagtc tagcgcttta cggaagacaa	10380
tgtatgtatt tcggttcctg gagaaactat tgcattctatt gcataggtaa tcttgcacgt	10440
cgcatccccg gttcattttc tgcgtttcca tcttgactt caatagcata tctttgttaa	10500
cgaagcatct gtgcttcatt ttgtagaaca aaaatgcaac gcgagagcgc taatttttca	10560
aacaaagaat ctgagctgca tttttacaga acagaaatgc aacgcgaaag cgctatttta	10620
ccaacgaaga atctgtgctt cttttttgta aaacaaaaat gcaacgcgag agcgctaatt	10680
tttcaacaa agaactctgag ctgcattttt acagaacaga aatgcaacgc gagagcgcta	10740
ttttaccaac aaagaatcta tacttctttt ttgttctaca aaaatgcac ccgagagcgc	10800
tatttttcta acaaagcatc ttagattact tttttctcc tttgtgcgt ctataatgca	10860
gtctcttgat aactttttgc actgtaggtc cgttaagggtt agaagaaggc tactttggtg	10920
tctattttct cttccataaa aaagcctga ctccacttcc cgcgtttact gattactagc	10980
gaagctgcgg gtgcattttt tcaagataaa ggcatccccg attatattct ataccgatgt	11040

ggattgcgca tactttgtga acagaaagtg atagcgttga tgattcttca ttggtcagaa 11100
aattatgaac ggtttcttct attttgtctc tatatactac gtataggaaa tgtttacatt 11160
ttcgtattgt ttctgattca ctctatgaat agttcttact acaatTTTTT tgtctaaaga 11220
gtaatactag agataaacat aaaaaatgta gaggtcagat ttagatgcaa gttcaaggag 11280

cgaaggtgg atgggtaggT tatataggga tatagcacag agatatatag caaagagata 11340
cttttgagca atgtttgtgg aagcggattt cgcaatattt tagtagctcg ttacagtccg 11400
gtgcgttttt ggttttttga aagtgcgtct tcagagcgct tttggttttc aaaagcgctc 11460
tgaagttcct atactttcta gagaatagga acttcggaat aggaacttca aagcgtttcc 11520
gaaaacgagc gcttccgaaa atgcaacgcg agctgcgcac atacagctca ctgttcacgt 11580
cgcacctata tctgcgtgtt gcctgtatat atatatacat gagaagaacg gcatagtgcg 11640
tgtttatgct taaatgcgta cttatatgcg tctatttatg taggatgaaa ggtagtctag 11700

tacctcctgt gatattatcc cattccatgc ggggtatcgt atgcttcctt cagcactacc 11760
ctttagctgt tctatatgct gccactctc aattggatta gtctcatcct tcaatgctat 11820
catttccttt gatattggat catatgcata gtaccgagaa actagaggat c 11871

<210> 177

<211> 52

<212> DNA

<213> Artificial sequence

<220><223> Primer OT1349

<400> 177

catcatcaca gtttaaacag tatgaaagtT ttctacgata aagactgcga cc 52

<210> 178

<211> 51

<212> DNA

<213> Artificial sequence

<220><223> Primer OT1318

<400>

> 178

gcacttgata ggctgcagg gccttagttc ttggctttgt cgacgatttt g 51

<210> 179

<211> 15

<212> DNA

<213> Artificial sequence

<220><223> Primer OT1383

<400> 179

ctagtcaccg gtggc 15

<210> 180

<211> 15

<212> DNA

<213> Artificial sequence

<220><223> Primer OT1384

<400> 180

ggccgccacc ggtga 15

<210> 181

<211> 9612

<212> DNA

<213> Artificial sequence

<220><223> pHR81::ILV5p-K9D3 (pLH702) plasmid

<400> 181

aaacagtatg gaagaatgta agatggctaa gatttactac caagaagact gtaacttgtc 60
cttgttggat ggtaagacta tcgccgttat cggttacggg tctcaaggtc acgctcatgc 120
cctgaatgct aaggaatccg gttgtaacgt tatcattggg ttatacgaag gtgctaagga 180
ttggaaaaga gctgaagaac aaggtttcga agtctacacc gctgctgaag ctgctaagaa 240
ggctgacatc attatgatct tgatcaacga tgaaaagcag gctaccatgt acaaaaacga 300
catcgaacca aacttggaag ccggtaacat gttgatgttc gtcacgggt tcaacatcca 360

tttcggttgt attgttcac caaaggacgt tgatgtcact atgatcgctc caaagggtcc 420
aggtcacacc gttagatccg aatacgaaga aggtaaagggt gtcccatgct tgggtgctgt 480
cgaacaagac gctactggca aggctttgga tatggctttg gcctacgctt tagccatcgg 540
tggtgctaga gccggtgtct tggaaactac cttcagaacc gaaactgaaa ccgacttgtt 600
cggatgaaca gctgttttat gtggtggtgt ctgcgctttg atgcaggccg gttttgaaac 660
cttgggtgaa gccggttacg acccaagaaa cgttacttc gaatgtatcc acgaaatgaa 720
gttgatcgtt gacttgatct accaatctgg ttctccggg atgcgttact ctatctccaa 780

cactgctgaa tacggtgact acattaccgg tccaaagatc attactgaag ataccaagaa 840
ggctatgaag aagattttgt ctgacattca agatgggtacc ttgccaagg acttcttggg 900
tgacatgtct gatgctggtt cccaggtcca cttcaaggct atgagaaagt tggcctccga 960

acaccagct gaagttgtcg gtgaagaaat tagatccttg tactcctggt cgcacgaaga 1020
caagttgatt aacaactgat attttcctct ggccctgcag gcctatcaag tgctggaaac 1080
tttttctctt ggaatttttg caacatcaag tcatagtcaa ttgaattgac ccaatttcac 1140
atttaagatt tttttttttt catccgacat acatctgtac actaggaagc cctgtttttc 1200

tgaagcagct tcaaatatat atatttttta catatttatt atgattcaat gaacaatcta 1260
attaaatcga aaacaagaac cgaaacgcga ataaataatt tatttagatg gtgacaagtg 1320
tataagtcct catcgggaca gctacgattt ctcttttcggt tttggctgag ctactggttg 1380
ctgtgacgca gcggcattag gcggcggtta tgagctaccc tcgtggcctg aaagatggcg 1440
ggaataaagc ggaactaaaa attactgact gagccatatt gaggtcaatt tgtcaactcg 1500
tcaagtcacg ttgggtggac ggcccccttc caacgaatcg tatatactaa catgcgcgcg 1560
cttctatat acacatatat atatatatat atatatatat gtgtgcgtgt atgtgtacac 1620

ctgtatttaa tticcttact cgcgggtttt tcttttttct caattcttgg cticctcttt 1680
ctcgagcgga ccggatcctc cgcggtgccg gcagatctat ttaaatggcg cgccgacgtc 1740
aggtggcact ttccggggaa atgtgcgcgg aaccctatt tgtttatttt tctaaataca 1800
ttcaaatatg tatccgctca tgagacaata accctgataa atgcttcaat aatattgaaa 1860
aaggaagagt atgagtattc aacatttccg tgcgccctt attccctttt ttgcggcatt 1920
ttgccttctt gtttttgctc acccagaaac gctggtgaaa gtaaaagatg ctgaagatca 1980
gttgggtgca cgagtgggtt acatcgaact ggatctcaac agcggtaaga tccttgagag 2040

ttttcgcgcc gaagaacgtt ttccaatgat gagcactttt aaagtcttgc tatgtggcgc 2100
ggtattatcc cgtattgacg ccgggcaaga gcaactcggg cgccgcatac actattctca 2160
gaatgacttg gttgagtact caccagtcac agaaaagcat cttacggatg gcatgacagt 2220
aagagaatta tgcagtgtcg ccataacat gagtgataac actgcggcca acttacttct 2280
gacaacgacg ggaggaccga aggagctaac cgcttttttg cacaacatgg gggatcatgt 2340
aactcgcctt gatcgttggg aaccggagct gaatgaagcc ataccaaagc acgagcgtga 2400
caccagatg cctgtagcaa tggcaacaac gttgcgcaaa ctattaactg gcgaactact 2460

tactctagct tcccggcaac aattaataga ctggatggag gcggataaag ttgcaggacc 2520
acttctgcgc tcggcccttc cggtggctg gtttattgct gataaatctg gagccgtga 2580
gcgtgggtct cgcggtatca ttgcagcact ggggccagat ggtaagccct cccgtatcgt 2640
agttatctac acgacgggga gtcaggcaac tatggatgaa cgaaatagac agatcgtga 2700
gataggtgcc tcaactgatta agcatttggt actgtcagac caagtttact catatatact 2760
ttagattgat ttaaaacttc atttttaatt taaaaggatc taggtgaaga tcctttttga 2820

taatctcatg accaaaaatcc cttaacgtga gttttcgttc cactgagcgt cagaccccgt	2880
agaaaagatc aaaggatctt cttgagatcc tttttttctg cgcgtaatct gctgcttgca	2940
aacaaaaaaa ccaccgtac cagcgggtgt ttttttgccg gatcaagagc taccaactct	3000
ttttccgaag gtaactggct tcagcagagc gcagatacca aatactgttc ttctagtgt	3060
gccgtagtta ggccaccact tcaagaactc ttagcaccg cctacatacc tcgctctgct	3120
aatcctgtta ccagtggctg ctgccagtgg cgataagtgc tgtcttacg ggttgactc	3180
aagacgatag ttaccggata aggcgcagcg gtcgggctga acggggggtt cgtgcacaca	3240
gcccagcttg gagcgaacga cctacaccga actgagatac ctacagcgtg agctatgaga	3300
aagcgccacg ctccccgaag ggagaaaggc ggacaggtat ccgtaagcg gcagggtcgg	3360
aacaggagag cgcacgaggg agcttccagg gggaaacgcc tggatatctt atagtctgt	3420
cgggtttcgc cacctctgac ttgagcgtcg atttttgtga tgcctgtcag gggggcggag	3480
cctatggaaa aacgccagca acgcggcctt ttacggttc ctggcctttt gctggccttt	3540
tgtcacatg ttctttcctg cgttatcccc tgattctgtg gataaccgta ttaccgcctt	3600
tgagtgagct gataccgtc gccgcagccg aacgaccgag cgcagcaggt cagtgagcga	3660
ggaagcggaa gagcgcccaa tacgcaaacc gcctctcccc gcgcgttggc cgattcatta	3720
atgcagctgg cagcacaggt tccccgactg gaaagcgggc agtgagcgca acgcaattaa	3780
tgtgagttag ctactcatt aggcacccca ggctttacac tttatgttc cggctcgtat	3840
gttgtgtgga attgtgagcg gataacaatt tcacacagga aacagctatg accatgatta	3900
cgccaagctt tttttttcca atttttttt ttctgtcatt ataaaaatca ttacgaccga	3960
gattccccgg taataactga tataattaa ttgaagctct aatttgtgag tttagtatac	4020
atgcatttac ttataataga gttttttagt ttgtctggcc gcattctctc aaatatgctt	4080
cccagcctgc ttttctgtaa cgttcacctt ctaccttagc atcccttccc ttgcaaata	4140
gtcctcttcc aacaataata atgtcagatc ctgtagagac cacatcatcc acggttctat	4200
actgttgacc caatgcgtct ccttgtcat ctaaaccac accgggtgtc ataatacaac	4260
aatcgtaacc ttcatctctt ccaccatgt ctctttgagc aataaagcgg ataacaaaat	4320
ctttgtcgt cttcgcaatg tcaacagtac ccttagtata ttctccagta gataggagc	4380
ccttgatga caattctgt aacatcaaaa ggctctagg ttctttgtt acttctctg	4440
ccgcctgctt caaacgcta acaatacctg ggcccaccac accgtgtgca ttcgtaatgt	4500
ctgccattc tgctattctg tatacaccg cagagtactg caatttgact gtattacaa	4560

tgtagcaaaa ttttctgtct tcgaagagta aaaaattgta cttggcggat aatgccttta	4620
gcggcttaac tgtgccctcc atggaaaaat cagtcaagat atccacatgt gtttttagta	4680
aacaaatfff gggacctaat gcttcaacta actccagtaa ttccttggtg gtacgaacat	4740
ccaatgaagc acacaagttt gtttgctttt cgtgcatgat attaaatagc ttggcagcaa	4800
caggactagg atgagtagca gcacgttcct tatatgtagc tttcgacatg atttatcttc	4860
gtttcctgca ggtttttgtt ctgtgcagtt gggttaagaa tactgggcaa tttcatgttt	4920
cttcaacact acatatgcgt atatatacca atctaagtct gtgctccttc cttcgttctt	4980
ccttctgttc ggagattacc gaatcaaaaa aatttcaagg aaaccgaaat caaaaaaag	5040
aataaaaaaa aaatgatgaa ttgaaaagct tgcatgcctg caggtcgact ctagtatact	5100
ccgtctactg tacgatacac ttccgctcag gtccttgtcc tttaacgagg cttaccact	5160
cttttgttac tctattgac cagctcagca aaggcagtg gatctaagat tctatcttcg	5220
cgatgtagta aaactagcta gaccgagaaa gagactagaa atgcaaaagg cacttctaca	5280
atggctgcca tcattattat ccgatgtgac gctgcatttt tttttttttt tttttttttt	5340
tttttttttt tttttttttt ttttttttgt acaaatatca taaaaaaga gaatcttttt	5400
aagcaaggat tttcttaact tcttcggcga cagcatcacc gacttcggtg gtactgttgg	5460
aaccacctaa atcaccagtt ctgatactg catccaaaac ctttttaact gcactttcaa	5520
tggttttacc ttttcaggc aagtccaatg acaatttcaa catcattgca gcagacaaga	5580
tagtggcgtat agggttgacc ttattctttg gcaaactctgg agcggaaacca tggcatggtt	5640
cgtacaaacc aaatgcggtg ttcttgtctg gcaaagaggc caaggacgca gatggcaaca	5700
aacccaagga gcctgggata acggaggcct catcggagat gatatacca aacatgttgc	5760
tgggtgattat aataccattt aggtgggttg ggttcttaac taggatcatg gcggcagaat	5820
caatcaattg atgttgaact ttcaatgtag ggaattcggt cttgatgggt tcctccacag	5880
tttttctcca taatcttgaa gaggccaaaa cattagcttt atccaaggac caaataggca	5940
atggtggctc atgtttagg gccatgaaag cgccattct tgtgattctt tgcacttctg	6000
gaacggtgta ttgttacta tccaagcga caccatcacc atcgtcttcc tttctcttac	6060
caaagtaaat acctccact aattctctaa caacaacgaa gtcagtacct ttagcaaatt	6120
gtggcttgat tggagataag tctaaaagag agtcggatgc aaagttacat ggtcttaagt	6180
tggcgtacaa ttgaagtctt ttacggattt ttagtaaacc ttgttcaggt ctaaacctac	6240
cggtacccca tttaggacca cccacagcac ctaacaaaac ggcatcagcc ttcttgaggg	6300
cttcagcgc ctcatctgga agtggaaacac ctgtagcatc gatagcagca ccaccaatta	6360
aatgattttc gaaatcgaac ttgacattgg aacgaacatc agaaatagct ttaagaacct	6420

taatggcttc ggctgtgatt tcttgaccaa cgtggtcacc tggcaaaacg acgatcttct	6480
taggggcaga cattacaatg gtatatcctt gaaatatata taaaaaaaaa aaaaaaaaaa	6540
aaaaaaaaaa atgcagcttc tcaatgatat tcgaatacgc tttgaggaga tacagcctaa	6600
tatccgacaa acigttttac agatttacga tcgtacttgt tacccatcat tgaattttga	6660
acatccgaac ctgggagttt tccctgaaac agatagtata tttgaacctg tataataata	6720
tatagtctag cgctttacgg aagacaatgt atgtatttcg gttcctggag aaactattgc	6780
atctattgca taggtaatct tgcacgtcgc atccccggtt cattttctgc gtttccatct	6840
tgcacttcaa tagcatactt ttgttaacga agcatctgtg cttcattttg tagaacaaaa	6900
atgcaacgcg agagcgctaa tttttcaaac aaagaatctg agctgcattt ttacagaaca	6960
gaaatgcaac gcgaaagcgc tattttacca acgaagaatc tgtgcttcat ttttgtaaaa	7020
caaaaatgca acgcgagagc gctaattttt caaacaaga atctgagctg catttttaca	7080
gaacagaaat gcaacgcgag agcgctatit taccaacaaa gaatctatac ttcttttttg	7140
ttctacaaaa atgcatcccg agagcgctat ttttctaaca aagcatctta gattactttt	7200
tttctccttt gtgcgtctta taatgcagtc tcttgataac tttttgact gtaggtccgt	7260
taagggtaga agaaggctac tttggtgtct attttctctt ccataaaaaa agcctgactc	7320
catttccgcg gtttactgat tactagcgaa gctgcgggtg cattttttca agataaaggc	7380
atccccgatt atattctata ccgatgtgga ttgcgcatac tttgtgaaca gaaagtata	7440
gcgttgatga ttcttcattg gtcagaaaat tatgaacggt ttcttctatt ttgtctctat	7500
atactacgta taggaaatgt ttacattttc gtattgtttt cgattcactc tatgaatagt	7560
tcttactaca atttttttgt ctaaagagta atactagaga taaacataaa aaatgtagag	7620
gtcgagttta gatgcaagtt caaggagcga aaggtggatg ggtaggttat atagggatat	7680
agcacagaga tatatagcaa agagatactt ttgagcaatg tttgtggaag cggtatctgc	7740
aataatttag tagctcgta cagtccggtg cgtttttggg tttttgaaag tgcgtcttca	7800
gagcgctttt ggttttcaaa agcgctctga agttcctata ctttctagag aataggaact	7860
tcggaatagg aacttcaaag cgtttccgaa aacgagcgct tccgaaaatg caacgcgagc	7920
tgcgcacata cagctcactg ttcacgtcgc acctatatct gcgtgttgcc tgtatatata	7980
tatacatgag aagaacggca tagtgctgtt ttatgcttaa atgcgtactt atatgcgtct	8040
atttatgtag gatgaaaggt agtctagtac ctctgtgat attatcccat tccatgcggg	8100
glatcgtatg cttccttcag cactaccctt tagctgttct atatgctgcc actcctcaat	8160
tggattagtc tcatccttca atgctatcat ttcccttgat attggatcat atgcatagta	8220
ccgagaaact agaggatctc ccattaccga catttgggcg ctatcgtgc atatgttcat	8280

glatgtatct gtatttaaaa cacttttgta ttatttttcc tcatatatgt gtataggttt 8340

atacggatga ttttaattatt acttcaccac cctttatttc aggctgatat cttagccttg 8400

ttactagtca ccggtggcgg ccgcacctgg taaaacctct agtggagtag tagatgtaat 8460

caatgaagcg gaagccaaaa gaccagagta gaggcctata gaagaaactg cgataccttt 8520

tgtgatggct aaacaaacag acatcttttt atatgttttt acttctgtat atcgtgaagt 8580

agtaagtgat aagcgaattt ggctaagaac gttgtaagtg aacaaggac ctcttttgcc 8640

tttcaaaaaa ggattaaatg gagttaatca ttgagattta gttttcgta gattctgtat 8700

ccctaaataa ctcccttacc cgacgggaag gcacaaaaga cttgaataat agcaaacggc 8760

cagtagccaa gaccaaataa tactagagtt aactgatggc cttaaacagg cattacgtgg 8820

tgaactccaa gaccaatata caaaatatcg ataagttatt cttgccacc aatttaagga 8880

gcctacatca ggacagtagt accattcttc agagaagagg tatacataac aagaaaatcg 8940

cgtgaacacc ttatataact tagcccgta ttgagctaaa aaaccttgca aaatttccta 9000

tgaataagaa tacttcagac gtgataaaaa ttactttct aactcttctc acgtgcccc 9060

tatctgttct tccgctctac cgtgagaaat aaagcatcga gtacggcagt tcgctgtcac 9120

tgaactaaaa caataaggct agttcgaatg atgaacttgc ttgctgtcaa acttctgagt 9180

tgccgtgat gtgacactgt gacaataaat tcaaaccggt tatagcggtc tcctccgta 9240

ccggttctgc cactccaat agagctcagt aggagtcaga acctctgcgg tggctgtcag 9300

tgactcatcc gcgtttcgta agttgtgcgc gtgcacattt cggccgttcc cgctcatctt 9360

gcagcaggcg gaaattttca tcacgtgta ggacgcaaaa aaaaaataat taatcgtaca 9420

agaatcttgg aaaaaaatt gaaaaatttt gtataaaagg gatgacctaa cttgactcaa 9480

tggcttttac acccagtatt ttccctttcc ttgtttgtta caattataga agcaagacaa 9540

aaacatatag acaacctatt cctaggagtt atattttttt acctaccag caatataagt 9600

aaaaaactgt tt 9612

<210> 182

<211> 12896

<212> DNA

<213> artificial sequence

<220><223> pBP915 plasmid

<400> 182

tcgcgcgttt cggatgatgac ggtgaaaacc tctgacacat gcagctcccg gagacggtea 60

cagcttgtct gtaagcggat gccgggagca gacaagcccg tcagggcgcg tcagcgggtg 120

ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc	180
accataaatt cccgttttaa gagcttgggtg agcgctagga gtcactgccca ggtatcgttt	240
gaacacggca ttagtcaggg aagtcataac acagtccttt cccgcaattt tctttttcta	300
ttactcttgg cctcctctag tacactctat atttttttat gcctcggtaa tgattttcat	360
tttttttttt ccacctagcg gatgactctt tttttttctt agcgattggc attatcacat	420
aatgaattat acattatata aagtaatgtg atttcttcga agaataact aaaaaatgag	480
caggcaagat aaacgaaggc aaagatgaca gagcagaaag ccctagtaaa gcgtattaca	540
aatgaaacca agattcagat tgcgatctct ttaaagggtg gtcccctagc gatagagcac	600
tcgatcttcc cagaaaaaga ggcagaagca gtagcagaac aggccacaca atcgcaagtg	660
attaacgtcc acacaggtat agggttttctg gaccatatga tacatgctct ggccaagcat	720
tccggtgggt cgtaatcgt tgagtgcatt ggtgacttac acatagacga ccatcacacc	780
actgaagact gcgggattgc tctcggtaaa gcttttaaag aggccttagg ggccgtgcgt	840
ggagtaaaaa ggtttggatc aggatttgcg cctttggatg aggcactttc cagagcgggtg	900
gtagatcttt cgaacaggcc gtacgcagtt gtcgaacttg gtttgcaaag ggagaaagta	960
ggagatctct ctgcgagat gatcccgcat tttcttgaag gctttgcaga ggctagcaga	1020
attaccctcc acgttgattg tctgcgagc aagaatgatc atcaccgtag tgagagtgcg	1080
ttcaaggctc ttgcggttgc cataagagaa gccacctcgc ccaatggtac caacgatgtt	1140
ccctccacca aaggtgttct tatgtagtga caccgattat ttaaagctgc agcatagat	1200
atataacat gtgtatatat gtataacctat gaatgtcagt aagtatgtat acgaacagta	1260
tgatactgaa gatgacaagg taatgcatca ttctatacgt gtcattctga acgaggcgcg	1320
ctttcctttt ttctttttgc tttttctttt tttttctctt gaactcgacg gatctatgcg	1380
gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcaggaaat tgtaagcgtt	1440
aataatttgt taaaattcgc gttaaatttt tgttaaatca gtcattttt taaccaatag	1500
gccgaaatcg gcaaaatccc ttataaatca aaagaataga ccgagatagg gttgagtgtt	1560
gttccagttt ggaacaagag tccactatta aagaacgtgg actccaacgt caaaggcgca	1620
aaaaccgtct atcaggcgca tggcccacta cgtgaacat caccctaact aagttttttg	1680
gggtcgaggt gccgtaaagc actaaatcgg aacctaaag ggagccccg atttagagct	1740
tgacggggaa agccggcgaa cgtggcgaga aaggaaggga agaaagcgaa aggagcgggc	1800
gctagggcgc tggcaagtgt agcggtcacg ctgcgcgtaa ccaccacacc cgccgcgctt	1860
aatgcgccgc tacaggcgcc gtccattcgc cattcaggct gcgcaactgt tgggaagggc	1920

gcggtgcggg ccctttcgct attacgccag ctggcgaaag ggggatgtgc tgcaaggcga 1980

ttaagttggg taacgccagg gttttcccag tcacgacgtt gtaaaacgac ggccagttag 2040

cgcgcgtaat acgactcact atagggcgaa ttgggtaccg ggccccccct cgaggtcgac 2100

ggcgcgccac tggtagagag cgactttgta tgccccaatt gcgaaaccg cgataatcctt 2160

ctcgattcct tagtaccga ccaggacaag gaaaaggagg tcgaaacgtt tttgaagaaa 2220

caagaggaac tacacggaag ctctaagat ggcaaccagc cagaaactaa gaaaatgaag 2280

ttgatggatc caactggcac cgctggcttg aacaacaata ccagccttc aacttctgta 2340

aataacggcg gtacgccagt gccaccagta ccgttacctt tcggtatacc tcctttcccc 2400

atgtttccaa tgccttcat gcctccaacg gctactatca caaatcctca tcaagctgac 2460

gcaagcccta agaaatgaat aacaatactg acagtactaa ataattgcct acttggttc 2520

acatacgttg catacgtcga tatagataat aatgataatg acagcaggat tatcgtaata 2580

cgtaatagct gaaaatcctca aaaatgtgtg ggtcattacg taaataatga taggaatggg 2640

attcttctat ttttctttt tccattctag cagccgtcgg gaaaacgtgg catcctctct 2700

ttcgggctca attggagtca cgctgccgtg agcatcctct cttccatat ctaacaactg 2760

agcacgtaac caatggaaaa gcatgagctt agcgttgctc caaaaaagta ttggatggtt 2820

aataccattt gtctgttctc ttctgacttt gactcctcaa aaaaaaaat ctacaatcaa 2880

cagatcgctt caattacgcc ctacaaaaa ctttttctt tcttcttcgc ccacgttaaa 2940

ttttatccct catgttgtct aacggatttc tgcacttgat ttattataaa aagacaaaga 3000

cataaactt ctctatcaat ttacgttatt gttcttctt gcgttattct tctgttcttc 3060

tttttcttt gtcatatata accataacca agtaatacat attcaacta gtatgactga 3120

caaaaaaact cttaaagact taagaaatcg tagttctgtt tacgattcaa tggtaaatac 3180

acctaactgt gctatgttgc gtgcaactgg tatgcaagat gaagactttg aaaaacctat 3240

cgtcggtgtc atttcaactt gggctgaaaa cacacctgt aatatccact tacatgactt 3300

tggtaacta gccaaagtgc gtgttaagga agctgggtgt tggccagttc agttcggaac 3360

aatcacggtt tctgatggaa tcgcatggg aaccaagga atgcgtttct ccttgacatc 3420

tcgtgatatt attgcagatt ctattgaagc agccatggga ggtcataatg cggatgcttt 3480

tgtagccatt ggcggttgtg ataaaaacat gcccggttct gttatcgta tggctaact 3540

ggatatccca gccatttttg cttacggcgg aacaattgca cctggtaatt tagacggcaa 3600

agatatcgat ttagtctctg tctttgaagg tgcggccat tggaaccacg gcgatatgac 3660

caaagaagaa gttaaagctt tggaatgtaa tgcttgctcc ggtcctggag gctgcggtgg 3720
 tatgtatact gctaacacaa tggcgacagc tattgaagtt ttgggactta gccttccggg 3780
 ttcatcttct cacccggtcg aatccgcaga aaagaaagca gatattgaag aagctggtcg 3840
 cgctgttgctc aaaatgctcg aaatgggctt aaaaccttct gacattttaa cgcgtgaagc 3900
 ttttgaagat gctattactg taactatggc tctgggaggt tcaaccaact caaccttca 3960
 cctcttagct attgcccatt ctgctaattg ggaattgaca cttgatgatt tcaatacttt 4020
 ccaagaaaaa gticctcatt tggctgattt gaaaccttct ggtcaatatg tattccaaga 4080

 cctttacaag gtcggagggg taccagcagt tatgaaatat ctcttataaa atggcttctt 4140
 tcatggtgac cgtatcactt gtactggcaa aacagtcgct gaaaatttga aggcttttga 4200
 tgatttaaca cctgggtcaa aggttattat gccgcttgaa aatcctaacc gtgaagatgg 4260
 tccgctcatt attctccatg gtaacttggc tccagacggt gccgttgcca aagtttctgg 4320
 tgtaaaagtg cgtcgtcatg tccgtcctgc taaggtcttt aattctgaag aagaagccat 4380
 tgaagctgct ttgaatgatg atattgttga tggatgatgt gttgtcgtac gttttgtagg 4440
 accaaagggc ggtcctggta tgctgaaat gctttccctt tcatcaatga ttgttggtaa 4500

 agggcaaggt gaaaaagttg ccttctgac agatggccgc ttctcaggtg gtacttatgg 4560
 tcttgcgtg ggtcatatcg ctctgaagc acaagatggc ggtccaatcg cctacctgca 4620
 aacaggagac atagtacta ttgaccaaga cactaaggaa ttacactttg atatctccga 4680
 tgaagagtta aaacatcgtc aagagacat tgaattgcca ccgctctatt cacgcggtat 4740
 ccttggtaaa tatgtcaca tcgtttctgc tgcttctagg ggagccgtaa cagacttttg 4800
 gaagcctgaa gaaactggca aaaaatgttg tctgggtgc tgtgggttaag cggccgcgtt 4860
 aattcaaatt aattgatata gttttttaat gattattgaa tctgtttaga aataatggaa 4920

 tattattttt attatattat ttatattatt ggtcggctct tttcttctga aggtcaatga 4980
 caaatgata tgaaggaaat aatgatttct aaaattttac aacgtaagat atttttacaa 5040
 aagcctagct catcttttgt catgcactat ttactcacg cttgaaatta acggccagtc 5100
 cactgcggag tcatttcaaa gtatccctaa tcgatctatc gtttttgata gctcattttg 5160
 gagttcgcga ttgtcttctg ttattcacia ctgttttaac ttttatttca ttctggaact 5220
 ctctgagttc tttgtaaagt ctttcatagt agcttacttt atctccaac atatttaact 5280
 tcatgtcaat ttcggctctt aaattttcca catcatcaag ttcaacatca tcttttaact 5340

 tgaatttatt ctctagctct tccaaccaag cctcattgct ccttgattta ctggtgaaaa 5400
 gtgatacact ttgcgcgcaa tccagggtcaa aactttctcg caaagaattc accaatttct 5460
 cgacatcata gtacaatttg tttgtttctc ccatcacaat ttaatatacc tgatggattc 5520

ttatgaagcg ctgggtaatg gacgtgtcac tctacttcgc ctttttcctt actcctttta	5580
gtacggaaga caatgctaataaataagagg gtaataataa tattattaat cggcaaaaaa	5640
gattaaacgc caagcgttta attatcagaa agcaaacgtc gtaccaatcc ttgaatgctt	5700
cccaattgta tattaagagt catcacagca acataattctt gttattaaat taattattat	5760
tgatttttga tattgtataa aaaaaccaa tatgtataaa aaaagtgaat aaaaaataacc	5820
aagtatggag aaatatatta gaagtctata cgtaaacca cccgggcccc ccctcgaggt	5880
cgacggtatc gataagcttg atatcgaatt cctgcagccc gggggatcca ctagtcttag	5940
agcggccgct ctagaactag taccacaggt gttgtcctct gaggacataa aatacacacc	6000
gagattcatc aactcattgc tggagttagc atatctacaa ttgggtgaaa tggggagcga	6060
tttgcaggca ttgtctcggc atgccggtag aggtgtggtc aataagagcg acctcatgct	6120
atactgaga aagcaacctg acctacagga aagagttact caagaataag aattttcgtt	6180
ttaaaaccta agagtcactt taaaatttgt atacacttat tttttttata acttatttaa	6240
taataaaaaat cataaatcat aagaaattcg ctactctta attaatcaaa aagttaaaat	6300
tgtacgaata gattcaccac ttcttaacaa atcaaacctt tcattgattt tctcgaatgg	6360
caatacatgt gtaattaaag gatcaagagc aaacttcttc gccataaagt cggcaacaag	6420
ttttggaaca ctatccttgc tcttaaaacc gccaaatata gctcccttcc atgtacgacc	6480
gcttagcaac agcataggat tcatcgacaa attttgtgaa tcaggaggaa cacctacgat	6540
cacactgact ccatatgcct cttgacagca ggacaacgca gttacatag tatcaagacg	6600
gcctataact tcaaaagaga aatcaactcc accgtttgac atttcagtaa ggacttcttg	6660
tattggtttc ttataatctt gaggtttaac acattcagta gccccgacct ccttagcttt	6720
tgcaaatattg tctttattga tgtctacacc tataatcttc gctgcgcctg cagctttaca	6780
ccccataata acgcttagtc ctactcctcc taaaccgaat actgcacaag tcgaacctg	6840
tgtaaccttt gcaactttta ctgcggaacc gtaaccggtg gaaaatccgc accctatcaa	6900
gcaaactttt tccagtggtg aagctgcac gatttttagcg acagatatct cgtccaccac	6960
tgtgtattgg gaaaatgtag aagtaccaag gaaatggtgt ataggtttcc ctctgcatgt	7020
aaatctgctt gtaccatcct gcatagtacc tctaggcata gacaaatcat ttttaaggca	7080
gaaattaccc tcaggatgtt tgcagactct acacttacca cattgaggag tgaacagtgg	7140
gatcacttta tcaccaggac gaacagtggg aacaccttca cctatggatt caacgattcc	7200
ggcagcctcg tgtcccgcga ttactggcaa aggagtaact agagtgcac tcaccacatg	7260
gtcgtcggat ctacagattc cgggtggcaac catcttgatt ctaacctcgt gtgcttttgg	7320
tggcgctact tctacttctt ctatgctaaa cggctttttc tcttcccaca aaactgccgc	7380

tttacactta ataactttac cggtgttga catcctcagc tagctattgt aatatgtgtg	7440
tttgtttgga ttattaagaa gaataattac aaaaaaatt acaaaggaag gtaattacaa	7500
cagaattaag aaaggacaag aaggaggaag agaatacagtt cattatttct tctttgttat	7560
ataacaaacc caagtagcga tttggccata cattaaaagt tgagaaccac cctccctggc	7620
aacagccaca actcgttacc attgttcac acgatcatga aactcgtgt cagctgaaat	7680
ttcacctcag tggatctctc tttttattct tcatcgttcc actaaccttt ttccatcagc	7740
tggcagggaa cggaaagtg aatcccatit agcgagcttc ctcttttctt caagaaaaga	7800
cgaagcttgt gtgtgggtgc gcgcgctagt atctttccac attaagaaat ataccataaa	7860
ggttacttag acatcactat ggctatatat atatatatat atatatgtaa cttagcacca	7920
tgcgcgtgc atcactgcat gtgttaaccg aaaagtttgg cgaacacttc accgacacgg	7980
tcatttagat ctgtcgtctg cattgcacgt ccttagcct taaatcctag gcgggagcat	8040
tctcgtgtaa ttgtgcagcc tgcgtagcaa ctcaacatag cgtagtctac ccagtttttc	8100
aagggtttat cgttagaaga ttctcccttt tcttctgct cacaatctt aaagtcatac	8160
attgcacgac taaatgaag catgcggatc ccccggtctg caggaattcg atatcaagct	8220
tatcgatacc gtcgactggc cattaatctt tcccatatta gatttcgcca agccatgaaa	8280
gttcaagaaa ggtctttaga cgaattaccc ttcatttctc aaactggcgt caagggatcc	8340
tggatatggtt ttatcgtttt atttctgggt cttatagcat cgttttggac ttctctgttc	8400
ccattaggcg gttcaggagc cagcgcagaa tcattctttg aaggatactt atcctttcca	8460
attttgattg tctgttacgt tggacataaa ctgtatacta gaaattggac ttigtatggtg	8520
aaactagaag atatggatct tgataccggc agaaaacaag tagatttgac tcttcgtagg	8580
gaagaaatga ggattgagcg agaaacatta gcaaaaagat ccttcgtaac aagattttta	8640
catttctggt gttgaaggga aagatatgag ctatacagcg gaatttccat atcactcaga	8700
ttttgtatc taatttttct cttcccacgt ccgcgggaat ctgtgtatat tactgcatct	8760
agatatatgt tatcttatct tggcgcgtac atttaatttt caacgtattc tataagaaat	8820
tgcgggagtt ttttcatgt agatgatact gactgcacgc aaatatagcg atgatttata	8880
ggcatgattt gatggctgta ccgataggaa cgctaagagt aacttcagaa tcgttatcct	8940
ggcggaaaaa attcatttgt aaactttaaa aaaaaagcc aatatccca aaattattaa	9000
gagcgctcc attattaact aaaatttcac tcagcatcca caatgtatca ggtatctact	9060
acagatatta catgtggcga aaaagacaag aacaatgcaa tagcgcatca agaaaaaca	9120

caaagctttc aatcaatgaa tcgaaaatgt cattaaaata gtatataaat tgaaactaag	9180
tcataaagct ataaaaagaa aattttattta aatcttggct ctcttgggct caaggtgaca	9240
aggtcctcga aaatagggcg cgccccaccg cggtaggagct ccagcttttg ttcccttttag	9300
tgagggttaa ttgcgcgctt ggcgtaatca tggatcatagc tgtttcctgt gtgaaattgt	9360
tatccgctca caattccaca caacatacga gccggaagca taaagtgtaa agcctggggt	9420
gcctaagtag tgagctaact cacattaatt gcgttgcgct cactgcccgc ttccagtcg	9480
ggaaacctgt cgtgccagct gcattaatga atcgccaac gcgcggggag aggcggtttg	9540
cgtattgggc gctcttccgc ttctcgtc actgactcgc tgcgtcgggt cgttcggctg	9600
cggcgagcgg tatcagctca ctcaaaggcg gtaatacgggt tatccacaga atcaggggat	9660
aacgcaggaa agaacatgtg agcaaaaggc cagcaaaagg ccaggaaccg taaaaaggcc	9720
gcgttctggt cgtttttcca taggctccgc cccctgacg agcatcaca aaatcgacgc	9780
tcaagtcaga ggtggcgaaa cccgacagga ctataaagat accaggcgtt tccccctgga	9840
agctccctcg tgcgtctcc tgttccgacc ctgccgtta ccggatacct gtccgccttt	9900
ctcccttcgg gaagcgtggc gctttctcat agctcacgt gtaggtatct cagttcgggtg	9960
taggtcgttc gctccaagct gggctgtgtg cacgaacccc ccgttcagcc cgaccgtgc	10020
gccttatccg gtaactatcg tcttgagtcc aaccggtaa gacacgactt atcgccactg	10080
gcagcagcca ctggtaacag gattagcaga gcgaggtatg taggcgggtgc tacagagttc	10140
ttgaagtgtt ggcctaacta cggctacact agaagaacag tatttggtat ctgcgtctg	10200
ctgaagccag ttaccttcgg aaaaagagtt gtagctctt gatccggcaa acaaaccacc	10260
gtggtagcgt gtggtttttt tgtttgcaag cagcagatta cgcgcagaaa aaaaggatct	10320
caagaagatc ctttgatctt ttctacgggg tctgacgtc agtggaacga aaactcacgt	10380
taagggtatt tggtcatgag attatcaaaa aggatcttca cctagatcct tttaaattaa	10440
aatgaagtt ttaaatcaat ctaaagtata tatgagtaaa cttggtctga cagttaccaa	10500
tgttaatca gtgaggcacc tatctcagcg atctgtctat ttcttctatc catagtggcc	10560
tgactccccg tctgttagat aactacgata cgggagggtt taccatctgg cccagtgct	10620
gcaatgatac cgcgagacc acgctcaccg gctccagatt tatcagcaat aaaccagcca	10680
gccggaaggg ccgagcgag aagtggctct gcaactttat ccgcctccat ccagtctatt	10740
aattgttgcc gggaagctag agtaagtagt tcgccagtta atagtgtgc caacgttgtt	10800
gccattgcta caggcatcgt ggtgtcacgc tcgtcgtttg gtatggcttc attcagctcc	10860
ggttcccaac gatcaaggcg agttacatga tccccatgt tgtgcaaaaa agcggtttagc	10920
tccttcggtc ctccgatcgt tgtcagaagt aagttggccg cagtgttatc actcatggtt	10980

atggcagcac tgcataattc tcttactgtc atgcatccg taagatgctt ttctgtgact	11040
gggtgagtact caaccaagtc attctgagaa tagtgtatgc ggcgaccgag ttgctcttgc	11100
ccggcgtcaa tacgggataa taccgcgcca catagcagaa ctttaaaagt gctcatcatt	11160
ggaaaacgtt cticggggcg aaaactctca aggatcttac cgctgttgag atccagttcg	11220
atgtaacca ctctgcacc caactgatct tcagcatctt ttactttcac cagcgtttct	11280
gggtgagcaa aaacaggaag gcaaaatgcc gcaaaaaagg gaataagggc gacacggaaa	11340
tgttgaatac tcatactctt cctttttcaa tattattgaa gcatttatca gggttattgt	11400
ctcatgagcg gatacatatt tgaatgtatt tagaaaaata aacaaatagg ggttcgcgc	11460
acatttcccc gaaaagtgcc acctgaacga agcatctgtg cttcattttg tagaacaaaa	11520
atgcaacgcg agagcgctaa tttttcaaac aaagaatctg agctgcattt ttacagaaca	11580
gaaatgcaac gcgaaagcg tattttacca acgaagaatc tgtgcttcat ttttgtaaaa	11640
caaaaatgca acgcgagagc gctaattttt caaacaaga atctgagctg catttttaca	11700
gaacagaaat gcaacgcgag agcgctatct taccaacaaa gaatctatac ttcttttttg	11760
ttctacaaaa atgcatcccg agagcgctat ttttcaaca aagcatctta gattactttt	11820
tttctccttt gtgcgctcta taatgcagtc tcttgataac ttttgcact gtaggtccgt	11880
taagggtaga agaaggctac tttggtgtct attttctctt ccataaaaaa agcctgactc	11940
cacttcccgc gtttactgat tactagcgaa gctgcgggtg cattttttca agataaaggc	12000
atccccgatt atattctata ccgatgtgga ttgcgcatac tttgtgaaca gaaagtata	12060
gcgttgatga ttcttcattg gtcagaaaat tatgaacggt ttcttctatt ttgtctctat	12120
atactacgta taggaaatgt ttacattttc gtattgtttt cgattcactc tatgaatagt	12180
tcttactaca atttttttgt ctaaagagta atactagaga taaacataaa aaatgtagag	12240
gtcgagtita gatgcaagtt caaggagcga aagggtgatg ggtaggttat ataggatat	12300
agcacagaga tatatagcaa agagatactt ttgagcaatg tttgtggaag cggatttcgc	12360
aatattttag tagctcgtaa cagtccgtg cgtttttggt ttttgaaag tgcgtcttca	12420
gagcgctttt ggttttcaaa agcgctctga agttcctata ctttctagag aataggaact	12480
tcggaatagg aacttcaaag cgtttccgaa aacgagcgct tccgaaaatg caacgcgagc	12540
tgcgacata cagctcactg ttacgctgc accatatact gcgtgttgcc tgtatatata	12600
tatacatgag aagaacggca tagtgctgtt ttatgcttaa atgcgtactt atatgcgtct	12660
atttatgtag gatgaaaggt agtctagtag ctctgtgat attatcccat tccatgcggg	12720
gtatcgtatg cttccttcag cactaccctt tagctgttct atatgctgcc actcctcaat	12780
tggattagtc tcaccttca atgctatcat ttctttgat attggatcat actaagaaac	12840

cattattatc atgacattaa cctataaaaa taggcgtatc acgaggccct ttcgtc 12896

<210> 183

<211> 22

<212> DNA

<213> artificial sequence

<220><223> primer oBP594

<400> 183

agctgtctcg tgttgtgggt tt 22

<210> 184

<211> 49

<212> DNA

<213> artificial sequence

<220><223> primer oBP595

<400> 184

cttaataata gaacaatata atcctttacg ggcatcttat agtgtcgtt 49

<210> 185

<211> 49

<212> DNA

<213> artificial sequence

<220><223> primer oBP596

<400> 185

gcgccaacga cactataaga tgcccgtaaa ggatgatatt gttctatta 49

<210> 186

<211> 49

<212> DNA

<213> artificial sequence

<220><223> primer oBP597

<400> 186

tatggaccct gaaaccacag ccacattgca acgacgacaa tgccaaacc 49

<210> 187

<211> 49

<212> DNA

<213> artificial sequence

<220><223> primer oBP598
 <400> 187
 tccttggttt ggcattgtcg tcgttgcaat gtggctgtgg tttcagggt 49
 <210> 188
 <211> 49
 <212> DNA
 <213> artificial sequence
 <220><223> primer oBP599
 <400> 188
 atcctctcgc ggagtcctg ttcagtaaag gccatgaagc tttttcttt 49
 <210> 189
 <211> 49
 <212> DNA
 <213> artificial sequence
 <220><223> primer oBP600
 <400> 189
 attggaaga aaaagcttca tggcctttac tgaacaggga ctccgcgag 49
 <210> 190
 <211> 22
 <212> DNA
 <213> artificial sequence
 <220><223> primer oBP601
 <400> 190
 tcataccaca atcttagacc at 22
 <210> 191
 <211> 21
 <212> DNA
 <213> artificial sequence
 <220><223> primer oBP602
 <400> 191
 tgttcaaacc cctaaccaac c 21
 <210> 192
 <211> 22

<212> DNA
 <213> artificial sequence
 <220><223> primer oBP603
 <400> 192
 tgttccaca atctattacc ta 22
 <210> 193
 <211> 21
 <212> DNA
 <213> artificial sequence
 <220><223> primer oBP605
 <400> 193
 tactgaacag ggactccgcg a 21
 <210> 194
 <211> 21
 <212> DNA
 <213> artificial sequence
 <220><223> primer oBP606
 <400> 194
 tcataccaca atcttagacc a 21
 <210> 195
 <211> 11856
 <212> DNA
 <213> Artificial sequence
 <220><223> Plasmid pYZ090
 <400> 195
 tccattacc gacatttggg cgctatacgt gcatatgttc atgtatgtat ctgtatttaa 60
 aacacttttg tattatTTTT cctcatatat gtgtataggt ttatcggat gatttaatta 120
 ttacttcacc accctttatt tcagcgtgat atcttagcct tggtactagt tagaaaaaga 180
 catttttgc tgcagtcact gtcaagagat tcttttgc tgcatttcttc tagaagcaaa 240
 aagagcgtg cgtcttttcc gctgaaccgt tccagcaaaa aagactacca acgcaatatg 300
 gattgtcaga atcatataaa agagaagcaa ataactcctt gtcttgtatc aattgcatta 360
 taatatcttc ttgttagtgc aatatcatat agaagtcac gaaatagata ttaagaaaaa 420
 caaactgtac aatcaatcaa tcaatcatcg ctgaggatgt tgacaaaagc aacaaaagaa 480

caaaaatccc ttgtgaaaaa cagaggggcg gagcttggtg ttgattgctt agtggagcaa 540
 ggtgtcacac atgtatttgg cattccaggt gcaaaaattg atgcggtatt tgacgcttta 600
 caagataaag gacctgaaat tatcgttgcc cggcacgaac aaaacgcagc attcatggcc 660
 caagcagtcg gccgtttaac tggaaaaccg ggagtcgtgt tagtcacatc aggaccgggt 720
 gcctctaact tggcaacagg cctgctgaca gcgaacactg aaggagaccc tgcgttgcg 780

cttgctggaa acgtgatccg tgcagatcgt ttaaaacgga cacatcaatc ttggataat 840
 gcggcgctat tccagccgat taaaaatac agtgtagaag ttcaagatgt aaaaaatata 900
 ccggaagctg ttacaaatgc atttaggata gcgtcagcag ggcaggctgg gcccgctttt 960
 gtgagctttc cgcaagatgt tgtgaatgaa gtcacaaata cgaaaaacgt gcgtgctgtt 1020
 gcagcgccaa aactcgggcc tgcagcagat gatgcaatca gtgcggccat agcaaaaatc 1080
 caaacagcaa aacttctgt cgttttggtc ggcatgaaag gcggaagacc ggaagcaatt 1140
 aaagcgggtc gcaagctttt gaaaaagggt cagcttccat ttgttgaaac atatcaagct 1200

gccggtaccc ttctagaga tttagaggat caatatattg gccgtatcgg ttgtttccgc 1260
 aaccagcctg gcgatttact gctagagcag gcagatgttg ttctgacgat cggctatgac 1320
 ccgattgaat atgatccgaa attctggaat atcaatggag accggacaat tatccattta 1380
 gacgagatta tcgtgacat tgatcatgct taccagcctg atcttgaatt gatcggtgac 1440
 attccgtcca cgatcaatca tatcgaacac gatgctgtga aagtgggaatt tgcagagcgt 1500
 gagcagaaaa tcctttctga tttaaacaa tatatgcatg aaggtgagca ggtgcctgca 1560
 gattggaaat cagacagagc gcaccctctt gaaatcgta aagagttgcg taatgcagtc 1620

gatgatcatg ttacagtaac ttgcgatac ggcttcgcacg ccatttggat gtcacgttat 1680
 ttccgcagct acgagccgtt aacattaatg atcagtaacg gtatgcaaac actcggcggtt 1740
 gcgcttccctt gggcaatcgg cgcttcattg gtgaaaccgg gagaaaaagt ggtttctgtc 1800
 tctggtgacg gcggtttctt attctcagca atggaattag agacagcagt tcgactaaaa 1860
 gcaccaattg tacacattgt atggaacgac agcacatatg acatggttgc attccagcaa 1920
 ttgaaaaaat ataaccgtac atctcgggtc gatttcggaa atatcgatat cgtgaaatat 1980
 gcggaaaagt tcggagcaac tggcttgccg gtagaatcac cagaccagct ggcagatgtt 2040

ctgcgtcaag gcatgaacgc tgaaggctct gtcacatcg atgtcccgtg tgactacagt 2100
 gataacatta atttagcaag tgacaagctt ccgaaagaat tcggggaact catgaaaacg 2160
 aaagctctct agttaattaa tcatgtaatt agttatgtca cgcttacatt caccgcctcc 2220
 cccacatcc gctctaaccg aaaaggaagg agttagacaa cctgaagtct aggtccctat 2280
 ttatTTTTTT atagtattgt tagtattaag aacgttattt atatttcaa tttttctttt 2340

ttttctgtac agacgcgtgt acgcatgtaa cattatactg aaaaccttgc ttgagaaggt	2400
tttgggacgc tcgaaggctt taatttgcgg gcggccgcac ctggtaaaac ctctagtga	2460
gtagtagatg taatcaatga agcggaagcc aaaagaccag agtagaggcc tatagaagaa	2520
actgcgatac cttttgtgat ggctaataca acagacatct ttttatatgt ttttacttct	2580
gtatatcgtg aagtagtaag tgataagcga atttggctaa gaacgttgta agtgaacaag	2640
ggacctcttt tgcctttcaa aaaaggatta aatggagtta atcattgaga tttagttttc	2700
gttagattct gtatccctaa ataactcct taccgcacgg gaaggcaca aagacttgaa	2760
taatagcaaa cggccagttag ccaagaccaa ataatactag agttaactga tggctctaaa	2820
caggcattac gtggtgaact ccaagaccaa tatacaaat atcgataagt tattcttgcc	2880
caccaattta aggagcctac atcaggacag tagtaccatt cctcagagaa gaggtataca	2940
taacaagaaa atcgcgtaga caccttatat aacttagccc gttattgagc taaaaaacct	3000
tgcaaaatct cctatgaata agaatacttc agacgtgata aaaatttact ttctaactct	3060
tctcacgtg cccctatctg ttcttccgct ctaccgtgag aaataaagca tcgagtacgg	3120
cagttcgtg tcaactgaact aaaacaataa ggctagtctg aatgatgaac ttgcttgctg	3180
tcaaacttct gatttgccgc tgatgtgaca ctgtgacaat aaattcaaac cggttatagc	3240
ggtctcctcc ggtaccggtt ctgccacctc caatagagct cagtaggagt cagaacctct	3300
gcggtggctg tcagtgactc atccgcgttt cgtaagtgt gcgcgtgcac atttcgccc	3360
ttcccgtca tcttgacga ggcggaatt ttcatcacgc tgtaggacgc aaaaaaaaa	3420
taattaatcg tacaagaatc ttggaaaaaa aattgaaaaa tttgtataa aagggatgac	3480
ctaacttgac tcaatggctt ttacaccag tattttcct ttccttggtt gttacaatta	3540
tagaagcaag acaaaaacat atagacaacc tattcctagg agttatattt ttttaccta	3600
ccagcaatat aagtaaaaaa ctgtttaaac agtatggcag ttacaatgta ttatgaagat	3660
gatgtagaag tatcagcact tgctggaaag caaattgcag taatcggta tggttcaca	3720
ggacatgctc acgcacagaa ttgctgtgat tctggtcaca acgttatcat tgggtgctgc	3780
cacggaaaat cttttgataa agcaaaagaa gatggctttg aaacatttga agtaggagaa	3840
gcagtagcta aagctgaigt tattatggtt ttggcaccag atgaacttca acaatccatt	3900
tatgaagagg acatcaaacc aaacttgaaa gcaggttcag cacttggttt tgctcacgga	3960
tttaatatcc attttggcta tattaaagta ccagaagacg ttgacgtctt tatggttgcg	4020
cctaaggctc caggtcacct tgtccgtcgg acttatactg aaggttttgg tacaccagct	4080

ttgtttgttt cacacaaaa tgcaagtggg catgcgcgtg aaatcgcaat ggattgggcc	4140
aaaggaattg gttgtgctcg agtgggaatt attgaaacaa cttttaaga agaaacagaa	4200
gaagatttgt ttggagaaca agctgttcta tgtggagggt tgacagcact tgttgaagcc	4260
ggttttgaaa cactgacaga agctggatac gctggcgaat tggttactt tgaagttttg	4320
cacgaaatga aattgattgt tgacctcatg tatgaagggt gttttactaa aatgcgtcaa	4380
tccatctcaa atactgctga gtttggcgat tatgtgactg gtccacggat tattactgac	4440
gaagttaaaa agaatatgaa gcttgttttg gctgatattc aatctggaaa atttgcctaa	4500
gatttcgttg atgacttcaa agcggggcgt ccaaaattaa tagcctatcg cgaagctgca	4560
aaaaatcttg aaattgaaaa aattggggca gagctacgtc aagcaatgcc attcacacaa	4620
tctggatgat acgatgcctt taaaatctat cagtaaggcc ctgcaggcca gagggaaata	4680
atatcaagtg ctggaaactt tttctcttgg aatttttgca acatcaagtc atagtcaatt	4740
gaattgaccc aatttcacat ttaagatttt ttttttttca tccgacatac atctgtacac	4800
taggaagccc tgtttttctg aagcagcttc aaatatatat attttttaca tatttattat	4860
gattcaatga acaatctaata taaatcgaaa acaagaaccg aaacgcgaat aaataattta	4920
tttagatggt gacaagtgtg taagtcctca tcgggacagc tacgatttct ctttcggttt	4980
tggtgagct actggttctg gtgacgcagc ggcatatagc cggcgttatg agctaccctc	5040
gtggcctgaa agatggcggg aataaagcgg aactaaaaat tactgactga gccatattga	5100
ggtcaatttg tcaactcgtc aagtcacgtt tgggtggacgg cccctttcca acgaatcgta	5160
tatactaaca tgcgcgcgtt tcctatatac acatatacat atatatatat atatatatgt	5220
gtgcgtgtat gtgtacacct gtatttaatt tccttactcg cgggtttttc ttttttctca	5280
attcttggct tcctctttct cgagcggacc ggatcctccg cgggtccggc agatctattt	5340
aatggcgcg ccgacgtcag gtggcacttt tcggggaaat gtgcgcggaa cccctatttg	5400
tttatttttc taaatacatt caaatatgta tccgtcatg agacaataac cctgataaat	5460
gcttcaataa tattgaaaaa ggaagagtat gattattcaa catttccgtg tcgcccttat	5520
tccttttttt gcggcatttt gccttcctgt ttttgctcac ccagaaacgc tggtgaaagt	5580
aaaagatgct gaagatcagt tgggtgcacg agtgggttac atcgaaactgg atctcaacag	5640
cggtaagatc cttgagagtt ttgcctccga agaacttttt ccaatgatga gcacttttaa	5700
agttctgcta tgtggcgcgg tattatcccg tattgacgcc gggcaagagc aactcggtcg	5760
ccgcatacac tattctcaga atgacttggg tgagtactca ccagtcacag aaaagcatct	5820

tacggatggc atgacagtaa gagaattatg cagtgtgcc ataaccatga gtgataacac 5880

tgcggccaac ttacttctga caacgatcgg aggaccgaag gagctaaccg cttttttgca 5940

caacatgggg gatcatgtaa ctgccttga tcgttgggaa ccggagctga atgaagccat 6000

accaaacgac gagcgtgaca ccacgatgcc tgtagcaatg gcaacaacgt tgcgcaaact 6060

attaactggc gaactactta ctctagcttc ccggcaacaa ttaatagact ggatggaggc 6120

ggataaagt tgcaggaccac ttctgcgtc ggcccttcg gctggctggt ttattgtga 6180

taaatctgga gccggtgagc gtgggtctcg cggtatcatt gcagcactgg ggccagatgg 6240

taagccctcc cgtatcgtag ttatctacac gacggggagt caggcaacta tggatgaacg 6300

aaatagacag atcgctgaga taggtgcctc actgattaag cattggtaac tgtcagacca 6360

agtttactca tatatacttt agattgattt aaaacttcat ttttaattta aaaggatcta 6420

ggtgaagatc ctttttgata atctcatgac caaaatccct taacgtgagt ttctgttcca 6480

ctgagcgtca gaccccgtag aaaagatcaa aggatcttct tgagatcctt tttttctgcg 6540

cgtaatctgc tgcttgcaaa caaaaaacc accgtacca gcggtggttt gtttgccgga 6600

tcaagagcta ccaactcttt ttccgaaggt aactggcttc agcagagcgc agatacaaaa 6660

tactgttctt ctagtgtagc cgtagttagg ccaccacttc aagaactctg tagcaccgcc 6720

tacatacttc gctctgctaa tctgttacc agtggctgct gccagtggcg ataagtcgtg 6780

tcttaccggg ttggactcaa gacgatagtt accggataag gcgcagcggg cgggctgaac 6840

gggggggttc tgacacagc ccagcttga gcgaacgacc tacaccgaac tgagatacct 6900

acagcgtgag ctatgagaaa gcgccacgt tcccgaagg agaaaggcgg acaggtatcc 6960

ggtaaagcgc aggttcggaa caggagagcg cagcaggag cttccagggg gaaacgcctg 7020

gtatctttat agtcctgtcg ggtttcgcca cctctgactt gagcgtcgat ttttgtgatg 7080

ctcgtcaggg ggcgggagcc tatggaaaaa cgccagcaac gcggcctttt tacggttcct 7140

ggccttttgc tggccttttg ctacatgtt ctttctgcg ttatcccctg attctgtgga 7200

taaccgtatt accgcctttg agtgagctga taccgtcgc gcgagccgaa cgaccgagcg 7260

cagcagatca gtgagcagg aagcggaaga gcgcccaata cgcaaaccgc ctctccccgc 7320

gcgttggcgg attcattaat gcagctggca cgacaggttt cccgactgga aagcgggcag 7380

tgagcgaac gcaattaat tgagtttagt cactcattag gcaccccagg ctttacattt 7440

tatgcttccg gctcgtatgt tgtgtggaat tgtgagcgga taacaatttc acacaggaaa 7500

cagctatgac catgattacg ccaagctttt tctttccaat ttttttttt tcgtcattat 7560

aaaaatcatt acgaccgaga ttcccgggta ataactgata taattaaatt gaagctctaa 7620

tttgtgagtt tagtatacat gcatttactt ataatacagt ttttagttt tgctggccgc 7680

atcttctcaa atatgcttcc cagcctgcctt ttctgtaacg ttcaccctct acccttagcat	7740
cccttccctt tgcaaatagt cctcttccaa caataataat gtcagatcct gtagagacca	7800
catcatccac ggttctatac tgttgacca atgcgtctcc cttgtcatct aaaccacac	7860
cgggtgtcat aatcaaccaa tcgtaacctt catctcttcc acccatgtct ctttgagcaa	7920
taaagccgat aacaaaatct ttgtcgctct tcgcaatgtc aacagtaccc ttagtatatt	7980
ctccagtaga tagggagccc ttgcatgaca attctgctaa catcaaaagg cctctaggtt	8040
cctttgttac ttcttctgcc gcctgcttca aaccgctaac aatacctggg cccaccacac	8100
cgtgtgcatt cgtaatgtct gccattctg ctattctgta tacacccgca gagtactgca	8160
atttgactgt attaccaatg tcagcaaatt ttctgtcttc gaagagtaaa aaattgtact	8220
tggcggataa tgcctttagc ggcttaactg tgccctccat ggaaaaatca gtcaagatat	8280
ccacatgtgt ttttagtaaa caaattttgg gacctaatgc ttcaactaac tccagtaatt	8340
ccttggtggt acgaacatcc aatgaagcac acaagtttgt ttgcttttcg tgcataatat	8400
taaatagctt ggcagcaaca ggactaggat gagtagcagc acgttcttca tatgtagctt	8460
tcgacatgat ttatcttctg ttectgcagg tttttgttct gtgcagttgg gttagaata	8520
ctgggcaatt tcatgtttct tcaacactac atatgcgtat atataccaat ctaagtctgt	8580
gtccttctct tcgttcttcc ttctgttcgg agattaccga atcaaaaaaa tttcaaggaa	8640
accgaaatca aaaaaaagaa taaaaaaaaa atgatgaatt gaaaagcttg catgcctgca	8700
ggtcgactct agtatactcc gtctactgta cgatacactt ccgctcaggt ccttgcctt	8760
taacgaggcc ttaccactct tttgttactc tatlgatcca gtcagcaaa ggcagtgtga	8820
tctaagattc tatcttcgcg atgtagttaa actagctaga ccgagaaaga gactagaaat	8880
gcaaaaggca ctctacaat ggctgccatc attattatcc gatgtgacgc tgcatttttt	8940
tttttttttt tttttttttt tttttttttt tttttttttt tttttgtac aaatatcata	9000
aaaaagaga atctttttta gcaaggattt tcttaacttc ttcggcgaca gcatcaccga	9060
cttcggtggt actgttgga ccacctaaat caccagtctt gatacctgca tccaaaacct	9120
ttttaactgc atcttcaatg gctttacctt cttcaggcaa gttcaatgac aatttcaaca	9180
tcattgcagc agacaagata gtggcgatag ggttgacctt attctttggc aaatctggag	9240
cggaaccatg gcatggttcg tacaaccaa atgcggtgtt cttgtctgac aaagaggcca	9300
aggacgcaga tggcaacaaa cccaaggagc ctgggataac ggaggcttca tcggagatga	9360
tatcaccaaa catgttgctg gtgattataa taccatttag gtgggttggg ttcttaacta	9420
ggatcatggc ggcagaatca atcaattgat gttgaacttt caatgtaggg aattcgttct	9480
tgatggtttc ctccacagt ttttccata atcttgaaga ggccaaaaca ttagctttat	9540

ccaaggacca aataggcaat ggtggctcat gttgtagggc catgaaagcg gccattcttg	9600
tgattctttg cacttctgga acgggtgtatt gttcactatc ccaagcgaca ccatcaccat	9660
cgtcttcttt tctcttacc aagtaaatac ctcccactaa ttctctaaca acaacgaagt	9720
cagtaccttt agcaaatgtt ggcttgattg gagataagtc taaaagagag tcggatgcaa	9780
agttacatgg tcttaagttg gcgtacaatt gaagttcttt acggattttt agtaaacctt	9840
gttcaggctt aacactaccg gtacccatt taggaccacc cacagcacct aacaaaacgg	9900
catcagcctt ctiggaggct tccagcgcct catctggaag tggaacacct gtagcatcga	9960
tagcagcacc accaattaaa tgattttcga aatcgaactt gacattggaa cgaacatcag	10020
aaatagcttt aagaacctta atggcttcgg ctgtgatttc ttgaccaacg tggtcacctg	10080
gcaaaacgac gatcttctta ggggcagaca ttacaatggg atctccttga aatatatata	10140
aaaaaaaaa aaaaaaaaaa aaaaaaaaaat gcagcttctc aatgatattc gaatacgctt	10200
tgaggagata cagcctaata tccgacaaac tgttttacag atttacgac gtacttgta	10260
cccatcattg aattttgaac atccgaacct gggagttttc cctgaaacag atagtatatt	10320
tgaacctgta taataatata tagtctagcg ctttacggaa gacaatgtat gtatttcggt	10380
tcctggagaa actattgcat ctattgcata ggtaatcttg cagtcgcat ccccggttca	10440
ttttctgcgt ttccatcttg cacttcaata gcatacttt gttaacgaag catctgtgct	10500
tcattttgta gaacaaaaat gcaacgcgag agcgctaatt tttcaacaa agaactctgag	10560
ctgcattttt acagaacaga aatgcaacgc gaaagcgcta tttaccaac gaagaatctg	10620
tgcttcattt ttgtaaaaca aaaatgcaac gcgagagcgc taatttttca aacaaagaat	10680
ctgagctgca tttttacaga acagaaatgc aacgcgagag cgctatttta ccaacaaaga	10740
atctatactt cttttttgtt ctacaaaaat gcaccccgag agcgctattt ttctaacaaa	10800
gcactctaga ttactttttt tctcttttgt gcgctctata atgcagtctc ttgataactt	10860
tttgactgt aggtccgta aggttagaag aaggctactt tgggtgtctat tttctcttc	10920
ataaaaaag cctgactcca ctccccgct ttactgatta ctagcgaagc tgcgggtgca	10980
ttttttcaag ataaaggcat ccccgattat attctatacc gatgtggatt gcgcatactt	11040
tgtgaacaga aagtgatagc gttgatgatt cttcattggg cagaaaatta tgaacggttt	11100
cttctatttt gtctctatat actacgtata ggaaatgttt acattttcgt attgttttcg	11160
attcactcta tgaatagttc ttactacaat tttttgtct aaagagtaat actagagata	11220
aacataaaaa atgiagaggt cgagtttaga tgcaagtcca aggagcgaaa ggtggatggg	11280

taggttatat agggatatag cacagagata tatagcaaag agatactttt gagcaatgtt 11340
 tgtggaagcg gtattcgcaa tattttagta gctcgttaca gtccggtgcg tttttggttt 11400
 tttgaaagtg cgtcttcaga gcgcttttgg ttttcaaag cgctctgaag ttcctatact 11460
 ttctagagaa taggaacttc ggaataggaa ctcaaagcg tttccgaaaa cgagcgcttc 11520
 cgaaaatgca acgcgagctg cgcacataca gctcactgtt cacgtcgac ctatatctgc 11580
 gtgttgctg tatatatata tacatgagaa gaacggcata gtgcgtgttt atgcttaaat 11640
 gcgtacttat atgcgtctat ttatgtagga tgaaggtag tctagtacct cctgtgatat 11700

tateccattc catgcggggt atcgtagct tccttcagca ctacccttta gctgttctat 11760
 atgctgccac tctcaattg gattagtctc atccttcaat gctatcattt cctttgatat 11820
 tggatcatat gcatagtacc gagaaactag aggatc 11856

<210> 196

<211> 338

<212> PRT

<213> Artificial sequence

<220><223> K9_David_SH

<400> 196

Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn Leu Ser Leu Leu Asp

1 5 10 15

Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser Gln Gly His Ala His

20 25 30

Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val Ile Ile Gly Leu Phe

35 40 45

Glu Gly Ala Glu Glu Trp Lys Arg Ala Glu Glu Gln Gly Phe Glu Val

50 55 60

Tyr Thr Ala Ala Glu Ala Ala Lys Lys Ala Asp Ile Ile Met Ile Leu

65 70 75 80

Ile Pro Asp Glu Met Gln Ala Thr Met Tyr Lys Asn Asp Ile Glu Pro

85 90 95

Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala His Gly Phe Asn Ile

100 105 110

His Phe Gly Cys Ile Val Pro Pro Lys Asp Val Asp Val Thr Met Ile

115 120 125

Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser Glu Tyr Glu Glu Gly

130 135 140

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

145 150 155 160

Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala Ile Gly Gly Ala Arg

165 170 175

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

180 185 190

Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val Cys Ala Leu Met Gln

195 200 205

Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr Asp Pro Arg Asn Ala

210 215 220

Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile Val Asp Leu Ile Tyr

225 230 235 240

Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile Ser Asn Thr Ala Glu

245 250 255

Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile Thr Glu Asp Thr Lys

260 265 270

Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln Asp Gly Thr Phe Ala

275 280 285

Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly Ser Gln Val His Phe

290 295 300

Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro Ala Glu Val Val Gly

305 310 315 320

Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp Glu Asp Lys Leu Ile

325 330 335

Asn Asn

<210> 197

<211> 34

<212

> DNA

<213> artificial sequence

<220><223> primer oBP562

<400> 197

aattgtttaa acatgtatac agtaggtgac tatc 34

<210> 198

<211> 49

<212> DNA

<213> artificial sequence

<220><223> primer oBP563

<400> 198

aatcataaat cataagaaat tcgcttatca gctcttgttt tgttctgca 49

<210> 199

<211> 49

<212> DNA

<213> artificial sequence

<220><223> primer oBP564

<400> 199

ttatttgcag aacaaaacaa gagctgataa gcgaatttct tatgattta 49

<210> 200

<211> 34

<212> DNA

<213> artificial sequence

<220><223> primer oBP565

<400> 200

aattggccgg ccaaaaaaag catgcacgta taca 34

<210> 201

<211> 32

<212> DNA

<213> artificial sequence

<220><223> primer oBP505

<400> 201

aattgagctc actgtagccc tagacttgat ag 32

<210> 202

<211> 35

<212> DNA
 <213> artificial sequence
 <220><223> primer oBP506
 <400> 202
 aattggcgcg cctgtatatg agatagttga ttgta 35

<210> 203
 <211> 35
 <212> DNA
 <213> artificial sequence
 <220><223> primer oBP507
 <400> 203
 aattttaatt aagtctagggt tctttggctg ttcaa 35

<210> 204
 <211> 33
 <212> DNA
 <213> artificial sequence
 <220><223> primer oBP508
 <400> 204
 aattgtcgac tttagaagtg tcaacaacgt atc 33

<210> 205
 <211> 34
 <212> DNA
 <213> artificial sequence
 <220><223> primer oBP674
 <400> 205
 aattggcgcg ccaattaccg tcgctcgtga ttg 34

<210> 206
 <211> 34
 <212> DNA
 <213> artificial sequence
 <220><223> primer oBP675
 <400> 206
 aattgtttaa acttgaatat gtattacttg gtta 34

<210> 207
 <211> 22
 <212> DNA
 <213> artificial sequence
 <220><223> primer oBP495
 <400> 207
 ggagatatac aatagaacag at 22
 <210> 208
 <211> 22
 <212> DNA
 <213> artificial sequence
 <220><223> primer oBP496
 <400> 208
 tagcaatggg gtttttttca gt 22
 <210> 209
 <211> 8366
 <212> DNA
 <213> Artificial Sequence
 <220><223> pRS316-UAS(PGK1)-PFBA1-GUS
 <400> 209
 tcgcgcgttt cggtgatgac ggtgaaaacc tctgacacat gcagctcccg gagacggtca 60
 cagcttgtct gtaagcggat gccgggagca gacaagcccg tcagggcgcg tcagcgggtg 120
 ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180
 accacgcttt tcaattcaat tcatcathtt ttttttattc ttttttttga tttcggtttc 240
 tttgaaatth ttttgattcg gtaatctccg aacagaagga agaacgaagg aaggagcaca 300
 gacttagatt ggtatatata cgcatatgta gtgttgaaga aacatgaaat tgcccagtat 360
 tcttaacca actgcacaga aaaaaaacct gcaggaaacg aagataaatc atgtcgaaag 420
 ctacatataa ggaacgtgct gctactcadc ctagtctgtg tgctgccaag ctatttaata 480
 tcatgcacga aaagcaaaca aacttgtgtg cttcattgga tgttcgtacc accaaggaat 540
 tactggagtt agttgaagca ttaggtccca aaatttgtht actaaaaaca catgtggata 600
 tcttgactga tttttccatg gagggcacag ttaagccgct aaaggcatta tccgccaagt 660
 acaatttttt actcttcgaa gacagaaaat ttgctgacat tggtaatata gtcaaattgc 720

agtactctgc ggggtgtatac agaatagcag aatgggcaga cattacgaat gcacacggtg 780

 tgggtggccc aggtattgtt agcggtttga agcaggcggc agaagaagta acaaaggaac 840
 ctagaggcct tttgatgtta gcagaattgt catgcaaggg ctcctatct actggagaat 900
 atactaaggg tactgttgac attgcgaaga gcgacaaaga ttttgttatc ggctttattg 960
 ctcaaagaga catgggtgga agagatgaag gttacgattg gttgattatg acaccggtg 1020
 tgggtttaga tgacaaggga gacgcattgg gtcaacagta tagaaccgtg gatgatgtgg 1080
 tctctacagg atctgacatt attattgttg gaagaggact atttgcaaag ggaagggatg 1140
 ctaaggtaga ggggtgaacgt tacagaaaag caggctggga agcatatttg agaagatgcg 1200

 gccagcaaaa ctaaaaaact gtattataag taaatgcatg tatactaaac tcacaaatta 1260
 gagcttcaat ttaattatat cagttattac cctgcggtgt gaaataccgc acagatgcgt 1320
 aaggagaaaa taccgatca ggaaattgta aacgttaata ttttgttaaa attcgcgtaa 1380
 aatTTTTgtt aaatcagctc atTTTTtaac caataggccg aaatcggcaa aatcccttat 1440
 aaatcaaaag aatagaccga gatagggttg agtgttgttc cagtttgaa caagagtcca 1500
 ctattaaaga acgtggactc caacgtcaaa gggcgaaaaa ccgtctatca gggcgatggc 1560
 ccactacgtg aaccatcacc ctaatcaagi tttttggggt cgagggtgccg taaagcacta 1620

 aatcggaacc ctaaaggag ccccgattt agagcttgac ggggaaagcc ggcgaacgtg 1680
 gcgagaaagg aagggaagaa agcgaagga gcgggcgcta gggcgctggc aagtgtagcg 1740
 gtcacgtgc gcgtaccac cacaccgcc gcgcttaatg cgccgctaca gggcgcgctcg 1800
 cgccattcgc cattcaggct gcgcaactgt tgggaagggc gatcgggtgcg ggcctcttcg 1860
 ctattacgcc agctggcgaa ggggggatgt gctgcaaggc gattaagttg ggtaacgcca 1920
 gggttttccc agtcacgacg ttgtaaaacg acggccagtg aattgtaata cgactcacta 1980
 tagggcgaat tggagctcca ccgcggtggt ttaacgtata gacttcta atatttctcc 2040

 atacttggtt tttttatc actttttta tacatatgtt gttttttat acaatatcaa 2100
 aaatcaataa taattaattt aataacaaga atatgttgct gtgatgactc ttaatataca 2160
 attgggaagc attcaaggat tggtagcagc tttgctttct gataattaaa cgcttggcgt 2220
 ttaatctttt ttgccgatta ataattat tattaccctc ttatttatta gcattgtctt 2280
 ccgtactaaa aggagtaggg aaaaaggcga agtagagtga cagtcatt acccagcgt 2340
 tcataagaat ccatcaggta tattaaattg tgatgggaga acaaaacaaa ttgtactatg 2400
 atgtcgagaa attggtgaat tctttgcagg aaagtttga cctggattgc gcgcaaagtg 2460

tatcactttt caccagtaaa tcaaggagca atgaggcttg gttggaagag ctagagaata	2520
aattcaagtt aaaagatgat gttgaacttg atgatgtgga aaatttaaga gccgaaattg	2580
acatgaagtt aaatatgttg gaggataaag taagctacta tgaaagactt tacaaagaac	2640
tcgaagagtt ccagaatgaa ataaaaatta aaacagttgt gaataacaga agacaatcgc	2700
gaactccaaa atgagctatc aaaaacgata gatcgattag gatgactttg aaatgactcc	2760
gcagtggact ggccgttaat ttcaagcgtg agtaaaatag tgcatgacaa aagatgagct	2820
aggcttttgt aaaaatatct tacgttgtaa aatttttagaa atcattatct cttcatatc	2880
attttgtcat tgaccttcag aagaaaagag ccgaccaata atataataa ataaataaaa	2940
ataatatcc attatttcta aacagattca atactatta aaaaactata tcaattaatt	3000
tgaattaacg cgcccgctct agttaattaa tcattgtttg cctccctgct gcggttttct	3060
accgaagttc atgccagtc agcggttttg cagcagaaaa gccgccgact tcggtttgcg	3120
gtcgcgagtg aagatccctt tcttgttacc gccaacgcgc aatatgcctt gcgaggtcgc	3180
aaaatcggcg aaattccata cctgttcacc gacgacggcg ctgacgcgat caaagacgcg	3240
gtgatacata tcagccatg cacactgata ctcttcactc cacatgtcgg tgtacattga	3300
gtgcagcccg gtaaacgtat ccacgccgta ttcggtgatg ataatcggt gatgcagttt	3360
ctctgccag gccagaagtt cttttccag taccttctct gccgtttcca aatcgccgt	3420
ttggacatac catccgtaat aacggttcag gcacagcaca tcaaagagat cgctgatggt	3480
atcggtgtga gcgtgcaga acattacatt gacgcaggtg atcggacgcg tcgggtcgag	3540
tttacgcgtt gcttccgcca gtggcgcgaa atattcccgt gcaccttgcg gacgggtatc	3600
cggttcgttg gcaatactcc acatcaccac gcttgggtgg tttttgtcac gcgtatcag	3660
ctctttaatc gcctgtaagt gcgcttgctg agtttccccg ttgactgcct ctctgctgta	3720
cagttctttc ggttgttgcc ccgcttcgaa accaatgcct aaagagaggt taaagccgac	3780
agcagcagtt tcatcaatca ccacgatgcc atgttcatct gcccagtcga gcatctcttc	3840
agcgtaaagg taatgcgagg tacggttagga gttggcccca atccagtcca ttaatgcgtg	3900
gtcgtgcacc atcagcagct tatcgaatcc ttggccacgc aagtccgcat ctatcatgacg	3960
accaaagcca gtaaagtaga acggtttgtg gttaatcagg aactgttcgc ctttactgc	4020
cactgaccgg atgccgacgc gaagcgggta gatatcacac tctgtctggc ttttggtgt	4080
gacgcacagt tcatagagat aaccttcacc cggttgccag aggtgcggat tcaccacttg	4140
caaagtcccg ctagtgcctt gtccagttgc aaccacctgt tgatccgcat cagcagttc	4200
aacgtgaca tcaccattgg ccaccactg ccagtcaaca gacgcgtggt tacagtcttg	4260
cgcgacatgc gtcaccacgg tgatatctc caccaggtg ttcggcgtgg tgtagagcat	4320

tacgctgcga tggattccgg catagttaaa gaaatcatgg aagtaagact gctttttctt	4380
gccgttttctg tcggtaatac ccattcccgg cgggatatgc tgccagttca gttcgttggt	4440
cacacaaacg gtgatacgta cacttttccc ggcaataaca tacggcgtga catcggttc	4500
aaatggagta tagccgcctt gatgtccat cacttctga ttattgacct acactttgcc	4560
gtaatgagt accgcatcga aacgcagcac gatacgtgg cctgcccac ctttcggtat	4620
aaagacttcg cgtgatacc agacgttgcc cgcataatta cgaatatctg catcggcgaa	4680
ctgatcgta aaactgcctg gcacagcaat tgcccggctt tcttgtaacg cgctttccca	4740
ccaacgtga tcaattccac agttttcgcg atccagactg aatgccaca ggccgtcgag	4800
tttttgatt tcacgggttg gggtttctac aggacgtacc atactagttt gaatatgtat	4860
tacttggtta tggttatata tgacaaaaga aaaagaagaa cagaagaata acgcaaggaa	4920
gaacaataac tgaaattgat agagaagtat tatgtctttg tctttttata ataatcaag	4980
tgcagaaatc cgtagacaa catgagggat aaaatttaac gtgggcgaag aagaaggaaa	5040
aaagtttttg tgaggcgta attgaagcga tctgttgatt gtagattttt ttttttgag	5100
gagtcaaagt cagaagagaa cagacaaatg gtattaacca tccaatactt ttttgagca	5160
acgctaaact catgcttttc cattgggttac gtgctcagtt gttagatatg gaaagagagg	5220
atgctcacgg cagcgtgact ccaattgagc ccgaaagaga ggatgccacg ttttcccgac	5280
ggctgctaga atggaaaaag gaaaaataga agaattccat tcctatcatt atttacgtaa	5340
tgaccacac atttttgaga ttttcaacta ttacgtatta cgataatcct gctgtcatta	5400
tcattattat ctatatcgac gtatgcaacg tatgtgaagc caagtaggcg tcgacaggac	5460
cttggtgtgt gacgaaattg gaagctgcaa tcaataggaa gacaggaagt cgagcgtgtc	5520
tgggtttttt cagttttgtt ctttttgcaa acaaatcacg agcgacggta attggtacct	5580
agcttttggt ccttttagtg agggttaatt ccgagcttgg cgtaatcatg gtcatactg	5640
tttctgtgt gaaattgtta tccgtcaca attccacaca acataggagc cggaagcata	5700
aagtgtaaag cctgggtgct ctaatgagt aggtaactca cattaattgc gttgcgtca	5760
ctgccgctt tccagtcggg aaacctgtcg tgccagctgc attaatgaat cggccaacgc	5820
gcggggagag gcggtttgct tattgggcgc tcttccgctt cctcgtcac tgactcgtg	5880
cgctcggtcg ttcggctcgc gcgagcggt ttagctcact caaaggcgtt aatacggtta	5940
tccacagaat caggggataa cgcaggaaag aacatgtgag caaaaggcca gcaaaaggcc	6000
aggaaccgta aaaaggccgc gttgctggcg tttttcata ggctcggccc cctgacgag	6060
catcacaaaa atcgacgtc aagtcagagg tggcgaaacc cgacaggact ataaagatac	6120
caggcgttcc cccctggaag ctccctcgtg cgctctctg ttcgaccct gccgcttacc	6180

ggatacctgt ccgcctttct cccttcggga agcgtggcgc tttctcaatg ctacacgtgt 6240

 aggtatctca gttcgggtga ggtcgttcgc tccaagctgg gctgtgtgca cgaaccccc 6300
 gttcagcccg accgctgcgc cttatccggt aactatcgtc ttgagtccaa cccggtaga 6360
 cagcacttat cgccactggc agcagccact ggtaacagga ttagcagagc gaggtatgta 6420
 ggcggtgcta cagagttctt gaagtgggtg cctaactacg gctacactag aaggacagta 6480
 tttggatatc gcgtctctgt gaagccagtt accttcggaa aaagagttgg tagctcttga 6540
 tccggcaaac aaaccaccgc tggtagcggg ggtttttttg tttgcaagca gcagattacg 6600
 cgcagaaaaa aaggatctca agaagatcct ttgatctttt ctacggggtc tgacgtcag 6660

 tggaaacgaaa acicacgtta agggattttg gtcattgagat tatcaaaaag gatcttcacc 6720
 tagatccttt taaattaaaa atgaagtttt aaatcaatct aaagtatata tgagtaaact 6780
 tggcttgaca gttaccaatg cttaatcagt gaggcaccta tctcagcgat ctgtctatct 6840
 cgttcatcca tagttgcctg actgcccgtc gtgtagataa ctacgatacg ggagggctta 6900
 ccatctggcc ccagtgtgc aatgataccg cgagaccac gtcaccggc tccagattta 6960
 tcagcaataa accagccagc cggaagggcc gagcgagaa gtggctctgc aactttatcc 7020
 gcctccatcc agtctattaa ttgttgccgg gaagctagag taagtagttc gccagttaat 7080

 agtttgcca acgttgttc cattgttaca ggcatcgtgg tgtcacgtc gtcgtttggt 7140
 atggcttcat tcagctccgg ttcccaacga tcaaggcgag ttacatgatc cccatgttg 7200
 tgaaaaaag cggttagctc cttcggtcct ccgatcgttg tcagaagtaa gttggccgca 7260
 gtgttatcac tcatggttat ggcagcactg cataattctc ttactgtcat gccatccgta 7320
 agatgctttt ctgtgactgg tgagtactca accaagtcac tctgagaata gtgtatgcgg 7380
 cgaccgagtt gctcttgccc ggcgtcaata cgggataata ccgcgccaca tagcagaact 7440
 ttaaaagtgc tcatcattgg aaaacgttct tcggggcgaa aactctcaag gatcttaccg 7500

 ctgttgagat ccagttcgat gtaaccact cgtgcacca actgatcttc agcatctttt 7560
 actttcacca gcgtttctgg gtgagcaaaa acaggaaggc aaaatgccgc aaaaaaggga 7620
 ataaggcgca cacgaaatg ttgaatactc atactcttc ttttcaata ttattgaagc 7680
 atttatcagg gttattgtct catgagcgga tacatatgtg aatgtattta gaaaaataaa 7740
 caaatagggg ttccgcgcac atttccccga aaagtccac ctgggtcctt ttcacacgt 7800
 gctataaaaa taattataat ttaattttt taatataaat atataaatta aaaatagaaa 7860
 gtaaaaaaag aaattaaaga aaaaatagtt ttgttttcc gaagatgtaa aagactctag 7920

ggggatcgcc aacaaatact accttttatac ttgctcttcc tgctctcagg tattaatgcc 7980
 gaattgtttc atcttgctcg tgtagaagac cacacacgaa aatcctgtga ttttacattt 8040
 tacttatcgt taatcgaatg tataatctatt taatctgctt ttcttgctta ataaatatat 8100
 atgtaaagta cgcttttttg tgaaattttt taaacctttg tttatttttt ttcttcatt 8160
 ccgtaactct tctaccttct ttatttactt tctaaaatcc aaatacaaaa cataaaaata 8220
 aataaacaca gagtaaattc ccaaattatt ccatcattaa aagatcagag gcgcgtgtaa 8280
 gtacaggca agcgatccgt cctaagaaac cattattatc atgacattaa cctataaaaa 8340

taggcgtatc acgaggccct ttcgtc 8366

<210> 210

<211> 22

<212> DNA

<213> Artificial sequence

<220><223> primer N160SeqF5

<400> 210

cctgaagtct aggtccctat tt 22

<210> 211

<211> 20

<212> DNA

<213> Artificial sequence

<220><223> N886 primer

<400> 211

caatgattgt tggtaaaggg 20

<210> 212

<211> 32

<212> DNA

<213> Artificial sequence

<220><223> Primer oBP622

<400> 212

aattggtacc ccaaaaggaa tattgggtca ga 32

<210> 213

<211> 49

<212> DNA

<213> Artificial sequence

<220><223> Primer oBP623

<400> 213

ccattgttta aacggcgcg cggatccttt gcgaaacct atgctctgt 49

<210> 214

<211> 49

<212> DNA

<213> Primer oBP624

<400> 214

gcaaaggatc cggcgcgccg tttaacaat ggaaggcgg gatgagcat 49

<210> 215

<211> 34

<212> DNA

<213> Artificial sequence

<220><223> Primer oBP625

<400> 215

aattggccgg cctacgtaac attctgtcaa ccaa 34

<210> 216

<211> 34

<212> DNA

<213> Artificial sequence

<220><223> Primer oBP626

<400> 216

aattgcggcc gttcatata tgacgtaata aaat 34

<210> 217

<211> 34

<212> DNA

<213> Artificial sequence

<220><223> Primer oBP627

<400> 217

aatTTtaatt aatTTTTttt ctTggaatca gtac 34

<210> 218

<211> 40

<212> DNA

<213> Artificial sequence

<220><223> Primer HY21

<400> 218

ttaaggcgcg cctatttgta atacgtatac gaattccttc 40

<210> 219

<211> 56

<212> DNA

<213> Artificial sequence

<220><223> Primer HY24

<400> 219

acttaataac ttaccggct gttgacattt tgttcttctt gttattgtat tgtgtt 56

<210> 220

<211> 56

<212> DNA

<213> Artificial sequence

<220><223> Primer HY25

<400> 220

aacacaatac aataacaaga agaacaaaat gtcaacagcc ggtaaagtta ttaagt 56

<210> 221

<211> 40

<212> DNA

<213> Artificial sequence

<220><223> Primer HY4

<400> 221

ggaagttaa acaccacagg tgttgctctc tgaggacata 40

<210> 222

<211> 30

<212> DNA

<213> Artificial sequence

<220><223> Primer URA3-end F

<400> 222

gcatatttga gaagatgcgg ccagcaaac 30

<210> 223

<211> 22

<212> DNA
 <213> Artificial sequence
 <220><223> Primer oBP636
 <400> 223
 catttttttc cctctaagaa gc 22

<210> 224
 <211> 22
 <212> DNA
 <213> Artificial sequence
 <220><223> Primer oBP637
 <400> 224
 tttttgcaca gttaaaactac cc 22

<210> 225
 <211> 40
 <212> DNA
 <213> Artificial sequence
 <220><223> Primer oBP691
 <400> 225
 aattggatcc gcgatcgca cgttctctcc gttgttcaaa 40

<210> 226
 <211> 41
 <212> DNA
 <213> Artificial sequence
 <220><223> Primer oBP692
 <400> 226
 aattggcgcg ccatttaa atatatgtat atatataaca c 41

<210> 227
 <211> 34
 <212> DNA
 <213> Artificial sequence
 <220><223> Primer oBP693
 <400> 227
 aattgtttaa acaaaggatg atattgttct atta 34

<210> 228
 <211> 33
 <212> DNA
 <213> Artificial sequence
 <220><223> Primer oBP694
 <400> 228
 aattggccgg ccgcaacgac gacaatgcca aac 33
 <210> 229
 <211> 34
 <212> DNA
 <213> Artificial sequence
 <220><223> Primer oBP695
 <400> 229
 aattgcggcc gcatgacagg tgaaagaatt gaaa 34
 <210> 230
 <211> 34
 <212> DNA
 <213> Artificial sequence
 <220><223> Primer oBP696
 <400> 230
 aattttaatt aaacgggcat cttatagtg t cgtt 34
 <210> 231
 <211> 40
 <212> DNA
 <213> Artificial sequence
 <220><223> Primer HY16
 <400> 231
 ttaaggcgcg ccccgacgc cgaaatgcat gcaagtaacc 40
 <210> 232
 <211> 56
 <212> DNA
 <213> Artificial sequence
 <220><223> Primer HY19
 <400> 232

acttaataac ttaccggct gttgacattt tgattgattt gactgtgtta ttttgc 56

<210> 233

<211> 56

<212> DNA

<213> Artificial sequence

<220><223> Primer HY20

<400> 233

gcaaaataac acagtcaaat caatcaaaat gtcaacagcc ggtaaagtta ttaagt 56

<210> 234

<211> 22

<212> DNA

<213> Artificial sequence

<220><223> Primer oBP730

<400> 234

ttgctccaaa gagatgtctt ta 22

<210> 235

<211> 22

<212> DNA

<213> Artificial sequence

<220><223> Primer oBP731

<400> 235

tgttcccaca atctattacc ta 22

<210> 236

<211> 34

<212> DNA

<213> Artificial sequence

<220><223> Primer HY31

<400> 236

gccgacttta tggcgaagaa gtttgctctt gatc 34

<210> 237

<211> 38

<212> DNA

<213> Artificial sequence

<220><223> N1179 primer
 <400> 237
 gcgaattctg atctgatgcg ctttgcataat ctcataatt 38
 <210> 238
 <211> 38
 <212> DNA
 <213> Artificial sequence
 <220><223> N1180 primer
 <400> 238
 gcggtaccgc cgtttcggct tggatacgcc atatgcaa 38
 <210> 239
 <211> 38
 <212> DNA
 <213> Artificial sequence
 <220><223> N1181 primer
 <400> 239
 gccgatcgtg taccaacctg catttctttc cgtcataat 38
 <210> 240
 <211> 38
 <212> DNA
 <213> Artificial sequence
 <220><223> N1182 primer
 <400> 240
 gcgcatacgt gtaccttttg attggaattg gttcgcag 38
 <210> 241
 <211> 30
 <212> DNA
 <213> Artificial sequence
 <220><223> N1212 primer
 <400> 241
 ttctatgtta aattttaacg atgtagacac 30
 <210> 242
 <211> 30

<212> DNA
 <213> Artificial sequence
 <220><223> N1213 primer
 <400> 242
 atggcgaaat ggcagtactc gggggaatgc 30
 <210> 243
 <211> 25
 <212> DNA
 <213> Artificial sequence
 <220><223> Primer N1239
 <400> 243
 gacgaggata atgtgcatga acggg 25
 <210> 244
 <211> 24
 <212> DNA
 <213> Artificial sequence
 <220><223> Primer N1242
 <400> 244
 gtttcgttgc gttggtggaa tggc 24

 <210> 245
 <211> 28
 <212> DNA
 <213> Artificial sequence
 <220><223> Primer N1249
 <400> 245
 gctttatgga ccctgaaacc acagccac 28
 <210> 246
 <211> 27
 <212> DNA
 <213> Artificial sequence
 <220><223> Primer N1250
 <400> 246
 ggaaagaaaa agcttcatgg ccttgcg 27
 <210> 247

<211> 23
 <212> DNA
 <213> Artificial sequence
 <220><223> Primer N1251
 <400> 247
 ccaatgtggc tgtggtttca ggg 23

<210> 248
 <211> 26
 <212> DNA
 <213> Artificial sequence
 <220><223> Primer N1252
 <400> 248
 gcaaggccat gaagcttttt ctttcc 26

<210> 249
 <211> 62
 <212> DNA
 <213> Primer N1253
 <400> 249
 ggaaagaaaa agcttcatgg ccttgcgctc atcaacacta aaattagagt cattctaatt 60
 gc 62

<210> 250
 <211> 29
 <212> DNA
 <213> Artificial sequence
 <220><223> Primer N1254
 <400> 250

ttatccacgg aagatatgat gaggtgacg 29

<210> 251
 <211> 64
 <212> DNA
 <213> Primer N1255
 <400> 251
 ggaccctgaa accacagcca cattggtacc ttccttgcca aagggattaa acctgacaat 60

tacc 64

<210> 252

<211> 29

<212> DNA

<213> Artificial sequence

<220><223> Primer N1256

<400> 252

acctagctaa actaagtaaa tctgtatac 29

<210> 253

<211> 29

<212> DNA

<213> Artificial sequence

<220><223> Primer N1257

<400> 253

gatgatgcta ttggtgcag agggatgatg 29

<210> 254

<211> 60

<212> DNA

<213> Artificial sequence

<220><223> Primer N1258

<400> 254

cagatttact tagtttagct aggtgtatca aaatacgttc tcaatgttct atttcccgcc 60

<210> 255

<211> 24

<212> DNA

<213> Artificial sequence

<220><223> Primer N1262

<400> 255

cacgtaaggg catgatagaa ttgg 24

<210> 256

<211> 26

<212> DNA

<213> Artificial sequence

<220><223> Primer N1263

<400> 256

ggatatagca gttgttgtac actagc 26

<210> 257

<211> 24

<212> DNA

<213> Artificial sequence

<220><223> primer N1223

<400> 257

caaggaatta ccatcacgt cacc 24

<210> 258

<211> 27

<212> DNA

<213> Artificial sequence

<220><223> primer N1225

<400> 258

cttgtgtata atgataaatt gggtggg 27

<210> 259

<211> 27

<212> DNA

<213> Artificial sequence

<220><223> Primer N1228

<400> 259

caacggttca tcattcatg gatctgc 27

<210> 260

<211> 26

<212> DNA

<213> Artificial sequence

<220><223> Primer N1229

<400> 260

gttacttggg tctggcgagg tattgg 26

<210> 261

<211> 25

<212> DNA

<213> Artificial sequence

<220><223> Primer N1230

<400> 261

gtaagaaggt tggattcca gctgg 25

<210> 262

<211> 26

<212> DNA

<213> Artificial sequence

<220><223> Primer N1231

<400> 262

ctgtttcgat accagaacct agaccg 26

<210> 263

<211> 1017

<212> DNA

<213> Artificial sequence

<220><223> K9_David_SH

<400> 263

atggctaaga ttactacca agaagactgt aacttgctct tgttggatgg taagactatc 60

gccgttatcg gttacgggtc tcaagggtcac gctcatgcc tgaatgctaa ggaatccggt 120

tgtaacgtta tcattgggtt attcgaaggt gcggaggagt ggaaaagagc tgaagaacaa 180

ggtttcgaag tctacaccgc tgctgaagct gctaagaagg ctgacatcat tatgatcttg 240

atcccagatg aaatgcaggc taccatgtac aaaaacgaca tcgaacaaaa cttggaagcc 300

ggtaacatgt tgatgttcgc tcacgggttc aacatccatt tcggttgat tgttcacca 360

aaggacgttg atgtcactat gatcgctcca aagggtccag gtcacaccgt tagatccgaa 420

tacgaagaag gtaaagggtg cccatgcttg gttgctgtcg aacaagacgc tactggcaag 480

gctttggata tggctttggc ctacgcttta gccatcgggtg gtgctagagc cgggtgtcttg 540

gaaactacct tcagaaccga aactgaaacc gacttggtcg gtgaacaagc tgttttatgt 600

ggtggtgtct gcgctttgat gcaggccggt tttgaaacct tgggtgaagc cggttacgac 660

ccaagaaacg cttacttcga atgtatccac gaaatgaagt tgatcgttga cttgatctac 720

caatctggtt tctccggtat gcgttactct atctccaaca ctgctgaata cggtgactac 780

attaccggtc caaagatcat tactgaagat accaagaagg ctatgaagaa gatattgtct 840

gacattcaag atggtacctt tgccaaggac ttcttggttg acatgtctga tgctggttcc 900

cagggtccact tcaaggctat gagaaagttg gcctccgaac acccagctga agttgtcggg 960
gaagaaatta gatccttgta ctcttggtcc gacgaagaca agttgattaa caactga 1017

<210> 264

<211> 31

<212> DNA

<213> Artificial sequence

<220><223> Primer pOLE1-NgoMI

<400> 264

cgtggccggc cgtgcaaaaa ctaactccga g 31

<210> 265

<211> 34

<212> DNA

<213> Artificial sequence

<220><223> Primer pOLE1-PmlI

<400> 265

gtactcacgt gttgttgta atgttttagt gctg 34

<210> 266

<211> 338

<212> PRT

<213> Artificial sequence

<220><223> K9_Eliza_SH

<400> 266

Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn Leu Ser Leu Leu Asp

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

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

325 330 335

Asn Asn

<210> 267

<211> 338

<212> PRT

<213> Artificial sequence

<220><223> K9_Frank_SH

<400> 267

Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn Leu Ser Leu Leu Asp

1 5 10 15

Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser Gln Gly His Ala His

20 25 30

Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val Ile Ile Gly Leu Phe

35 40 45

Glu Gly Ala Glu Glu Trp Lys Arg Ala Glu Glu Gln Gly Phe Glu Val

50 55 60

Tyr Thr Ala Ala Glu Ala Ala Lys Lys Ala Asp Ile Ile Met Ile Leu

65 70 75 80

Ile Pro Asp Glu Lys Leu Ala Thr Met Tyr Lys Asn Asp Ile Glu Pro

85 90 95

Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala His Gly Phe Asn Ile

100 105 110

His Phe Gly Cys Ile Val Pro Pro Lys Asp Val Asp Val Thr Met Ile

115 120 125

Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser Glu Tyr Glu Glu Gly

130 135 140

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

145 150 155 160

Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala Ile Gly Gly Ala Arg

165 170 175

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

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

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

290 295 300
Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro Ala Glu Val Val Gly
305 310 315 320
Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp Glu Asp Lys Leu Ile
325 330 335
Asn Asn

<210> 268

<211> 4236

<212> DNA

<213> artificial sequence

<220><223> Synthetic construct

<400> 268

gatccgcatt gcggattacg tattctaagt ttcagataac ttcgtatagc atacattata 60

cgaagtatg cagattgtac tgagagtgc ccataccaca gcttttcaat tcaattcatc 120

atTTTTTTTT tattctTTTT ttTgatttcg gtttctttga aatTTTTttg attcggtaat 180

ctccgaacag aaggaagaac gaaggaagga gcacagactt agattgggat atatacgcat 240

atgtagtgtt gaagaaacat gaaattgcc agtattctta acccaactgc acagaacaaa 300

aacctgcagg aaacgaagat aaatcatgtc gaaagctaca tataaggaac gtgctgctac	360
tcacacctagt cctgttgctg ccaagctatt taatatcatg cacgaaaagc aaacaaactt	420
gtgtgcttca ttggatgttc gtaccaccaa ggaattactg gagttagtgt aagcattagg	480
tcccaaaatt tgtttactaa aaacacatgt ggatatcttg actgattttt ccatggaggg	540
cacagttaag ccgctaaagg cattatccgc caagtacaat tttttactct tcgaagacag	600
aaaatttgct gacattggta atacagtcaa attgcagtac tctgcgggtg tatacagaat	660
agcagaatgg gcagacatta cgaatgcaca cgggtgtggtg ggcccaggtg ttgttagcgg	720
tttgaagcag gcggcagaag aagtaacaaa ggaacctaga ggccttttga tgtagcaga	780
attgtcatgc aagggtccc tatctactgg agaataact aagggtactg ttgacattgc	840
gaagagcgac aaagattttg ttatcggttt tattgtcaa agagacatgg gtggaagaga	900
tgaaggttac gattggttga ttatgacacc cgggtgtgggt ttagatgaca agggagacgc	960
attgggtcaa cagtatagaa ccgtggatga tgtggtctct acaggatctg acattattat	1020
tgttggaaga ggactatttg caaagggaag ggatgctaag gtagagggtg aacgttacag	1080
aaaagcaggc tgggaagcat atttgagaag atgcggccag caaaactaaa aaactgtatt	1140
ataagtaaat gcatgtatac taaactcaca aattagagct tcaatttaat tatatcagtt	1200
attaccctat gcggtgtgaa ataccgcaca gatgcgtaag gagaaaatac cgcacagga	1260
aattgtaaac gttaatattt tggttaaaatt cgcgttaaat ttttgtaaa tcagtcatt	1320
ttttaaccaa taggccgaaa tcggcaaaat cccttataaa tcaaaagaat agaccgagat	1380
agggttgagt gttgttccag ttggaacaa ggtccacta ttaaagaacg tggactcaa	1440
cgtcaaaggg cgaaaaaccg tctatcaggg cgatggccca ctactgaac catcaccta	1500
atcaagataa cttcgtatag catacattat acgaagtat ccagtgatga tacaacgagt	1560
tagccaaggt gaattcactg gccgtcggtt tacaacgtcg tgactgggaa aaccctggcg	1620
ttaccaact taatcgctt gcagcacatc ccctttcgc cagctggcgt aatagcgaag	1680
aggcccgcac cgatcgccct tccaacagt tgcgcagcct gaatggcgaa tggcgcctga	1740
tgcggtatth tctccttacg catctgtgcg gtatttcaca ccgcatatgg tgcactctca	1800
gtacaatctg ctctgatgcc gcatagttaa gccagccccg acaccgcca acaccgctg	1860
acgcgccctg acgggcttgt ctgctcccg catccgtta cagacaagct gtgaccgtct	1920
ccgggagctg catgtgtcag aggttttcac cgtcatcacc gaaacgcgcg agacgaaagg	1980
gcctcgtgat acgcctatth ttataggtha atgtcatgat aataatggtt tcttagacgt	2040
caggtggcac ttttcgggga aatgtgcgcg gaaccctat ttgtttatth ttctaaatac	2100

attcaaatat gtatccgctc atgagacaat aaccctgata aatgcttcaa taatattgaa	2160
aaaggaagag tatgagtatt caacatttcc gtgtcgccct tattcccttt tttgcggcat	2220
tttgccttcc tgtttttgct caccagaaaa cgttggtgaa agtaaaagat gctgaagatc	2280
agttgggtgc acgagtgggt tacatcgaac tggatctcaa cagcggtaag atccttgaga	2340
gttttcgccc cgaagaacgt tttccaatga tgagcacttt taaagtcttg ctatgtggcg	2400
cggattatc ccgtattgac gccgggcaag agcaactcgg tcgccgata cactattctc	2460
agaatgactt ggttagtac tcaccagtca cagaaaagca tcttacggat ggcatgacag	2520
taagagaatt atgcagtgt gccataacca tgagtataa cactgcggcc aacttacttc	2580
tgacaacgat cggaggaccg aaggagctaa ccgctttttt gcacaacatg ggggatcatg	2640
taactcgctt tgatcgttgg gaaccggagc tgaatgaagc cataccaaac gacgagcgtg	2700
acaccacgat gctgtagca atggcaacaa cgttcgcaa actattaact ggcaactac	2760
ttactctagc ttcccgcaa caattaatag actggatgga ggcgataaa gttgcaggac	2820
cacttctgcg ctcgccctt ccggttggtt ggtttattgc tgataaatct ggagccggtg	2880
agcgtgggtc tcgcggtatc attgcagcac tggggccaga tggtaagccc tcccgtatcg	2940
tagttatcta cagacgggg agtcaggcaa ctatggatga acgaaataga cagatcgctg	3000
agataggtgc ctactgatt aagcattggt aactgtcaga ccaagtttac tcatatatac	3060
tttagattga tttaaaactt catttttaat ttaaaaggat ctaggtgaag atcctttttg	3120
ataatctcat gacaaaaac ccttaacgtg agttttcgtt ccactgagcg tcagaccccg	3180
tagaaaagat caaaggatct tcttgagatc ctttttttct gcgcgtaac tgctgcttgc	3240
aaacaaaaa accaccgcta ccagcgggtg tttgtttgcc ggatcaagag ctaccaactc	3300
tttttcgcaa ggtaactggc ttcagcagag cgcagatacc aaatactgtc cttctagtgt	3360
agccgtagtt aggccaccac ttcaagaact ctgtagcacc gcctacatac ctgctctgc	3420
taatcctgtt accagtggct gctgccagtg gcgataagtc gtgtcttacc gggttggact	3480
caagacgata gttaccggtt aaggcgcagc ggtcgggctg aacggggggt tcgtgcacac	3540
agcccagctt ggagcgaacg acctacaccg aactgagata cctacagcgt gagctatgag	3600
aaagcggcac gcttccgaa gggagaaagg cggacaggta tccgtaagc ggcagggtcg	3660
gaacaggaga gcgcacagg gagcttccag ggggaaacgc ctggtatctt tatagtcttg	3720
tcgggtttcg ccacctctga cttgagcgtc gattttttgt atgctcgtca ggggggcgga	3780
gcctatggaa aaacgccagc aacgcggcct ttttacggtt cctggccttt tgctggcctt	3840

ttgctcacat gtcttttctt gcgttatccc ctgattctgt ggataaccgt attaccgcct 3900
 ttgagtgagc tgataccgct cgccgcagcc gaacgaccga gcgcagcgag tcagtgagcg 3960
 aggaagcggg agagcgccca atacgaaac cgctctccc cgcgcttgg ccgattcatt 4020
 aatgcagctg gcacgacagg tttcccagct ggaaagcggg cagtgcgcgc aacgcaatta 4080
 atgtgagtta gctcactcat taggcacccc aggctttaca ctttatgctt ccggctcgta 4140
 tgttgtgtgg aattgtgagc ggataacaat ttcacacagg aaacagctat gaccatgatt 4200
 acgccaagct tgcattgcctg caggctgact ctagag 4236

<210> 269

<211> 17

<212> DNA

<213> Artificial sequence

<220><223> M13 forward primer

<400> 269

gtaaaacgac ggccagt 17

<210> 270

<211> 22

<212> DNA

<213> Artificial sequence

<220><223> M13 reverse primer

<400> 270

cacacaggaa acagctatga cc 22

<210> 271

<211> 7523

<212> DNA

<213> Artificial sequence

<220><223> Construct

<400> 271

ccagcttttg ttcccttttag tgagggttaa ttgcgcgctt ggcgtaatca tggatcatagc 60

tgtttcctgt gtgaaattgt tatccgctca caattccaca caacatagga gccggaagca 120

taaagtgtaa agcctggggg gcctaattgag tgaggtaact cacattaatt gcgttgcgct 180

cactgcccgc ttccagctcg ggaaacctgt cgtgccagct gcattaatga atcgccaac 240

gcgcggggag aggcggtttg cgtattgggc gctcttcgc ttcctcgctc actgactcgc 300

tgcgctcggt cgttcggctg cggcgagcgg tatcagctca ctcaaaggcg gtaatacgg 360

tatccacaga atcaggggat aacgcaggaa agaacaatgtg agcaaaaaggc cagcaaaaagg	420
ccaggaaccg taaaaaggcc gcgttgctgg cgtttttcca taggctccgc cccctgacg	480
agcatcacia aaatcgacgc tcaagtcaga ggtggcgaaa cccgacagga ctataaagat	540
accaggcggtt tccccctgga agctccctcg tgcgctctcc tgttccgacc ctgccgtta	600
ccggatacct gtccgccttt ctcccttcgg gaagcgtggc gctttctcat agctcacgt	660
gtaggtatct cagttcgggtg taggtcggtc gtcceaagct gggctgtgtg cacgaacccc	720
ccgttcagcc cgaccgtgc gccttatccg gtaactatcg tcttgagtcc aaccggtaa	780
gacacgactt atcgccactg gcagcagcca ctgtaacag gattagcaga gcgaggtatg	840
taggcgggtgc tacagagttc ttgaagtggg ggcctaacta cggctacact agaaggacag	900
tatttggat ctgcgctctg ctgaagccag ttaccttcgg aaaaagagtt gtagctctt	960
gatccggcaa acaaaccacc gctggtagcg gtggtttttt tgtttgcaag cagcagatta	1020
cgcgcaaaa aaaaggatct caagaagatc ctttgatctt ttctacgggg tctgacgctc	1080
agtggaacga aaactcacgt taagggattt tggatcatgag attatcaaaa aggatcttca	1140
cctagatcct tttaaattaa aaatgaagtt ttaaatcaat ctaaagtata tatgagtaaa	1200
cttggctga cagttaccaa tgcttaatca gtgaggcacc tatctcagcg atctgtctat	1260
ttcgttcac catagttgcc tgactccccg tcgtgtagat aactacgata cgggagggt	1320
taccatctgg cccagtgct gcaatgatac cgcgagacc acgctcacg gctccagatt	1380
tatcagcaat aaaccagcca gccgaaggc ccgagcgag aagtggctct gcaactttat	1440
ccgcctccat ccagttctatt aattgttgcc gggaagctag agtaagtagt tcgccagtta	1500
atagtttgcg caacgttgtt gccattgcta caggcatcgt ggtgtcacgc tcgtcgtttg	1560
gtatggcttc attcagctcc ggttcccaac gatcaaggcg agttacatga tccccatgt	1620
tgtgcaaaaa agcggttagc tccttcggtc ctccgatcgt tgtcagaagt aagttggccg	1680
cagtgttacc actcatgggt atggcagcac tgcataattc tcttactgtc atgcatccg	1740
taagatgctt tctgtgact ggtgagtact caaccaagtc attctgagaa tagtgtatgc	1800
ggcgaccgag ttgtctttgc ccggcgtcaa tacgggataa taccgcgcca catagcagaa	1860
ctttaaaagt gctcatcatt ggaaaacgtt cttcggggcg aaaactctca aggatcttac	1920
cgctgttag atccagttcg atgtaacca ctctgtcacc caactgatct tcagcatctt	1980
ttactttcac cagcggtttct gggtagagaa aacaggaag gcaaaatgcc gcaaaaaagg	2040
gaataagggc gacacggaaa tgttgaatac tcatactctt ctttttcaa tattattgaa	2100

gcatttatca gggttattgt ctcatgagcg gatacatatt tgaatgtatt tagaaaaata 2160

aacaaatagg ggttccgcgc acatttcccc gaaaagtgcc acctgaacga agcatctgtg 2220

cttcattttg tagaacaaaa atgcaacgcg agagcgctaa tttttcaaac aaagaatctg 2280

agctgcattt ttacagaaca gaaatgcaac gcgaaagcgc tatttacca acgaagaatc 2340

tgtgcttcat ttttgtaaaa caaaaatgca acgcgagagc gctaattttt caaacaaga 2400

atctgagctg catttttaca gaacagaaat gcaacgcgag agcgctattt taccaacaaa 2460

gaatctatac ttcttttttg ttctacaaaa atgcatcccg agagcgctat ttttctaaca 2520

aagcatctta gattactttt tttctccttt gtgcgctcta taatgcagtc tcttgataac 2580

ttttgcact gtaggtccgt taaggttaga agaaggctac tttggtgtct attttctctt 2640

ccataaaaaa agcctgactc cacttcccgc gtttactgat tactagcgaa gctgcgggtg 2700

cattttttca agataaaggc atccccgatt atattctata ccgatgtgga ttgcgcatac 2760

tttgtgaaca gaaagtata gcgttgatga ttcttcattg gtcagaaaaat tatgaacggt 2820

ttcttctatt ttgtctctat atactacgta taggaaatgt ttacattttc gtattgtttt 2880

cgattcactc tatgaatagt tcttactaca atttttttgt ctaaagagta atactagaga 2940

taaacataaa aaatgtagag gtcgagttta gatgcaagtt caaggagcga aagggtggatg 3000

ggtaggttat atagggatat agcacagaga tatatagcaa agagatactt ttgagcaatg 3060

tttgtggaag cggtattcgc aatattttag tagctcgtaa cagtcgggtg cgtttttggt 3120

tttttgaaag tgcgtcttca gacgctttt ggttttcaaa agcgctctga agttcctata 3180

ctttctagag aataggaact tcggaatagg aacttcaaag cgtttccgaa aacgagcgct 3240

tccgaaaatg caacgcgagc tgcgcacata cagctcactg ttcacgtcgc acctatatct 3300

gcgtgttgcc tgtatatata tatacatgag aagaacggca tagtgcgtgt ttatgcttaa 3360

atgcgtactt atatgcgtct atttatgtag gatgaaaggt agtctagtag ctcctgtgat 3420

attatcccat tccatgcggg gtatcgtag ctctcttcag cactaccctt tagctgttct 3480

atatgtgcc actcctcaat tggattagtc tcatecttca atgctatcat ttcttttgat 3540

attggatcat ctaagaaacc attattatca tgacattaac ctataaaaat aggcgtatca 3600

cgaggccctt tcgtctcgcg cgtttcgggtg atgacgggtga aaacctctga cacatgcagc 3660

tcccggagac ggtcacagct tgtctgtaag cggatgccgg gagcagacaa gcccgtcagg 3720

gcgcgtcagc ggggtgttggc ggggtgcggg gctggcttaa ctatgcggca tcagagcaga 3780

ttgtactgag agtgcacat aaattcccgt ttttaagagct tggtagcgc taggagtcac 3840

tgccaggat cgtttgaaca cggcattagt caggggaagtc ataacacagt cttttccgc	3900
aattttcttt ttctattact cttggcctcc tctagtagac tctatatttt ttatgcctc	3960
ggtaatgatt ttcatttttt tttttccct agcgatgac tcttttttt tcttagcgat	4020
tggcattatc acataatgaa ttatacatta tataaagtaa tgtgatttct tcgaagaata	4080
tactaaaaa tgagcaggca agataaacga aggcaaagat gacagagcag aaagccctag	4140
taaagcgtat tacaaatgaa accaagattc agattgcgat ctctttaag ggtggcccc	4200
tagcgataga gcactcgatc ttcccagaaa aagaggcaga agcagtagca gaacaggcca	4260
cacaatcgca agtgattaac gtccacacag gtatagggtt tctggaccat atgatacatg	4320
ctctggccaa gcattccggc tggctcgctaa tcgttgagtg cattgggtgac ttacacatag	4380
acgaccatca caccactgaa gactgcggga ttgctctcgg tcaagctttt aaagaggccc	4440
tactggcgcg tggagtaaaa aggtttggat caggatttgc gcctttggat gaggcacttt	4500
ccagagcggg ggtagatctt tcgaacaggc cgtacgcagt tgtcgaactt ggtttgcaaa	4560
gggagaaaagt aggagatctc tcttgcgaga tgatcccgca ttttcttgaa agctttgcag	4620
aggctagcag aattaccctc cacgttgatt gtctgcgagg caagaatgat catcaccgta	4680
gtgagagtgc gttcaaggct cttgcggttg ccataagaga agccacctcg cccaatggta	4740
ccaacgatgt tccctccacc aaagggttgc ttatgtagt acaccgatta tttaaagctg	4800
cagcatagca tatatatata tgtgtatata tgtataccta tgaatgtcag taagtatgta	4860
tacgaacagt atgatactga agatgacaag gtaatgcac attctatacg tgcattctg	4920
aacgaggcgc gctttccttt tttctttttg ctttttcttt tttttctct tgaactcgac	4980
ggatctatgc ggtgtgaaat accgcacaga tgcgtaagga gaaaataccg catcaggaaa	5040
ttgtaaacgt taatattttg ttaaaattcg cgttaaattt ttgttaaac agctcatttt	5100
ttaaccaata ggccgaaatc ggcaaaatcc cttataaatc aaaagaatag accgagatag	5160
ggttgagtgt tgttccagtt tggaacaaga gtccactatt aaagaacgtg gactccaacg	5220
tcaaaggcgc aaaaaccgtc tatcagggcg atggccact acgtgaacca tcaccctaat	5280
caagtttttt ggggtcgagg tgccgtaaag cactaaatcg gaaccctaaa gggagcccc	5340
gatttagagc ttgacgggga aagccggcga acgtggcgag aaaggaaggg aagaaagcga	5400
aaggagcggg cgctagggcg ctggcaagt tagcgggtcac gctgcgcta accaccacac	5460
ccgccgcgt taatgcgccg ctacagggcg cgtcgcgcca ttcgccattc aggctgcgca	5520
actgttggga agggcgatcg gtgcgggcct cttcgtatt acgccagctg gcgaaagggg	5580
gatgtgctgc aaggcgatta agttgggtaa cgccagggtt ttccagtcga cgacgttgta	5640
aaacgacggc cagtgagcgc gcgtaatacg actcactata gggcgaattg ggtaccgggc	5700

ccccccctga ggiattagaa gccgccgagc gggcgacagc cctccgacgg aagactctcc 5760
 tccgtgcgtc ctgctcttca ccggtcgcgt tcctgaaacg cagatgtgcc tcgcgccgca 5820
 ctgctccgaa caataaagat tctacaatac tagcttttat ggttatgaag aggaaaaatt 5880
 ggcagtaacc tggccccaca aaccttcaaa ttaacgaatc aaattaacaa ccataggatg 5940

 ataatgcgat tagtttttta gccttatttc tggggttaatt aatcagcgaa gcgatgattt 6000
 ttgatctatt aacagatata taaatggaaa agctgcataa ccactttaac taatactttc 6060
 aacattttca gtttgtatta cttcttattc aaatgtcata aaaglatcaa caaaaaattg 6120
 ttaataatacc tctatacttt aacgtcaagg agaaaaatgt ccaatttact gcccgtaac 6180
 caaaatttgc ctgcattacc ggtcgatgca acgagtgatg aggttcgcaa gaacctgatg 6240
 gacatgttca gggatcgcca ggcgttttct gagcatacct ggaaaatgct tctgtccgtt 6300
 tgccggtcgt gggcggtcatg gtgcaagttg aataaccgga aatggtttcc cgcagaacct 6360

 gaagatgttc gcgattatct tctatatctt caggcgcgcg gtctggcagt aaaaactatc 6420
 cagcaacatt tgggccagct aaacatgctt catcgtcggc cggggtgcc acgaccaagt 6480
 gacagcaatg ctgtttcact ggttatgcgg cggatccgaa aagaaaacgt tgatgccggt 6540
 gaacgtgcaa aacaggctct agcggtcgaa cgcactgatt tcgaccaggt tcgttactc 6600
 atggaaaata gcgatcgctg ccaggatata cgtaatctgg ctttctggg gattgcttat 6660
 aacacctgt tacgtatagc cgaaattgcc aggatcaggg ttaaagatat ctcacgtact 6720
 gacggtggga gaatgttaat ccatattggc agaacgaaa cgctgggttag caccgcaggt 6780

 gtagagaagg cacttagcct gggggttaact aaactggtcg agcgatggat ttccgtctct 6840
 ggtgtagctg atgatccgaa taactacctg ttttgccggg tcagaaaaaa tgggtgtgcc 6900
 gcgccatctg ccaccagcca gctatcaact cgcgccctgg aagggtttt tgaagcaact 6960
 catcgattga tttacggcgc taaggatgac tctggtcaga gatacctggc ctggtctgga 7020
 cacagtgccg gtgtcggagc cgcgcgagat atggcccgcg ctggagtttc aataccggag 7080
 atcatgcaag ctggtggctg gaccaatgta aatattgtca tgaactatat ccgtaacctg 7140
 gatagtgaag caggggcaat ggtgcgcctg ctggaagatg gcgattagga gtaagcgaat 7200

 ttcttatgat ttatgatttt tattattaaa taagttataa aaaaaataag tgtatacaaa 7260
 ttttaaagtg actcttaggt tttaaaacga aaattcttat tcttgagtaa ctctttcctg 7320
 taggtcaggt tgcttttcca ggtatagcat gaggtcgtc ttattgacca cacctctacc 7380
 ggcatgccga gcaaatgcct gcaaatcgct cccatttca cccaattgta gatatgctaa 7440
 ctccagcaat gagttagatg atctcggtgt gtattttatg tcctcagagg acaacacctg 7500
 tggtcgccca ccgcggtgga gct 7523

<210> 272

<211> 47

<212> DNA

<213> Artificial sequence

<220><223> upper primer for pYZ058 construction

<400> 272

catcatcaca gtttaaacag tatgttgaag caaatcaact tcggtgg 47

<210> 273

<211> 48

<212> DNA

<213> Artificial sequence

<220><223> lower primer for pYZ058 construction

<400> 273

ggacgggccc tgcaggcctt attggttttc tggctcgaac tttctgac 48

<210> 274

<211> 1005

<212> DNA

<213> Clostridium beijerinckii

<400> 274

atgccaaaga tgtactacga aaaggacact gatttgaact tgttgagagg taagaaggtc 60

gccattatcg gttacggttc ccaaggtcac gctcatgcct tgaacttgca cgaatctggt 120

gtcgacgttg tcgttggttt gtacaacggc tctaagtcct gggctaaggc tgaagctgct 180

ggtctgcaag tcgccaccgt tgctgacgct gccaaaggctg ccgatttgat catgattttg 240

ctgccagatg aaaagcaagc taaaatctac aacgaagaaa ttgctccaaa ctggaagaa 300

ggtaatgcct tggttttcgc tcacggtttc aacatccatt tcggtcaagt cgttccacca 360

tcctacgtcg acgttttcat ggtcgtccca aagggtccag gtcacttggt tagacgtacc 420

tacactgaag gtgctgggtg cccatgtttg atcgtgttc atcaagatgc cactggtaaa 480

gctaaacagt acgttttggc ctacgctaac ggtattggtg gtgccagagc tgggtgtctg 540

gaaaccacct tcaaggacga aaccgaaacc gacttggtcg gtgaacaagc tgtcttatgt 600

ggtggtgttt ctgaattgat caaggtggt tttgaaacct tggtcgaagc cggttacgct 660

ccagaaaacg cttacttcga atgcatgcac gaaatgaagt tgattgttga ctgtttgtac 720

caaggtgggt tatccatgat gagatactct atctccgaca ctgctgaata cgggtgactac 780

cagatttgta gaagaatcat taccgacgaa accaagaagg aaatgaagaa ggttttgacc 840
gaaatccaag atggtacttt cgctaagaac tggctgttgg aaaaccagac caatagacca 900

ggtttcaacg ccagaagaag aatggaagct gaacacccaa tcgaaaaggt tggtaaggaa 960
ttgagaggta tgatgtcctg gatcgacacc gctaagggtg attaa 1005

<210> 275

<211> 334

<212> PRT

<213> Clostridium beijerinckii

<400> 275

Met Pro Lys Met Tyr Tyr Glu Lys Asp Thr Asp Leu Asn Leu Leu Arg

1 5 10 15

Gly Lys Lys Val Ala Ile Ile Gly Tyr Gly Ser Gln Gly His Ala His

20 25 30

Ala Leu Asn Leu His Glu Ser Gly Val Asp Val Val Val Gly Leu Tyr

35 40 45

Asn Gly Ser Lys Ser Trp Ala Lys Ala Glu Ala Ala Gly Leu Gln Val

50 55 60

Ala Thr Val Ala Asp Ala Ala Lys Ala Ala Asp Leu Ile Met Ile Leu

65 70 75 80

Leu Pro Asp Glu Lys Gln Ala Lys Ile Tyr Asn Glu Glu Ile Ala Pro

85 90 95

Asn Leu Glu Glu Gly Asn Ala Leu Val Phe Ala His Gly Phe Asn Ile

100 105 110

His Phe Gly Gln Val Val Pro Pro Ser Tyr Val Asp Val Phe Met Val

115 120 125

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

130 135 140

Ala Gly Val Pro Cys Leu Ile Ala Val His Gln Asp Ala Thr Gly Lys

145 150 155 160

Ala Lys Gln Tyr Ala Leu Ala Tyr Ala Asn Gly Ile Gly Gly Ala Arg

165 170 175

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

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

290 295 300
 Arg Arg Arg Met Glu Ala Glu His Pro Ile Glu Lys Val Gly Lys Glu
 305 310 315 320
 Leu Arg Gly Met Met Ser Trp Ile Asp Thr Ala Lys Val Asp
 325 330

<210> 276

<211> 1020

<212> DNA

<213> *Zymomonas mobilis*

<400> 276

atgaaagttt actacgattc tgacgccgac ctgggcctca tcaaattctaa gaaaattgcc 60

atcctcggct acggttccca gggccacgct catgcacaga acctgcgcga ttccggcgta 120

gcggaagtgg ctattgcctt gcgtccagac tcggcttccg ttaagaaggc tcaagacgcc 180

ggcttcaaag tgctcaccaa tgctgaagcc gctaagtggg ccgacatcct catgatcctg 240

gccccggacg agcaccagge cgcatctac gcggaggacc tcaaggacaa cctccgcccc 300

ggctccgcta tcgccttcgc ccacggcctc aacatccact tcggcctgat tgagcctcgt 360

aaggacatcg acgtgttcat gatcgcgccg aagggtccag gccacaccgt acgttccgag 420

tacgtccgcg gcggtgggtg tccttgccctg gtcgcggtcg accaggacgc ttccggcaac 480
 gcccacgaca ttgccctggc ctacgcctca ggcatcggcg gcggccgttc cggcgtcac 540

gaaaccacct tccgcgagga gggtgaaacc gacctgttcg gtgagcaggc tgttctgtgt 600
 ggcggtctca ccgcgtgat cacagccggt ttcgaaacc tgactgaagc tggctacgt 660
 ccagaaatgg ccttcttcga gtgcatgcac gaaatgaagc tgatcgttga cctcatctac 720
 gaagccggta tcgccaacat gcgctactcg atctccaaca ccgctgaata cggcgacatc 780
 gtgtctggtc cacgcgtcat caacgaggaa tcaaaaaagg ccatgaaggc tattctggac 840
 gacatccaga gcggccgctt cgtgtcgaag ttcgtcctcg acaaccgcgc cggccagcca 900
 gagttgaagg ccgcgcgtaa acgcatggct gctcacccta tcgaacaggt cggcgcgcgc 960

ctgcgcaaga tgatgccttg gatcgttcc aacaaactcg tcgacaaagc ccgcaactaa 1020

<210> 277

<211> 339

<212> PRT

<213> *Zymomonas mobilis*

<400> 277

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

1 5 10 15

Lys Lys Ile Ala Ile Leu Gly Tyr Gly Ser Gln Gly His Ala His Ala

20 25 30

Gln Asn Leu Arg Asp Ser Gly Val Ala Glu Val Ala Ile Ala Leu Arg

35 40 45

Pro Asp Ser Ala Ser Val Lys Lys Ala Gln Asp Ala Gly Phe Lys Val

50 55 60

Leu Thr Asn Ala Glu Ala Ala Lys Trp Ala Asp Ile Leu Met Ile Leu

65 70 75 80

Ala Pro Asp Glu His Gln Ala Ala Ile Tyr Ala Glu Asp Leu Lys Asp

85 90 95

Asn Leu Arg Pro Gly Ser Ala Ile Ala Phe Ala His Gly Leu Asn Ile

100 105 110

His Phe Gly Leu Ile Glu Pro Arg Lys Asp Ile Asp Val Phe Met Ile

115 120 125

Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser Glu Tyr Val Arg Gly

130 135 140

Gly Gly Val Pro Cys Leu Val Ala Val Asp Gln Asp Ala Ser Gly Asn

145 150 155 160

Ala His Asp Ile Ala Leu Ala Tyr Ala Ser Gly Ile Gly Gly Gly Arg

165 170 175

Ser Gly Val Ile Glu Thr Thr Phe Arg Glu Glu Val Glu Thr Asp Leu

180 185 190

Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Leu Thr Ala Leu Ile Thr

195 200 205

Ala Gly Phe Glu Thr Leu Thr Glu Ala Gly Tyr Ala Pro Glu Met Ala

210 215 220

Phe Phe Glu Cys Met His Glu Met Lys Leu Ile Val Asp Leu Ile Tyr

225 230 235 240

Glu Ala Gly Ile Ala Asn Met Arg Tyr Ser Ile Ser Asn Thr Ala Glu

245 250 255

Tyr Gly Asp Ile Val Ser Gly Pro Arg Val Ile Asn Glu Glu Ser Lys

260 265 270

Lys Ala Met Lys Ala Ile Leu Asp Asp Ile Gln Ser Gly Arg Phe Val

275 280 285

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

290 295 300

Ala Arg Lys Arg Met Ala Ala His Pro Ile Glu Gln Val Gly Ala Arg

305 310 315 320

Leu Arg Lys Met Met Pro Trp Ile Ala Ser Asn Lys Leu Val Asp Lys

325 330 335

Ala Arg Asn

<210> 278

<211> 267

<212> PRT

<213> Saccharomyces cerevisiae

<400> 278

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> 279
<211> 500
<212> PRT
<213> *Saccharomyces cerevisiae*
<400> 279
Met Thr Lys Leu His Phe Asp Thr Ala Glu Pro Val Lys Ile Thr Leu
1 5 10 15
Pro Asn Gly Leu Thr Tyr Glu Gln Pro Thr Gly Leu Phe Ile Asn Asn
20 25 30
Lys Phe Met Lys Ala Gln Asp Gly Lys Thr Tyr Pro Val Glu Asp Pro
35 40 45
Ser Thr Glu Asn Thr Val Cys Glu Val Ser Ser Ala Thr Thr Glu Asp
50 55 60
Val Glu Tyr Ala Ile Glu Cys Ala Asp Arg Ala Phe His Asp Thr Glu
65 70 75 80
Trp Ala Thr Gln Asp Pro Arg Glu Arg Gly Arg Leu Leu Ser Lys Leu
85 90 95
Ala Asp Glu Leu Glu Ser Gln Ile Asp Leu Val Ser Ser Ile Glu Ala
100 105 110
Leu Asp Asn Gly Lys Thr Leu Ala Leu Ala Arg Gly Asp Val Thr Ile
115 120 125
Ala Ile Asn Cys Leu Arg Asp Ala Ala Ala Tyr Ala Asp Lys Val Asn
130 135 140
Gly Arg Thr Ile Asn Thr Gly Asp Gly Tyr Met Asn Phe Thr Thr Leu
145 150 155 160
Glu Pro Ile Gly Val Cys Gly Gln Ile Ile Pro Trp Asn Phe Pro Ile
165 170 175

Met Met Leu Ala Trp Lys Ile Ala Pro Ala Leu Ala Met Gly Asn Val

180 185 190
Cys Ile Leu Lys Pro Ala Ala Val Thr Pro Leu Asn Ala Leu Tyr Phe

195 200 205
Ala Ser Leu Cys Lys Lys Val Gly Ile Pro Ala Gly Val Val Asn Ile

210 215 220
Val Pro Gly Pro Gly Arg Thr Val Gly Ala Ala Leu Thr Asn Asp Pro
225 230 235 240

Arg Ile Arg Lys Leu Ala Phe Thr Gly Ser Thr Glu Val Gly Lys Ser

245 250 255
Val Ala Val Asp Ser Ser Glu Ser Asn Leu Lys Lys Ile Thr Leu Glu

260 265 270
Leu Gly Gly Lys Ser Ala His Leu Val Phe Asp Asp Ala Asn Ile Lys

275 280 285
Lys Thr Leu Pro Asn Leu Val Asn Gly Ile Phe Lys Asn Ala Gly Gln

290 295 300
Ile Cys Ser Ser Gly Ser Arg Ile Tyr Val Gln Glu Gly Ile Tyr Asp

305 310 315 320
Glu Leu Leu Ala Ala Phe Lys Ala Tyr Leu Glu Thr Glu Ile Lys Val

325 330 335
Gly Asn Pro Phe Asp Lys Ala Asn Phe Gln Gly Ala Ile Thr Asn Arg

340 345 350
Gln Gln Phe Asp Thr Ile Met Asn Tyr Ile Asp Ile Gly Lys Lys Glu

355 360 365
Gly Ala Lys Ile Leu Thr Gly Gly Glu Lys Val Gly Asp Lys Gly Tyr

370 375 380
Phe Ile Arg Pro Thr Val Phe Tyr Asp Val Asn Glu Asp Met Arg Ile

385 390 395 400
Val Lys Glu Glu Ile Phe Gly Pro Val Val Thr Val Ala Lys Phe Lys

405 410 415
Thr Leu Glu Glu Gly Val Glu Met Ala Asn Ser Ser Glu Phe Gly Leu

420 425 430
Gly Ser Gly Ile Glu Thr Glu Ser Leu Ser Thr Gly Leu Lys Val Ala

435 440 445
Lys Met Leu Lys Ala Gly Thr Val Trp Ile Asn Thr Tyr Asn Asp Phe

450 455 460
Asp Ser Arg Val Pro Phe Gly Gly Val Lys Gln Ser Gly Tyr Gly Arg

465 470 475 480
Glu Met Gly Glu Glu Val Tyr His Ala Tyr Thr Glu Val Lys Ala Val

485 490 495
Arg Ile Lys Leu

500

<210> 280

<211> 804

<212> DNA

<213> *Saccharomyces cerevisiae*

<400> 280

atgtccaag gtagaaaagc tgcagaaaga ttggctaaga agactgtcct cattacaggt 60
gcatctgctg gtattggtaa ggcgaccgca ttagagtact tggaggcatc caatggtgat 120
atgaaactga tcttggctgc tagaagatta gaaaagctcg aggaattgaa gaagaccatt 180
gatcaagagt ttccaaacgc aaaagtcat gtggcccagc tggatatcac tcaagcagaa 240
aaaatcaagc ccttcattga aaacttgcca caagagtca aggatattga cattctggtg 300
aacaatgccg gaaaggctct tggcagtgac cgtgtgggcc agatcgcaac ggaggatata 360

caggacgtgt ttgacaccaa cgtcacggct ttaatcaata tcacacaagc tgtactgccc 420
atattccaag ccaagaattc aggagatatt gtaaatttgg gttcaatcgc tggcagagac 480
gcatacccaa caggttctat ctattgtgcc tctaagttg ccgtgggggc gttcactgat 540
agtttgagaa aggagctcat caacactaaa attagagtca ttctaattgc accagggcta 600
gtcgagactg aattttcact agttagatac agaggtaacg aggaacaagc caagaatgtt 660
tacaaggata ctacccatt gatggctgat gacgtggctg atctgatcgt ctatgcaact 720
tccagaaaac aaaatactgt aattgcagac actttaatct ttccaacaaa ccaagcgtca 780

cctcatcata tcttcggtgg ataa 804

<210> 281

<211> 30
 <212> DNA
 <213> Artificial sequence
 <220><223> N1214 primer
 <400> 281
 aaaaaggggg tagaagcgtt ttgaagctat 30
 <210> 282
 <211> 62
 <212> DNA
 <213> Artificial sequence
 <220><223> N1176 primer
 <400> 282
 gcatagcaat ctaatctaag ttccagctga ggatgacaac agattactca tcaccagcat 60
 at 62
 <210> 283
 <211> 62
 <212> DNA
 <213> Artificial sequence
 <220><223> N1117 primer
 <400> 283
 atcaacacac aaacactaaa tcaaagctga ggatggattt attttagtca ttagcacaaa 60
 aa 62
 <210> 284
 <211> 800
 <212> DNA
 <213> *Saccharomyces cerevisiae*
 <400> 284
 ccactagtca gatgccgcgg gcacttgagc acctcatgca cagcaataac acaacacaat 60
 ggtagtagc aacctgaatt cggtcattga tgcattgatg tgccgtgaag cgggacaacc 120
 agaaaagtcg tctataaatg ccggcacgtg cgatcatcgt ggccggggtt taagagtgc 180
 tatcacaaat tgtgcatta ccgcggaacc gccagatatt cattacttga cgcaaaagcg 240
 ttgaaataa tgacgaaaaa gaaggaagaa aaaaaaagaa aaataccgct tctaggcggg 300
 ttatctactg atccgagctt ccactaggat agcacccaaa cacctgcata ttggacgac 360

ctttacttac accacaaaa accactttcg cctctcccg cctgataac gtccactaat 420
 tgagcgatta cctgagcggg cctcttttgt ttgcagcatg agacttgcat actgcaaate 480
 gtaagtagca acgtctcaag gtcaaaactg tatggaaacc ttgtcacctc acttaattct 540
 agctagccta cctgcaagt caagaggtct ccgtgattcc tagccacctc aaggatatgcc 600

tctccccgga aactgtggcc ttttctggca cacatgatct ccacgatttc aacatataaa 660
 tagcttttga taatggcaat attaataaaa tttattttac ttctttcttg taacatctct 720
 cttgtaatcc cttattcctt ctagctattt ttcataaaaa accaagcaac tgcttatcaa 780
 cacacaaaca ctaaataaaa 800

<210> 285
 <211> 44
 <212> DNA
 <213> Artificial sequence
 <220><223> N822 primer
 <400> 285

cgccctcagct ttgatttagt gtttgtgtgt tgataagcag ttgc 44

<210> 286
 <211> 63

<212> DNA
 <213> Artificial sequence
 <220><223> N1178 primer
 <400> 286

ggtatcgata agcttgatat cgaattcctg cgccccgggc actagtcaga tgccgcgggc 60
 act 63

<210> 287
 <211> 38
 <212> DNA
 <213> Artificial sequence
 <220><223> N821 primer
 <400> 287

cgccccgggc actagtcaga tgccgcgggc acttgagc 38

<210> 288
 <211> 30
 <212> DNA

<213> Artificial sequence

<220><223> N1115 primer

<400> 288

tttttgtgct aatgactcaa ataaatccat 30

<210> 289

<211> 30

<212> DNA

<213> Artificial sequence

<220><223> N1110 primer

<400> 289

gcgatttaat ctctaattat tagttaaagt 30

<210> 290

<211> 30

<212> DNA

<213> Artificial sequence

<220><223> N1114 primer

<400> 290

atatgctggg gatgagtaat ctgttgatcat 30

<210> 291

<211> 9220

<212> DNA

<213> Artificial sequence

<220><223> pUC19-URA3::pdc1::TEF(M4)-xpk1::kan Plasmid

<400> 291

ccgcattgcg gattacgtat tctaattgtc agataacttc gtataatgta tgctatacga 60

agttatcgaa cagagaaact aaatccacat taattgagag ttctatctat tagaaaatgc 120

aaactccaac taaatgggaa aacagataac ctcttttatt tttttttaat gtttgatatt 180

cgagtctttt tcttttgta gggtttatatt catcatttca atgaataaaa gaagcttctt 240

attttggttg caaagaatga aaaaaaagga ttttttcata cttctaaagc ttcaattata 300

accaaaaatt ttataaatga agagaaaaaa tctagtagta tcaagttaaa cttagaaaaa 360

ctcatcgagc atcaaatgaa actgcaattt attcatatca ggattatcaa taccatattt 420

ttgaaaaagc cgtttctgta atgaaggaga aaactcaccg aggcagttcc ataggatggc 480

aagatcctgg tatcggctcg cgattccgac tcgtccaaca tcaatacaac ctattaattt 540
ccctcgtca aaaataaggt tatcaagtga gaaatcacca tgagtgacga ctgaatccgg 600
tgagaatggc aaaagcttat gcatttcttt ccagacttgt tcaacaggcc agccattacg 660
ctcgtcatca aaatcactcg catcaaccaa accgttattc attcgtgatt gcgcctgagc 720
gagacgaaat acgcgatcgc tgttaaaagg acaattacaa acaggaatcg aatgcaaccg 780

gcgcaggaac actgccagcg catcaacaat attttcacct gaatcaggat attcttctaa 840
tacctggaat gctgttttgc cggggatcgc agtggtgagt aacctgcat catcaggagt 900
acggataaaa tgcttgatgg tcggaagagg cataaattcc gtcagccagt ttagtctgac 960
catctcatct gtaacatcat tggcaacgct acctttgcca tgtttcagaa acaactctgg 1020
cgcatcgggc ttcccataca atcgatagat tctcgcacct gattgcccgga cattatcgcg 1080
agcccattta taccatata aatcagcatc catgttggaa tttaatcgcg gcctcgaaac 1140
gtgagtcctt tccttaccca tctcgagttt taatgttact tctcttgagc ttagggaact 1200

ataatgtaac tcaaaataag attaaacaaa ctaaaataaa aagaagttat acagaaaaac 1260
ccatataaac cagtactaat ccataataat aatacacaaa aaaactatca aataaaacca 1320
gaaaacagat tgaatagaaa aattttttcg atctcctttt atattcaaaa ttcgatatat 1380
gaaaaagga actctcagaa aatcaccaaa tcaatttaat tagatttttc ttttccttct 1440
agcgttggaa agaaaaattt tttttttttt ttttagaaat gaaaaatttt tgccgtagga 1500
atcacggtat aaacctgta taaacgtac tctgttcacc tgtgtaggct atgattgacc 1560
cagtgttcat tgttattgag agagagcggg agaaaagaac cgatacaaga gatccatgct 1620

ggtatagttg tctgtccaac actttgatga actttagga cgatgatgtg tatttagacg 1680
agtacgtgtg tgactattaa gtagttatga tagagaggtt tgtacggtgt gttctgtgta 1740
attcgattga gaaaatggtt atgaatccct agataacttc gtataatgta tgctatacga 1800
agttatccag tgatgataca acgagttagc caaggtgggg gatcctctag agtcttaagg 1860
ccgccgcaa attaaagcct tcgagcgtcc caaaaccttc tcaagcaagg ttttcagtat 1920
aatgttacat gcgtacacgc gttgttacag aaaaaaaga aaaatttgaa atataaataa 1980
cgttcttaat actaacataa ctattaaaaa aaataaatag ggacctagac ttcaggttgt 2040

ctaactcctt ccttttcggt tagagcggat gtgggaggag ggcgatgaatg taagcgtgac 2100
ataactaatt acatgattaa ttaattattt taaaccttc cattgccaat cattaacttc 2160
tggaagtca gttccggcat cccgatata ggcatgtgt ttagcaagca tattatccat 2220
ggattgaacg aaggccgcac cagtgttttc cattgctggt tgcgccgcaa ttgccgactt 2280
agctaagtcg aagcgggtcca tctggttcat gacctgacg tcgaatggtg tggtaatatc 2340

accattttca cggtaaccgt ggacgtataa gttatggttg tgacgatcaa agaagatgtc	2400
acgaactaag tcttcgtaac cgtggaaagc aaagaccact ggtttgcct tagtaaagta	2460
atgggtcaaac tcagcatctg acaagccccg cggatccttt tcaggactac gtaacttcaa	2520
gatgtcgacc acgttcacga aacgaatctt catctctggg aaactgtcgt gtagtaattg	2580
gatggcagcc aacgtttcaa gcgttgggtc cgtcccagca gctgcaaaga caatgtctgg	2640
ttcgctacct tggctccgtac ttgcccaatc aatgatacca agaccattgt caactaattg	2700
cttagcttct tcaatgctga accattgttg acgtgggtgt tttgacgtaa ccacgtagtt	2760
gatcttttct tggctccgga aaatgacgtc accgacagct aataacgtgt tggcatcggc	2820
tggtaaaat tcaacgaatgt attctggttt ctttccggcc aaatgagtta atgcacctgg	2880
atcttgggtg gtataacat tatggtcttg ttggaataca gttgaagccg cgataatgtt	2940
aagtgatggg tactttttac gccaatcaag ttcatgggt ttacgtaacc acttgaagtg	3000
ttgcgtcaac attgagtcca caacgcgtag gaaggcttca taactggcaa ataaccatg	3060
acgtccagtt aagacgtaac cttctaacca accttcagct tgggtttcag ataactgagc	3120
atctaagacc cggccagctg gtgcttcata ttggtcacta tctggatgaa tgtcttccat	3180
ccattgacga ttagtgggtt cgaagacacc atataaacgg ttagacatgg tttcatcagg	3240
tccgaacaac cggaagtat caggattttt cttgatgaca tcccgcaaat agtctgacca	3300
aacgatcata tcttgcttaa cattcgcgcc ttctttggac gtatcgaccg cataatcacg	3360
gaagtttggg aagtccaagg ctttcggatc gacccacca ttggtgattg ggtagcagc	3420
catccgactg tccccagtag gaataatttc ttaatatca tcttcaaag agccatcttc	3480
attgaagagt tcttttgggt gatatgattc gagccaatca actaaagcat ccgcatgttc	3540
catgtcattt tgatcaacag gaatcggaat ttgatgagca cggaatgaac cttcgatctt	3600
atcacgtcc catgacttcg gaccagtcca gcccttaggt gcgcggaaga cgatcattgg	3660
ccatactggc aatgttgcat cgttattttc gcgagcatgc ttctggattg cttgatctt	3720
ttcaacggct tcattcatgg ccttagctaa ggtgggtga accttttcag gatcgtcacc	3780
ttcaacgaag attggttccc aattcatgct ttcgaagtat tccttaatct tagcatcaga	3840
agtcgacca aaaatcgttg gattagaaat cttaaaacca ttttaagtca agattggtaa	3900
aacagccccg tcgttgattg ggtaaatgaa ctctgttgat tgccatgaag ttgctaattg	3960
acccgtttcg gattcccat caccaacaac aaccggcg atttcgtcag gattgtcaag	4020
aattgcccc accccgtgtg aaattgagta accaagttcg ccaccttcgt ggattgaacc	4080

gggtgtttca ggtgccgcat gggaagcaac cccacctggg aatgagaatt gcttgaagag 4140

cttttgcac ccttcaacat cctgcgtaat ttctggataa atatcggtgt aagtaccgtc 4200

aaggtaagag tttagaaacca tcaattgacc accatgacct ggaccttcaa cgtagaacat 4260

cttcaaaccg tacttgttga tgaccgggtt aagatgagca tagataaagt tttagaccggc 4320

aatcgteccc cagtaccaa ttggatgaac cttaacgtca ctggccttca atggccgttg 4380

taatagtga ttatctttta aataaagttg accaactgat aagtagttgg cagcacgcca 4440

gtacttatca actttttgca aatatgctgg tgatgagtaa tctgttgtca tcctcagctg 4500

gaacttagat tagattgcta tgctttctct ctaacgagca agaagtaaaa aaagttgtaa 4560

tagaacaaga aaaatgaaac tgaagcttga gaaattgaag accgtttatt agcttaaata 4620

tcaatgggag gtcacgaaa gagaaaaaaa tcaagaaaga aactctcaag aaaaagaaac 4680

gtgataaaaa tttttattgc ctctctcgac gaagagaaag aaacgaggcg gtcccttttt 4740

tcttttccaa accttttagta cgggtaattt gcgacaccct agaggaagaa agaggggaaa 4800

tttagtatgc tgtgcttggg tgtcttgaag tggtagggcg atgcgcggag tccgagaaaa 4860

tctggaagag taaaaggggg gtagaagcgt tttagaagcta tccgcggtgg ttaagcctaa 4920

ccaggccaat tcaacagact gtcggcaact tcttgtctgg tctttccatg gtaagtgaca 4980

gtgcagtaat aatatgaacc aattttttt tcgttacata aaaatgctta taaaacttta 5040

actaataatt agagattaaa tcgcaaaccg ccggccaatg tggctgtggt ttcagggtcc 5100

ataaagcttt tcaattcatc ttttttttt ttgttctttt ttttgattcc ggtttctttg 5160

aaattttttt gattcggtaa tctccgagca gaaggaagaa cgaaggaagg agcacagact 5220

tagattggta tatatacgca tatgtgggtg tgaagaaaca tgaattgcc cagtattctt 5280

aacccaactg cacagaacaa aaacctgcag gaaacgaaga taaatcatgt cgaaagctac 5340

atataaggaa cgtgctgcta ctcatcctag tctgttgct gccaaagctat ttaatatcat 5400

gcacgaaaag caaacaact tgtgtgcttc attggatgtt cgtaccacca aggaattact 5460

ggagttagtt gaagcattag gtcccaaat ttgtttacta aaaacacatg tggatatctt 5520

gactgatttt tccatggagg gcacagttaa gccgctaaag gcattatccg ccaagtacaa 5580

ttttttactc ttggaagaca gaaaatttgc tgacattggg aatacagtca aattgcagta 5640

ctctgcgggt gtatacagaa tagcagaatg ggacagacatt acgaatgcac acggtgtggt 5700

gggccagggt attgttagcg gtttgaagca ggcggcggaa gaagtaacaa aggaacctag 5760

aggccttttg atgttagcag aattgtcatg caagggtcc ctagctactg gagaatatac 5820

taagggtact gttgacattg cgaagagcga caaagatfff gttatcggtt ttattgctca	5880
aagagacatg ggtggaagag atgaagggtta cgattgggtg attatgacac ccggtgtggg	5940
tttagatgac aaggagagacg cattgggtca acagtataga accgtggatg atgtggtctc	6000
tacaggatct gacattatta ttgttgggaag aggactatff gcaaaggga gggatgctaa	6060
ggtagagggt gaacgtttaca gaaaagcagg ctgggaagca tttttgagaa gatgcggcca	6120
gcaaaactaa aaaactgtat tataagtaaa tgcatgtata ctaaactcac aaattagagc	6180
ttcaatttaa ttatatcagt tattaccggg gaatctcggt cgtaatgatt tctataatga	6240
cgaaaaaaa aaaattggaa agaaaaagct tcatggcctt gcggccgctt aattaatcta	6300
gagtcgacct gcaggcatgc aagcttggcg taatcatggt catagctgtt tcctgtgtga	6360
aattgttate cgctcacaat tccacacaac atacgagccg gaagcataaa gtgtaaagcc	6420
tgggtgtcct aatgagttag ctaactcaca ttaattgcgt tgcgtcact gcccgctttc	6480
cagtcgggaa acctgtcgtg ccagctgcat taatgaatcg gccaacgcgc ggggagaggc	6540
ggtttgcgta ttgggcgctc ttccgcttcc tcgtcactg actcgtcgc ctcggtcgtt	6600
cggtcgggc gagcgggtac agctcactca aaggcggtaa tacggttatc cacagaatca	6660
ggggataacg caggaaagaa catgtgagca aaaggccagc aaaaggccag gaaccgtaaa	6720
aaggccgctg tctgtggcgtt ttcccatagg ctccgcccc ctgacgagca tcacaaaaat	6780
cgacgtcaa gtcagagggt gcgaaacccg acaggactat aaagatacca ggcgtttccc	6840
cctggaagct cctcgtcgc ctctcctgtt ccgacctgc cgcttaccgg atacctgtcc	6900
gcctttctcc cttcgggaag cgtggcgtt tctcatagct cacgtgttag gtatctcagt	6960
tcggtgtagg tcgttcgctc caagctgggc tgtgtgcacg aacccccgt tcagcccgac	7020
cgctgcgct tatccggtaa ctatcgtctt gagtccaacc cggtagaca cgacttatcg	7080
ccactggcag cagccactgg taacaggatt agcagagcga ggtatgtagg cgggtctaca	7140
gagttcttga agtgggtggc taactacggc tacactagaa ggacagtatt tggatctgc	7200
gctctgctga agccagttac cttcggaaaa agagttggtg gctcttgatc cggcaaaaa	7260
accaccgtg gtagcgggtg ttttttgtt tgcaagcagc agattacgcg cagaaaaaaa	7320
ggatctcaag aagatccttt gatcttttct acggggtctg acgctcagt gaacgaaaac	7380
tcacgttaag ggattttgt catgagatta tcaaaaagga tcttcaccta gatcctttta	7440
aattaaaaat gaagttttta atcaatctaa agtatatatg agtaaacttg gtctgacagt	7500
taccaatgct taatcagtga ggcacctatc tcagcgatct gtctatttcg ttcattcata	7560
gttgctgac tccccgtcgt gtagataact acgatacggg agggcttacc atctggcccc	7620
agtgtgcaa tgataccgcg agaccacgc tcaccggctc cagatttatc agcaataaac	7680

cagccagccg gaagggccga ggcgagaagt ggtcctgcaa ctttatccgc ctccatccag 7740
tctattaatt gttgccggga agctagagta agtagttcgc cagttaatag ttgcgcaac 7800
gttgttgcca ttgctacagg catcgtggtg tcacgctcgt cgtttggtat ggcttcattc 7860
agctccggtt cccaacgata aaggcgagtt acatgatccc ccatgttgtg caaaaaagcg 7920

gtagctcct tcggtcctcc gatcgttgtc agaagtaagt tggccgcagt gttatcactc 7980
atggttatgg cagcactgca taattctctt actgtcatgc catccgtaag atgcttttct 8040
gtgactgggt agtactcaac caagtcattc tgagaatagt gtatgcggcg accgagttgc 8100
tcttgcccg cgtcaatacg ggataatacc gcgccacata gcagaacttt aaaagtgtc 8160
atcattggaa aacgttcttc ggggcgaaaa ctctcaagga tcttaccgct gttgagatcc 8220
agttcgatgt aaccactcg tgcaccaac tgatcttcag catcttttac tttcaccagc 8280
gtttctgggt gagcaaaaac aggaaggcaa aatgccgcaa aaaagggaat aaggcgaca 8340

cggaaatgtt gaatactcat actcttctt tttcaatatt attgaagcat ttatcagggt 8400
tattgtctca tgagcggata catatttgaa tgtatttaga aaaataaaca aataggggtt 8460
ccgcgcacat ttccccgaaa agtgccacct gacgtctaag aaaccattat tatcatgaca 8520
ttaacctata aaaataggcg tatcacgagg ccctttcgtc tcgcgcgttt cggatgatgac 8580
ggtgaaaacc tctgacacat gcagctcccg gagacggtca cagcttgtct gtaagcggat 8640
gccgggagca gacaagcccc tcagggcgcg tcagcgggtg ttggcgggtg tcggggctgg 8700
cttaactatg cggcatcaga gcagattgta ctgagagtgc accatatgcg gtgtgaaata 8760

ccgcacagat gcgtaaggag aaaataccgc atcaggcgcc attcgccatt caggctgcgc 8820
aactgttggg aaggcgcatc ggtgcgggcc tcttcgtat tacgccagct ggcgaaaggg 8880
ggatgtgctg caaggcgatt aagtgggtg acgccagggt tttccagtc acgacgttgt 8940
aaaacgacgg ccagtgaatt cgagctcggg acccggggat ccggcgcgcc gttttatttg 9000
tatcgagggt tctagtcttc tattacacta atgcagtttc agggttttgg aaaccacact 9060
gtttaaacag tgttctttaa tcaaggatac ctcttttttt ttccttggtt ccactaatc 9120
atcggttttt tttttggaag acatcttttc caacgaaaag aatatacata tcgtttaaga 9180

gaaattctcc aaatttgtaa agaagcggac ccagacttaa 9220

<210> 292
<211> 38
<212> DNA
<213> artificial sequence
<220><223> primer BK468

<400> 292
gcctcgagtt ttaatgttac ttctcttgca gttaggga 38

<210> 293

<211> 10934

<212> DNA

<213> Artificial sequence

<220><223> pRS423::TEF(M4)-xpk1+EN01-eutD

<400> 293

ggtggagctc cagcttttgt tccctttagt gagggttaat tgcgcgcttg gcgtaatcat 60
ggtcatagct gtttcctgtg tgaaattgtt atccgctcac aattccacac aacataggag 120

ccggaagcat aaagtgtaaa gcctggggtg cctaagtgt gaggttaact acattaattg 180
cgttgcgctc actgcccgtt ttccagtcgg gaaacctgtc gtgccagctg cattaatgaa 240
tcggccaacg cgcggggaga ggcggtttgc gtattgggcg ctcttcgctt tctcgtctca 300
ctgactcgct gcgctcggtc gttcggctgc ggcgagcggt atcagctcac tcaaaggcgg 360
taatacgggt atccacagaa tcaggggata acgcaggaaa gaacatgtga gcaaaaggcc 420
agcaaaaggc caggaaccgt aaaaaggccg cgttgctggc gttttccat aggctccgcc 480
cccctgacga gcatcacaaa aatcgacgtt caagtcagag gtggcgaaac ccgacaggac 540

tataaagata ccaggcgttt ccccttgaa gctccctcgt gcgctctcct gttccgacct 600
tgccgcttac cggatacctg tccgcctttc tcccttcggg aagcgtggcg cttttcata 660
gtcacgctg taggtatctc agttcgggtg aggtcgttcg ctccaagctg ggctgtgtgc 720
acgaaccccc cgttcagccc gaccgtcgc ccttatccgg taactatcgt cttgagtcca 780
accggtaag acacgactta tcgccactgg cagcagccac tggtaacagg attagcagag 840
cgaggatagt aggcgggtgt acagagttct tgaagtgggt gcctaactac ggctacacta 900
gaaggacagt atttggtatc tgcgctctgc tgaagccagt taccttcgga aaaagagttg 960

gtagctcttg atccggcaaa caaaccaccg ctggtagcgg tggttttttt gtttgcaagc 1020
agcagattac gcgcagaaaa aaaggatctc aagaagatcc tttgatcttt tctacggggt 1080
ctgacgtca gtggaacgaa aactcacgtt aagggttttt ggtcatgaga ttatcaaaaa 1140
ggatcttcac ctgatctctt ttaaattaaa aatgaagttt taaatcaatc taaagtatat 1200
atgagtaaac ttggtctgac agttaccaat gcttaatcag tgaggcacct atctcagcga 1260
tctgtctatt tcgttcaccc atagttgcct gactccccgt cgtgtagata actacgatac 1320
gggagggttt accatctggc cccagtgtg caatgatacc gcgagacca cgctcaccgg 1380

ctccagatatt atcagcaata aaccagccag ccggaagggc cgagcgcaga agtggtcctg	1440
caactttatc cgcctccatc cagtctatta attgttgccg ggaagctaga gtaagtagtt	1500
cgccagttaa tagtttgccg aacgttggtg ccattgctac aggcacgtg gtgtcacgct	1560
cgctgtttgg taaggcttca ttacgtccg gttcccaacg atcaaggcga gttacatgat	1620
cccccatgtt gtgcaaaaaa gcggttagct ccttcgggcc tccgatcgtt gtcagaagta	1680
agttggccgc agtgttatca ctcatggtta tggcagcact gcataattct ctactgtca	1740
tgccatccgt aagatgcttt tctgtgactg gtgagtactc aaccaagtca ttctgagaat	1800
agtgtatgcg gcgaccgagt tgctcttgcc cggcgtcaat acgggataat accgcgccac	1860
atagcagaac tttaaaagtg ctcatcattg gaaaacgttc ttccggggcga aaactctcaa	1920
ggatcttacc gctgttgaga tccagttcga tgtaaccac tcgtgcacc aactgatctt	1980
cagcatcttt tactttcacc agcgtttctg ggtgagcaaa aacaggaagg caaatgccg	2040
caaaaaaggg aataagggcg acacggaaat gttgaatact catactcttc ctttttcaat	2100
attattgaag catttatcag ggttattgtc tcatgagcgg atacatattt gaatgtattt	2160
agaaaaataa acaaataggg gttccgcgca catttcccg aaaagtgcc cctgaacgaa	2220
gcatctgtgc ttcatattgt agaacaaaaa tgcaacgcga gagcgctaatt tttcaaaca	2280
aagaatctga gctgcatttt tacagaacag aaatgcaacg cgaaagcgt attttaccaa	2340
cgaagaatct gtgcttcatt ttgtaaaaac aaaaatgcaa cgcgagagcg ctaatttttc	2400
aaacaagaa tctgagctgc atttttacag aacagaaatg caacgcgaga gcgctatatt	2460
accaacaaag aatctatact tcttttttgt tctacaaaaa tgcatcccga gagcgctatt	2520
tttctaaca agcatcttag attacttttt ttctccttg tgcgctctat aatgcagtct	2580
cttgataact tttgactg taggtccgtt aaggtagaa gaaggctact ttggtgtcta	2640
ttttctcttc cataaaaaaa gcctgactcc acttcccgcg ttactgatt actagcgaag	2700
ctgcgggtgc attttttcaa gataaaggca tccccgatta tattctatac cgatgtggat	2760
tgcgcatact ttgtgaacag aaagtgatag cgttgatgat tcttcattgg tcagaaaatt	2820
atgaacggtt tcttctattt tgtctctata tactacgtat aggaaatgtt tacattttcg	2880
tattgttttc gattcactct atgaatagtt ctactacaa tttttttgtc taaagagtaa	2940
tactagagat aaacataaaa aatgtagagg tcgagtttag atgcaagttc aaggagcgaa	3000
aggtggatgg gtaggttata tagggatata gcacagagat atatagcaaa gagatacttt	3060
tgagcaatgt ttgtggaagc ggtattcgca atattttagt agctcgttac agtccgggtgc	3120
gtttttgggt ttttgaaagt gcgtcttcag agcgtttttg gttttcaaaa gcgctctgaa	3180
gttcctatac ttctagaga ataggaactt cggaatagga acttcaaagc gtttccgaaa	3240

acgagcgctt ccgaaaatgc aacgcgagct gcgcacatac agctcactgt tcacgtcgca 3300

cctatatctg cgtgttgctt gtatatatat atacatgaga agaacggcat agtgcgtgtt 3360

tatgcttaaa tgcgtactta tatgcgtcta tttatgtagg atgaaaggta gtctagtacc 3420

tcctgtgata ttatcccatt ccatgcgggg tatcgtatgc ttccttcagc actacccttt 3480

agctgttcta tatgctgcca ctctcaatt ggattagtct catccttcaa tgctatcatt 3540

tcctttgata ttggatcatc taagaaacca ttattatcat gacattaacc tataaaaata 3600

ggcgtatcac gaggcccttt cgtctcgcgc gtttcggtga tgacggtgaa aacctctgac 3660

acatgcagct cccggagacg gtcacagctt gtctgtaagc ggatgccggg agcagacaag 3720

cccgtcaggg cgcgtcagcg ggtgttggcg ggtgtcgggg ctggcttaac tatgcggcat 3780

cagagcagat tgiactgaga gtgcaccata aattcccgtt ttaagagctt ggtgagcgt 3840

aggagtcact gccaggtatc gtttgaacac ggcattagtc agggaagtca taacacagtc 3900

ctttcccgca attttctttt tctattactic ttggcctcct ctagtacact ctatatTTTT 3960

ttatgcctcg gtaatgattt tcatTTTTTT ttttcccta gcggatgact cttttTTTT 4020

cttagcgatt ggcattatca cataatgaat tatacattat ataaagtaat gtgatttctt 4080

cgaagaatat actaaaaaat gagcaggcaa gataaacgaa ggcaaagatg acagagcaga 4140

aagccctagt aaagcgtatt acaaatgaaa ccaagattca gattgcgac tctttaaagg 4200

gtggtccct agcgatagag cactcgatct tccagaaaa agaggcagaa gcagtagcag 4260

aacaggccac acaatcgcaa gtgattaacg tccacacagg tatagggttt ctggaccata 4320

tgatacatgc tctggccaag cattccggct ggtcgctaata cgttgagtgc attggtgact 4380

tacacataga cgaccatcac accactgaag actgcgggat tgctctcggt caagctttta 4440

aagaggccct actggcgctt ggagtaaaaa ggtttggatc aggatttgcg cttttggatg 4500

aggcactttc cagagcggtg gtagatcttt cgaacaggcc gtacgcagtt gtcgaacttg 4560

gtttgcaaag ggagaaagta ggagatctct ctgcgagat gatcccgcat tttcttgaaa 4620

gctttgcaga ggctagcaga attaccctcc acgttgattg tctgcgagc aagaatgac 4680

atcacgtag tgagagtgcg ttcaaggctc ttgcggttgc cataagagaa gccacctgc 4740

ccaatggtac caacgatgtt ccctccacca aaggtgttct tatgtagtga caccgattat 4800

ttaaagctgc agcatacgtat atatatatat gtgtatatat gtatactat gaatgtcagt 4860

aagtatgtat acgaacagta tgatactgaa gatgacaagg taatgcatca ttctatcgt 4920

gtcattctga acgaggcgcg ctttcctttt ttctttttgc tttttcttt tttttcttt 4980

gaactcgacg gatctatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc 5040

atcaggaaat tgtaaactgt aatatTTTgt taaaattcgc gttaaatttt tgtaaataca 5100

gtcattttt taaccaatag gccgaaatcg gcaaaatccc ttataaatca aaagaataga 5160

ccgagatagg gttgagtgtt gttccagttt ggaacaagag tccactatta aagaacgtgg 5220

actccaacgt caaagggcga aaaaccgtct atcagggcga tggcccaacta cgtgaacat 5280

caccctaatac aagttttttg gggtcgaggt gccgtaaagc actaaatcgg aaccctaaag 5340

ggagcccccg atttagagct tgacggggaa agccggcgaa cgtggcgaga aaggaaggga 5400

agaaagcgaa aggagcgggc gctagggcgc tggcaagtgt agcggtcacg ctgcgcgtaa 5460

ccaccacacc cgccgcgtt aatgcgccgc tacagggcgc gtcgcgccat tcgccattca 5520

ggctgcgcaa ctgttgggaa gggcgatcgg tgcgggcctc ttcgctatta cgccagctgg 5580

cgaaaggggg atgtgctgca aggcgattaa gttgggtaac gccagggttt tcccagtcac 5640

gacgttgtaa aacgacggcc agtgagcgcg cgtaatacga ctactatag ggcgaattgg 5700

gtaccgggcc cccctcgag gtcgacggtg tcgataagct tgatatcgaa ttcctgcgcc 5760

cgggccacta gtcagatgcc gcgggcactt gagcacctca tgcacagcaa taacacaaca 5820

caatggttag tagcaacctg aattcgggtc ttgatgcatg catgtgccgt gaagcgggac 5880

aaccagaaaa gtcgtctata aatgccggca cgtgcgatca tcgtggcggg gttttaagag 5940

tgcatatcac aaattgtcgc attaccgcgg aaccgccaga tttcattac ttgacgcaa 6000

agcgtttgaa ataatgacga aaaagaagga agaaaaaaaa agaaaaatac cgcttctagg 6060

cgggttatct acigatccga gcttccacta ggatagcacc caaacacctg catatttgga 6120

cgacctttac ttacaccacc aaaaaccact ttcgcctctc ccgccctga taacgtccac 6180

taattgagcg attacctgag cggtcctctt ttgtttgcag catgagactt gcatactgca 6240

aatcgtaagt agcaacgtct caaggtcaaa actgtatgga aaccttgtca cctacttaa 6300

ttctagctag cctacctgc aagtcaagag gtctccgtga ttcctagcca cctcaaggta 6360

tgcctctccc cgaaaactgt ggccttttct ggcacacatg atctccacga tttcaacata 6420

taaatagctt ttgataatgg caatattaat caaatattatt ttacttcttt ctgtaacat 6480

ctctcttgta atcccttatt cttctagct atttttcata aaaaaccaag caactgtta 6540

tcaacacaca aacactaaat caaagctgag gatggattta tttgagtcac tagcacaaaa 6600

aattactggt aaagatcaaa caattgtttt ccctgaagga actgaacccc gaattgtcgg 6660

tgcggcagcg cgattagctg cagacggctt ggttaagccg attgttttag gtgcaacgga 6720

caaagttcag gcgtggctc acgatttgaa tgcggattta acaggcgttc aagtccttga 6780

tcctgcgaca taccggctg aagataagca agcaatgctt gatgccctcg ttgaacggcg 6840

gaaaggtgaag aatagccag aacaagcggc taaaatgctg gaagatgaaa actactttgg	6900
cacgatgctc gtttataatg gcaaagcggg tgggatgggt tcaggtgcaa tccatccaac	6960
tggatgatac gtacggccag cgttacaaat tattaagacc aagcccgggt cacaccgaat	7020
ctcgggtgca tttatcatgc aaaaggggtga ggaacgctac gtctttgctg actgtgccat	7080
caatatgatg cccgatgccg atacgttagc ggaaattgcc actcagagtg cggctactgc	7140
taaggctctc gatattgacc cgaaagtgtc gatgctcagc ttctcaacta aggggttcggc	7200
taagggtgaa atggctacta aagtgaaga agcaacggcc aaggcgcaag ctgctgaacc	7260
ggaattggct atcgatgggt aacttcaatt tgacgcggcc ttcgttgaaa aagttgggtt	7320
gcaaaaggct cctgggtcca aagtagctgg tcatgccaat gtctttgtat ttccagagct	7380
tcagtctggt aatattggct ataagattgc gcaacgattt ggctattttg aagcgggtggg	7440
tcctgtcttg caaggcctga acaagccggt ctccgacttg tcacgtggat gcagtgaaga	7500
agacgtttat aaggttgcga ttattacagc agcccaagga ttagcttaat taattaagag	7560
taagcgaatt tcttatgatt tatgattttt attattaaat aagttataaa aaaaataagt	7620
gtatacaaat tttaaagtga ctcttaggtt ttaaaacgaa aattcttatt cttagtaac	7680
tctttcctgt aggtcaggtt gctttctcag gtatagcatg aggtcgtctt tattgaccac	7740
acctctaccg gcatgccgag caaatgcctg caaatcgtc cccatttcac ccaattgtag	7800
atatgctaac tcagcaatg agttgatgaa tctcgggtgtg tattttatgt cctcagagga	7860
caacacctgt ggtactagtt cttagcggc cgcccgaat taaagcctt cgagcgtccc	7920
aaaaccttct caagcaaggt ttccagtata atgttacatg cgtacacgcg ttgtacaga	7980
aaaaaaaaaa aaatttgaaa tataaataac gttcttaata ctaacataac tattaaaaaa	8040
aataaatagg gacctagact tcaggttgct taactccttc ctttcgggtt agagcggatg	8100
tgggaggagg gcgtgaatgt aagcgtgaca taactaatta catgattaat taattatttt	8160
aaacctttcc attgccaatc attaatctct ggcaagtcag ttccggcatc ccgatatag	8220
gcattgtgtt tagcaagcat attatccatg gattgaacga aggccgcacc agtgttttcc	8280
attgctgggt gcgcccgaat tgccgactta gctaagtcga agcgggccat ctgggtcatg	8340
accgtacgt cgaatgggtg ggtaatatca ccattttcac ggtaaccgtg gacgtataag	8400
ttatgggtgt gacgatcaaa gaagatgtca cgaactaagt cttcgtaac gtggaaagca	8460
aagaccactg gttgtcctt agtaaagtaa tggtaaaact cagcatctga caagccccgc	8520
ggatcctttt caggactacg taacttcaag atgtcgacca cgttcacgaa acgaatcttc	8580
atctctggga aactgtcgtg tagtaattgg atggcagcca acgtttcaag cgttggttcc	8640
gtcccagcag ctgcaaagac aatgtctggt tcgctacctt ggtccgtact tgcccaatca	8700

atgataccaa gaccattgtc aactaattgc ttagcttctt caatgctgaa ccattgttga	8760
cgtgggtgtt ttgacgtaac cacgtagtgt atcttttctt ggctccgaa aatgacgtca	8820
ccgacagcta ataacgtgtt ggcatcggtt ggtaaatatt cacgaatgta ttctgttttc	8880
ttttcgcca aatgagttaa tgcacctgga tcttgggtgg tataaccatt atggtcttgt	8940
tggaatacag ttgaagccgc gataatgtta agtgaagggt actttttacg ccaatcaagt	9000
tcattggctt tacgtaacca cttgaagtgt tgcgtcaaca ttgagtcac aacgcgtagg	9060
aaggcttcat aactggcaaa taacctatga cgtccagtta agacgtaacc ttctaacca	9120
ccttcagctt ggtgttcaga taactgagca tctaagacct ggccagctgg tgcctcatat	9180
tggctactat ctggatgaat gtcttccatc cattgacgat tagtgggttc gaagacacca	9240
tataaacggt tagacatggt ttcatcaggt ccgaacaacc ggaagttatc aggatttttc	9300
ttgatgacat cccgcaaata gtctgaccaa acgatcatat cttgcttaac attcgcgcct	9360
tctttggacg tatcgaccgc ataatcacgg aagtttggta agttcaaggc ttccgcatcg	9420
acccaccat tggtgattgg gttagcagcc atccgactgt cccagtagg aataatttct	9480
ttaatatcat cttcaaaga gccatcttca ttgaagagtt ctttgggttg atatgattcg	9540
agccaatcaa ctaaagcatc cgcattgtcc atgtcatttt gatcaacagg aatcggaatt	9600
tgatgagcac ggaatgaacc ttcatctta tcaccgtccc atgacttcgg accagtccag	9660
cccttaggtg cgcggaagac gatcattggc catactggca atgttgcatc gttattttcg	9720
cgagcatgct tctggattgc cttgatcttt tcaacggctt catccatggc cttagctaag	9780
gctgggtgaa ctttttcagg atcgtcacct tcaacgaaga ttggttccca attcatgctt	9840
tcgaagtatt cttaatctt agcatcagaa gtccgaccaa aaatcggttg attagaaatc	9900
ttaaaacat ttaagttcaa gattggtaaa acagccccgt cgttgattgg gttaatgaac	9960
ttcgttgatt gccatgaagt tgctaattga cccgtttcgg attccccatc accaacaaca	10020
accgaggcga ttctgtcagg attgtcaaga attgccccaa cccgtgtga aattgagtaa	10080
ccaagtgcg caccttcgtg gattgaaccg ggtgtttcag gtgccgatg ggaagcaacc	10140
ccacctggga atgagaattg cttgaagagc ttttgcatcc cttcaacatc ctgcgtaatt	10200
tctggataaa tatcgggtga agtaccgtca agtaagagt ttgaaacat cacttgacca	10260
ccatgacctg gaccttcaac gtagaacatc ttcaaaccgt acttgttgat gaccgggtta	10320
agatgagcat agataaagtt ttgaccggca atcgtcccc agtgaccaat tggatgaacc	10380
ttaacgtcac tggccttcaa tggccgttgt aatagtggat tatcttttaa ataaagttga	10440
ccaactgata agtagttggc agcacgccag tacttatcaa ctttttgcaa atatgctggt	10500
gatgagtaat ctgttgtcat cctcagctgg aacttagatt agattgctat gctttctctc	10560

taacgagcaa gaagtaaaaa aagttgtaat agaacaagaa aaatgaaact gaagcttgag 10620

aaattgaaga ccgtttatta gcttaaatat caatgggagg tcatcgaaag agaaaaaaat 10680

caagaaagaa actctcaaga aaaagaaacg tgataaaaat ttttattgcc tctctcgacg 10740

aagagaaaga aacgaggcgg tccctttttt cttttccaaa cctttagtac gggttaattag 10800

cgacacccta gaggaagaaa gaggggaaat ttagtatgct gtgcttgggt gtcttgaagt 10860

ggtacggcga tgcgcggagt ccgagaaaat ctggaagagt aaaaaggggg tagaagcgtt 10920

ttgaagctat ccgc 10934

<210> 294

<211> 80

<212> DNA

<213> Artificial sequence

<220><223> Primer

<400> 294

ttccggtttc ttgaaattt ttttgattcg gtaatctccg agcagaagga gcattgcgga 60

ttacgtattc taatgttcag 80

<210> 295

<211> 81

<212> DNA

<213> Artificial sequence

<220><223> Primer

<400> 295

gggtaataac tgatataatt aaattgaagc tctaatttgt gagtttagta caccttggt 60

aactcgttgt atcatcactg g 81

<210> 296

<211> 38

<212> DNA

<213> Artificial sequence

<220

><223> Primer

<400> 296

gcctcgagtt ttaatgttac ttctcttgca gttaggga 38

<210> 297

<211> 31

<212> DNA
 <213> Artificial sequence
 <220><223> Primer
 <400> 297
 gctaaattcg agtgaaacac aggaagacca g 31
 <210> 298
 <211> 24
 <212> DNA
 <213> Artificial sequence
 <220><223> Primer
 <400> 298
 ttctcgacgt gggccttttt cttg 24
 <210> 299
 <211> 49
 <212> DNA
 <213> Artificial sequence
 <220><223> Primer
 <400> 299
 tgcagcttta aataatcggg gtcactactt tgccttcgtt tatcttgcc 49
 <210> 300
 <211> 49
 <212> DNA
 <213> Artificial sequence
 <220><223> Primer
 <400> 300
 gagcaggcaa gataaacgaa ggcaaagtag tgacaccgat tatttaaag 49
 <210> 301
 <211> 49
 <212> DNA
 <213> Artificial sequence
 <220><223> Primer
 <400> 301
 tatggaccct gaaaccacag ccacattgta accaccacga cggttgttg 49
 <210> 302

<211> 49

<212> DNA

<213> Artificial sequence

<220><223> Primer

<400> 302

ttttagcaaca accgtcgtgg tggttacaat gtggctgtgg tttcagggt 49

<210> 303

<211> 49

<212> DNA

<213> Artificial sequence

<220><223> Primer

<400> 303

ccagaaaccc tatacctgtg tggacgtaag gccatgaagc tttttcttt 49

<210> 304

<211> 49

<212> DNA

<213> Artificial sequence

<220><223> Primer

<400> 304

attggaaaga aaaagcttca tggccttacg tccacacagg tatagggtt 49

<210> 305

<211> 22

<212> DNA

<213> Artificial sequence

<220><223> Primer

<400> 305

cataagaaca cctttggtgg ag 22

<210> 306

<211> 22

<212> DNA

<213> Artificial sequence

<220><223> Primer

<400> 306

aggattatca ttcataagtt tc	22
<210> 307	
<211> 23	
<212> DNA	
<213> Artificial sequence	
<220><223> Primer	
<400> 307	
ttcttggagc tgggacatgt ttg	23
<210> 308	
<211> 22	
<212> DNA	
<213> Artificial sequence	
<220><223> Primer	
<400> 308	
tgatgatatt tcataaataa tg	22
<210> 309	
<211> 23	
<212> DNA	
<213> Artificial sequence	
<220><223> Primer	
<400> 309	
atgcgtccat ctttacagtc ctg	23
<210> 310	
<211> 22	
<212> DNA	
<213> Artificial sequence	
<220><223> Primer	
<400> 310	
tacgtacgga ccaatcgaag tg	22
<210> 311	
<211> 49	
<212> DNA	
<213> Artificial sequence	

<220><223> Primer
 <400> 311
 aattcgtttg agtacactac taatggcttt gttggcaata tgtttttgc 49
 <210> 312
 <211> 49
 <212> DNA
 <213> Artificial sequence
 <220><223> Primer
 <400> 312
 atatagcaaa aacatattgc caacaaagcc attagtagtg tactcaaac 49
 <210> 313
 <211> 49
 <212> DNA
 <213> Artificial sequence
 <220><223> Primer
 <400> 313
 tatggaccct gaaaccacag ccacattctt gttatttata aaaagacac 49
 <210> 314
 <211> 49
 <212> DNA
 <213> Artificial sequence
 <220><223> Primer
 <400> 314
 ctcccggtgc tttttataaa taacaagaat gtggctgtgg tttcagggt 49
 <210> 315
 <211> 49
 <212> DNA
 <213> Artificial sequence
 <220><223> Primer
 <400> 315
 taccgtaggc gtccttagga aagatagaag gccatgaagc tttttcttt 49
 <210> 316
 <211> 49
 <212> DNA

<213> Artificial sequence

<220><223> Primer

<400> 316

attggaaaga aaaagcttca tggccttcta tctttcctaa ggacgccta 49

<210> 317

<211> 21

<212> DNA

<213> Artificial sequence

<220><223> Primer

<400> 317

ttattgtttg gcattttag c 21

<210> 318

<211> 22

<212> DNA

<213> Artificial sequence

<220><223> Primer

<400> 318

ccaagcatct cataaaccta tg 22

<210> 319

<211> 22

<212> DNA

<213> Artificial sequence

<220><223> Primer

<400> 319

tgtgcagatg cagatgtgag ac 22

<210> 320

<211> 17

<212> DNA

<213> Artificial sequence

<220><223> Primer

<400> 320

agttattgat accgtac 17

<210> 321

<211> 19

<212> DNA

<213> Artificial sequence

<220><223> Primer

<400> 321

cgagataccg taggcgtcc 19

<210> 322

<211> 2237

<212> DNA

<213> Artificial sequence

<220><223> Fragment

<400> 322

ttatgtatgc tcttctgact tttcgtgtga tgaggctcgt ggaaaaaatg aataatttat 60

gaatttgaga acaattttgt gttgttacgg tattttacta tggaataatc aatcaattga 120

ggattttatg caaatatcgt ttgaatattt ttccgaccct ttgagtactt ttcttcataa 180

ttgcataata ttgtccgctg cccctttttc tgtagacgg tgtcttgatc tacttgctat 240

cgttcaacac caccattatt tctaactatt ttttttttag ctcatittgaa tcagcttatg 300

gtgatggcac atttttgcat aaacctagct gtccctcgtt aacataggaa aaaaaatat 360

ataaacaagg ctctttcact ctccctgcaa tcagatttgg gtttgttccc tttattttca 420

tatttcttgt catattcctt tctcaattat tattttctac tcataacctc acgcaaaata 480

acacagtcaa atcaatcaaa atgactgaca aaaaaactct taaagactta agaaatcgta 540

gttctgttta cgattcaatg gttaaatac ctaatcgtgc tatgttgcgt gcaactggta 600

tgcaagatga agactttgaa aaacctatcg tcggtgtcat ttcaacttgg gctgaaaaca 660

caccttgtaa tatccactta catgactttg gtaaactagc caaagtcggt gttaaggaag 720

ctggtgcttg gccagttcag ttcggaacaa tcacggtttc tgatggaatc gccatgggaa 780

cccaaggaat gcgtttctcc ttgacatctc gtgatattat tgcagattct attgaagcag 840

ccatgggagg tcataatgcg gatgcttttg tagccattgg cggttgtgat aaaaacatgc 900

ccggttctgt tatcgctatg gctaacatgg atatccagc catttttgct tacggcggaa 960

caattgcacc tggtaattta gacggcaaag atatcgattt agtctctgtc tttgaagggtg 1020

tcggccattg gaaccacggc gatatgacca aagaagaagt taaagctttg gaatgtaatg 1080

cttgtcccgg tcttgagggc tgcggtggta tgtatactgc taacacaatg gcgacagcta 1140

ttgaagtttt gggacttagc cttccgggtt catcttctca cccggctgaa tccgcagaaa 1200

agaaagcaga tattgaagaa gctggtcgcg ctgttgtaaa aatgctcgaa atgggcttaa 1260
 aaccttctga cattttaacg cgtgaagctt ttgaagatgc tattactgta actatggctc 1320
 tgggaggttc aaccaactca acccttcacc tcttagctat tgcccatgct gctaattgtg 1380

 aattgacact tgatgatttc aatactttcc aagaaaaagt tcctcatttg gctgatttga 1440
 aaccttctgg tcaatatgta ttccaagacc tttaacaagg cggaggggta ccagcagtta 1500
 tgaaatatct ccttaaaaaat ggcttccttc atggtgaccg taccatttgt actggcaaaa 1560
 cagtcgctga aaatttgaag gcttttgatg atttaacacc tgggtcaaaag gttattatgc 1620
 cgcttgaaaa tcctaaacgt gaagatggtc cgctcattat tctccatggt aacttggctc 1680
 cagacggatg cgcttgcaaa gtttctgggtg taaaagtgcg tcgtcatgct ggtcctgcta 1740
 aggtctttta ttctgaagaa gaagccattg aagctgtctt gaatgatgat attgttgatg 1800

 gtgatgttgt tgctgtacgt tttgtaggac caaaggcggtg tcctgggatg cctgaaatgc 1860
 tttccctttc atcaatgatt gttggtaaag ggcaaggatg aaaagttgcc cttctgacag 1920
 atggccgctt ctcagggtgt acttatggtc ttgtcgtggg tcatatcgct cctgaagcac 1980
 aagatggcgg tccaatcgcc tacctgcaaa caggagacat agtcactatt gaccaagaca 2040
 ctaaggaatt acactttgat atctccgatg aagagttaaa acatcgtaaa gagaccattg 2100
 aattgccacc gctctattca cgcggtatcc ttggtaaata tgctcacatc gtttcgtctg 2160
 cttctagggg agccgtaaca gacttttga agcctgaaga aactggcaaa aaatgagcga 2220

 tttaattctt aattatt 2237
 <210> 323
 <211> 4420
 <212> DNA
 <213> Artificial sequence
 <220><223> Cassette
 <400> 323

 ttatgtatgc tcttctgact tttcgtgtga tgaggctcgt ggaaaaaatg aataatttat 60
 gaatttgaga acaattttgt gttgttacgg tattttacta tggaataatc aatcaattga 120
 ggattttatg caaatatcgt ttgaatattt ttccgacctt ttgagtactt ttcttcataa 180
 ttgcataata ttgtccgctg cccctttttc tgtagacagg tgtcttgatc tacttgcctat 240
 cgttcaacac caccttattt tctaactatt ttttttttag ctcatgtgaa tcagcttatg 300

 gtgatggcac atttttgcat aaacctagct gtccctcgtt aacataggaa aaaaaaatat 360
 ataaacaagg ctctttcact ctccctgcaa tcagatttgg gtttgttccc tttattttca 420

tattttctgt catattcctt tctcaattat tttttctac tcataacctc acgcaaaata	480
acacagtcaa atcaatcaaa atgactgaca aaaaaactct taaagactta agaaatcgta	540
gttctgttta cgattcaatg gttaaatcac ctaatcgtgc tatgttgcgt gcaactggta	600
tgcaagatga agactttgaa aaacctatcg tcggtgtcat ttcaacttgg gctgaaaaca	660
caccttgtaa tatccactta catgactttg gtaaaactagc caaagtcggt gttaaggaag	720
ctggtgcttg gccagttcag ttcggaacaa tcacggtttc tgatggaatc gccatgggaa	780
cccaaggaat gcgtttctcc ttgacatctc gtgatatatt tgcagattct attgaagcag	840
ccatgggagg tcataatcgc gatgcttttg tagccattgg cggttgtgat aaaaacatgc	900
ccggttctgt tatcgctatg gctaacatgg atatcccagc ctttttgcgt tacggcggaa	960
caattgcacc tggtaattta gacggcaaag atatcgattt agtctctgtc tttgaagggtg	1020
tcggccattg gaaccacggc gatatgacca aagaagaagt taaagctttg gaatgtaatg	1080
cttgtcccgg tcctggaggc tgcggtggta tgtatactgc taacacaatg gcgacagcta	1140
ttgaagtttt gggacttagc cttccgggtt catcttctca cccggtgaa tccgcagaaa	1200
agaaagcaga tattgaagaa gctggtcgcg ctgttgtcaa aatgctcgaa atgggcttaa	1260
aaccttctga ctttttaacg cgtgaagctt ttgaagatgc tattactgta actatggctc	1320
tgggagggtc aaccaactca acccttcacc tcttagctat tgcccatgct gctaattgtg	1380
aattgacact tgatgatttc aatactttcc aagaaaaagt tcctcatttg gctgatttga	1440
aaccttctgg tcaatatgta ttccaagacc tttaacaagg cggaggggta ccagcagtta	1500
tgaaatatct ccttaaaaat ggcttccttc atggtgaccg tatcacttgt actggcaaaa	1560
cagtcgctga aaatttgaag gcttttgaatg atttaacacc tggtcaaaag gttattatgc	1620
cgcttgaaaa tcctaaacgt gaagatggtc cgctcattat tctccatggt aacttggtc	1680
cagacgggtc cggtgccaaa gtttctgggtg taaaagtgcg tcgtcatgtc ggtcctgcta	1740
aggctcttaa ttctgaagaa gaagccattg aagctgtctt gaatgatgat attgttgatg	1800
gtgatgttgt tgtcgtacgt tttgtaggac caaaggcggt tcctggatg cctgaaatgc	1860
tttccttttc atcaatgatt gttggtaaag ggcaagggtg aaaagtggcc ctctgacag	1920
atggccgctt ctcagggtgt acttatggtc ttgtcgtggg tcatatcgct cctgaagcac	1980
aagatggcgg tccaatcgcc tacctgcaaa caggagacat agtcactatt gaccaagaca	2040
ctaaggaatt acactttgat atctccgatg aagagttaaa acatcgtaaa gagaccattg	2100
aattgccacc gctctattca cgcggtatcc ttggtaaata tgctcacatc gtttcgtctg	2160
cttctagggg agccgtaaca gacttttggg agcctgaaga aactggcaaa aaatgagcga	2220
tttaatctct aattattagt taaagtttta taagcatttt tatgtaacga aaaataaatt	2280

ggttcatatt attactgcac tgtcacttac catggaaaga ccagacaaga agttgccgac	2340
agtctgttga attggcctgg ttaggcttaa gtctgggtcc gcttctttac aaatttggag	2400
aatttctctt aaacgatatg tatattcttt tcgttggaaa agatgtcttc caaaaaaaaa	2460
accgatgaat tagtggaaacc aaggaaaaaa aaagaggtat ccttgattaa ggaacactgt	2520
ttaaacagtg tggtttccaa aacctgaaa ctgcattagt gtaatagaag actagacacc	2580
tcgatacaaa taatggttac tcaattcaaa actgccagcg aattcgactc tgcaattgct	2640
caagacaagc tagttgtcgt agatttctac gccacttggg gcggtccatg taaaatgatt	2700
gctccaatga ttgaaaaatg tggctgtggg ttcagggtcc ataaagcttt tcaattcatc	2760
tttttttttt ttgttctttt ttttgattcc gggttctttg aaattttttt gattcggtaa	2820
tctccgagca gaaggaagaa cgaaggaagg agcacagact tagatttgta tatatacgca	2880
tatgtggtgt tgaagaaaca tgaaattgcc cagtattctt aaccaactg cacagaacaa	2940
aaacctgcag gaaacgaaga taaatcatgt cgaagctac atataaggaa cgtgctgcta	3000
ctcatcctag tctgtttgct gccaaagctat ttaatatcat gcacgaaaag caaacaact	3060
tgtgtgcttc attggatgtt cgtaccacca aggaattact ggagttagtt gaagcattag	3120
gtcccaaat ttgtttacta aaaacacatg tggatatctt gactgatttt tccatggagg	3180
gcacagttaa gccgctaaag gcattatccg ccaagtacaa ttttttactc ttcgaagaca	3240
gaaaatttgc tgacatttgt aatacagtca aattgcagta ctctgcgggt gtatacagaa	3300
tagcagaatg ggcagacatt acgaatgcac acggtgttgt gggcccaggt attgttagcg	3360
gtttgaagca ggcggcggaa gaagtaacaa aggaacctag aggccttttg atgttagcag	3420
aattgtcatg caagggctcc ctagctactg gagaatatac taagggtact gttgacattg	3480
cgaagagcga caaagatttt gttatcggtt ttattgtcga aagagacatg ggtggaagag	3540
atgaaggtta cgattggttg attatgacac ccggtgtggg tttagatgac aaggagacg	3600
cattgggtca acagtataga accgtggatg atgtggtctc tacaggatct gacattatta	3660
ttgttgaag aggactattt gcaaagggaa gggatgctaa ggtagagggt gaacgttaca	3720
gaaaagcagg ctgggaagca tattttagaa gatgcggcca gcaaaactaa aaaactgtat	3780
tataagtaaa tgcatgtata ctaaactcac aaattagagc ttcaatttaa ttatatcagt	3840
tattaccgg gaatctcggg ctgaatgatt tctataatga cgaaaaaaaa aaaattggaa	3900
agaaaaagct tcatggcctt ccactttccc aaacaacacc tacggtatct ctcaagtctt	3960
atggggttcc attggtttca ccaactggtgc taccttgggt gctgctttcg ctgctgaaga	4020

aattgatcca aagaagagag ttatcttatt cattgggtgac ggttctttgc aattgactgt 4080

tcaagaaatc tccacatga tcagatgggg cttgaagcca tacttggtcg tcttgaacaa 4140

cgatggttac accattgaaa agttgattca cgggccaaag gctcaatata acgaaattca 4200

aggttgggac cacctatcct tgttgccaac ttctgggtgct aaggactatg aaaccacag 4260

agtcgtacc accggtgaat gggacaagt gacccaagac aagtctttca acgacaactc 4320

taagatcaga atgattgaaa tcatgttgcc agtcttcgat gctccacaaa acttggttga 4380

acaagctaag ttgactgctg ctaccaacgc taagcaataa 4420

<210> 324

<211> 3521

<212> DNA

<213> Artificial sequence

<220><223> Cassette

<400> 324

aaggaaataa agcaaataac aataacacca ttattttaat tttttttcta ttactgtcgc 60

taacacctgt atggttgcaa ccagggtgaga atccttctga tgcatacttt atgcgtttat 120

gcgttttgcg ccccttgga aaaaattgat tctcatcgta aatgcatact acatgcgttt 180

atgggaaaag cctccatata caaagggtcgc gtttctttta gaaaaactaa tacgtaaacc 240

tgcattaagg taagattata tcagaaaatg tgttgcaaga aatgcattat gcaatttttt 300

gattatgaca atctctcgaa agaaatttca tatgatgaga cttgaataat gcagcggcgc 360

ttgctaaaag aacttgtata taagagctgc cattctcgat caatatactg tagtaagtcc 420

tttctctctt ttcttattac acttatttca cataatcaat ctcaaagaga acaacacaat 480

acaataacaa gaagaacaaa atgaaagctc tggttttatca cggtgaccac aagatctcgc 540

ttgaagacaa gcccaagccc acccttcaaa agcccacgga tgtagtagta cgggttttga 600

agaccacgat ctgcggcacg gatctcggca tctacaaagg caagaatcca gaggtcgccg 660

acgggcgcgt cctgggccat gaagggttag gcgtcatcga ggaagtgggc gagagtgtca 720

cgcagttcaa gaaaggcgac aaggctcctga ttctctgcgt cacttcttgc ggctcgtgcg 780

actactgcaa gaagcagctt tactccatt gccgcgacgg cgggtggatc ctgggttaca 840

tgatcgatgg cgtgcaggcc gaatacgtcc gcatcccgca tgccgacaac agcctctaca 900

agatccccc gacaattgac gacgaaatcg cgtcctgct gagcgacatc ctgccaccg 960

gccacgaaat cggcgtccag tatgggaatg tccagccggg cgatgcggtg gctattgtcg 1020

gcgcggggccc cgtcggcatg tccgtactgt tgaccgcca gttctactcc cctcagacca 1080

tcacgtgat cgacatggac gagaatcgcc tccagctcgc caaggagctc ggggcaacgc	1140
acaccatcaa ctccggcagc gagaacgttg tcgaagccgt gcataggatt gcggcagagg	1200
gagtcgatgt tgcgatcgag gcggtgggca taccggcgac ttgggacatc tgccaggaga	1260
tcgtcaagcc cggcgcgcac atcgccaacg tcggcgtgca tggcgtcaag gttgacttcg	1320
agattcagaa gctctggatc aagaacctga cgatcaccac gggactgttg aacacgaaca	1380
cgacgcccac gctgatgaag gtgcctcga ccgacaagct tccgttgaag aagatgatta	1440
cccatcgctt cgagctggcc gagatcgagc acgcctatca ggtattcctc aatggcgcca	1500
aggagaaggc gatgaagatc atcctctcga acgcaggcgc tgcctgagct aattaacata	1560
aaactcatga ttcaacgttt gtgtatTTTT ttacttttga aggttataga tgtttaggta	1620
aataattggc atagatatag ttttagtata ataaatttct gatttggttt aaaatatcaa	1680
ctatTTTTTT tcacatatgt tcttgtaatt acttttctgt cctgtcttcc aggttaaaga	1740
ttagcttcta atatTTTAGG tggtttatta tttaatTTTA tgcctgattaa tttatttact	1800
tgtttaaacg gccggccaat gtggctgtgg tttcagggtc cataaagctt ttcaattcat	1860
ctTTTTTTTT tttgttcttt tttttgattc cggtttcttt gaaatTTTT tgattcggtta	1920
atctccgagc agaaggaaga acgaaggaag gacacagac ttagattggt atatatcgc	1980
atatgtggtg ttgaagaaac atgaaattgc ccagtattct taaccaact gcacagaaca	2040
aaaacctgca ggaaacgaag ataaatcatg tcgaaagcta catataagga acgtgctgct	2100
actatccta gtctgttgc tgccaagcta tttaatatca tgcacgaaaa gcaaacaac	2160
ttgtgtgctt cattggatgt tcgtaccacc aaggaattac tggagttagt tgaagcatta	2220
ggtcccaaaa tttgtttact aaaaacacat gtggatatct tgactgattt ttccatggag	2280
ggcacagtta agccgctaaa ggcatatcc gccaaagtaca atTTTTtact cttcgaagac	2340
agaaaatttg ctgacattgg taatacagtc aaattgcagt actctgcggg tgtatacaga	2400
atagcagaat gggcagacat tacgaatgca cacggtgtgg tgggccagg tattgttagc	2460
ggtttgaagc aggcggcgga agaagtaaca aaggaaccta gaggcctttt gatgttagca	2520
gaattgtcat gcaagggtc cctagctact ggagaatata ctaagggtac tgttgacatt	2580
gcgaagagcg acaaagattt tgttatcggc tttattgctc aaagagacat ggggtgaaga	2640
gatgaaggtt acgattggtt gattatgaca ccggtgtgg gtttagatga caaggagac	2700
gcattgggtc aacagtatag aaccgtggat gatgtggtct ctacaggatc tgacattatt	2760
attgttggaagaggactatt tgcaaaggga agggatgcta aggttagagg tgaacgttac	2820
agaaaagcag gctgggaagc atatttgaga agatgcggcc agcaaaacta aaaaactgta	2880

ttataagtaa atgcatgtat actaaactca caaattagag cttcaattta attatatcag 2940

ttattacccg ggaatctcgg tcgtaatgat ttctataatg acgaaaaaa aaaaattgga 3000

aagaaaaagc ttcatggcct tctactttcc caacagatgt atacgctatc gtccaagtct 3060

tgtggggttc cattggtttc acagtcggcg ctctattggg tgctactatg gccgctgaag 3120

aacttgatcc aaagaagaga gttattttat tcattgggtga cggttctcta caattgactg 3180

ttcaagaaat ctctaccatg attagatggg gtttgaagcc atacattttt gtcttgaata 3240

acaacggtta caccattgaa aaattgattc acggctcctca tgccgaatat aatgaaattc 3300

aaggttggga ccacttggcc ttattgccaa cttttggtgc tagaaactac gaaaccacac 3360

gagttgctac cactggtgaa tgggaaaagt tgactcaaga caaggacttc caagacaact 3420

ctaagattag aatgattgaa gttatgttgc cagtctttga tgctccacaa aacttggtta 3480

aacaagctca attgactgcc gctactaacg ctaacaata a 3521

<210> 325

<211> 2686

<212> DNA

<213> Artificial sequence

<220><223> Plasmid

<400> 325

tcgcgcgttt cggatgatgac ggtgaaaacc tctgacacat gcagctcccg gagacgggtca 60

cagcttgtct gtaagcggat gccgggagca gacaagcccg tcagggcgcg tcagcgggtg 120

ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180

accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcaggcgcc 240

attcgccatt caggctgcgc aactgttggg aaggcgatc ggtgcgggcc tcttcgctat 300

tacgccagct ggcgaaaggg ggatgtgctg caaggcgatt aagtgggta acgccagggt 360

tttcccagtc acgacgttgt aaaacgacgg ccagtgaatt cgagctcggt acccggggat 420

cctctagagt cgacctgcag gcatgcaagc ttggcgtaat catggtcata gctgtttcct 480

gtgtgaaatt gttatccgt cacaattcca cacaacatac gagccggaag cataaagtgt 540

aaagcctggg gtgcctaata agtgagctaa ctacattaa ttgcgttgcg ctcaactgcc 600

gctttccagt cgggaaacct gtctgtccag ctgcattaat gaatcgcca acgcgcgggg 660

agaggcgggt tgcgtattgg gcgctcttcc gcttctctgc tcactgactc gctgcgtcgc 720

gtcgttcggc tgcggcgagc ggtatcagct cactcaaagg cggtaatacg gttatccaca 780

gaatcagggg ataacgcagg aaagaacatg tgagcaaaag gccagcaaaa ggccaggaac 840

cgtaaaaagg ccgcgttgct ggcgtttttc cataggctcc gccccctga cgagcatcac 900
 aaaaatcgac gctcaagta gaggtggcga aacccgacag gactataaag ataccaggcg 960
 ttccccctg gaagctccct cgtgcgctct cctgttccga ccctgccgct taccgatac 1020

ctgtccgctt ttctcccttc gggaagcgtg gcgctttctc atagctcacg ctgtaggtat 1080
 ctcagttcgg tgtaggtcgt tcgctccaag ctgggctgtg tgcacgaacc ccccgttcag 1140
 cccgaccgct gcgccttacc cggttaactat cgtcttgagt ccaaccgggt aagacacgac 1200
 ttatcgccac tggcagcagc cactggtaac aggattagca gagcgaggta ttaggcgggt 1260
 gctacagagt tcttgaagtg gtggcctaac tacggctaca ctagaaggac agtatttggt 1320
 atctgcgctc tgcgaagcc agttaccttc ggaaaaagag ttggtagctc ttgatccggc 1380
 aaacaaacca ccgctggtag cgggtggtttt ttgtttgca agcagcagat tacgcgcaga 1440

aaaaaaggat ctcaagaaga tcctttgatc ttttctacgg ggtctgacgc tcagtggaac 1500
 gaaaactcac gtttaaggat ttgtgtcatg agattatcaa aaaggatctt cacctagatc 1560
 cttttaaat aaaaatgaag ttttaaatca atctaaagta tatatgagta aacttggtct 1620
 gacagttacc aatgcttaat cagtgaggca cctatctcag cgatctgtct atttcgttca 1680
 tccatagtgt cctgactccc cgtcgtgtag ataactacga tacgggaggg cttaccatct 1740
 ggccccagtg ctgcaatgat accgcgagac ccacgctcac cggctccaga ttatcagca 1800
 ataaaccagc cagccggaag ggccgagcgc agaagtggtc ctgcaacttt atccgcctcc 1860

atccagtcta ttaattgttg ccggaagct agagtaagta gttcgccagt taatagtttg 1920
 cgcaacgttg ttgccattgc tacaggcatc gtgggtgtcac gctcgtcgtt tggtaggct 1980
 tcattcagct ccggttccca acgatcaagg cgagttacat gatccccat gttgtgcaaa 2040
 aaagcggtta gtccttcgg tcctccgatc gttgtcagaa gtaagttggc cgcagtgtta 2100
 tcaactcatg ttatggcagc actgcataat tctcttactg tcatgccatc cgtaagatgc 2160
 ttttctgtga ctggtagta ctcaaccaag tcattctgag aatagtgtat gcggcgaccg 2220
 agttgtcttt gcccgcgctc aatacgggat aataccgcgc cacatagcag aactttaaaa 2280

gtgctcatca ttggaaaacg ttcttcgggg cgaaaactct caaggatctt accgctgttg 2340
 agatccagtt cgatgtaacc cactcgtgca cccaactgat cttcagcatc ttttactttc 2400
 accagcgttt ctgggtgagc aaaaacagga agcgaataatg ccgcaaaaaa gggaataagg 2460
 gcgacacgga aatgttgaat actcatactc ttctttttc aatattattg aagcatttat 2520
 cagggttatt gtctcatgag cggatacata ttggaatgta tttagaaaaa taaacaaata 2580
 ggggttccgc gcacatttcc ccgaaaagtg ccacctgacg tctaagaaac cattattatc 2640

atgacattaa cctataaaaa taggcgtatc acgaggccct ttcgtc 2686

<210> 326

<211> 24

<212> DNA

<213> Artificial sequence

<220><223> Primer

<400> 326

ttatgtatgc tcttctgact tttc 24

<210> 327

<211> 49

<212> DNA

<213> Artificial sequence

<220><223> Primer

<400> 327

aataattaga gattaaatcg ctcatTTTTT gccagtttct tcaggcttc 49

<210> 328

<211> 49

<212> DNA

<213> Artificial sequence

<220><223> Primer

<400> 328

agcctgaaga aactggcaaa aaatgagcga tttaattctt aattattag 49

<210> 329

<211> 49

<212> DNA

<213> Artificial sequence

<220><223> Primer

<400> 329

tatggaccct gaaaccacag ccacattttt caatcattgg agcaatcat 49

<210> 330

<211> 49

<212> DNA

<213> Artificial sequence

<220><223> Primer
 <400> 330
 taaaatgatt gctccaatga ttgaaaaatg tggctgtggt ttcagggtc 49
 <210> 331
 <211> 49
 <212> DNA
 <213> Artificial sequence
 <220><223> Primer
 <400> 331
 accgtaggtg ttgtttggga aagtggaagg ccatgaagct tttctttc 49
 <210> 332
 <211> 49
 <212> DNA
 <213> Artificial sequence
 <220><223> Primer
 <400> 332
 ttggaaagaa aaagcttcat ggccttccac ttcccaaac aacacctac 49
 <210> 333
 <211> 23
 <212> DNA
 <213> Artificial sequence
 <220><223> Primer
 <400> 333
 ttattgctta gcgttggtag cag 23
 <210> 334
 <211> 57
 <212> DNA
 <213> Artificial sequence
 <220><223> Primer
 <400> 334
 aattgcatcc ggcgcgccgt ttaaaggcc ggccaatgtg gctgtggttt cagggtc 57
 <210> 335
 <211> 49

<212> DNA

<213> Artificial sequence

<220><223> Primer

<400> 335

aatttctaga ttaattaagc ggccgcaagg ccatgaagct ttttctttc 49

<210> 336

<211> 21

<212> DNA

<213> Artificial sequence

<220><223> Primer

<400> 336

tttttgggtgg ttccggttc c 21

<210> 337

<211> 22

<212> DNA

<213> Artificial sequence

<220><223> Primer

<400> 337

aaagttggca tagcggaac tt 22

<210> 338

<211> 16

<212> DNA

<213> Artificial sequence

<220><223> Primer

<400> 338

gtcattgaca ccatct 16

<210> 339

<211> 19

<212> DNA

<213> Artificial sequence

<220><223> Primer

<400> 339

agagataccg tagtgttg 19

<210> 340

<211> 90
 <212> DNA
 <213> Artificial sequence
 <220><223> Primer
 <400> 340
 gtattttggt agattcaatt ctctttccct ttccttttcc ttcgtcctcc ttccttatca 60

 gcattgcgga ttacgtattc taatgttcag 90
 <210> 341
 <211> 90
 <212> DNA
 <213> Artificial sequence
 <220><223> Primer
 <400> 341
 ttggttgggg gaaaaagagg caacaggaaa gatcagaggg ggaggggggg ggagagtgtc 60
 accttggtta actcgttgta tcatcactgg 90
 <210> 342
 <211> 21
 <212> DNA
 <213> Artificial sequence
 <220><223> Primer
 <400> 342
 tcggtgcggg cctcttcgct a 21
 <210>
 > 343
 <211> 21
 <212> DNA
 <213> Artificial sequence
 <220><223> Primer
 <400> 343
 aatgtgagtt agctcactca t 21
 <210> 344
 <211> 33
 <212> DNA
 <213> Artificial sequence

<220><223> Primer
 <400> 344
 aattggcgcg ccatgaaagc tctggtttat cac 33
 <210> 345
 <211> 49
 <212> DNA
 <213> Artificial sequence
 <220><223> Primer
 <400> 345
 tgaatcatga gttttatgtt aattagctca ggcagcgcct gcgttcgag 49
 <210> 346
 <211> 49
 <212> DNA
 <213> Artificial sequence
 <220><223> Primer
 <400> 346
 atcctctcga acgcaggcgc tgcctgagct aattaacata aaactcatg 49
 <210> 347
 <211> 34
 <212> DNA
 <213> Artificial sequence
 <220><223> Primer
 <400> 347
 aattgtttaa acaagtaaataa aaattaatca gcat 34
 <210> 348
 <211> 49
 <212> DNA
 <213> Artificial sequence
 <220><223> Primer
 <400> 348
 acacaatata ataacaagaa gaacaaaatg aaagctctgg tttatcacg 49
 <210> 349
 <211> 49

<212> DNA
 <213> Artificial sequence
 <220><223> Primer
 <400> 349
 agcgtataca tctgttggga aagtagaagg ccatgaagct ttttctttc 49
 <210> 350
 <211> 49
 <212> DNA
 <213> Artificial sequence
 <220><223> Primer
 <400> 350
 ttggaaagaa aaagcttcat ggccttctac tttcccaaca gatgtatac 49
 <210> 351
 <211> 22
 <212> DNA
 <213> Artificial sequence
 <220><223> Primer
 <400> 351
 ttattgttta gcgttagtag cg 22
 <210> 352
 <211> 72
 <212> DNA
 <213> Artificial sequence
 <220><223> Primer
 <400> 352
 cataatcaat ctcaaagaga acaacacaat acaataacaa gaagaacaaa atgaaagctc 60
 tgggtttatca cg 72
 <210> 353
 <211> 21
 <212> DNA
 <213> Artificial sequence
 <220><223> Primer
 <400> 353
 taggcataat caccgaagaa g 21

<210> 354	
<211> 21	
<212> DNA	
<213> Artificial sequence	
<220><223> Primer	
<400> 354	
aaaatggtaa gcagctgaaa g	21
<210> 355	
<211> 17	
<212> DNA	
<213> Artificial sequence	
<220><223> Primer	
<400> 355	
agttgttaga actgttg	17
<210> 356	
<211> 19	
<212> DNA	
<213> Artificial sequence	
<220><223> Primer	
<400> 356	
gacgatagcg tatacatct	19
<210> 357	
<211> 22	
<212> DNA	
<213> Artificial sequence	
<220><223> Primer	
<400> 357	
cttagcctct agccatagcc at	22
<210> 358	
<211> 23	
<212> DNA	
<213> Artificial sequence	
<220><223> Primer	

<400> 358	
ttagttttgc tggccgcac ttc	23
<210> 359	
<211> 15456	
<212> DNA	
<213> Artificial sequence	
<220><223> Template	
<400> 359	
aaagagtaat actagagata aacataaaaa atgtagaggt cgagtttaga tgcaagttca	60
aggagcgaaa ggtggatggg taggttataat agggatatag cacagagata tatagcaaag	120
agatactttt gagcaatggt tgtggaagcg gtattcgcaa tattttagta gctcgttaca	180
gtccggtgcg tttttggttt ttgaaagtgc cgtcttcaga gcgcttttgg ttttcaaag	240
cgctctgaag ttctatact ttctagagaa taggaacttc ggaataggaa cttcaaagcg	300
tttccgaaaa cgagcgcttc cgaaaatgca acgcgagctg cgcacataca gtcactgtt	360
cacgtcgac ctatatctgc gtgttgccgt tatatatata tacatgagaa gaacggcata	420
gtgcgtgttt atgcttaaat gcgtacttat atgcgtctat ttatgtagga tgaaaggtag	480
tctagtaacct cctgtgatat tatccattc catgcggggg atcgtatgct tccttcagca	540
ctacccttta gctgttctat atgctgccac tcctcaattg gattagtctc atccttcaat	600
gctatcattt cctttgatat tggatcacac taagaaacca ttattatcat gacattaacc	660
tataaaaaata ggcgatcac gaggcctttt cgtctcgcgc gtttcggtga tgacggtgaa	720
aacctctgac acatgcagct cccggagacg gtcacagctt gctgtgaagc ggatgccggg	780
agcagacaag cccgtcaggg cgcgtcagcg ggtgttggcg ggtgtcgggg ctggcttaac	840
tatgcggcat cagagcagat tgtactgaga gtgcaccata aattcccgtt ttaagagctt	900
ggtgagcgt aggagtcact gccaggtatc gtttgaacac ggcatatgc agggaagtca	960
taacacagtc ctttcccgca attttctttt tctattactc ttggcctcct ctagtacact	1020
ctatatTTTT ttatgcctcg gtaatgattt tcattttttt tttccacct agcggatgac	1080
tctttttttt tcttagcgat tggcattatc acataatgaa ttatacatTA tataaagtaa	1140
tgtgatttct tcgaagaata tactaaaaaa tgagcaggca agataaacga aggcaaagat	1200
gacagagcag aaagccctag taaagcgtat tacaatgaa accaagattc agattgcgat	1260
ctctttaaag ggtggctccc tagcgataga gcactcgatc ttcccagaaa aagaggcaga	1320
agcagtagca gaacaggcca cacaatcgca agtgattaac gtccacacag gtatagggtt	1380

tctggaccat atgatacatg ctctggccaa gcattccggc tggtcgctaa tcgttgagtg	1440
cattggtgac ttacacatag acgacatca caccactgaa gactgcggga ttgctctcgg	1500
tcaagctttt aaagaggccc taggggccgt gcgtggagta aaaaggtttg gatcaggatt	1560
tgcgcctttg gatgaggcac tttccagagc ggtggtagat ctttcgaaca ggccgtacgc	1620
agttgtcgaa cttggtttgc aaaggagaa agtaggagat ctctcttgcg agatgatccc	1680
gcattttctt gaaagctttg cagaggctag cagaattacc ctccacgttg attgtctcgc	1740
aggcaagaat gatcatcacc gtagtgagag tgcgttcaag gctcttgctg ttgccataag	1800
agaagccacc tcgccaatg gtaccaacga tgttcctcc accaaagtg ttcttatgta	1860
gtgacaccga ttatttaaag ctgcagcata cgatatatat acatgtgtat atatgtatac	1920
ctatgaatgt cagtaaglat gtatacgaac agtatgatac tgaagatgac aaggtaatgc	1980
atcattctat acgtgtcatt ctgaacgagg cgcgctttcc ttttttcttt ttgctttttc	2040
tttttttttc tcttgaactc gacggatcta tgcggtgtga aataccgcac agatgcgtaa	2100
ggagaaaata ccgcatcagg aaattgtaag cgtaaatatt ttgttaaaat tcgcgttaaa	2160
tttttgtaa atcagctcat tttttaacca ataggccgaa atcggcacaaa tcccttataa	2220
atcaaaagaa tagaccgaga tagggttgag tgttgtcca gtttgaaca agagtccact	2280
attaaagaac gtggactcca acgtcaaagg gcgaaaaacc gtctatcagg gcgatggccc	2340
actacgtgaa ccatcacctt aatcaagttt ttgggggtcg aggtgccgta aagcactaaa	2400
tcggaacctt aaaggagccc ccgatttag agcttgacgg ggaaagccgg cgaacgtggc	2460
gagaaaggaa gggaagaaag cgaaaggagc gggcgctagg gcgctggcaa gtgtagcgg	2520
cacgtcgcgc gtaaccacca caccgccgc gcttaatgcg ccgctacagg gcgcgtccat	2580
tcgccattca ggctgcgcaa ctgttgggaa gggcgcggtg cgggcctctt cgctattacg	2640
ccagctggcg aaagggggat gtgctgcaag gcgattaagt tgggtaacgc cagggttttc	2700
ccagtcacga cgttgtaaaa cgacggccag tgagcgcgcg taatagact cactataggg	2760
cgaattgggt accgggcccc ccctcgaggt cgacggcgcg ccaactgtag agagcgactt	2820
tgtatgcccc aattgcgaaa cccgcgatat ccttctcgat tctttagtag ccgaccagga	2880
caaggaaaag gaggtcgaaa cgtttttgaa gaaacaagag gaactacacg gaagctctaa	2940
agatggcaac cagccagaaa ctaagaaaat gaagttgatg gatccaactg gcaccgtgg	3000
cttgaacaac aataccagcc ttccaacttc tgtaaataac ggcggtacgc cagtgcacc	3060
agtaccgtta ctttcggta tacctccttt ccccatgttt ccaatgcctt tcatgcctcc	3120
aacggtact atcacaaatc ctcatcaagc tgacgcaagc cctaagaaat gaataacaat	3180
actgacagta ctaaataatt gcctacttgg cttcacatac gttgcatacg tcgatataga	3240

taataatgat aatgacagca ggattatcgt aatacgtaat agctgaaaat ctcaaaaatg	3300
tgtgggtcat tacgtaaata atgataggaa tgggattctt ctatttttcc tttttccatt	3360
ctagcagccg tcgggaaaac gtggcatcct ctctttcggg ctcaattgga gtcacgctgc	3420
cgtgagcatc ctctctttcc atatctaaca actgagcacg taaccaatgg aaaagcatga	3480
gcttagcggtt gctccaaaaa agtattggat ggtaataacc atttgtctgt tctcttctga	3540
ctttgactcc tcaaaaaaaaa aaatctacaa tcaacagatc gttcaatta cgccctcaca	3600
aaaacttttt tcttcttctt tcgcccacgt taaattttat cctcatgtt gtctaacgga	3660
tttctgcact tgatttatta taaaaagaca aagacataat acttctctat caatttcagt	3720
tattgttctt ccttgcggtt ttcttctgtt ctcttttttc tttgtcata tataaccata	3780
accaagtaat acatattcaa actagtatga ctgacaaaaa aactcttaaa gacttaagaa	3840
atcgtagttc tgtttacgat tcaatgggtt aatcacctaa tcgtgctatg ttgcgtgcaa	3900
ctggtatgca agatgaagac tttgaaaaac ctatcgtcgg tgtcatttca acttgggctg	3960
aaaacacacc ttgtaatatc cacttacatg actttggtaa actagccaaa gtcggtgtta	4020
aggaagctgg tgcttgccca gttcagttcg gaacaatcac ggtttctgat ggaatcgcca	4080
tgggaaccca aggaatgcgt ttctccttga catctcgtga tattattgca gattctattg	4140
aagcagccat gggagggtcat aatgcggatg cttttgtagc cattggcggg tgtgataaaa	4200
acatgcccgg ttctgttacc gctatggcta acatggatat cccagccatt ttgtcttacg	4260
gcggaacaat tgcacctggt aatttagacg gcaaagatat cgatttagtc tctgtctttg	4320
aagggtgcgg ccattggaac cacggcgata tgaccaaaga agaagttaaa gctttggaat	4380
gtaatgcttg tcccggctct ggaggctcgc gtggtatgta tactgctaac acaatggcga	4440
cagctattga agttttggga cttagccttc cgggttcac tctcacccg gctgaatccg	4500
cagaaaagaa agcagatatt gaagaagctg gtcgcgctgt tgtcaaatg ctcgaaatgg	4560
gcttaaaacc ttctgacatt ttaacgcgtg aagcttttga agatgctatt actgtaacta	4620
tggctctggg aggttcaacc aactcaacc ttcacctctt agctattgcc catgctgcta	4680
atgtggaatt gacacttgat gatttcaata cttccaaga aaaagtctct catttggctg	4740
atttgaacc ttctggtaaa tatgtattcc aagaccttta caaggtcgga ggggtaccag	4800
cagttatgaa atatctctt aaaaatggct tcttcatgg tgaccgtatc acttgtactg	4860
gcaaaacagt cgctgaaaat ttgaaggctt ttgatgattt aacacctggt caaaaggtta	4920
ttatgccgct tgaaaatcct aaacgtgaag atggccgct cattattctc catggttaact	4980
tggctccaga cggtgccgtt gccaaagttt ctgggtgtaaa agtgcgctgt catgtcggtc	5040
ctgctaaggt ctttaattct gaagaagaag ccattgaagc tgtcttgaat gatgatattg	5100

ttgatggatga tgttgttgc gtacgttttg taggacaaa gggcggctct ggtatgcctg	5160
aaatgctttc cctttcatca atgattgttg gtaaaggga aggtgaaaaa gttgcccttc	5220
tgacagatgg ccgtttctca ggtggtactt atggtcttgt cgtgggtcat atcgctcctg	5280
aagcacaaga tggcgggtcca atcgcttacc tgcaaacagg agacatagtc actattgacc	5340
aagacactaa ggaattacac ttgatatact ccgatgaaga gttaaaacat cgtcaagaga	5400
ccattgaatt gccaccgttc tattcacgcg gtatccttgg taaatatgct cacatcgttt	5460
cgtctgcttc taggggagcc gtaacagact ttggaagcc tgaagaaact ggcaaaaaat	5520
gttgtcctgg ttgctgtggt taagcggccg cgttaattca aattaattga tatagttttt	5580
taatgagtat tgaatctgtt tagaaataat ggaatattat ttttatttat ttatttatat	5640
tattggtcgg ctcttttctt ctgaaggcca atgacaaaat gatatgaagg aaataatgat	5700
ttctaaaatt ttacaacgta agatatTTTT acaaaagcct agtcactctt ttgtcatgca	5760
ctattttact cacgcttgaa attaacggcc agtccactgc ggagtcattt caaagtcac	5820
ctaactgac tctggtttt gatagctcat ttggagttc gcgattgtct tctgttattc	5880
acaactgttt taatttttat ttattcttgg aactcttcga gttctttgta aagtctttca	5940
tagtagctta ctttatctc caacatattt aacttcatgt caatttcggc tcttaaattt	6000
tccacatcat caagtccaac atcatctttt aacttgaatt tattctctag ctcttccaac	6060
caagcctcat tgctcctga ttactggig aaaagtata cactttgcgc gcaatccagg	6120
tcaaaacttt cctgcaaaga attaccaat ttctcgacat catagtacaa ttgttttgt	6180
tctccatca caatttaata tacctgatgg attcttatga agcgtgggt aatggacgtg	6240
tcactctact tcgcttttt cctactcct tttagtacgg aagacaatgc taataataa	6300
gagggtata ataataattt taatcgcaa aaaagattaa acgccaagcg ttaattatc	6360
agaaagcaaa cgtcgtacca atccttgaat gcttccaat tgtatattaa gagtcacac	6420
agcaacatat tctgttatt aaattaatta ttattgattt ttgatattgt ataaaaaac	6480
caaatatgta taaaaaaagt gaataaaaaa taccaagtat ggagaaatat attagaagtc	6540
tatacgtaa accacccggg cccccctcg aggtcgacgg tatcgataag cttagatcgc	6600
aattcctgca gcccggggga tccactagt cttagcggc cgtctagaa ctagtaccac	6660
aggtgttgc ctctgaggac ataaaataca caccgagatt catcaactca ttgctggagt	6720
tagcatatct acaattgggt gaaatgggga gcgatttgca ggcatgtgct cggcatgccg	6780
gtagaggtgt ggtcaataag agcgacctca tgctatact gagaaagcaa cctgacctac	6840

aggaaagagt tactcaagaa taagaatttt cgttttaaaa cctaagagtc actttaaaat	6900
ttgtatacac ttatTTTTTT tataacttat ttaataataa aaatcataaa tcataagaaa	6960
ttcgcttact ctttaattaat caggcagcgc ctgcgttcga gaggatgata ttcacgcct	7020
tctccttggc gccattgagg aataacctgat aggcgtgctc gatctcggcc agctcgaagc	7080
gatgggtaat catcttcttc aacggaagct tgcggtcga ggcgaccttc atcagcatgg	7140
gcgtcgtgtt cgtgttcacc agtcccggtg tgatcgtcag gttcttgatc cagagcttct	7200
gaatctcgaa gtcaaccttg acgccatgca cgccgacgtt ggcgatgtgc gcgccgggct	7260
tgacgatctc ctggcagatg tcccaagtcg ccggtatgcc caccgcctcg atcgcaacat	7320
cgactccctc tgcgcaatc ctatgcacgg cttcgacaac gttctccgtg ccggagtga	7380
tgggtgtcgt tgcgccgagc tccttggcga gctggaggcg attctcgtcc atgtcgatca	7440
cgatgatggt cgagggggag tagaactggg cggatcaacag tacggacatg ccgacggggc	7500
ccgcgccgac aatagccacc gcatcgcccg gctggacatt cccatactgg acgccgattt	7560
cgtggccggt gggcaggatg tcgctcagca ggacggcgat ttcgtcgtca attgtctggg	7620
ggatcttgta gaggctgttg tcggcatgcg ggatgcggac gtattcggcc tgcacgccat	7680
cgatcatgta acccaggatc caccgcctg cgcggcaatg ggagtaaagc tgcttcttgc	7740
agtagtcga cgagccgcaa gaagtacgc aggaatcag gacctgtcg ctttcttga	7800
actgcgtgac actctgccc acttctcga tgacgcctac ccttcatgg cccaggatgc	7860
gcccgtcggc gacctctgga ttcttgctt ttagatgcc gagatccgtg ccgcagatcg	7920
tggcttcaa aaccgtact actacatccg tgggcttttg aagggtgggc ttgggcttgt	7980
cttcaagca gatcttggg tcacgtgat aaaccagagc tttcactc agctattgta	8040
atatgtgtgt ttgtttggat tattaagaag aataattaca aaaaaatta caaaggaagg	8100
taattacaac agaattaaga aaggacaaga aggaggaaga gaatcagttc attatttctt	8160
ctttgtata taacaaacc aagtagcgt ttggccatac attaaaagt gagaaccacc	8220
ctccctggca acagccaca ctcgttacca ttgttcatca cgatcatgaa actcgtgtc	8280
agctgaaatt tcacctcagt ggatctctt tttattctt catcgttcca ctaacctttt	8340
tccatcagct ggcaggaac ggaaagtga atccattta gcgagcttc tctttcttc	8400
aagaaaagac gaagcttgtg tgtgggtgcg cgcgctagta tctttccaca ttaagaata	8460
taccataaag gtiacttaga catcactatg gctatatata tatatatata tatatatgta	8520
acttagcacc atcgcgctg catcactgca tgtgttaacc gaaaagttg gcgaacactt	8580
caccgacag gtcatttaga tctgtcgtc gattgcagc tcccttagcc ttaaatccta	8640
ggcgggagca ttctcgtgta attgtgcagc ctgcgtagca actcaacata gcgtagtcta	8700

cccagttttt caagggttta tcgttagaag attctccctt ttcttcctgc tcacaaatct	8760
taaagtcata cattgcacga ctaaagcaa gcatgcggat ccccgggct gcaggaattc	8820
gatatcaagc ttatcgatac cgtcgactgg ccattaatct ttcccatatt agatttcgcc	8880
aagccatgaa agttcaagaa aggtcttttag acgaattacc cttcatttct caaactggcg	8940
tcaagggatc ctggtatggt ttatcgttt ttttctggt tcttatagca tcgttttgga	9000
cttctctgtt ccatttaggc gggtcaggag ccagcgaga atcattcttt gaaggatact	9060
tatcctttcc aattttgatt gtctgttacg ttggacataa actgtatact agaaattgga	9120
ctttgatggt gaaactagaa gatatggatc ttgataccgg cagaaaacaa gtagatttga	9180
ctcttcgtag ggaagaaatg aggattgagc gagaaacatt agcaaaaaga tccttcgtaa	9240
caagattttt acatttctgg tgttgaaggg aaagatatga gctatacagc ggaatttcca	9300
tatcactcag attttgttat ctaatttttt ccttcccacg tccgcgggaa tctgtgtata	9360
ttactgcac tagataatg ttatcttacc ttggcgcgta catttaattt tcaacgtatt	9420
ctataagaaa ttgcgggagt tttttcatg tagatgatac tgactgcacg caaatatagg	9480
catgatttat aggcattgatt tgatggctgt accgatagga acgctaagag taacttcaga	9540
atcgttatcc tggcggaataa aattcatttg taaactttaa aaaaaaacg caatatcccc	9600
aaaattatta agagcgctc cattattaac taaaatttca ctcagcatcc acaatgtatc	9660
aggtaatctac tacagatatt acatgtggcg aaaaagacaa gaacaatgca atagcgcatc	9720
aagaaaaaac acaaagcttt caatcaatga atcgaaaatg tcattaaaat agtatataaa	9780
ttgaaactaa gtcataaagc tataaaaaga aaatttatat aaatgcaaga tttaaagtaa	9840
attcacggcc ctgcaggcct cagctcttgt ttgttctgc aaataactta cccatctttt	9900
tcaaaacttt aggtgcaccc tcttttgcta gaataagttc tatccaatac atcctatttg	9960
gatctgcttg agcttcttc atcacggata cgaattcatt ttctgttctc acaattttgg	10020
acacaactct gtcttccgtt gcccgaac tttctggcag ttttgagtaa ttccacatag	10080
gaatgtcatt ataactctgg ttcggacat gaatttcct ctcaaccgtg taaccatcgt	10140
tattaatgat aaagcagatt gggtttatct tctctctaag ggctagtcct aattcttgg	10200
cagtcagttg caatgatcca tctccgataa acaataaatg tctagattct ttatctgcaa	10260
tttgctgccc tagagctcgc gggaaagtgt atcctataga tccccacaag ggttgaccaa	10320
taaaatgtga tticgatttc agaaatatag atgaggcacc gaagaaagaa gtgccttgtt	10380
cagccacgat cgtctcatta ctttgggtca aattttcgac agcttgccac agtctatctt	10440
gtgacaacag cgcgttagaa ggtacaaaat cttcttgctt ttatctatg tacttgccct	10500
tatatccaat ttcggacaag tcaagaagag atgatatcag ggattcgaag tcgaaatttt	10560

ggattctttc gttgaaaatt ttaccttcat cgatattcaa ggaaatcatt ttattttcat	10620
taagatggtg agtaaatgca cccgtactag aatcggtgtaag ctttacaccc aacataagaa	10680
taaaatcagc agattccaca aattccttca agtttggctc tgacagagta cgtttgtaaa	10740
tccccaaaaa tgagggaat gcttcatcaa cagatgattt accaaagtgc aaagtagtaa	10800
taggtaactt agtctttgaa ataaactgag taacagtcct ctctaggccg aacgatataa	10860
tttcatggcc tgtgattaca attggtttct tggcattctt cagactttcc tgtattttgt	10920
tcagaatctc ttgatcagat gtattcgacg tggaaatttc cttcttaaga ggcaaggatg	10980
gtttttcagc cttagcggca gctacatcta caggtaaatt gatgtaaacc ggctttcttt	11040
ccttttagtaa ggcagacaac actctatcaa tttcaacagt tgcattctcg gctgtcaata	11100
aagtcttggc agcagtaacc ggttcgtgca tcttcataaa gtgcttgaaa tcaccatcag	11160
ccaacgtatg gtgaacaaac ttaccttctg tctgcacttt cgaggtagga gatcccacga	11220
tctcaacaac aggcagggtc tcagcatagg agcccgttaa gccattaact gcggataatt	11280
cgccaacacc aaatgtagtc aagaatgccg cagccttttt cgttcttgcg taccgctcgg	11340
ccatatagga ggcatttaac tcattagcat tttccacca tttcatatct ttgtgtgaaa	11400
taatttgatc tagaaattgc aaattgtagt cacctggtac tccgaatatt tcttctatac	11460
ctaattcgtg taatctgtcc aacagatagt cacctactgt atacattttg tttactagtt	11520
tatgtgtgtt tattcgaaac taagtcttctg gtgtttttaa actaaaaaa agactaacta	11580
taaaagtaga atttaagaag tttaagaaat agatttacag aattacaatc aatacctacc	11640
gtctttatat acttattagt caagtagggg aataatttca gggaaactgt ttcaaccttt	11700
tttttcagct ttttccaaat cagagagagc agaaggtaat agaagggtga agaaaatgag	11760
atagatacat gcgtgggtca attgccttgt gtcacatctt actccaggca ggttgcatca	11820
ctccatttag gttgtgcccc ttttttgctt gtttgtgccc ctgttctctg tagttgcgct	11880
aagagaatgg acctatgaac tgatggttgg tgaagaaaac aatatttttg tgctgggatt	11940
cttttttttt ctggatgcca gcttaaaaag cgggctccat tataatttagt ggatgccagg	12000
aataaactgt tcaccagac acctacgatg ttatatattc tgtgtaaccc gccccctatt	12060
ttgggcatgt acgggttaca gcagaattaa aaggctaatt ttttgactaa ataaagttag	12120
gaaaatcact actattaatt atttacgtat tctttgaaat ggacgtattg ataataataa	12180
actcgaactg aaaaagcgtg ttttttattc aaaatgattc taactccctt acgtaatcaa	12240
ggaatctttt tgccttggcc tccgcgtcat taaacttctt gttgttgacg ctaacattca	12300

acgctagtat atattcgttt ttttcaggta agttcttttc aacgggtctt actgatgagg	12360
cagtcgcgtc tgaacctgtt aagaggtaa atatgtcttc ttgaccgtac gtgtcttgca	12420
tgttattagc tttgggaatt tgcataagt cataggaaaa tttaaactctt ggctctcttg	12480
ggctcaaggt gacaaggfcc tcgaaaatag ggcgcgcccc accgcggtgg agctccagct	12540
tttgttcctt ttagtgaggg ttaattgcgc gcttggcgta atcatggtca tagctgtttc	12600
ctgtgtgaaa ttgttatccg ctacacaattc cacacaacat acgagccgga agcataaagt	12660
gtaaagcctg ggggtgcctaa tgagttagct aactcacatt aattgcgttg cgctcactgc	12720
ccgctttcca gtcgggaaac ctgtcgtgcc agctgcatta atgaatcggc caacgcgcgg	12780
ggagaggcgg tttagctatt gggcgctctt ccgcttcctc gctcactgac tcgctgcgct	12840
cggctgttcg gctgcggcga gcggtatcag ctactcaaa ggcggttaata cggttatcca	12900
cagaatcagg ggataacgca ggaaagaaca tgtgagcaaa aggccagcaa aaggccagga	12960
accgtaaaaa ggccgcgttg ctggcgtttt tccataggct ccgccccct gacgagcatc	13020
acaaaaatcg acgtcaagt cagaggtggc gaaacccgac aggactataa agataccagg	13080
cgtttcccc tggaagctcc ctctgctgct ctctgttcc gacctgccc ctaccggat	13140
acctgtccgc ctttctccct tcgggaagcg tggcgctttc tcatagctca cgctgtaggt	13200
atctcagttc ggtgtaggtc gttcgtcca agctgggctg tgtgcacgaa cccccgttc	13260
agcccgaccg ctgcgcctta tccgtaact atcgtcttga gtccaaccg gtaagacacg	13320
acttatcgcc actggcagca gccactggta acaggattag cagagcgagg tatgtaggcg	13380
gtgtacaga gttcttgaag tggtagccta actacggcta cactagaaga acagtatttg	13440
gtatctgcgc tctgtgaag ccagttacct tcggaaaaag agttggtagc tcttgatccg	13500
gcaaacaaac caccgtggt agcgggtggt tttttgttg caagcagcag attacgcgca	13560
gaaaaaaagg atctcaagaa gatcctttga tctttctac ggggtctgac gctcagtgga	13620
acgaaaactc acgttaaggg atttttgtca tgagattatc aaaaaggatc ttcacctaga	13680
tcctttttaa ttaaaaatga agttttaaat caatctaaag tataatagag taaacttgg	13740
ctgacagtta ccaatgctta atcagttagg cacctatctc agcgatctgt ctatttcgtt	13800
catccatagt tgctgactc ccgctcgtgt agataactac gatacgggag ggcttaccat	13860
ctggccccag tctgcaatg ataccgcgag acccacgctc accggtcca gatttatcag	13920
caataaacca gccagccgga agggccgagc gcagaagtgg tcttgcaact ttatccgct	13980
ccatccagtc tattaattgt tgccgggaag ctagagtaag tagttcgcca gttaatagtt	14040
tgcgcaacgt tgttgccatt gctacaggca tcttggtgtc acgtcgtcg tttgggtatgg	14100
cttcattcag ctccggttcc caacgatcaa ggcgagttac atgatcccc atgttgtgca	14160

aaaaagcggg tagctccttc ggtcctccga tcgtttgcag aagtaagttg gccgcagtgt	14220
tatcactcat ggttatggca gcactgcata attctcttac tgtcatgcca tccgtaagat	14280
gcttttctgt gactggtag tactcaacca agtcattctg agaatagtgt atgcggcgac	14340
cgagttgctc ttgccccggc tcaatacggg ataataccgc gccacatagc agaactttaa	14400
aagtgtcat cattgaaaa cgttcttcgg ggcgaaaact ctcaaggatc ttaccgtgt	14460
tgagatccag ttcatgtaa cccactctg caccgaactg atcttcagca tcttttactt	14520
tcaccagcgt ttctgggtga gcaaaaacag gaaggcaaaa tgccgcaaaa aagggaataa	14580
ggcgacacg gaaatgttg atactcatac tcttctttt tcaatattat tgaagcattt	14640
atcagggtta ttgtctcatg agcggataca tatttgaatg tatttagaaa aataa caaaa	14700
taggggttcc gcgcacattt ccccgaaaag tgccacctga acgaagcatc tgtgcttcat	14760
ttttagaac aaaaatgcaa cgcgagagcg ctaattttt aaacaagaa tctgagctgc	14820
atttttacag aacagaaatg caacgcgaaa gcgctatttt accaacgaag aatctgtgct	14880
tcatttttgt aaaacaaaa tgcaacgcga gagcgctaatt ttttcaaaca aagaatctga	14940
gtgcatttt tacagaacag aaatgcaacg cgagagcgct attttacc aaagaatct	15000
atacttcttt ttgtttctac aaaaatgcat cccgagagcg ctatttttct aacaaagcat	15060
cttagattac tttttttct ctttgtgcgc tctataatgc agtctcttga taactttttg	15120
cactgtaggc ccgttaaggt tagaagaagg ctacttttgt gtctatttt tcttcataa	15180
aaaaagcctg actccacttc ccgcgtttac tgattactag cgaagctgcg ggtgcatttt	15240
ttcaagataa aggcacccc gatttatattc tataccgatg tggattgcgc atactttgtg	15300
aacagaaagt gatagcgttg atgattcttc attggtcaga aaattatgaa cggtttcttc	15360
tattttgtct ctatatacta cgtataggaa atgtttacat tttcgtattg ttttcgattc	15420
actctatgaa tagttcttac tacaattttt ttgtct	15456
<210> 360	
<211> 1559	
<212> DNA	
<213> Artificial sequence	
<220><223> Template	
<400> 360	
gcattgcgga ttacgtattc taatgttcag taccgttcgt ataatgtatg ctatacgaag	60
ttatgcagat tgtactgaga gtgcaccata ccaccttttc aattcatcat tttttttta	120
ttcttttttt tgatttcggg ttcccttgaaa tttttttgat tcggtaatct ccgaacagaa	180

ggaagaacga aggaaggagc acagacttag attggtatat atacgcataat gtagtggtga 240

agaaacatga aattgcccag tattcttaac ccaactgcac agaacaaaaa cctgcaggaa 300

acgaagataa atcatgtcga aagctacata taaggaacgt gctgctactc atcctagttc 360

tgttgctgcc aagctattta atatcatgca cgaagagcaa acaaaactgt gtgcttcatt 420

ggatgttcgt accaccaagg aattactgga gttagttgaa gcattaggtc ccaaaatttg 480

tttactaaaa acacatgtgg atatcttgac tgatttttcc atggagggca cagttaagcc 540

gctaaaggca ttatccgcca agtacaattt ttactcttc gaagacagaa aatttgctga 600

cattggtaat acagtcaaat tgcagtactc tgcgggtgta tacagaatag cagaatgggc 660

agacattacg aatgcacacg gtgtggtggg cccagggtatt gttagcgggt tgaagcaggc 720

ggcagaagaa gtaacaaagg aacctagagg ccttttgatg ttagcagaat tgtcatgcaa 780

gggctcccta tctactggag aatatactaa gggtactgtt gacattgcga agagcgacaa 840

agattttgtt atcggtctta ttgctcaaag agacatgggt ggaagagatg aaggttacga 900

ttggttgatt atgacacccg gtgtgggttt agatgacaag ggagacgcat tgggtcaaca 960

gtatagaacc gtggatgatg tggctcttac aggatctgac attattattg ttggaagagg 1020

actatttgca aagggaagggt atgctaaggt agagggtgaa cgttacagaa aagcaggctg 1080

ggaagcatat ttgagaagat gcggccagca aaactaaaaa actgtattat aagtaaatgc 1140

atgtatacta aactcacaaa ttagagcttc aatttaatta tatcagttat taccctatgc 1200

ggtgtgaaat accgcacaga tgcgtaagga gaaaataccg catcaggaaa ttgtaaacgt 1260

taatattttg ttaaaattcg cgtaaattt ttgttaaate agtcattttt ttaaccaata 1320

ggccgaaatc ggcaaaatcc cttataaatc aaaagaatag accgagatag ggttgagtgt 1380

tgttccagtt tggaacaaga gtccactatt aaagaacgtg gactccaacg tcaaagggcg 1440

aaaaaccgtc tatcagggcg atggccact acgtgaacca tcaccctaat caagataact 1500

tcgtataatg tatgctatag gaacggtagc agtgatgata caacgagttg gccaaagggtg 1559

<210> 361

<211> 1685

<212> DNA

<213> Artificial sequence

<220><223> Cassette

<400> 361

gtattttggt agattcaatt ctctttccct ttcttttcc ttctgtcccc ttctttatca 60

gcattgcgga ttacgtattc taatgttcag ataacttcgt atagcatata ttatacgaag 120

ttatgcagat tgiactgaga gtgcaccata ccacagcttt tcaattcaat tcatcatctt 180

ttttttattc ttttttttga tttcggtttc tttgaaatct ttttgattcg gtaatctccg 240

aacagaagga agaacgaagg aaggagcaca gacttagatt ggtatatata cgcataatga 300

gtgttgaaga aacatgaaat tgcccagtat tcttaacca actgcacaga acaaaaacct 360

gcaggaaacg aagataaatc atgtcgaaag ctacatatata ggaacgtgct gctactcatc 420

ctagtcctgt tgcgtccaag ctatttaata tcatgcacga aaagcaaaca aacttgtgtg 480

cttcattgga tgttcgtacc accaaggaat tactggagtt agttgaagca ttaggtccca 540

aaatttgctt actaaaaaca catgtggata tcttgactga tttttccatg gagggcacag 600

ttaagccgct aaaggcatta tccgccaagt acaatttttt actcttcgaa gacagaaaat 660

ttgctgacat tggtaatata gtcaaatgac agtactctgc gggtgtatag agaatagcag 720

aatgggcaga cattacgaat gcacacggtg tgggtggccc aggtattgtt agcgggttga 780

agcaggcggc agaagaagta acaaaggaac ctaggagcct tttgatgtta gcagaattgt 840

catgcaaggg ctccctatct actggagaat atactaaggg tactgttgac attgcaaga 900

gcgacaaaga ttttgttacc ggctttattg ctcaaagaga catgggtgga agagatgaag 960

gttacgattg gttgattatg acacccggtg tgggtttaga tgacaaggga gacgcattgg 1020

gtcaacagta tagaacctg gatgatgtgg tctctacagg atctgacatt attattgttg 1080

gaagaggact atttgcaaag ggaagggatg ctaaggtaga gggtgaacgt tacagaaaag 1140

caggctggga agcatatttg agaagatgcg gccagcaaaa ctaaaaaact gtattataag 1200

taaatgcatg tataactaac tcacaaatta gagcttcaat ttaattatat cagttattac 1260

cctatgcggt gtgaaatacc gcacagatgc gtaaggagaa aataccgcat caggaaattg 1320

taaacgttaa tattttgtta aaattcgcgt taaatttttg ttaaatcagc tcatttttta 1380

accaataggg cgaaatcggc aaaatccctt ataaatcaaa agaataagacc gagatagggt 1440

tgagtgttgt tccagtttgg aacaagagtc cactattaaa gaacgtggac tccaacgtca 1500

aagggcgaaa aaccgtctat caggcgatg gccactacg tgaacatca ccctaataca 1560

gataacttcg tatagcatat attatacgaa gttatccagt gatgatacaa cgagttagcc 1620

aaggtgacac tctcccccc cctccccctc tgatctttcc tgttgccctt ttttccccca 1680

accaa 1685

<210> 362

<211> 21

<212> DNA

<213> Artificial sequence

<220><223> Primer oBP591
 <400> 362
 cccattaata tactattgag a 21

<210> 363
 <211> 20
 <212> DNA
 <213> Artificial sequence
 <220><223> primer N98SeqF1
 <400> 363
 cgtgttagtc acatcaggac 20
 <210> 364
 <211> 20
 <212> DNA
 <213> Artificial sequence
 <220><223> primer N98SeqF4
 <400> 364
 ggtttctgtc tctggtgacg 20
 <210> 365
 <211> 22
 <212> DNA
 <213> Artificial sequence
 <220><223> primer N99SeqR2
 <400> 365
 catcgactgc attacgcaac tc 22

<210> 366
 <211> 30
 <212> DNA
 <213> Artificial sequence
 <220><223> primer N1111
 <400> 366
 tatttgtatc gaggtgtcta gtcttctatt 30
 <210> 367
 <211> 30

<212> DNA
 <213> Artificial sequence
 <220><223> primer N1110
 <400> 367
 gcgatttaat ctctaattat tagttaagt 30
 <210> 368
 <211> 38
 <212> DNA
 <213> Artificial sequence
 <220><223> BK468 Primer
 <400> 368
 gcctcgagtt ttaatgttac ttctcttgca gttaggga 38
 <210> 369
 <211> 7589
 <212> DNA
 <213> Artificial sequence
 <220><223> pUC19-kan::pdc1::FBA-alsS::TRX1
 <400> 369
 tatttgatc gaggtgtcta gtcttctatt acactaatgc agtttcaggg ttttgaaaac 60
 cacactgttt aaacagtgtt ccttaatacaa ggatacctct ttttttttcc ttggttccac 120
 taattcatcg gttttttttt tggaagacat cttttccaac gaaaagaata tacatatcgt 180
 ttaagagaaa ttctccaaat ttgtaaagaa gcggacccag acttaaggcc gcccgcaaat 240
 taaagccttc gagcgtccca aaaccttctc aagcaagggtt ttcagtataa tgttacatgc 300
 gtacacgcgt ctgtacagaa aaaaaagaaa aatttgaaat ataaataacg ttcttaatac 360
 taacataact ataaaaaat aaataggac ctagacttca ggttgtctaa ctcttctctt 420
 ttcggttaga gcggatgtgg ggggagggcg tgaatgtaag cgtgacataa ctaattacat 480
 gattaattaa ctagagagct ttcgttttca tgagttcccc gaattcttcc ggaagcttgt 540
 cacttgctaa attaatgtta tcaactgtagt caaccgggac atcgatgatg acaggacctt 600
 cagcgttcat gccttgacgc agaacatctg ccagctggtc tggtgattct acgcgcaagc 660
 cagttgtctc gaagctttcc gcataattca cgatatcgat atttccgaaa tcgaccgcag 720
 atgtacgggt atattttttc aattgctgga atgcaacat gtcatatgtg ctgtcgttcc 780
 atacaatgtg tacaattggt gcttttagtc gaactgctgt ctctaattcc attgctgaga 840

ataagaaacc gccgtcacca gagacagaaa ccactttttc tcccggtttc accaatgaag	900
cgccgattgc ccaaggaagc gcaacgccga gtgtttgcat accgttactg atcattaatg	960
ttaacggctc gtagctgcgg aaataacgtg acatccaaat ggctgtcgaa ccgatatcgc	1020
aagttactgt aacatgatca tcgactgcat tacgcaactc tttaacgatt tcaagagggt	1080
gcgctctgtc tgattttcaa tctgcaggca cctgctcacc ttcatgcata tattgtttta	1140
aatcagaaag gattttctgc tcacgtctcg caaattccac tttcacagca tcgtgttcga	1200
tatgattgat cgtggacgga atgtcaccga tcaattcaag atcaggctgg taagcatgat	1260
caatgtcagc gataatctcg tctaaatgga taattgtccg gtctccattg atattccaga	1320
atttcggatc atattcaatc gggtcatagc cgatcgtcag aacaacatct gcctgctcta	1380
gcagtaaate gccaggctgg ttgcggaaca aaccgatacg gccaaaatat tgatcctcta	1440
aatctctaga aagggtaccg gcagcttgat atgtttcaac aaatggaagc tgaacctttt	1500
tcaaaagctt gcgaaccgct ttaattgctt ccggctcttc gcctttcatg ccgaccaaaa	1560
cgacaggaag ttttctgtt tggatttttg ctatggccgc actgattgca tcattctgtg	1620
caggaccgag ttttggcgt gcaacagcac gcacgttttt cgtatttttg acttcattca	1680
caacatcttg cgaaagctc acaaaagcgg cccagcctg cctgctgac gctatcctaa	1740
atgcatttgt aacagcttcc ggtatatttt ttacatcttg aacttctaca ctgtattttg	1800
taatcggtg gaatagcgc gcattatcca aagattgatg tgtccgtttt aaacgatctg	1860
cacggatcac gttccagca agcgcaacga cagggtctcc ttcagtgttc gctgtcagca	1920
ggcctgttgc caagttagag gcacccggtc ctgatgtgac taacacgact cccggttttc	1980
cagttaaagc gccgactgct tgggccatga atgctgcgtt ttgttcgtgc cgggcaacga	2040
taatttcagg tcctttatct tgtaaagcgt caaataccgc atcaattttt gcacctggaa	2100
tgccaaatac atgtgtgaca ctttctcca ctaagcaatc aacaacaagc tccgcccctc	2160
tgtttttcac aagggaattt tgttcttttg ttgcttttgt caacatcctc agctctagat	2220
ttgaatatgt attacttggg tatggttata tatgacaaaa gaaaaagaag aacagaagaa	2280
taacgcaagg aagaacaata actgaaattg atagagaagt attatgtctt tgtcttttta	2340
taataaatca agtgcagaaa tccgttagac aacatgaggg ataaaattta acgtgggcga	2400
agaagaagga aaaaagtttt tgtgaggcg taattgaagc gatctgttga ttgtagattt	2460
tttttttttg aggagtcaaa gtcagaagag aacagacaaa tggatattaac catccaatac	2520
ttttttggag caacgctaag ctcatgcttt tcatttggtt acgtgctcag ttgttagata	2580
tggaaagaga ggatgctcac ggcagcgtga ctccaattga gcccgaaaga gaggatgcc	2640
cgttttcccg acggctgcta gaatggaaaa aggaaaaata gaagaatccc attcctatca	2700

ttatttacgt aatgaccac acatttttga gattttcaac tattacgtat tacgataatc	2760
ctgctgtcat tatcattatt atctatatcg acgtatgcaa cgtatgtgaa gccaaagtagg	2820
caattatitta gtactgtcag tattgttatt catttcagat cttaaagccta accaggccaa	2880
ttcaacagac tgtcggcaac ttcttgtctg gtctttccat ggtaagtgaac agtgcagtaa	2940
taatatgaac caattttatt ttctgttacct aaaaatgctt ataaaacttt aactaataat	3000
tagagattaa atcgcaaacg gccgactcta gaggatcccc caccttggtt aactcgttgt	3060
atcatcactg gataacttcg tatagcatac attatacgaa gttatctagg gattcataac	3120
cattttctca atcgaattac acagaacaca ccgtacaaac ctctctatca taactactta	3180
atagtccacac acgtactcgt ctaaatacac atcatcgtcc tacaagttca tcaaagtggt	3240
ggacagacaa ctataccagc atggatctct tgtatcggtt cttttctccc gctctctcgc	3300
aataacaatg aacactgggt caatcatagc ctacacaggt gaacagagta gcgtttatac	3360
agggtttata cgggtgattcc tacggcaaaa atttttcatt tctaaaaaaa aaaagaaaaa	3420
tttttcttc caacgctaga aggaaaagaa aaatctaatt aaattgattt ggtgattttc	3480
tgagagtcc ctttttcata tatcgaattt tgaatataaa aggagatcga aaaaattttt	3540
ctattcaatc tgttttctgg ttttatttga tagttttttt gtgtattatt attatggatt	3600
agtactgggt tatatgggtt tttctgtata acttcttttt attttagttt gtttaattct	3660
attttgagtt acattatagt tccctaactg caagagaagt aacattaaaa ctcgagatgg	3720
gtaaggaaaa gactcacgtt tcgaggccgc gattaaattc caacatggat gctgatttat	3780
atgggtataa atgggctcgc gataatgtcg ggcaatcagg tgcgacaatc tatcgattgt	3840
atgggaagcc cgatgcgcca gatttgtttc tgaacatgg caaaggtagc gttgccaatg	3900
atgttacaga tgagatggtc agactaaact ggctgacgga atttatgcct cttccgacca	3960
tcaagcattt tatccgtact cctgatgatg catggttact caccactgcg atccccggca	4020
aaacagcatt ccaggtatta gaagaatac ctgattcagg tgaaaatatt gttgatgcgc	4080
tggcagtgtt cctgcgccgg ttgcattcga ttctgtttg taattgtcct tttaacagcg	4140
atcgctgatt tcgtctcgct caggcgcaat cacgaatgaa taacggtttg gttgatgcga	4200
gtgattttga tgacgagcgt aatggctggc ctgttgaaca agtctggaaa gaaatgcata	4260
agcttttgcc attctaccg gattcagtcg tcaactatgg tgatttctca ctigataacc	4320
ttatttttga cgaggggaaa ttaatagggt gtattgatgt tggacgagtc ggaatcgag	4380
accgatacca ggatcttgcc atcctatgga actgcctcgg tgagttttct ctttcattac	4440

agaaacggct ttttcaaaa tatggtattg ataatcctga tatgaataaa ttgcagtttc 4500

 atttgatgct cgatgagttt ttctaagttt aacttgatac tactagattt ttctctttca 4560
 ttataaaaat ttttggttat aattgaagct ttagaagtat gaaaaaatcc ttttttttca 4620
 ttctttgcaa ccaaaataag aagcttcttt tattcattga aatgatgaat ataaacctaa 4680
 caaaagaaaa agactcgaat atcaaacatt aaaaaaaaaat aaaagagggtt atctgttttc 4740
 ccatttagtt ggagtttgca ttttctaata gatagaactc tcaattaatg tggatttagt 4800
 ttctctgttc gataacttcg tatagcatac attatacgaa gttatctgaa cattagaata 4860
 cgtaatccgc aatgcggggc cgcttaatta atctagagtc gacctgcagg catgcaagct 4920

 tggcgtaatc atggtcatag ctgtttccig tgtgaaattg ttatccgctc acaattccac 4980
 acaacatacg agccggaagc ataaagtga aagcctgggg tgcctaata gtagagctaac 5040
 tcacattaat tgcgttgccg tcaactgccg ctttccagtc gggaaacctg tctgtccagc 5100
 tgcattaatg aatcgcccaa cgcgcgggga gaggcgggtt gcgtattggg cgctcttccg 5160
 cttcctcgct cactgactcg ctgcgctcgg tcttccggct gcggcgagcg gtatcagctc 5220
 actcaaaggc ggtaatacgg ttatccacag aatcagggga taacgcagga aagaacatgt 5280
 gagcaaaagg ccagcaaaag gccaggaacc gtaaaaaggc cggttgctg gcgtttttcc 5340

 ataggtccg cccccctgac gagcatcaca aaaatcgacg ctcaagtcag aggtggcgaa 5400
 acccgacagg actataaaga taccaggcgt tccccctgg aagctccctc gtgcgctctc 5460
 ctgttccgac cctgccgtt accggatacc tgcgcctt tctcccttcg ggaagcgtgg 5520
 cgctttctca tagctcacgc tgtaggtatc tcagttcggg gtaggtcgtt cgctccaagc 5580
 tgggctgtgt gcacgaacct cccgttcagc ccgaccgtg cgccttatcc ggtaactatc 5640
 gtcttgatgc caaccggta agacacgact tatcgccact ggcagcagcc actggttaaca 5700
 ggattagcag agcgaggat gtagcggtg ctacagagtt cttgaagtgg tggcctaact 5760

 acggctacac tagaaggaca gtatttggta tctgcgctct gctgaagcca gttaccttcg 5820
 gaaaaagagt tggtagctct tgatccggca aacaaaccac cgctggtagc ggtggttttt 5880
 ttgtttgcaa gcagcagatt acgcgagaa aaaaaggatc tcaagaagat cctttgatct 5940
 tttctacggg gtctgacgt cagtggaacg aaaactcacg ttaagggtt ttggtcatga 6000
 gattatcaaa aaggatcttc acctagatcc ttttaatta aaaatgaagt tttaatcaa 6060
 tctaaagtat atatgagtaa acttggctcg acagttacca atgcttaatc agtgaggcac 6120
 ctatctcagc gatctgtcta tttcgttcat ccatagttgc ctgactcccc gtcgtgtaga 6180

taactacgat acgggagggc ttaccatctg gccccagtgc tgcaatgata ccgcgagacc 6240
 cagctcacc ggctccagat ttatcagcaa taaaccagcc agccggaagg gccgagcgca 6300
 gaagtggtec tgcaacttta tccgcctcca tccagtctat taattgttgc cgggaagcta 6360
 gagtaagtag ttcgccagtt aatagtttgc gcaacgttgt tgccattgct acaggcatcg 6420
 tgggtgcacg ctctctgttt ggtatggctt cattcagctc cggttcccaa cgatcaaggc 6480
 gagttacatg atcccccatg ttgtgcaaaa aagcggttag ctcttcgggt cctccgatcg 6540
 ttgtcagaag taagtgggcc gcagtgttat cactcatggt tatggcagca ctgcataatt 6600

ctcttactgt catgccatcc gtaagatgct tttctgtgac tggtagtac tcaaccaagt 6660
 cattctgaga atagtgtatg cggcgaccga gttgtctttg cccggcgctca atacgggata 6720
 ataccgcgcc acatagcaga actttaaaag tgctcatcat tggaaaacgt tcttcggggc 6780
 gaaaactctc aaggatctta ccgctgttga gatccagttc gatgtaaccc actcgtgcac 6840
 ccaactgac ttcagcatct tttactttca ccagcgtttc tgggtgagca aaaacaggaa 6900
 ggcaaatgc cgcaaaaaag ggaataaggg cgacacggaa atgttgaata ctcatactct 6960
 tcctttttca atattattga agcattttatc agggttattg tctcatgagc ggatacatat 7020

ttgaatgtat ttagaaaaat aaacaaatag gggttccgcg cacatttccc cgaaaagtgc 7080
 cacctgacgt ctaagaaacc attattatca tgacattaac ctataaaaat aggcgtatca 7140
 cgaggccctt tcgtctcgcg cgtttcgggtg atgacggtga aaacctctga cacatgcagc 7200
 tccccggagac ggtcacagct tgtctgtaag cggatgccgg gagcagacaa gcccgtcagg 7260
 gcgcgtcagc ggggtgttggc ggggtgtcggg gctggcttaa ctatgcggca tcagagcaga 7320
 ttgtactgag agtgcacat atgcgggtgtg aaataccgca cagatgcgta aggagaaaaat 7380
 accgcatcag gcgccattcg ccattcaggc tgcgcaactg ttgggaaggg cgatcggtgc 7440

gggcctcttc gctattacgc cagctggcga aagggggatg tgctgcaagg cgattaagtt 7500
 gggtaacgcc agggttttcc cagtcacgac gttgtaaaac gacggccagt gaattcgagc 7560
 tcggtaccgc gggatccggc gcgccgttt 7589

<210> 370

<211> 38

<212> DNA

<213> Artificial sequence

<220><223> primer N191

<400> 370

atccgcggat agatctccca ttaccgacat ttgggcgc 38

<210> 371

<211> 9593

<212> DNA

<213> Artificial sequence

<220><223> Plasmid pYZ090DalsS

<400> 371

ggccgcacct ggtaaacct ctagtggagt agtagatgta atcaatgaag cggaagccaa	60
aagaccagag tagaggccta tagaagaaac tgcgatacct tttgtgatgg ctaaacaaac	120
agacatcttt ttatatgttt ttacttctgt atatcgtgaa gtagtaagtg ataagcgaat	180
ttggctaaga acgttgtaag tgaacaaggg acctcttttg cctttcaaaa aaggattaaa	240
tggagttaat cattgagatt tagttttcgt tagattctgt atccctaaat aactccctta	300
cccgacggga aggcacaaaa gacttgaata atagcaaacg gccagtagcc aagaccaaat	360
aatactagag ttaactgatg gtcttaaaca ggcattacgt ggtgaactcc aagaccaata	420
tacaaaatat cgataagtta ttcttgccca ccaatttaag gaggctacat caggacagta	480
gtaccattcc tcagagaaga ggtatacata acaagaaaat cgcgtaaca cttatataa	540
cttagcccggt tattgagcta aaaaaccttg caaaatttcc tatgaataag aatacttcag	600
acgtgataaa aatttacttt ctaactcttc tcacgctgcc cctatctgtt cttccgctct	660
accgtgagaa ataaagcatc gagtacggca gttcgctgtc actgaactaa aacaataagg	720
ctagttcgaa tgatgaactt gcttgctgtc aaacttctga gttccgctg atgtgacact	780
gtgacaataa attcaaaccg gttatagcgg tctcctccgg taccggttct gccacctcca	840
atagagctca gtaggagtca gaacctctgc ggtggctgtc agtgactcat ccgcgtttcg	900
taagttgtgc gcgtgcacat ttcgcccgtt cccgctcacc ttgcagcagg cggaaatfff	960
catcacgctg taggacgcaa aaaaaaata attaatcgta caagaatctt ggaaaaaaa	1020
ttgaaaaatt ttgtataaaa gggatgacct aacttgactc aatggctttt acaccagta	1080
ttttcccttt cttgtttgt tacaattata gaagcaagac aaaaacatat agacaaccta	1140
ttcctaggag ttatattttt ttaccctacc agcaatataa gtaaaaaact gtttaaacag	1200
tatggcagtt acaatgtatt atgaagatga tgtagaagta tcagcacttg ctggaaagca	1260
aattgcagta atcggttatg gttcacaagg acatgctcac gcacagaatt tgcgtgattc	1320
tggtcacaac gttatcattg gtgtgcgcca cggaaaatct tttgataaag caaaagaaga	1380
tggctttgaa acatttgaag taggagaagc agtagctaaa gctgatgtta ttatggtttt	1440
ggcaccagat gaacttcaac aatccattta tgaagaggac atcaaaccac acttgaaagc	1500
aggttcagca ctiggttttg ctcacggatt taatatccat tttggctata ttaaaatacc	1560

agaagacgtt gacgtcttta tggttgcgcc taaggctcca ggtcaccttg tccgtcggac	1620
ttatactgaa ggttttggta caccagcttt gtttgtttca caccaaaatg caagtgggtca	1680
tgcgcgtgaa atcgcaatgg attgggccaa aggaattggt tgtgctcgag tgggaattat	1740
tgaacaact tttaaagaag aaacagaaga agatttgttt ggagaacaag ctgttctatg	1800
tggaggtttg acagcacttg ttgaagccgg ttttgaaca ctgacagaag ctggatacgc	1860
tggcgaattg gcttactttg aagttttgca cgaatgaaa ttgattgttg acctcatgta	1920
tgaaggtggt ttactaaaa tgcgtcaatc catctcaaat actgctgagt ttggcgatta	1980
tgtgactggt ccacggatta ttactgacga agtataaaag aatatgaagc ttgttttggc	2040
tgatatccaa tctggaaaat ttgctcaaga tttcgttgat gacttcaaag cggggcgtcc	2100
aaaattaata gcctatcgcg aagctgcaaa aaatcttgaa attgaaaaaa ttggggcaga	2160
gctacgtcaa gcaatgccat tcacacaatc tggatgatgac gatgccttta aaatctatca	2220
gtaaggccct gcaggccaga ggaaaataat atcaagtgtt ggaaactttt tctcttggaa	2280
tttttgaac atcaagtcac agtcaattga attgacccaa tttcacattt aagatttttt	2340
ttttttcacc cgacatacat ctgtacacta ggaagccctg tttttctgaa gcagcttcaa	2400
atataatata tttttacata tttattatga ttcaatgaac aatctaatta aatcgaaaac	2460
aagaaccgaa acgcgaataa ataatttatt tagatggtga caagtgtata agtcctcatc	2520
gggacagcta cgatttctct ttcggttttg gctgagctac tggttgctgt gacgcagcgg	2580
cattagcgcg gcgttatgag ctacctcgt ggctgaaag atggcgggaa taaagcggaa	2640
ctaaaaatta ctgactgagc catattgagg tcaatttgc aactcgtcaa gtcacgtttg	2700
gtggacggcc cttttcaac gaatcgtata tactaacatg cgcgcgttc ctatatacac	2760
atatacatat atataatata atatatgtgt gcgtgtatgt gtacacctgt atttaatttc	2820
cttactcgcg ggtttttctt ttttctcaat tcttggttc ctctttctcg agcggaccgg	2880
atcctccgcg gtgccggcag atctatttaa atggcgcgc gacgtcaggt ggcacttttc	2940
ggggaaatgt gcgcggaacc cctatttgtt tatttttcta aataattca aatatgtatc	3000
cgctcatgag acaataaccc tgataaatgc ttcaataata ttgaaaaagg aagagtatga	3060
gtattcaaca tttccgtgtc gcccttattc ctttttttgc ggcattttgc cttcctgttt	3120
ttgctcaccg agaaacgtg gtgaaagtaa aagatgtga agatcagttg ggtgcacgag	3180
tgggttacat cgaactggat ctcaacagcg gtaagatcct tgagagtttt cgtcccgaag	3240
aacgttttcc aatgatgagc acttttaaag ttctgctatg tggcgcggta ttatcccgtg	3300

ttgacgccgg gcaagagcaa ctccgtcgcc gcatacacta ttctcagaat gacttggttg 3360

agtactcacc agtcacagaa aagcatctta cggatggcat gacagtaaga gaattatgca 3420

gtgctgcat aaccatgagt gataaacctg cggccaactt acttctgaca acgatcggag 3480

gaccgaagga gctaaccgct tttttgcaca acatggggga tcatgtaact cgccttgatc 3540

gttgggaacc ggagctgaat gaagccatac caaacgacga gcgtgacacc acgatgcctg 3600

tagcaatggc aacaacgttg cgcaaaactat taactggcga actacttact ctagcttccc 3660

ggcaacaatt aatagactgg atggaggcgg ataaagttgc aggaccactt ctgcgctcgg 3720

cccttcgggc tggttggttt attgctgata aatctggagc cggtagcgt gggtctcgcg 3780

glatcattgc agcactgggg ccagatggta agccctcccg tatcgtagtt atctacacga 3840

cggggagtca ggcaactatg gatgaacgaa atagacagat cgctgagata ggtgcctcac 3900

tgattaagca ttggttaactg tcagaccaag ttactcata tatactttag attgatttaa 3960

aacttcattt ttaattttaa aggatctagg tgaagatcct ttttgataat ctcagacca 4020

aaatccctta acgtgagttt tcgttccact gagcgtcaga cccgtagaa aagatcaaag 4080

gatcttcttg agatcctttt tttctgcgcg taatctgctg cttgcaaaca aaaaaaccac 4140

cgctaccagc ggtggtttgt ttgccggatc aagagctacc aactcttttt ccgaaggtaa 4200

ctggcttcag cagagcgcag ataccaaata ctgttcttct agttagccg tagttaggcc 4260

accacttcaa gaactctgta gcaccgccta catacctcgc tctgctaata ctgttaccag 4320

tggctgctgc cagtggcgat aagtctgtc ttaccgggtt ggactcaaga cgatagttac 4380

cggataaggc gcagcggtcg ggctgaacgg ggggttcgtg cacacagccc agcttggagc 4440

gaacgaccta caccgaactg agatacctac agcgtgagct atgagaaagc gccacgcttc 4500

ccgaaggag aaaggcggac aggtatccgg taagcggcag ggtcggaaca ggagagcgca 4560

cgaggagct tccaggggga aacgcctggt atctttatag tctgtcggg tttcgccacc 4620

tctgacttga gcgtcgattt ttgtgatgct cgtcaggggg gcggagccta tggaaaaacg 4680

ccagcaacgc ggctttttta cggttcctgg ctttttctg gccttttct cactgttct 4740

ttctgcgtt atccccgat tctgttgata accgtattac cgcctttgag tgagctgata 4800

ccgctgcgcg cagccgaacg accgagcgca gcgagtcagt gagcgaggaa gcggaagagc 4860

gccaataacg caaacgcct ctccccgcgc gttggccgat tcattaatgc agctggcacg 4920

acaggtttcc cgactggaaa gcgggcagtg agcgcaacgc aattaatgtg agttagctca 4980

ctcattagge accccaggt ttacacttta tgcttcggc tcgtatgtt tgtggaattg 5040

tgagcggata acaatttcac acaggaaaca gctatgacca tgattacgcc aagctttttc	5100
tttccaattt ttttttttc gtcattataa aaatcattac gaccgagatt cccgggtaat	5160
aactgatata attaaattga agctctaat tgtgagtta gtatacatgc atttacttat	5220
aatacagttt ttiagttttg ctggccgcat cttctcaaat atgcttccca gcctgctttt	5280
ctgtaacgtt caccctctac cttagcatcc cttccctttg caaatagtcc tcttccaaca	5340
ataataatgt cagatcctgt agagaccaca tcatccacgg ttctatactg ttgacceaat	5400
gcgtctccct tgcacatcaa acccacaccg ggtgtcataa tcaaccaatc gtaaccttca	5460
tctcttccac ccatgtctct ttgagcaata aagccgataa caaaatcttt gtcgctcttc	5520
gcaatgtcaa cagtaccctt agtatattct ccagtagata gggagccctt gcatgacaat	5580
tctgctaaca tcaaaaggcc tctaggttcc ttgttactt cttctgccgc ctgcttcaaa	5640
ccgctaacaa tacctgggcc caccacaccg tgtgcattcg taatgtctgc ccattctgct	5700
attctgtata ccccgcaga gtactgcaat ttgactgtat taccaatgtc agcaaatttt	5760
ctgtcttcca agagtaaaaa attgtacttg gcggataatg cctttagcgg cttactgtg	5820
ccctccatgg aaaaatcagt caagatatcc acatgtgttt ttagtaacaa aattttggga	5880
cctaattgctt caactaactc cagtaattcc ttggtggtac gaacatcaa tgaagcacac	5940
aagtttgttt gcttttcgtg catgatatta aatagcttgg cagcaacagg actaggatga	6000
gtagcagcac gttccttata tgtagctttc gacatgattt atcttcgttt cctgcaggtt	6060
tttgttctgt gcagttgggt taagaatact gggcaatttc atgtttcttc aacactacat	6120
atgcgtatat ataccaatct aagtctgtgc tccttccttc gttcttctt ctgttcggag	6180
attaccgaat caaaaaaatt tcaaggaaac cgaatcaaa aaaaagaata aaaaaaaat	6240
gatgaattga aaagcttgca tgcctgcagg tcgactctag tatactccgt ctactgtacg	6300
atacacttcc gtcagggtcc ttgtccttta acgaggcctt accactcttt tgttactcta	6360
ttgatccagc tcagcaaagg cagtgtgatc taagattcta tcttcgcgat gtagtaaac	6420
tagctagacc gagaaagaga ctagaatgc aaaaggcact tctacaatgg ctgccatcat	6480
tattatccga tgtgacgtg catttttttt ttttttttt ttttttttt ttttttttt	6540
ttttttttt tttgtacaa atatcataaa aaaagagaat ctttttaagc aaggattttc	6600
ttaacttctt cggcgacagc atcaccgact tcggtggtac tgttgaacc acctaaatca	6660
ccagttctga tacttgcatc caaaaccttt ttaactgcat cttcaatggc ttaccttct	6720
tcaggcaagt tcaatgacaa tttcaacatc attgcagcag acaagatagt ggcgataggg	6780
ttgaccttat tctttggcaa atctggagcg gaacctggc atggttcgta caaaccaaat	6840
gcggtgttct tgtctggcaa agaggccaag gacgcagatg gcaacaaacc caaggagcct	6900

gggataacgg aggccttcac ggagatgata tcaccaaaca tgttgctggt gattataata 6960
 ccatttaggt gggttgggtt cttaactagg atcatggcgg cagaatcaat caattgatgt 7020
 tgaactttca atgtaggga ttcgttcttg atggtttcct ccacagtttt tctccataat 7080
 ctigaagagg caaaacatt agctttatcc aaggacaaa taggcaatgg tggctcatgt 7140

 ttagggcca tgaagcggc cattcttggt attctttgca cttctggaac ggtgtattgt 7200
 tcactatccc aagcgacacc atcaccatcg tcttctttc tcttaccaaa gtaaatacct 7260
 cccactaatt ctctaacaac aacgaagtca gtacctttag caaattgtgg ctgtattgga 7320
 gataagtcta aaagagagtc ggatgcaaag ttacatggtc ttaagttggc gtacaattga 7380
 agttctttac ggatttttag taaaccttgt tcaggtctaa cactaccggt accccattta 7440
 ggaccacca cagcacctaa caaacggca tcagccttct tggaggcttc cagcgctca 7500
 tctggaagtg gaacacctgt agcatcgata gcagcaccac caattaaatg attttcgaaa 7560

 tcgaacttga cattggaacg aacatcagaa atagctttaa gaaccttaat ggcttcggct 7620
 gtgatttctt gaccaacgtg gtcacctggc aaaacgacga tcttcttagg ggcagacatt 7680
 acaatgggtat atccttgaaa tatatataaa aaaaaaaaaa aaaaaaaaaa aaaaaaatgc 7740
 agcttctcaa tgatattcga atacgctttg aggagataca gcctaataatc cgacaaactg 7800
 ttttacagat ttacgatcgt acttgttacc catcattgaa ttttgaacat ccgaacctgg 7860
 gagttttccc tgaacagat agtatatttg aacctgtata ataatatata gtctagcgct 7920
 ttacggaaga caatgtatgt atttcggctc ctggagaaac tattgcatct attgcatagg 7980

 taatcttgca cgtcgatcc cgggttcatt tctgcgttt ccatcttgca ctccaatagc 8040
 atatctttgt taacgaagca tctgtgcttc attttgtaga acaaaaatgc aacgcgagag 8100
 cgctaatttt tcaaacaaag aatctgagct gcatttttac agaacagaaa tgcaacgca 8160
 aagcgtatatt ttaccaacga agaatctgtg ctctattttt gtaaaacaaa aatgcaacgc 8220
 gagagcgcta atttttcaa caaagaatct gagctgcatt tttacagaac agaaatgcaa 8280
 cgcgagagcg ctattttacc aacaaagaat ctatacttct ttttgttct acaaaaatgc 8340
 atcccagagag cgctattttt ctaacaaagc atcttagatt acttttttc tctttgtgc 8400

 gctctataat gcagtctctt gataactttt tgcactgtag gtccgttaag gttagaagaa 8460
 ggctactttg gtgtctattt tctcttccat aaaaaagcc tgactccact tcccgcgttt 8520
 actgattact agcgaagctg cgggtgcatt tttcaagat aaaggcatcc ccgattatat 8580
 tctataccga tgtggattgc gcatactttg tgaacagaaa gtgatagcgt tgatgattct 8640
 tcattggtea gaaaattatg aacggtttct tctattttgt ctctatatac tacgtatagg 8700
 aaatgtttac attttcgtat tgttttcgat tcactctatg aatagttctt actacaattt 8760

ttttgtctaa agagtaatac tagagataaa cataaaaaat gtagaggtcg agtttagatg 8820

caagttcaag gagcgaaagg tggatgggta ggttatatag ggatatagca cagagatata 8880

tagcaaagag atacttttga gcaatgtttg tggaagcggg attcgcaata ttttagtagc 8940

tcgttacagt ccggtgctgt tttgggtttt tgaaagtgcg tcttcagagc gcttttggtt 9000

ttcaaaagcg ctctgaagtt cctatacttt ctagagaata ggaacttcgg aataggaact 9060

tcaaagcgtt tccgaaaacg agcgcttccg aaaatgcaac gcgagctgcg cacatacagc 9120

tcactgttca cgctgcacct atatctgcgt gttgcctgta tatatatata catgagaaga 9180

acggcatagt gcgtgtttat gcttaaagtc gtacttatat gcgtctatct atgtaggatg 9240

aaaggtagtc tagtacctcc tgtgatattt tccattcca tgcgggggtat cgtatgcttc 9300

cttcagcact acccttttagc tgttctatat gctgccactc ctcaattgga ttagtctcat 9360

ccttcaatgc tatcatttcc tttgatattg gatcatatgc atagtaccga gaaactagag 9420

gatctcccat taccgacatt tgggcgctat acgtgcatat gttcatgtat gtatctgtat 9480

ttaaaacact tttgtattat ttttcctcat atatgtgtat aggtttatac ggatgattta 9540

attattactt caccaccctt tatttcaggc tgatatctta gccttggtac tag 9593

<210> 372

<211> 22

<212> DNA

<213> Artificial sequence

<220><223> primer oBP556

<400> 372

tattttcgag gaccttgta cc 22

<210> 373

<211> 25

<212> DNA

<213> Artificial sequence

<220><223> oBP561

<400> 373

tggccattaa tctttcccat attag 25

<210> 374

<211> 13114

<212> DNA

<213> artificial sequence

<220><223> constructed plasmid

<400> 374

tcgcgcgttt cggatgatgac ggtgaaaacc tctgacacat gcagctcccg gagacgggtca	60
cagcttgtct gtaagcggat gccgggagca gacaagcccg tcagggcgcg tcagcgggtg	120
ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc	180
accataaatt cccgttttaa gagcttgggtg agcgctagga gtcactgcca ggtatcggtt	240
gaacacggca ttagtcaggg aagtcataac acagtccttt cccgcaattt tctttttcta	300
ttactcttgg cctcctctag tacactctat atttttttat gcctcggtaa tgattttcat	360
tttttttttt ccacctagcg gatgactctt tttttttctt agcgattggc attatcacat	420
aatgaattat acattatata aagtaatgtg atttcttcga agaataact aaaaaatgag	480
caggcaagat aaacgaaggc aaagatgaca gagcagaaag ccctagtaaa gcgtattaca	540
aatgaaacca agattcagat tgcgatctct ttaaagggtg gtcccctagc gatagagcac	600
tcgatcttcc cagaaaaaga ggcagaagca gtagcagaac aggccacaca atcgcaagtg	660
attaacgtcc acacaggat agggtttctg gaccatatga tacatgctct ggccaagcat	720
tccggctggt cgtaatcgt tgagtgcatt ggtgacttac acatagacga ccatcacacc	780
actgaagact gcgggattgc tctcggtcaa gcttttaaag aggccctagg ggccgtgcgt	840
ggagtaaaaa ggtttggatc aggtattgctg cctttggatg aggcactttc cagagcgggtg	900
glagatcttt cgaacaggcc gtacgcagtt gtcgaacttg gtttgcaaag ggagaaagta	960
ggagatctct cttgcgagat gatcccgcat tttcttgaaa gctttgcaga ggctagcaga	1020
attaccttcc acgttgattg tctgcgaggc aagaatgac atcacctag tgagagtgcg	1080
ttcaaggctc ttgcggttgc cataagagaa gccacctgc ccaatggtac caacgatgtt	1140
ccctccacca aagggtttct tatgtagtga caccgattat ttaaagctgc agcatacgat	1200
atataacat gtgtatata gtatacctat gaatgtcagt aagtatgtat acgaacagta	1260
tgatactgaa gatgacaagg taatgcatca ttctatacgt gtcattctga acgaggcgcg	1320
ctttcttttt ttctttttgc tttttctttt tttttctctt gaactcgacg gatctatgcg	1380
gltgaaata ccgcacagat gcgtaaggag aaaataccgc atcaggaaat tgtaagcgtt	1440
aataatttgt taaaattcgc gttaaatttt tgttaaata gctcatTTTT taaccaatag	1500
gccgaaatcg gcaaaatccc ttataaatca aaagaataga ccgagatagg gttgagtgtt	1560
gttccagttt ggaacaagag tccactatta aagaacgtgg actccaacgt caaaggcgca	1620
aaaaccgtct atcaggcgca tggcccacta cgtggccggc ttacatacgt ttgcatacgt	1680
cgatatagat aataatgata atgacagcag gattatcgta atacgtaata gctgaaaatc	1740

tcaaaaatgt gtgggtcatt acgtaaataa tgataggaat gggattcttc tatttttcct	1800
ttttccattc tagcagccgt cgggaaaacg tggcatcctc tctttcgggc tcaattggag	1860
tcacgtgcc gtgagcatcc tctctttcca tatctaaca ctgagcacgt aaccaatgga	1920
aaagcatgag cttagcgttg ctccaaaaa gtattggatg gttaatacca tttgtctgtt	1980
ctctctgac tttgactcct caaaaaaaaa aatctacaat caacagatcg cttcaattac	2040
gccctcaca aaactttttt cttctctctt cgcccacgtt aaattttatc cctcatgttg	2100
tctaacggat tctgacctt gatttattat aaaaagaca agacataata cttctctatc	2160
aatttcagtt attgttcttc cttgcgttat tctctgttc tttttttct tttgtcatat	2220
ataaccataa ccaagtaata catattcaaa cacgtgagta tgactgacaa aaaaactctt	2280
aaagacttaa gaaatcgtag ttctgtttac gattcaatgg ttaaatacacc taatcgtgct	2340
atgttgcgtg caactggat gcaagatgaa gactttgaaa aacctatcgt cgggtgcatt	2400
tcaacttggg ctgaaaacac accttgtaat atccacttac atgactttgg taaactagcc	2460
aaagtcggtg ttaaggaagc tgggtccttg ccagttcagt tcggaacaat cacggtttct	2520
gatggaatcg ccatgggaac ccaaggaatg cgtttctct tgacatctcg tgatattatt	2580
gcagattcta ttgaagcagc catgggaggt cataatgcgg atgcttttgt agccattggc	2640
ggttgtgata aaaacatgcc cggttctgtt atcgctatgg ctaacatgga tatcccagcc	2700
atttttgctt acggcggaac aattgcacct ggtaatttag acggcaaaga tatcgattta	2760
gtctctgtct ttgaagggtg cggccattgg aaccacggcg atatgaccaa agaagaagtt	2820
aaagctttgg aatgtaatgc ttgtcccgtt cctggaggct gcggtggtat gtatactgct	2880
aacacaatgg cgacagctat tgaagttttg ggacttagcc ttccgggttc atcttctcac	2940
ccggctgaat ccgcagaaaa gaaagcagat attgaagaag ctggtcgcgc tgttgtcaaa	3000
atgctcgaaa tgggcctaaa acctctgac attttaacgc gtgaagcttt tgaagatgct	3060
attactgtaa ctatggctct gggaggttca accaactcaa ccttcacct cttagctatt	3120
gcccagctg ctaatgtgga attgacactt gatgatttca atactttcca agaaaaagtt	3180
cctcatttgg ctgatttgaa acctctggt caatatgtat tccaagacct ttacaaggtc	3240
ggaggggtac cagcagttat gaaatatct cttaaaaatg gcttccttca tggtgaccgt	3300
atcacttgta ctggcaaac agtcgctgaa aatttgaagg cttttgatga tttaacacct	3360
ggtcaaaagg ttattatgcc gcttgaaaat cctaaacgtg aagatgggcc gctcattatt	3420
ctccatggta acttggctcc agacggtgcc gttgccaaag tttctggtgt aaaagtgcgt	3480

cgtcatgtcg gtccgtgctaa ggtctttaat tctgaagaag aagccattga agctgtcttg 3540
 aatgatgata ttgttgatgg tgatgttggt gtcgtacgtt ttgtaggacc aaagggcgggt 3600
 cctggatgac ctgaaatgct ttccctttca tcaatgattg ttggtaaagg gcaaggtgaa 3660
 aaagttgccc ttctgacaga tggccgcttc tcaggtggta cttatggtct tgtcgtgggt 3720
 catatcgctc ctgaagcaca agatggcgggt ccaatcgctt acctgcaaac aggagacata 3780
 gtcactattg accaagacac taaggaatta cactttgata tctccgatga agagttaaaa 3840
 catcgtaag agaccattga attgccaccg ctctattcac gcggtatcct tggtaaatat 3900

 gctcacatcg tttcgtctgc ttctagggga gccgtaacag acttttgaa gcctgaagaa 3960
 actggcaaaa aatgttgctc tggttgctgt ggttaagcgg ccgcttaat tcaaattaat 4020
 tgatatagtt ttttaatgag tattgaatct gtttagaaat aatggaatat tatttttatt 4080
 tattttatta tattattggt cggtcttttt ctctgaagg tcaatgacaa aatgatatga 4140
 aggaaataat gattttctaa attttacaac gtaagatatt ttacaaaag cctagctcat 4200
 cttttgtcat gcaactatctt actcacgctt gaaattaacg gccagtccac tgcggagtca 4260
 tttcaaagtc atcctaactg atctatcgtt ttgatagct cattttgag ttcgcgagga 4320

 tccactagtt ctagagcggc cgctctagaa ctagtaccac aggtgttgct ctctgaggac 4380
 ataaaataca caccgagatt catcaactca ttgctggagt tagcatatct acaattgggt 4440
 gaaatgggga gcgatttgca ggcatttgct cggcatgccg gtagagggtt ggtcaataag 4500
 agcgacctca tgctatactt gagaagcaa cctgacctac aggaaagagt tactcaagaa 4560
 taagaatctt cgtttttaaa cctaagagtc acttttaaat ttgtatacac ttattttttt 4620
 tataacttat ttaataataa aaatcataaa tcataagaaa ttcgcttact ctttaattaat 4680
 caaaaagtta aaattgtacg aatagattca ccacttctta acaaatcaaa cccttcattg 4740

 attttctcga atggcaatac atgtgtaatt aaaggatcaa gagcaaactt cttcgccata 4800
 aagtcggcaa caagttttgg aacactatcc ttgctcttaa aaccgcaaaa tatagctccc 4860
 ttccatgtac gaccgcttag caacagcata ggattcatcg acaaatcttg tgaatcagga 4920
 ggaacaccta cgatcacact gactccatat gcctcttgac agcaggacaa cgcagttacc 4980
 atagtatcaa gacggcctat aacttcaaaa gagaaatcaa ctccaccgtt tgacatttca 5040
 gtaaggactt cttgtattgg tttcttataa tcttgagggt taacacattc agtagccccg 5100
 acctccttag cttttgcaaa ttgttcctta ttgatgtcta cacctataat cctcgtgcg 5160

 cctgcagctt tacaccccat aataacgctt agtctactc ctctaaacc gaatactgca 5220
 caagtgaac cctgtgtaac ctttgcaact ttaactgcgg aaccgtaacc ggtggaaaaat 5280
 ccgcacccta tcaagcaaac tttttccagt ggtgaagctg catcgatttt agcgacagat 5340

atctctgcca ccactgtgta ttgggaaaat gtagaagtac caaggaaatg gtgtataggt	5400
ttccctctgc atglaaatct gcttgtacca tcttgcatag tacctctagg catagacaaa	5460
tcatttttaa ggcagaaatt accctcagga tgtttgcaga ctctacactt accacattga	5520
ggagtgaaca gtgggatcac tttatcacca ggacgaacag tggtaacacc ttcacctatg	5580
gattcaacga ttccggcagc ctctgtgtcc gcgattactg gcaaaggagt aactagagt	5640
ccactcacca catggctgct ggatctacag attccggtgg caaccatctt gattctaacc	5700
tcgtgtgctt ttgggtggcg tacttctact tcttctatgc taaacggctt tttctcttc	5760
cacaaaactg ccgctttaca cttaataact ttaccgctg ttgacatcct cagctagcta	5820
ttgtaatatg tgtgtttgtt tggattatta agaagaataa ttacaaaaa aattacaaag	5880
gaaggttaatt acaacagaat taagaaagga caagaaggag gaagagaatc agttcattat	5940
ttcttctttg ttatataaca aaccaagta gcgatttggc catacattaa aagttgagaa	6000
ccaccctccc tggcaacagc cacaactcgt taccattgtt catcacgac atgaaactcg	6060
ccgtcagctg aaatttcacc tcagtggatc tctcttttta ttcttcacg ttccactaac	6120
ctttttccat cagctggcag ggaacggaaa gtggaatccc atttagcgag ctctctctt	6180
tcttcaagaa aagacgaagc ttgtgtgtgg gtgcgcgcgc tagtatcttt ccacattaag	6240
aaatatacca taaaggttac ttagacatca ctatggctat atatatatat atatatatat	6300
glaacttagc accatcgcg gtgcatcact gcatgtgtta accgaaaagt ttggcgaaca	6360
cttcaccgac acggtcattt agatctgtcg tctgcattgc acgtccctta gccttaaatc	6420
ctaggcggga gcattctcgt gtaattgtgc agcctgcgta gcaactcaac atagcgtagt	6480
ctaccagatt ttcaagggt ttatcgtag aagattctcc ctttctctcc tgctcacaaa	6540
tcttaaagtc atacattgca cgactaaatg caagcgacgt cagggaaga tatgagctat	6600
acagcggaat ttccatatca ctacagattt gttatctaata ttttctctc ccacgtccgc	6660
gggaatctgt gtatattact gcacttagat atatgttata ttatcttggc gcgtacattt	6720
aattttcaac gtattctata agaaattgcg ggagtttttt tcatgtagat gatactgact	6780
gcacgcaaat ataggcatga tttataggca tgatttgatg gctgtaccga taggaacgct	6840
aagagtaact tcagaatcgt tatcctggcg gaaaaaatc atttgtaaac tttaaaaaa	6900
aaagccaata tcccaaaat tattaagagc gctccatta ttaactaaaa tttaactcag	6960
catccacaat gtatcaggta tctactacag atattacatg tggcgaaaaa gacaagaaca	7020
atgcaatagc gcatcaagaa aaaacacaaa gctttcaatc aatgaatcga aaatgtcatt	7080
aaaatagtat ataaattgaa actaagtcac aaagtataa aaagaaaatt tatttaaatg	7140
caagatttaa aglaaattca cggccctgca ggcctcagct cttgttttgt tctgcaata	7200

acttacccat ctttttcaaa actttagggtg caccctcctt tgctagaata agttctatcc	7260
aatacatcct atttgatct gcttgagctt ctttcacac ggatacgaat tcattttctg	7320
ttctcacaat ttggacaca actctgtctt ccgttgcccc gaaactttct ggcagttttg	7380
agtaattcca cataggaatg tcattataac tctgggtcgg accatgaatt tcctctcaa	7440
ccgtgtaacc atcgttatta atgataaagc agattgggtt tatcttctct ctaatggcta	7500
gtcctaattc ttggacagtc agttgcaatg atccatctcc gataaacaat aaatgtctag	7560
attctttatc tgcaatttgg ctgcctagag ctgcggggaa agtgtatcct atagatcccc	7620
acaagggttg accaataaaa tgtgatttcg atttcagaaa tatagatgag gcaccgaaga	7680
aagaagtgcc ttgttcagcc acgatcgtct cattactttg ggtcaaattt tcgacagctt	7740
gccacagtct atcttgtgac aacagcgcgt tagaaggtag aaaatcttct tgctttttat	7800
ctatgtactt gcctttatat tcaatttcgg acaagtcaag aagagatgat atcagggatt	7860
cgaagtcgaa attttggtt ctttcgttga aaattttacc ttcatcgata ttcaaggaaa	7920
tcattttatt ttcattaaga tggtagtaaa atgcaccctg actagaatcg gtaagcttta	7980
caccaacat aagaataaaa tcagcagatt ccacaaatc cttcaagttt ggctctgaca	8040
gagtaccgtt gtaaatcccc aaaaatgagg gcaatgcttc atcaacagat gatttaccaa	8100
agttcaaagt agtaataggt aacttagtct ttgaaataaa ctgagtaaca gtcttctcta	8160
ggccgaacga tataatttca tggcctgtga ttacaattgg tttcttgga ttcttcagac	8220
tttctgtat ttgttcaga atctcttgat cagatgtatt cgacgtggaa ttttcttct	8280
taagaggcaa ggatggtttt tcagccttag cggcagctac atctacaggt aaattgatgt	8340
aaaccggctt tctttcttt agtaaggcag acaacactct atcaatttca acagttgcat	8400
tctcgctgt caataaagtc ctggcagcag taaccggttc gtgcatcttc ataaagtgt	8460
tgaaatcacc atcagccaac gtatggtgaa caaacttacc ttcgttctgc actttcgagg	8520
taggatcc cagatctca acaacaggca ggttctcagc ataggagccc gctaagccat	8580
taactcgga taattcgcca acaccaatg tagtcaagaa tgccgcagcc tttttcgttc	8640
ttgcgtacc gtcggccata taggaggcat ttaactcatt agcatttccc acccatttca	8700
tatctttgtg tgaataaatt tgatctagaa attgcaaatt gtagtcacct ggtactccga	8760
atattttctc tatacctaatt tcgtgtaatc tgccaacag atagtcacct actgtataca	8820
tgtttaaact ttgtttacta gtttatgtgt gtttattcga aactaagttc ttggtgtttt	8880
aaaactaaaa aaaagactaa ctataaaagt agaatttaag aagttaaga aatagattta	8940

cagaattaca atcaataacct accgtcttta tatacttatt agtcaagtag gggaataatt	9000
tcagggaact ggittcaacc ttttttttca gctttttcca aatcagagag agcagaaggt	9060
aatagaaggt gtaagaaat gagatagata catgcgtggg tcaattgcct tgtgtcatca	9120
tttactccag gcaggttgca tcactccatt gaggtttgtgc ccgttttttg cctgtttgtg	9180
ccctgtttct ctgtagtgc gctaagagaa tggacctatg aactgatggt tggagaagaa	9240
aacaatattt tgggtctggg attctttttt tttctggatg ccagcttaaa aagcgggctc	9300
cattatattt agtggatgcc aggaataaac tgttcacca gacacctacg atgttatata	9360
ttctgtgtaa cccgccccct attttgggca tgtacgggtt acagcagaat taaaaggcta	9420
attttttgac taaataaagt taggaaaatc actactatta attatttacg tattctttga	9480
aatggcagta ttggagctcc agcttttgtt cccttttagtg agggtaatt gcgcgcttgg	9540
cgtaatcatg gtcatagctg tttcctgtgt gaaattgtta tccgctcaca attccacaca	9600
acatacgagc cggaagcata aagtgtaaag cctgggggtgc ctaatgagtg agctaaactca	9660
cattaattgc gttgcgtca ctgcccgtt tccagtcggg aaacctgtcg tgccagctgc	9720
attaatgaat cggccaacgc gcggggagag gcggtttgcg tattgggcgc tttccgctt	9780
ctcgcctcac tgactcgtg cgctcggctg ttcggctgcg gcgagcggtc tcagctcact	9840
caaaggcggg aatacggta tccacagaat caggggataa cgcaggaaag aacatgtgag	9900
caaaaggcca gcaaaaggcc aggaaccgtc aaaaggccgc gttgctggcg tttttccata	9960
ggctccgccc cctgacgag catcacaaaa atcgacgtc aagtcagagg tggcgaaacc	10020
cgacaggact ataaagatac caggcgtttc cccctggaag ctcctcgtg cgtctcctg	10080
ttccgacctt gccgcttacc ggatacctgt ccgcttttct ccttcggga agcgtggcgc	10140
tttctcatag ctacgcgtg aggtatctca gttcgggtga ggtcgttcgc tccaagctgg	10200
gctgtgtgca cgaaccccc gttcagcccc accgtgcgc cttatccgtt aactatcgtc	10260
ttgagtcgaa cccggtaaga cagcacttat cgccactggc agcagccact ggtaacagga	10320
ttagcagagc gaggtatgta ggcggtgcta cagagtctt gaagtgttg cctaactacg	10380
gtacactag aagaacagta tttggtatct gcgctctgct gaagccagtt accttcggaa	10440
aaagagttgg tagctcttga tccggcaaac aaaccaccgc tggtagcggg ggtttttttg	10500
tttgcaagca gcagattacg cgcagaaaaa aaggatctca agaagatcct ttgatctttt	10560
ctacggggtc tgacgtcag tggaacgaaa actcacgtta agggattttg gtcagagat	10620
tatcaaaaag gatcttcacc tagatccttt taaattaaaa atgaagtttt aaatcaatct	10680
aaagtatata tgagtaaaat tggctcgaca gttaccaatg cttaatcagt gaggcaccta	10740
tctcagcgat ctgtctattt cgttcatcca tagttgcctg actccccgtc gtgtagataa	10800

ctacgatacg ggagggctta ccatctggcc ccagtgtgc aatgataccg cgagaccac	10860
gtcaccggc tccagattta tcagcaataa accagccagc cggaagggcc gagcgagaa	10920
gtggtcctgc aactttatcc gcctccatcc agtctattaa ttgttgccgg gaagctagag	10980
taagtagttc gccagttaat agtttgcga acgttgttgc cattgctaca ggcatcgtgg	11040
tgtcacgtc gtcgttttgt atggcttcat tcagctccgg ttccaacga tcaaggcgag	11100
ttacatgac ccccatgttg tgcaaaaaag cggtagctc cttcggtcct ccgatcgttg	11160
tcagaagtaa gtggccgca gtgttatcac tcattggtat ggagcactg cataattctc	11220
ttactgtcat gccatccgta agatgctttt ctgtgactgg tgagtactca accaagtcat	11280
tctgagaata gtgtatgcgg cgaccgagtt gctcttgccc gggtcaata cgggataata	11340
ccgcccaca tagcagaact taaaagtgc tcattattgg aaacgttct tcggggcgaa	11400
aacttcaag gatcttaccg ctgttgagat ccagttcgat gtaacccact cgtgcacca	11460
actgatcttc agcatctttt actttcacca gcgtttctgg gtgagcaaaa acaggaaggc	11520
aaaatgccgc aaaaaaggga ataaggcgca cacggaaatg ttgaatactc atactcttc	11580
tttttcaata ttattgaagc atttatcagg gttattgtct catgagcgga tacatatgtg	11640
aatgtattta gaaaaataa caaatagggg ttccgcgcac atttccccga aaagtccac	11700
ctgaacgaag catctgtgtc tcatTTTgta gaacaaaaat gcaacgcgag agcgctaatt	11760
tttcaaaaca agaattcgag ctgcatTTTT acagaacaga aatgcaacgc gaaagcgcta	11820
ttttaccaac gaagaatctg tgcttcattt ttgtaaaaca aaaatgcaac gcgagagcg	11880
taatttttca aacaaagaat ctgagctgca tttttacaga acagaaatgc aacgcgagag	11940
cgctatttta ccaacaaaga atctatactt cttttttgtt ctacaaaaat gcatcccgag	12000
agcgctattt ttctaacaa gcatcttaga ttactttttt tctcctttgt gcgctctata	12060
atgcagtctc ttgataactt ttgactgt aggtccgtta aggttagaag aaggctactt	12120
tggtgtctat ttctcttcc ataaaaaag cctgactcca cttcccgctt ttactgatta	12180
ctagcgaagc tgcgggtgca ttttttcaag ataaaggcat ccccgattat attctatacc	12240
gatgtggatt gcgcatactt tgtgaacaga aagtgatagc gttgatgatt cttcattgg	12300
cagaaaatta tgaacggitt cttctatttt gtctctatat actacgtata ggaaatgttt	12360
acattttcgt attgttttcg attcactcta tgaatagttc ttactacaat tttttgtct	12420
aaagagtaat actagagata aacataaaaa atgtagaggt cgagtttaga tgcaagttca	12480
aggagcgaaa ggtggatggg taggttatat agggatatag cacagagata tatagcaaag	12540
agatactttt gagcaatgtt tgtggaagcg gtattcgcaa tattttagta gctcgttaca	12600
gtccggtgcg tttttggttt ttgtaaagtg cgtcttcaga gcgcttttgg ttttcaaag	12660

cgctctgaag ttctataact ttctagagaa taggaacttc ggaataggaa cttcaaagcg 12720

tttccgaaaa cgagcgcttc cgaaaatgca acgcgagctg cgacataca gtcactgtt 12780

cacgtcgac ctatatctgc gtgttgctg tatatatata tacatgagaa gaacggcata 12840

gtgcgtgttt atgcttaaat gcgtacttat atgcgtctat ttatgtagga tgaaaggtag 12900

tctagtacct cctgtgatat tatcccatc catgcggggt atcgtatgct tccttcagca 12960

ctaccttta gctgttctat atgctgccac tctcaattg gattagtctc atccttcaat 13020

gctatcattt cctttgatat tggatcatac taagaaacca ttattatcat gacattaacc 13080

tataaaaata ggcttatcac gaggccttt cgtc 13114

<210> 375

<211> 987

<212> DNA

<213> Enterococcus gallinarum

<400> 375

atggttaagg tctactacaa cgaagttaact gacaacttgg gtgacaagac tatcgtatc 60

atcggttacg gtctcaagg tcacgccac gctcaaaact tgagagacac cggtcataat 120

gttgtcatcg gtatcagagc tggtaaactc gccgaaaacg ctaagaacga tggtttact 180

gttttgccag tctccgaagc tgttcaacag gccaacgttg tcatgatctt ggctccagac 240

gaaattcagg gtgaattgta cgaaaaggaa atcgtcctaa acttgtctgc tggtagcgt 300

ttagccttcg ctcacggttt caacatccat ttcgaagtca ttacccacc aaaagaagtt 360

gacgttttct tggttgctcc aaagggtcca ggtcacttgg tcagacgtac cttctctgaa 420

ggtttcgccg ttccagcttt gttcgccgtc taccaagatg ctaccggtaa cgcccaggaa 480

actgctttgt cctacgcaa gggtatcggg gctaccagag ttggtgtctt ggaaaccacc 540

ttcaaggaag aaaccgaaac cgacttggtc ggtgaacaag ctgttttgtg tggtaggtctg 600

acttctatga ttgaaaccgg ttctgaaact ctggtcgaag ctggttacca accagaattg 660

gcctacttcg aagtttgcca cgaaatgaag ttgatcgtcg attgatcta cgaaggtgg 720

ttcgctaaga tgagagactc tatctcaac accgccgaat acggtgacta cgtttctggt 780

tctagaatca ttaccgaaca aactaaggct aacatgaaaa atgtcttgaa ggacatccag 840

aacggtaact ttgctaaggg ttctattgct gataaccaag ccggtttccc agaattcaag 900

aagatgagag aacagaacgg taatcacca atcgaaaagg tcggttctga attgagaaag 960

atgatgcat tcgttactaa ggactga 987

<210> 376

<211> 328

<212> PRT

<213> Enterococcus gallinarum

<400> 376

Met Val Lys Val Tyr Tyr Asn Glu Val Thr Asp Asn Leu Gly Asp Lys

1 5 10 15

Thr Ile Ala Ile Ile Gly Tyr Gly Ser Gln Gly His Ala His Ala Gln

20 25 30

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

35 40 45

Lys Ser Ala Glu Asn Ala Lys Asn Asp Gly Phe Thr Val Leu Pro Val

50 55 60

Ser Glu Ala Val Gln Gln Ala Asn Val Val Met Ile Leu Ala Pro Asp

65 70 75 80

Glu Ile Gln Gly Glu Leu Tyr Glu Lys Glu Ile Ala Pro Asn Leu Ser

85 90 95

Ala Gly Asp Ala Leu Ala Phe Ala His Gly Phe Asn Ile His Phe Glu

100 105 110

Val Ile Thr Pro Pro Lys Glu Val Asp Val Phe Leu Val Ala Pro Lys

115 120 125

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

130 135 140

Pro Ala Leu Phe Ala Val Tyr Gln Asp Ala Thr Gly Asn Ala Gln Glu

145 150 155 160

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

165 170 175

Leu Glu Thr Thr Phe Lys Glu Glu Thr Glu Thr Asp Leu Phe Gly Glu

180 185 190

Gln Ala Val Leu Cys Gly Gly Leu Thr Ser Met Ile Glu Thr Gly Phe

195 200 205

Glu Thr Leu Val Glu Ala Gly Tyr Gln Pro Glu Leu Ala Tyr Phe Glu

210 215 220
Val Cys His Glu Met Lys Leu Ile Val Asp Leu Ile Tyr Glu Gly Gly
225 230 235 240
Phe Ala Lys Met Arg Asp Ser Ile Ser Asn Thr Ala Glu Tyr Gly Asp
245 250 255
Tyr Val Ser Gly Ser Arg Ile Ile Thr Glu Gln Thr Lys Ala Asn Met
260 265 270

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

<210> 377

<211> 1023

<212> DNA

<213> Lactococcus lactis

<400> 377

atggctgtca ctatgtacta cgaagaagac gttgaagtcg ccgccttggc cggtaagaag	60
atcgccgtta tcggttacgg ttcccaaggt cacgctcatg cccaaaactt gagagactct	120
ggtcacgatg ttatcattgg tgttagacag ggtaaacttt tcgacaaggc taaagaagat	180
ggtttcgaaa cctttgaagt tgggtgaagct gtcgccaagg ctgacgttat catggctctg	240
gtccagatg aattgcaaca gtctatctac gaagaagaca ttaagccaaa ctgaaggct	300
ggttctgctt tgggtttcgc tcacggtttc aacatccatt tcggttacat tgaagttcca	360
gaagacgttg atgtcttcat ggtcgcccca aaggctccag gtcacttggt tagacgtacc	420
ttcacgaag gtttcggtac ccagctttg ttcgtctctc accaaaacgc cactggtcat	480
gctagagaaa tcgcatgga ctgggctaag ggtattggtt gtgctagagt tggatcatt	540
gaaaccactt tcaaggaaga aaccgaagaa gatitgttcg gtgaacaagc tgtcctgtgc	600
ggtaggttaa ctgccttggg tgaagctggg ttcgaaacct tgactgaagc cggttacgct	660
ggtagaattgg cttacttcga agttttacac gaaatgaagt tgatcgttga cttgatgtac	720

gaagggtggtt tcaccaagat gagacaatct atttccaaca ctgctgaatt cggtagactac 780
gtcacccggtc caagaatcat cactgatgaa gttaagaaga acatgaagtt ggtcctggcc 840

gacatccaat ccggtaaatt tgctcaggat ttcggtgacg attttaaggc cggtagacca 900
aagttgaccg cttaccgtga agctgccaag aacttggaag tcgaaaagat cggtagctgaa 960
ttgagaaagg ctatgccatt cacccaatcc ggtgacgacg acgctttcaa gatttaccaa 1020
taa 1023

<210> 378

<211> 340

<212> PRT

<213> Lactococcus lactis

<400> 378

Met Ala Val Thr Met Tyr Tyr Glu Glu Asp Val Glu Val Ala Ala Leu
1 5 10 15

Ala Gly Lys Lys Ile Ala Val Ile Gly Tyr Gly Ser Gln Gly His Ala
20 25 30

His Ala Gln Asn Leu Arg Asp Ser Gly His Asp Val Ile Ile Gly Val
35 40 45

Arg Gln Gly Lys Ser Phe Asp Lys Ala Lys Glu Asp Gly Phe Glu Thr
50 55 60

Phe Glu Val Gly Glu Ala Val Ala Lys Ala Asp Val Ile Met Val Leu
65 70 75 80

Ala Pro Asp Glu Leu Gln Gln Ser Ile Tyr Glu Glu Asp Ile Lys Pro
85 90 95

Asn Leu Lys Ala Gly Ser Ala Leu Gly Phe Ala His Gly Phe Asn Ile
100 105 110

His Phe Gly Tyr Ile Glu Val Pro Glu Asp Val Asp Val Phe Met Val
115 120 125

Ala Pro Lys Ala Pro Gly His Leu Val Arg Arg Thr Phe Thr Glu Gly
130 135 140

Phe Gly Thr Pro Ala Leu Phe Val Ser His Gln Asn Ala Thr Gly His
145 150 155 160

Ala Arg Glu Ile Ala Met Asp Trp Ala Lys Gly Ile Gly Cys Ala Arg
165 170 175
Val Gly Ile Ile Glu Thr Thr Phe Lys Glu Glu Thr Glu Glu Asp Leu
180 185 190
Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Leu Thr Ala Leu Val Glu
195 200 205

Ala Gly Phe Glu Thr Leu Thr Glu Ala Gly Tyr Ala Gly Glu Leu Ala
210 215 220
Tyr Phe Glu Val Leu His Glu Met Lys Leu Ile Val Asp Leu Met Tyr
225 230 235 240
Glu Gly Gly Phe Thr Lys Met Arg Gln Ser Ile Ser Asn Thr Ala Glu
245 250 255
Phe Gly Asp Tyr Val Thr Gly Pro Arg Ile Ile Thr Asp Glu Val Lys
260 265 270

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

Lys Ile Tyr Gln
340

<210> 379

<211> 1023

<212> DNA

<213> Lactococcus lactis

<400> 379

atggcagtta caatgtatta tgaagatgat gtagaagtat cagcacttgc tggaaagcaa	60
attgcagtaa tcggttatgg ttcaacaagga catgctcacg cacagaattt gcgtgattct	120
ggtcacaacg ttatcattgg tgtgcgccac ggaaaatctt ttgataaagc aaaagaagat	180

ggctttgaaa catttgaagt aggagaagca gtagctaaag ctgatgttat tatggttttg 240
gcaccagatg aacttcaaca atccatttat gaagaggaca tcaaaccaaa cttgaaagca 300
ggttcagcac ttggttttgc tcacggattt aatatccatt ttggctatat taaagtacca 360

gaagacgttg acgtctttat ggttgcgcct aaggctccag gtcaccttgt ccgtcggact 420
tatactgaag gttttgttac accagctttg ttgttttcac accaaaatgc aagtggatcat 480
gcgcgtgaaa tcgcaatgga ttgggccaaa ggaattgggt gtgctcgagt gggaattatt 540
gaaacaactt ttaaagaaga aacagaagaa gatttgtttg gagaacaagc tgttctatgt 600
ggaggtttga cagcacttgt tgaagccggt ttgaaacac tgacagaagc tggatacgt 660
ggcgaattgg cttactttga agttttgcac gaaatgaaat tgattgttga cctcatgtat 720
gaaggtgggt ttactaaaat gcgtcaatcc atctcaaata ctgctgagtt tggcgattat 780

gtgactggtc cacgatttat tactgacgaa gttaaaaaga atatgaagct tgttttggt 840
gatattcaat ctggaaaatt tgctcaagat ttcgttgatg acttcaaagc ggggcgtcca 900
aaattaatag cctatcgca agctgcaaaa aatcttgaaa ttgaaaaaat tggggcagag 960
ctacgtcaag caatgccatt cacacaatct ggtgatgacg atgcctttaa aatctatcag 1020
taa 1023

<210> 380

<211> 340

<212> PRT

<213> Lactococcus lactis

<400> 380

Met Ala Val Thr Met Tyr Tyr Glu Asp Asp Val Glu Val Ser Ala Leu

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

65	70	75	80
Ala Pro Asp Glu Leu Gln Gln Ser Ile Tyr Glu Glu Asp Ile Lys Pro			

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

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

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

 260 265 270
 Lys Asn Met Lys Leu Val Leu Ala Asp Ile Gln Ser Gly Lys Phe Ala
 275 280 285
 Gln Asp Phe Val Asp Asp Phe Lys Ala Gly Arg Pro Lys Leu Ile Ala
 290 295 300
 Tyr Arg Glu Ala Ala Lys Asn Leu Glu Ile Glu Lys Ile Gly Ala Glu
 305 310 315 320
 Leu Arg Gln Ala Met Pro Phe Thr Gln Ser Gly Asp Asp Asp Ala Phe

 325 330 335

Lys Ile Tyr Gln

340

<210> 381

<211> 1023

<212> DNA

<213> Streptococcus thermophilus

<400> 381

atggccgttc aaatggaata cgaaaaggac gttaaggtcc cagccttgga tggtaagaag	60
atcgccgtta tcggttacgg ttctcaaggt cacgctcatt ctcaaaactt gagagacacc	120
ggtcacgatg ttatcattgg tgttagacca ggtaaatctt tcgacaaggc taaagaagac	180
ggttttgata cctacactgt cgctgaagcc actaagttgg cgcacgttat catgatattg	240
gctccagatg aaatccaaca ggaattgtac gaagctgaaa ttgccccaaa ctggaagct	300
ggtaatgccg tcggtttcgc tcacggtttc aacatccatt tcgaattcat caaggttcca	360
gccgacgttg atgtcttcat gtgtgctcca aagggtccag gtcacttggc cagacgtacc	420
ttcgaagaag gttttgggtg ccagcctttg tacgccgttt accaagatgc tactggtaac	480
gccaaaggaca tcgctatgga ttggtgtaaa ggtattgggtg ctgccagagt tggtttggtg	540
gaaaccactt acaaggaaga aaccgaagaa gacttggttc gtgaacaggc tgtcttatgc	600
ggtgggttaa ccgctttgat cgaaactggt ttcgaagttt tgactgaagc tggttacgcc	660
ccagaattgg cttacttcga agtcctgcac gaaatgaagt tgatcgttga cttgatctac	720
gaaggtggtt tcaagaagat gagacaatct atctccaaca ccgctgaatt cggtgactac	780
gtttctggtc caagagtcac cactgaacaa gtttaaggaaa acatgaaggc tgttttggct	840
gacattcaga atggtaaatt cgctaacgac ttcggttaacg attacaaggc cggtagacca	900
aagttgaccg ctiaccgtga agaagctgcc aacttggaag tcgaaaaggc cgggtgctgaa	960
ttgagaaagg ctatgccatt cgttggtcaa aacgacgacg acgctttcaa gatttacaac	1020
tga	1023

<210> 382

<211> 340

<212> PRT

<213> Streptococcus thermophilus

<400> 382

Met Ala Val Gln Met Glu Tyr Glu Lys Asp Val Lys Val Pro Ala Leu

1	5	10	15

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

 Tyr Thr Val Ala Glu Ala Thr Lys Leu Ala Asp Val Ile Met Ile Leu
 65 70 75 80
 Ala Pro Asp Glu Ile Gln Gln Glu Leu Tyr Glu Ala Glu Ile Ala Pro
 85 90 95
 Asn Leu Glu Ala Gly Asn Ala Val Gly Phe Ala His Gly Phe Asn Ile
 100 105 110
 His Phe Glu Phe Ile Lys Val Pro Ala Asp Val Asp Val Phe Met Cys
 115 120 125

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

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

 Phe Gly Asp Tyr Val Ser Gly Pro Arg Val Ile Thr Glu Gln Val Lys

260 265 270
 Glu Asn Met Lys Ala Val Leu Ala Asp Ile Gln Asn Gly Lys Phe Ala
 275 280 285
 Asn Asp Phe Val Asn Asp Tyr Lys Ala Gly Arg Pro Lys Leu Thr Ala
 290 295 300
 Tyr Arg Glu Glu Ala Ala Asn Leu Glu Ile Glu Lys Val Gly Ala Glu
 305 310 315 320

Leu Arg Lys Ala Met Pro Phe Val Gly Gln Asn Asp Asp Asp Ala Phe
 325 330 335
 Lys Ile Tyr Asn
 340

<210> 383

<211> 11017

<212> DNA

<213> Artificial sequence

<220><223> pRS426::GPD-xpk1+ADH-eutD

<400> 383

gatcccccgg gctgcaggaa ttcgatatca agcttatcga taccgtcgac ctcgaggggg 60
 ggccccgtac ccaattcgcc ctatagttag tctgtattacg cgcgctcact ggccgtcggt 120
 ttacaacgtc gtgactggga aaacctggc gttaccaac ttaatgcct tgcagcacat 180

ccccctttcg ccagctggcg taatagcgaa gaggcccgca ccgatcgccc ttcccaacag 240
 ttgcgcagcc tgaatggcga atggcgcgac ggcacctgta gcggcgcat aagcgcgcg 300
 ggtgtggtgg ttacgcgcag cgtgaccgct acacttgcca gcgcctagc gcccgtcct 360
 ttcgctttct tcccttcctt tctcgccacg ttcgceggct ttccccgtca agctctaaat 420
 cgggggctcc ctttagggtt ccgatttagt gctttacggc acctcgacce caaaaaactt 480
 gattaggggtg atggttcacg tagtgggcca tcgccctgat agacggtttt tcgccctttg 540
 acgttggagt ccacgttctt taatagtga ctctgttcc aaactggaac aacactcaac 600

cctatctcgg tctattcttt tgatttataa gggattttgc cgatttcggc ctattggtta 660
 aaaaatgagc tgatttaaca aaaatttaac gcgaatttta acaaaatatt aacgtttaca 720
 atttcctgat gcggtatfff ctctttacgc atctgtgcgg tatttcacac cgcatagggt 780
 aataactgat ataattaaat tgaagctcta atttgtgagt ttagtataca tgcatttact 840
 tataatacag ttttttagtt ttgctggcgg catcttctca aatatgttc ccagcctgct 900

tttctgtaac gttcaccctc taccttagca tcccttcctt ttgcaaatag tcctcttcca	960
acaataataa tgtcagatcc tgtagagacc acatcatcca cggttctata ctgttgaccc	1020
aatgcgtctc ccttgtcacc taaaccacaca ccgggtgtca taatcaacca atcgtaacct	1080
tcatctcttc caccatgtc tctttgagca ataaagccga taacaaaatc tttgtcgctc	1140
ttcgcaatgt caacagtacc cttagtatat tctccagtag ataggagacc ctgcatgac	1200
aattctgcta acatcaaaag gcctctaggt tcttttgta cttcttctgc cgcctgcttc	1260
aaaccgctaa caatacctgg gccaccacaca ccgtgtgcat tcgtaatgtc tgccattct	1320
gctattctgt atacaccgc agagtactgc aatttgactg tattaccaat gtcagcaaat	1380
tttctgtctt cgaagagtaa aaaattgtac ttggcggata atgcctttag cggcttaact	1440
gtgccctcca tggaaaaatc agtcaagata tccacatgtg tttttagtaa acaaattttg	1500
ggacctaatg cttcaactaa ctccagtaat tctttggtgg tacgaacatc caatgaagca	1560
cacaagtttg tttgcttttc gtgcatgata ttaaatagtc tggcagcaac aggactagga	1620
tgagtagcag cagtttctt atatgtagct ttcgacatga tttatcttcg tttcctgcag	1680
gtttttgttc tgtgcagttg ggttaagaat actgggcaat ttcattgttc ttcaacacta	1740
catatgcgta tatatacaa tctaagtctg tgctccttc ttcgttcttc cttctgttcg	1800
gagattaccg aatcaaaaaa atttcaaaga aaccgaaatc aaaaaaaga ataaaaaaa	1860
aatgatgaat tgaattgaaa agctgtggta tgggtgactc tcagtacaat ctgctctgat	1920
gccgcatagt taagccagcc ccgacacccg ccaacacccg ctgacgcgcc ctgacgggct	1980
tgtctgctcc cggcatccgc ttacagacaa gctgtgaccg tctccgggag ctgcatgtgt	2040
cagaggtttt caccgtcacc accgaaacgc gcgagacgaa agggcctcgt gatacgccta	2100
tttttatagg ttaatgtcat gataataatg gtttcttagt atgatccaat atcaaaggaa	2160
atgatagcat tgaaggatga gactaatcca attgaggagt ggcagcatat agaacagcta	2220
aagggtagt ctgaaggaag catacgtac cccgcatgga atgggataat atcacaggag	2280
gtactagact acctttcacc ctacataaat agacgcatat aagtacgcat ttaagcataa	2340
acacgacta tgccgttctt ctcatgtata tatatatata ggcaacacgc agatataggt	2400
gcgacgtgaa cagttagctg tatgtgcgca gctcgcgttg cattttcgga agcgtcgtt	2460
ttcggaacg ctttgaagtt cctattccga agttcctatt ctctagaaag tataggaact	2520
tcagagcgtt ttigaaaacc aaaagcgtc tgaagacgca ctttcaaaaa accaaaaacg	2580
caccggactg taacgagcta ctaaaatatt gcgaataccg cttccacaaa cattgtctaa	2640

aagtatctct ttgctatata tctctgtgct atatccctat ataacctacc catccacctt	2700
tcgctccttg aacttgcac taaactcgac ctctacattt tttatgttta tctctagtat	2760
tactcttttag acaaaaaaat tgtagtaaga actattcata gagtgaatcg aaaacaatac	2820
gaaaatgtaa acatttccta tacgtagtat atagagacaa aatagaagaa accgttcata	2880
atcttctgac caatgaagaa tcatcaacgc taccactttc tgttcacaaa gtatgcgcaa	2940
tccacatcgg tatagaatat aatcggggat gcctttatct tgaaaaaatg caccgcgagc	3000
ttcgctagta atcagtaaac gcgggaagtg gagtcaggct tttttatgg aagagaaat	3060
agacacccaa gtagccttct tctaacctta acggacctac agtgcaaaaa gttatcaaga	3120
gactgcatta tagagcgac aaaggagaaa aaaagtaatc taagatgctt tgtagaaaa	3180
atagcgtctc cgggatgcat tttttagaa caaaaaagaa gtatagattc tttgttgga	3240
aaatagcgtc ctgcgttgc atttctgttc tgtaaaaatg cagctcagat tctttgtttg	3300
aaaaattagc gcctcgcgt tgcatttttg tttacaaaa atgaagcaca gattcttcgt	3360
tggtaaaata gcgctttgc gttgcatttc tgttctgtaa aaatgcagct cagattcttt	3420
gtttgaaaaa ttagcgtctc cgcggtgcat tttgttcta caaaatgaag cacagatgct	3480
tcgttcaggt ggcacttttc ggggaaatgt gcgcggaacc cctatttgtt tatttttcta	3540
aatacattca aatatgtatc cgctcatgag acaataaccc tgataaatgc ttcaataata	3600
ttgaaaaagg aagagtatga gtattcaaca tttccgtgtc gcccttattc ccttttttgc	3660
ggcattttgc ctctctgttt ttgtctaccc agaaacgctg gtgaaagtaa aagatgctga	3720
agatcagttg ggtgcacgag tgggttacat cgaactggat ctcaacagcg gtaagatcct	3780
tgagagtttt cgccccgaag aacgttttcc aatgatgagc acttttaaag ttctgctatg	3840
tggcgcggtt ttatcccgtt ttgacgccgg gcaagagcaa ctgcgtcgcc gcatacacta	3900
ttctcagaat gacttggttg agtactcacc agtcacagaa aagcatctta cggatggcat	3960
gacagtaaga gaattatgca gtgctgcat aacctagat gataaactg cggccaactt	4020
acttctgaca acgatcggag gaccgaagga gctaaccgct tttttgcaca acatggggga	4080
tcatgtaaat cgcttgatc gttgggaacc ggagctgaat gaagccatac caaacgacga	4140
gcgtgacacc acgatgcctg tagcaatggc aacaacgttg cgcaactat taactggcga	4200
actacttact ctacttccc ggcaacaatt aatagactgg atggaggcgg ataaagtgc	4260
aggaccactt ctgcgctcgg ccttccggc tggctggttt attgctgata aatctggagc	4320
cggtagcgtt ggttctcgcg gtatcattgc agcactgggg ccagatggta agccctccc	4380

tatcgtagtt atctacacga cggggagtca ggcaactatg gatgaacgaa atagacagat	4440
cgctgagata ggtgcctcac tgattaagca ttggttaactg tcagaccaag ttactcata	4500
tatacttttag attgatttaa aacttcattt ttaatttaaa aggatctagg tgaagatcct	4560
ttttgataat ctcatgacca aaatccctta acgtgagttt tcgttccact gagcgtcaga	4620
ccccgtagaa aagatcaaag gatcttcttg agatcctttt tttctgcgcg taatctgctg	4680
cttgcaaaca aaaaaaccac cgctaccagc ggtggtttgt ttgccggatc aagagctacc	4740
aactcttttt ccgaaggtaa ctggcttcag cagagcgcag ataccaaata ctgtccttct	4800
agtgtagccg tagttaggcc accacttcaa gaactctgta gcaccgccta catacctcgc	4860
tctgctaate ctgttaccag tggctgctgc cagtggcgat aagtcgtgic ttaccgggtt	4920
ggactcaaga cgatagttac cggataaggc gcagcggctg ggctgaacgg ggggttcgtg	4980
cacacagccc agcttggagc gaacgacctt caccgaactg agatacctac agcgtgagct	5040
atgagaaagc gccacgcttc ccgaagggag aaaggcggac aggtatccgg taagcggcag	5100
ggtcggaaca ggagagcgca cgagggagct tccaggggga aacgcctggt atctttatag	5160
tcctgtcggg ttctgccacc tctgacttga gcgtcgatct ttgtgatgct cgtcaggggg	5220
gcggagccta tggaaaaacg ccagcaacgc ggccttttta cggttcctgg ccttttgctg	5280
gccttttgct cacatgttct ttctgcgtt atccccgat tctgtggata accgtattac	5340
cgcttttag tgagctgata ccgctgcgcg cagccgaacg accgagcgca gcgagtcagt	5400
gagcgaggaa gcggaagagc gcccaatacg caaacgcct ctccccgcgc gttggccgat	5460
tcattaatgc agctggcacg acaggtttcc cgactggaaa gcgggcagtg agcgcaacgc	5520
aattaatgtg agttacctca ctcattaggc acccaggct ttacacttta tgcttcggc	5580
tcctatgttg tgtggaattg tgagcggata acaatttcac acaggaaaca gctatgacca	5640
tgattacgcc aagcgcgcaa ttaaccctca ctaaaggga caaaagctgg agctccaccg	5700
cggatagatc tagttcgagt ttatcattat caatactgcc atttcaaaga atacgtaaat	5760
aattaatagt agtgattttc ctaactttat ttagtcaaaa aattagcctt ttaattctgc	5820
tgtaacccgt acatgcccaa aatagggggc gggttacaca gaatatataa catcgtaggt	5880
gtctgggtga acagtttatt cctggcatcc actaaatata atggagcccg ctttttaagc	5940
tggcatccag aaaaaaaaaa aatcccagca ccaaaatatt gttttcttca ccaaccatca	6000
gttcataggt ccatttcttt agcgcaacta cagagaacag gggcacaac aggcaaaaaa	6060
cgggcacaac ctcaatggag tgatgcaacc tgcttgaggt aatgatgac acaaggcaat	6120
tgaccacgc atgtatctat ctcatcttct tacaccttct attaccttct gctctctctg	6180
atttgaaaa agctgaaaaa aaaggttgaa accagttccc tgaaattatt ccctacttg	6240

actaataagt atataaagac ggtaggtatt gattgtaatt ctgtaaatct atttcttaaa	6300
cttcttaaat tctactttta tagttagtct ttttttagt tttaaaacac caagaactta	6360
gtttcgaata aacacacata aacgctgagg atgacaacag attactcatc accagcatat	6420
ttgcaaaaag ttgataagta ctggcgtgct gccaaactact tatcagttgg tcaactttat	6480
ttaaaagata atccactatt acaacggcca ttgaaggcca gtgacgttaa ggttcatcca	6540
attggtcact gggggacgat tgccggtcaa aactttatct atgctcatct taaccgggtc	6600
atcaacaagt acggtttgaa gatgttctac gttgaaggtc caggtcattg tggtcaagtg	6660
atggtttcaa actcttacct tgacgtact tacaccgata tttatccaga aattacgcag	6720
gatgttgaag ggatgcaaaa gctcttcaag caattctcat tcccaggagg ggttgcttcc	6780
catgcggcac ctgaaacacc cggttcaatc cacgaagggt gcgaacttgg ttactcaatt	6840
tcacacgggg ttggggcaat tcttgacaat cctgacgaaa tcgccgcggt tgttgttgt	6900
gatggggaat ccgaaacggg tccattagca acttcatggc aatcaacgaa gttcattaac	6960
ccaatcaacg acggggctgt tttaccaatc ttgaacttaa atggttttaa gatttctaat	7020
ccaacgattt ttggtcgacg tcttgatgct aagattaagg aatacttcca aagcatgaat	7080
tggaaccaa tcttcgtga aggtgacgat cctgaaaagg ttcaccacgc ctagctaag	7140
gccatggatg aagccgttga aaagatcaag gcaatccaga agcatgctcg cgaaaataac	7200
gatgcaacat tgccagtatg gccaatgac gtcttcgcg cacctaaggg ctggactggt	7260
ccgaagtcat gggacgggtga taagatcgaa ggttcattcc gtgctcatca aattccgatt	7320
cctgttgatc aaaatgacat ggaacatgcg gatgcttag ttgattggct cgaatcatat	7380
caacaaaag aactcttcaa tgaagatggc tctttgaagg atgatattaa agaaattatt	7440
cctactgggg acagtcggat ggctgctaac ccaatcacca atggtggggc cgatccgaaa	7500
gccttgaact taccaaactt ccgtgattat gcggtcgata cgtccaaaga aggcgcgaat	7560
gttaagcaag atatgatcgt ttggtcagac tatttgcggg atgcatcaa gaaaaatcct	7620
gataacttcc ggttgttcgg acctgatgaa accatgtcta accgtttata tgggtgtctc	7680
gaaaccacta atcgtcaatg gatggaagac attcatccag atagtacca atatgaagca	7740
ccagctggcc gggctcttaga tgctcagtt tctgaacacc aagctgaagg ttggttagaa	7800
ggttacgtct taactggacg tcatgggtta tttgccagtt atgaagcctt cctacgcgtt	7860
gtggactcaa tgttgacgca acacttcaag ttggttacgt aagccaatga acttgattgg	7920
cgtaaaaagt accatcact taacattatc gcggcttcaa ctgtattcca acaagaccat	7980
aatggttata cccaccaaga tccagggtgca ttaactcatt tggccgaaaa gaaaccagaa	8040
tacattcgtg aatatttacc agccgatgcc aacacgttat tagctgtcgg tgacgtcatt	8100

ttccggagcc aagaaaagat caactacgtg gttacgtcaa aacacccacg tcaacaatgg	8160
ttcagcattg aagaagctaa gcaattagtt gacaatggtc ttggtatcat tgattgggca	8220
agtacggacc aaggtagcga accagacatt gtctttgcag ctgctgggac ggaaccaacg	8280
cttgaacgt tggctgcat ccaattacta cagcacagtt tcccagagat gaagattcgt	8340
ttcgtgaacg tggctgacat cttgaagtta cgtagtcctg aaaaggatcc gcggggcttg	8400
tcagatgctg agtttgacca ttactttact aaggacaaac cagtggctctt tgctttccac	8460
ggttacgaag acttagttcg tgacatcttc ttgatcgtc acaaccataa cttatacgtc	8520
cacggttacc gtgaaaatgg tgatattacc acaccattcg acgtacgggt catgaaccag	8580
atggaccgct tcgacttagc taagtcggca attgcggcgc aaccagcaat ggaaaacact	8640
ggtgcggcct tcgttcaatc catggataat atgcttgcta aacacaatgc ctatatccgg	8700
gatgccggaa ctgacttgcc agaagttaat gattggcaat ggaagggttt aaaataatta	8760
attaatcatg taattagtta tgtcacgctt acattcacgc cctcctcca catccgctct	8820
aaccgaaaag gaaggagtta gacaacctga agtctaggtc cctatttatt tttttaata	8880
gttatgttag tattaagaac gttatttata ttcaaatTT tcttttttt tctgtacaaa	8940
cgcgtgtacg catgtaacat tatactgaaa accttgcttg agaagggtttt gggacgctcg	9000
aaggctttaa tttgcgggcg gccgctctag aactagtacc acagggttg tcctctgagg	9060
acataaaata cacaccgaga ttcatcaact cattgctgga gttagcatat ctacaattgg	9120
gtgaaatggg gagcgatttg caggcatttg ctcgcatgc cggtagaggt gtggtcaata	9180
agagcgacct catgctatac ctgagaaagc aacctgacct acaggaaaga gttactcaag	9240
aataagaatt ttcgttttaa aacctagag tcactttaaa atttgtatac acttattttt	9300
tttataactt atttaataat aaaaatcata aatcataaga aattcgctta ctcttaatta	9360
attaagctaa tccttgggct gctgtaataa tcgcaacctt ataaacgtct tcttcaactgc	9420
atccacgtga caagtcggag accggcttgi tcaggccttg caagacagga cccaccgctt	9480
caaaatgacc aaatcgttgc gcaatcttat agccaatatt accagactga agctctggaa	9540
atacaaagac attggcatga ccagctactt tggaaccagg agccttttgc aaaccaactt	9600
tttcaacgaa ggccgcgtca aattgaagt caccatcgat agccaattcc ggttcagcag	9660
cttgccctt ggccgttgc tcttgcaact tagtgacctt ttcacctta gccgaacctt	9720
tagttgagaa gctgagcatc gcaactttcg ggccaatata gaagacctta gcagtagccg	9780
cactctgagt ggcaatttcc gctaacgtat cggcatcggg atcaatattg atggcacagt	9840

cagcaaagac gtagcggtcc tcaccccttt gcatgataaa tgcacccgag attcgggtgtg 9900
aaccgggctt ggtcttaata atttgaacg ctggccgtac cgtatcacca gttggatgga 9960
ttgcacctga aaccatccca tccgctttgc ccatataaac gagcatcgtg ccaaagtagt 10020
tttcatcttc cagcatttta gccgcttggt ctggcgtatt cttacctttc cgccgttcaa 10080
cgagggcatc aagcattgct tgccttatctt cagccgggta tgtcgcagga tcaaggactt 10140
gaacgcctgt taaatccgca ttcaaatcgt tagccacagc ctgaactttg tccgttgac 10200
ctaaaacaat cggcttaacc aagccgtctg cagctaatacg cgctgccgca cgcacaattc 10260

ggggttcagt tccttcaggg aaaacaattg ttgatcttt accagtaatt ttttgtgcta 10320
atgactcaaa taaatccatc ctacgcgaga tagttgattg tatgcttggt atagcttgaa 10380
atattgtgca gaaaaagaaa caaggaagaa agggaaacgag aacaatgacg aggaaacaaa 10440
agattaataa ttgcaggtct atttatactt gatagcaaga cagcaaatct ttttttattt 10500
caaattcaag taactggaag gaaggccgta taccgttgct cattagagag tagtgtgcgt 10560
gaatgaagga aggaaaaagt ttcgtgtgct tcgagatacc cctcatcagc tctggaacaa 10620
cgacatctgt tgggtgctgct tttgtcgta attttttctt ttagtgtctt ccatcatctt 10680

tttgtcattg cggatatggt gagacaacaa cgggggagag agaaaagaaa aaaaaagaaa 10740
agaagttgca tgcgcctatt attacttcaa tagatggcaa atggaaaaag gtagtgaaa 10800
cttcgatatg atgatggcta tcaagtctag ggctacagta ttagttcgtt atgtaccacc 10860
atcaatgagg cagtgttaatt ggtgtagtct tgtttagccc attatgtctt gtctggtatc 10920
tgttctattg tatatctccc ctccgccacc tacatgttag ggagaccaac gaaggtatta 10980
taggaatccc gatgtatggg ttgggttgcc agaaaag 11017

<210> 384

<211> 1862

<212> DNA

<213> Artificial sequence

<220><223> 1.8kb KpnI-HincII geneticin resistant cassette

<400> 384

gtcgactcta gaggatcccc caccttggct aactcgttgt atcatcactg gataacttcg 60
tatagcatac attatacgaa gttatctagg gattcataac cattttctca atcgaattac 120
acagaacaca ccgtacaaac ctctctatca taactactta atagtcacac acgtactcgt 180
ctaaatacac atcatcgtcc tacaagttca tcaaagtgtt ggacagacaa ctataccagc 240
atggatctct tgiatcgggt cttttctccc gctctctcgc aataacaatg aacactgggt 300

caatcatagc ctacacaggt gaacagagta gcgtttatac agggtttata cggtgattcc 360

tacggcaaaa atttttcatt tctaaaaaaa aaaagaaaaa tttttctttc caacgctaga 420

aggaaaagaa aaatctaatt aaattgattt ggtgattttc tgagagttcc ctttttcata 480

tatcgaattt tgaatataaa aggagatcga aaaaattttt ctattcaatc tgttttctgg 540

ttttatttga tagttttttt gtgtattatt attatggatt agtactggtt tatatgggtt 600

tttctgtata acttcttttt attttagttt gtttaacttt attttgagtt acattatagt 660

tccctaactg caagagaagt aacattaaaa ctcgagatgg gtaaggaaaa gactcacgtt 720

tcgaggccgc gattaaattc caacatggat gctgatttat atgggtataa atgggctcgc 780

gataatgtcg ggcaatcagg tgcgacaatc tatcgattgt atgggaagcc cgatgcgcca 840

gagttgtttc tgaacatgg caaaggtagc gttgccaatg atgttacaga tgagatggtc 900

agactaaact ggctgacgga atttatgcct cttccgacca tcaagcattt tatccgtact 960

cctgatgatg catggttact caccactgcg atccccggca aaacagcatt ccaggtatta 1020

gaagaatatc ctgattcagg tgaaaatatt gttgatgcgc tggcagtgtt cctgcgccgg 1080

ttgcattcga ttctgtttg taattgtcct tttaacagcg atcgcgtatt tcgtctcgct 1140

caggcgcaat cacgaatgaa taacggtttg gttgatgcga gtgattttga tgacgagcgt 1200

aatggctggc ctgttgaaca agtctggaaa gaaatgcata agcttttgcc attctcaccg 1260

gattcagtcg tactcatgg tgatttctca cttgataacc ttatttttga cgaggggaaa 1320

ttaataggtt gtattgatgt tggacgagtc ggaatcgcag accgatacca ggatcttgcc 1380

atcctatgga actgcctcgg tgagttttct ctttcattac agaaacggct ttttcaaaaa 1440

tatggatttg ataatcctga tatgaataaa ttgcagtttc atttgatgct cgatgagttt 1500

ttctaagttt aacttgatac tactagattt tttctcttca tttataaaat ttttggttat 1560

aattgaagct ttagaagtat gaaaaaatcc ttttttttca ttctttgcaa ccaaaataag 1620

aagcttcttt tattcattga aatgatgaat ataaacctaa caaaagaaaa agactcgaat 1680

atcaaacatt aaaaaaaaaa aaaagaggtt atctgttttc ccatttagtt ggagtttgca 1740

ttttctaata gatagaactc tcaattaatg tggatttagt ttctctgttc gataacttcg 1800

tatagcatac attatacgaa gttatctgaa cattagaata cgtaatccgc aatgcgggta 1860

cc 1862

<210> 385

<211> 7938

<212> DNA

<213> Artificial sequence

<220><223> YZ067DkivDDhADH

<400> 385

tcgcgcgttt cggatgatgac ggtgaaaacc tctgacacat gcagctcccg gagacgggtca	60
cagcttgtct gtaagcggat gccgggagca gacaagcccg tcaggcgcg tcagcgggtg	120
ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc	180
accataaatt cccgttttaa gagcttgggtg agcgctagga gtcactgcca ggtatcggtt	240
gaacacggca ttagtcaggg aagtcataac acagtccttt cccgcaattt tctttttcta	300
ttactcttgg ctcctcttag tacactctat atttttttat gcctcggtaa tgattttcat	360
tttttttttt ccacctagcg gatgactctt tttttttctt agcgattggc attatcacat	420
aatgaattat acattatata aagtaatgtg atttcttcga agaataact aaaaaatgag	480
caggcaagat aaacgaaggc aaagatgaca gagcagaaag ccctagtaaa gcgtattaca	540
aatgaaacca agattcagat tgcgatctct ttaaagggtg gtcccctagc gatagagcac	600
tcgatcttcc cagaaaaaga ggcagaagca gtagcagaac aggccacaca atcgcaagtg	660
attaacgtcc acacaggtat agggtttctg gaccatatga tacatgctct ggccaagcat	720
tccggctggt cgtaatcgt tgagtgcatt ggtgacttac acatagacga ccatcacacc	780
actgaagact gcgggattgc tctcggtcaa gcttttaaag aggccctagg ggccgtgcgt	840
ggagtaaaaa ggtttggatc aggatgtgcg cttttggatg aggcactttc cagagcgggtg	900
gtagatcttt cgaacaggcc gtacgcagtt gtcgaacttg gtttgcaaag ggagaaagta	960
ggagatctct cttgcgagat gatcccgcat tttcttgaag gctttgcaga ggctagcaga	1020
attaccctcc acgttgattg tctgcgaggc aagaatgac atcaccttag tgagagtgcg	1080
ttcaaggctc ttgcggttgc cataagagaa gccacctcgc ccaatggtac caacgatgtt	1140
ccctccacca aagggtttct tatgtagtga caccgattat ttaaagctgc agcatacgat	1200
atataacat gtgtatatat gtatacctat gaatgtcagt aagtatgtat acgaacagta	1260
tgatactgaa gatgacaagg taatgcatca ttctatacgt gtcattctga acgaggcgcg	1320
ctttcttttt ttctttttgc tttttctttt tttttctctt gaactcgacg gatctatgcg	1380
gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcaggaaat tgtaagcgtt	1440
aatatattgt taaaattcgc gttaaatttt tgttaaata gctcatTTTT taaccaatag	1500
gccgaaatcg gcaaaatccc ttataaatca aaagaataga ccgagatagg gttgagtgtt	1560
gttccagttt ggaacaagag tccactatta aagaacgtgg actccaacgt caaaggcgca	1620
aaaaccgtct atcaggcgca tggcccaacta cgtggccggc ttacatacgt ttgcatacgt	1680

cgatatagat aataatgata atgacagcag gattatcgta atacgtaata gctgaaaatc	1740
tcaaaaatgt gtgggtcatt acgtaaataa tgataggaat gggattcttc tatttttcct	1800
ttttccattc tagcagccgt cgggaaaacg tggcatcctc tctttcgggc tcaattggag	1860
tcacgtgcc gtgagcatcc tctctttcca tatctaaca ctagcacgt aaccaatgga	1920
aaagcatgag cttagcgttg ctccaaaaaa gtattggatg gttaatacca ttgtctgtt	1980
ctcttctgac ttgactcct caaaaaaaaa aatctacaat caacagatcg cttcaattac	2040
gccctcaca aaactttttt ccttcttctt cgcccacgtt aaattttatc cctcatgttg	2100
tctaacggat ttctgcactt gatttattat aaaaagacaa agacataata cttctctatc	2160
aatttcagtt atgtttcttc cttgcgttat tcttctgttc ttctttttct ttgtcatat	2220
ataaccataa ccaagtaata catattcaaa cacgtgagta tgactgacaa aaaaactctt	2280
aaagacttaa gaaatcgtag ttctgtttac gattcaatgg ttaaatcacc taatcgtgct	2340
atgttgcgtg caactgglat gcaagatgaa gactttgaaa aacctatcgt cgggtgcatt	2400
tcaacttggg ctgaaaacac accttgtaat atccacttac atgactttgg taaactagcc	2460
aaagtcggtg ttaaggaagc tgggtccttg ccagttcagt tcggaacaat cacggtttct	2520
gatggaatcg ccatgggaac ccaaggaatg cgtttctcct tgacatctcg tgatattatt	2580
gcagattcta ttgaagcagc catgggaggt cataatgcgg atgcttttgt agccattggc	2640
ggtttgata aaaacatgcc cggttctgtt atcgctatgg ctaacatgga tatcccagcc	2700
atttttgctt acggcggaac aattgcacct ggtaatttag acggcaaaga tatcgattta	2760
gtctctgtct ttgaaggtgt cggccattgg aaccacggcg atatgaccaa agaagaagtt	2820
aaagcttttg aatgtaatgc ttgtcccgtt cctggaggct gcggtggat gtatactgct	2880
aacacaatgg cgacagctat tgaagttttg ggacttagcc ttccgggttc atcttctcac	2940
ccggtgaat ccgcagaaaa gaaagcagat attgaagaag ctggtcgcgc tgttgtcaaa	3000
atgctcgaat tgggcttaaa accttctgac attttaacgc gtgaagcttt tgaagatgct	3060
attactgtaa ctatggctct gggagggttca accaactcaa ccttcacct cttagctatt	3120
gcccatgctg ctaatgtgga attgacactt gatgatttca atactttcca agaaaaagtt	3180
cctcatttgg ctgatttgaa accttctggt caatatgtat tccaagacct ttacaaggtc	3240
ggagggttac cagcagttat gaaatatctc cttaaaaatg gcttcttca tgggtaccgt	3300
atcacttgta ctggcaaac agtcgctgaa aatttgaagg cttttgatga tttaacacct	3360
ggtcaaaagg ttattatgcc gcttgaaaat cctaaacgtg aagatggtcc gctcattatt	3420

ctccatggta acttggctcc agacggtgcc gttgccaaag tttctggtgt aaaagtgcgt	3480
cgtcatgtcg gtctgtctaa ggtctttaat tctgaagaag aagccattga agctgtcttg	3540
aatgatgata ttgttgatgg tgatgttgtt gtcgtacgtt ttgtaggacc aaagggcgg	3600
cctggatgac ctgaaatgct ttccctttca tcaatgattg ttggtaaagg gcaaggtgaa	3660
aaagttgccc ttctgacaga tggccgcttc tcaggtggta cttatggtct tgcgtgggt	3720
catatcgctc ctgaagcaca agatggcggg ccaatcgctt acctgcaaac aggagacata	3780
gtcactattg accaagacac taaggaattt cactttgata tctccgatga agagttaaaa	3840
catcgtcaag agaccattga attgccaccg ctctattcac gcggtatcct tggtaaatat	3900
gtcacatcgt tttcgtctgc ttctagggga gccgtaacag acttttgaa gcctgaagaa	3960
actggcaaaa aatgttgctc tggttgctgt ggttaagcgg ccgcttaat tcaaataat	4020
tgatatagtt ttttaatgag tattgaatct gtttagaaat aatggaatat tatttttatt	4080
tatttattta tattattggt cggctctttt cttctgaagg tcaatgacaa aatgatatga	4140
aggaaataat gatttctaaa attttacaac gtaagatatt ttacaaaag cctagctcat	4200
cttttgtcat gcaactattt actcacgctt gaaattaacg gccagtccac tgcggagtca	4260
tttcaaagtc atcctaactg atctatcgtt ttgatagct cattttgag ttcgcgagga	4320
tcccagcttt tgttcccttt agtgagggtt aattgcgcgc ttggcgtaat catggtcata	4380
gtgttttctt gtgtgaaatt gttatccgct cacaattcca cacaacatac gagccggaag	4440
cataaagtgt aaagcctggg gtgcctaata agtgagctaa ctacattaa ttgcgttgcg	4500
ctcactgccc gctttccagt cgggaaacct gtcgtgccag ctgcattaat gaatcggcca	4560
acgcgcgggg agaggcgggt tgcgtattgg gcgctcttcc gcttctctgc tcaactgactc	4620
gctgcgctcg gtcgttcggc tgcggcgagc ggtatcagct cactcaaagg cgtaatacg	4680
gttatccaca gaatcagggg ataacgcagg aaagaacatg tgagcaaaag gccagcaaaa	4740
ggccaggaac cgtaaaaagg ccgcttgctt ggcgtttttc cataggctcc gccccctga	4800
cgagcatcac aaaaatcgac gctcaagtca gaggtggcga aaccgcagac gactataaag	4860
ataccaggcg tttccccctg gaagctccct cgtgcgtctt cctgttccga cctgcccgt	4920
taccggatac ctgtccgctt ttctcccttc gggaagcgtg gcgctttctc atagctcacg	4980
ctgtaggtat ctcaattcgg tgtaggtcgt tcgctccaag ctgggctgtg tgcacgaacc	5040
ccccgttcag cccgaccgct gcgccttata cggttaactat cgtcttgagt ccaaccgggt	5100
aagacacgac ttatcgccac tggcagcagc cactggtaac aggattagca gagcgaggta	5160
tgtagcgggt gctacagagt tcttgaagtg gtggcctaac tacggctaca ctagaagaac	5220
agtatttggt atctgcgctc tgctgaagcc agttaccttc ggaaaaagag ttggtagctc	5280

ttgatccggc aaacaaacca ccgctggtag cggtggtttt tttgtttgca agcagcagat 5340
 tacgcgcaga aaaaaaggat ctcaagaaga tcctttgatc tttctacgg ggtctgacgc 5400
 tcagtggac gaaaactcac gttaagggat tttggtcatg agattatcaa aaaggatctt 5460
 cacctagatc cttttaaat aaaaatgaag ttttaaatca atctaaagta tatatgagta 5520

 aacttggctt gacagttacc aatgcttaat cagtgaggca cctatctcag cgatctgtct 5580
 atttcgttca tccatagtgt cctgactccc cgtcgtgtag ataactacga tacgggaggg 5640
 ctaccatct ggccccagtg ctgcaatgat accgcgagac ccacgctcac cggtccaga 5700
 tttatcagca ataaaccagc cagccggaag ggccgagcgc agaagtgttc ctgcaacttt 5760
 atccgcctcc atccagtcta ttaattgttg ccgggaagct agagtaagta gttcgccagt 5820
 taatagtttg cgcaacgttg ttgccattgc tacaggcatc gtggtgtcac gctcgtcgtt 5880
 tggtaggct tcattcagct ccggttccca acgatcaagg cgagttacat gatccccat 5940

 gttgtcaaaa aaagcggtta gctccttcgg tcctccgatc gttgtcagaa gtaagttggc 6000
 cgcagtgtta tcactcatgg ttatggcagc actgcataat tctcttactg tcatgccatc 6060
 cgtaagatgc ttttctgtga ctggtagta ctcaaccaag tcattctgag aatagtgtat 6120
 gcggcgaccg agttgctctt gcccgcgctc aatacgggat aataccgcgc cacatagcag 6180
 aactttaaaa gtgctcatca ttggaaaacg ttcttcgggg cgaaaactct caaggatctt 6240
 accgctgttg agatccagtt cgatgtaacc cactcgtgca cccaactgat cttcagcatc 6300
 ttttactttc accagcgttt ctgggtgagc aaaaacagga aggcaaatg ccgcaaaaaa 6360

 gggaataagg gcgacacgga aatgttgaat actcatactc ttctttttc aatattattg 6420
 aagcatttat cagggttatt gtctcatgag cggatacata tttgaatgta tttagaaaaa 6480
 taaacaaata ggggttccgc gcacatttcc ccgaaaagtg ccacctgaac gaagcatctg 6540
 tgcttcattt tgtagaacaa aaatgcaacg cgagagcgct aatttttcaa acaaagaatc 6600
 tgagctgcat ttttacagaa cagaaatgca acgcgaaagc gctattttac caacgaagaa 6660
 tctgtgttc atttttgtaa aacaaaaatg caacgcgaga gcgctaattt ttcaaacaaa 6720
 gaatctgagc tgcattttta cagaacagaa atgcaacgcg agagcgctat ttaccaaca 6780

 aagaatctat acttctttt tgttctacaa aaatgcatcc cgagagcgct atttttctaa 6840
 caaagcatct tagattactt tttttctct tttgtgcgtc tataatgcag tctcttgata 6900
 actttttgca ctgtaggctc gttaagggtt gaagaaggct actttggtgt ctattttctc 6960
 ttccataaaa aaagcctgac tccacttccc gcgtttactg attactagcg aagctgcggg 7020
 tgcattttt caagataaag gcatccccga ttatattcta taccgatgtg gattgcgcat 7080
 actttgtgaa cagaaagtga tagcgttgat gattcttcat tggtcagaaa attatgaacg 7140

gtttcttcta ttttgtctct atatactacg tataggaaat gtttacattt tcgtattgtt 7200

ttcgattcac tctatgaata gttcttacta caattttttt gtctaaagag taataactaga 7260

gataaacata aaaaatgtag aggtcgagtt tagatgcaag ttcaaggagc gaaagggtga 7320

tgggtagggtt atatagggat atagcacaga gatatatagc aaagagatac ttttgagcaa 7380

tgtttgtgga agcgggtattc gcaatatattt agtagctcgt tacagtccgg tgcgtttttg 7440

gttttttgaa agtgcgtctt cagagcgctt ttggttttca aaagcgctct gaagttccta 7500

tactttctag agaataggaa cttcggaata ggaacttcaa agcgtttccg aaaacgagcg 7560

cttccgaaaa tgcaacgcga gctgcgcaca tacagctcac tgttcacgtc gcacctatat 7620

ctgcgtgttg ccigtatata tatatacatg agaagaacgg catagtgcgt gtttatgctt 7680

aaatgcgtac ttatatgcgt ctatttatgt aggatgaaag gtagtctagt acctcctgtg 7740

atattatccc attccatcgc gggatcgtgta tgccttcttc agcactaccc tttagctgtt 7800

ctatatgctg ccactcctca attggattag tctcatcctt caatgctatc atttcctttg 7860

atattggatc atactaagaa accattatta tcatgacatt aacctataaa aataggcgta 7920

tcacgaggcc ctttcgtc 7938

<210> 386

<211> 4586

<212> DNA

<213> Artificial sequence

<220><223> pLA54 template DNA

<400> 386

gggtaccgag ctcgattca ctggccgtcg ttttacaacg tcgtgactgg gaaaaccctg 60

gcgttaccca acttaatcgc cttgcagcac atcccccttt cgccagctgg cgtaatagcg 120

aagaggcccc caccgatcgc ccttcccaac agttgcgcag cctgaatggc gaatggcgcc 180

tgatgcggta ttttctcctt acgcatctgt gcggtatttc acaccgcata tgggtgactc 240

tcagtacaat ctgctctgat gccgcatagt taagccagcc ccgacaccg ccaacaccg 300

ctgacgcgcc ctgacgggct tgtctgctcc cggcatccgc ttacagacaa gctgtgaccg 360

tctccgggag ctgcatgtgt cagaggtttt caccgtcatc accgaaacgc gcgagacgaa 420

agggcctcgt gatagccta tttttatagg ttaatgtcat gataataatg gtttcttaga 480

cgtcagggtgg cacttttcgg ggaaatgtgc gcggaacccc tatttgttta tttttctaaa 540

tacattcaaa tatgtatccg ctcatgagac aataaccctg ataaatgctt caataatatt 600

gaaaaaggaa gagtatgagt attcaacatt tccgtgtcgc ccttattccc ttttttcggg 660

cattttgcct tctgttttt gctcaccag aaacgctggt gaaagtaaaa gatgctgaag	720
atcagttggg tgcacgagtg ggttacatcg aactggatct caacagcggg aagatccttg	780
agagttttcg ccccgaaaga cgtttttcaa tgatgagcac ttttaaagtt ctgctatgtg	840
gcgcggtatt atcccgtatt gacgccgggc aagagcaact cggtcgccgc atacactatt	900
ctcagaatga cttggttgag tactcaccag tcacagaaaa gcatcttacg gatggcatga	960
cagtaagaga attatgcagt gctgccataa ccatgagtga taacactgcg gccaacttac	1020
ttctgacaac gatcggagga ccgaaggagc taaccgcttt ttgacacaac atgggggagc	1080
atgtaactcg cttgatcgt tgggaaccgg agctgaatga agccatacca aacgacgagc	1140
gtgacaccac gatgcctgta gcaatggcaa caacgttgcg caaactatta actggcgaac	1200
tacttactct agcttcccg caacaattaa tagactggat ggaggcggat aaagttgcag	1260
gaccacttct gcgctcgcc cttccggctg gctggtttat tgcgtataaa tctggagccg	1320
gtgagcgtgg gtctcgcggt atcattgcag cactggggcc agatggtaag ccctcccgta	1380
tcgtagttat ctacacgacg gggagtcagg caactatgga tgaacgaaat agacagatcg	1440
ctgagatagg tgcctcactg attaagcatt ggtaactgtc agaccaagtt tactcatata	1500
tactttagat tgatttaaaa cttcattttt aatttaaaag gatctaggtg aagatccttt	1560
ttgataatct catgacaaa atcccttaac gtgagttttc gttccactga gcgtcagacc	1620
ccgtagaaaa gatcaaagga tcttcttgag atcctttttt tctgcgcgta atctgctgct	1680
tgcaaacaaa aaaaccaccg ctaccagcgg tggtttggtt gccgatcaa gagctaccaa	1740
ctctttttcc gaaggttaact ggcttcagca gagcgcagat accaaatact gtccttctag	1800
tgtagccgta gtiaggccac cacttcaaga actctgtagc accgcctaca tacctcgctc	1860
tgctaactct gttaccagtg gctgctgcca gtggcgataa gtcgtgtctt accgggttgg	1920
actcaagacg atagttaccg gataaggcgc agcggtcggg ctgaacgggg ggttcgtgca	1980
cacagcccag cttggagcga acgacctaca ccgaactgag atacctacag cgtgagctat	2040
gagaaagcgc cacgcttccc gaaggagaga aggcggacag gtatccggtg agcggcaggg	2100
tcggaacagg agagcgcacg agggagcttc cagggggaaa cgcctggtat ctttatagtc	2160
ctgtcgggtt tcgccacctc tgacttgagc gtcgattttt gtgatgctcg tcaggggggc	2220
ggagcctatg gaaaaacgcc agcaacgcgg cttttttacg gttcctggcc ttttctggc	2280
cttttctca catgttcttt cctgcgttat ccctgatctc tgttgataac cgtattaccg	2340
cctttgagtg agctgatacc gctcgccgca gccgaacgac cgagcgcagc gagtcatga	2400
gcgaggaagc ggaagagcgc ccaatacgca aaccgcctct ccccgcgctg tggccgattc	2460

attaatgcag ctggcacgac aggtttcccg actggaaagc gggcagtgag cgcaacgcaa	2520
ttaatgtgag ttagctcact cattaggcac ccaggtttt acactttatg cttccggctc	2580
gtatgttgtg tggaattgtg agcggataac aatttcacac aggaaacagc tatgaccatg	2640
attacgcaa gcttgcacgc ctgcaggtcg actctagagg atccccgcat tgcggattac	2700
gtattctaataa gttcagataa cttcgtatag catacattat acgaagttaa ctagggtatc	2760
ataaccattt tctcaatcga attacacaga acacaccgta caaacctctc tatcataact	2820
acttaatagt cacacacgta ctgctctaaa tacacatcat cgtcctacaa gttcatcaaa	2880
gtgttgga gacaactata ccagcatgga tctcttgat cggttctttt ctcccgtct	2940
ctcgcaataa caatgaacac tgggtcaatc atagcctaca caggtgaaca gagtagcgtt	3000
tatacagggt ttataggtg attcctacgg caaaaatttt tcatttctaa aaaaaaaag	3060
aaaaattttt cttccaacg ctagaaggaa aagaaaaatc taattaaatt gatttggtga	3120
ttttctgaga gtccctttt tcatatatcg aattttgaat ataaaaggag atcgaaaaaa	3180
tttttctatt caatctgtt tctggtttta ttgatagtt ttttgtgta ttattattat	3240
ggattagtagc tggtttatat gggtttttct gtataacttc tttttatttt agtttgttta	3300
atcttatttt gagttacatt atagtccct aactgcaaga gaagtaacat taaaactcga	3360
gatgggtaag gaaaagactc acgtttcgag gccgcgatta aattccaaca tggatgctga	3420
tttatatggg tataaatggg ctgcgataa tgcgggcaa tcaggtgcga caatctatcg	3480
attgtatggg aagcccgatg cgccagagtt gtttctgaaa catggcaaag gtagcgttgc	3540
caatgatgtt acagatgaga tggtcagact aaactggctg acggaattta tgcctcttc	3600
gaccatcaag cattttatcc gtactcctga tgatgcatgg ttactacca ctgcgatccc	3660
cggcaaaaca gcattccagg tattagaaga atatcctgat tcaggtgaaa atattgttga	3720
tgcgctggca gtgttctgc gccggttgca ttcgattcct gtttgaatt gtccttttaa	3780
cagcgatcg gtatttcgtc tcgtcaggc gcaatcacga atgaataacg gtttggttga	3840
tgcgagtgt tttgatgacg agcgtaatgg ctggcctgtt gaacaagtct ggaaagaaat	3900
gcataagctt ttgccattct caccggattc agtcgtcact catggtgatt tctcattga	3960
taaccttatt ttgacgagg ggaaattaat aggttgtatt gatgttgac gtagcggat	4020
cgcagaccga taccaggatc ttgccatcct atggaactgc ctcggtgagt tttctccttc	4080
attacagaaa cggctttttc aaaaatatgg tattgataat cctgatatga ataaattgca	4140
gtttcatttg atgctcgatg agtttttcta agtttaactt gataactacta gattttttct	4200

cttcatttat aaaatttttg gttataattg aagctttaga agtatgaaaa aatccttttt 4260
 ttccattctt tgcaacaaaa ataagaagct tcttttattc attgaaatga tgaatataaa 4320
 cctaacaaaa gaaaaagact cgaatatcaa acattaaaaa aaaataaaag aggttatctg 4380
 ttttccatt tagttggagt ttgcattttc taatagatag aactctcaat taatgtggat 4440
 ttagtttctc tgttcgtttt tttttgtttt gtctctactg tatttacatt tctatttagt 4500
 atttagttat tcatataatc tataacttcg tatagcatac attatacgaa gttatccagt 4560
 gatgatacaa cgagtttagcc aagggtg 4586

<210> 387

<211> 7589

<212> DNA

<213> Artificial sequence

<220><223> pUC19-kan::pdc1::FBA-alsS::TRX1

<400> 387

tatttgatc gaggtgtcta gtcttctatt acactaatgc agtttcaggg ttttgaaac 60
 cacactgttt aaacagtgtt ccttaataca ggatacctct tttttttcc ttggttccac 120
 taattcatcg gttttttttt tggaagacat cttttccaac gaaaagaata tacatatcgt 180
 ttaagagaaa ttctccaaat ttgtaagaa gcggaccag acttaaggcc gcccgcaaat 240
 taaagccttc gagcgtccca aaaccttcic aagcaagggt ttcagtataa tgttacatgc 300

gtacacgct ctgtacagaa aaaaaagaaa aatttgaaat ataaataacg ttcttaatac 360
 taacataact ataaaaaaat aaataggac ctgacttca ggttgtctaa ctcttcctt 420
 ttcggttaga gcggatgtgg ggggagggcg tgaatgtaag cgtgacataa ctaattacat 480
 gattaattaa ctagagagct ttctgtttca tgagttcccc gaattctttc ggaagcttgt 540
 cacttgctaa attaatgtta tcactgtagt caaccgggac atcgatgatg acaggacctt 600
 cagcgttcat gccttgacgc agaacatctg ccagctggtc tgggtattct acgcgcaagc 660
 cagttgctcc gaagctttcc gcataattca cgatatcgat atttccgaaa tcgaccgcag 720

atgtacggtt atattttttc aattgctgga atgcaacat gtcatatgtg ctgtcgttcc 780
 atacaatgtg tacaattggt gcttttagtc gaactgtgt ctctaattcc attgctgaga 840
 ataagaaacc gccgtcacca gagacagaaa ccactttttc tcccggtttc accaatgaag 900
 cgccgattgc ccaaggaagc gcaacgccga gtgtttgcat accgttactg atcattaatg 960
 ttaacggctc gtagctgagg aaataacgtg acatccaaat ggctgtcgaa ccgatatcgc 1020
 aagttactgt aacatgatca tcgactgcat tacgcaactc tttaacgatt tcaagagggt 1080

gcgctctgtc tgatttccaa tctgcaggca cctgtcacc ttcattgata tattgtttta 1140

aatcagaaag gattttctgc tcacgtctg caaattccac tttcacagca tcgtgttcga 1200

tatgattgat cgtggacgga atgtcacgga tcaattcaag atcaggctgg taagcatgat 1260

caatgtcagc gataatctcg tctaaatgga taattgtccg gtctccattg atattccaga 1320

atttcggatc atattcaatc ggggtcatagc cgatcgtcag aacaacatct gcctgtctta 1380

gcagtaaate gccaggctgg ttgcggaaca aaccgatacg gccaaaatat tgatectcta 1440

aatctctaga aagggtaccg gcagcttgat atgtttcaac aaatggaagc tgaacctttt 1500

tcaaaagctt gcgaaccgct ttaattgtct cgggtcttcc gcctttcatg ccgaccaaaa 1560

cgacaggaag ttttgcgtgt tggatttttg ctatggccgc actgattgca tcatctgctg 1620

caggaccgag ttttggcgtt gcaacagcac gcacgttttt cgtatttttg acttcattca 1680

caacatcttg cgaaaagctc acaaaagcgg cccagcctg cctgtgtgac gctatcctaa 1740

atgcatttgt aacagcttcc ggtatatttt ttacatcttg aacttctaca ctgtattttg 1800

taatcggctg gaatagcgcc gcattatcca aagattgatg tgtccgtttt aaacgatctg 1860

cacggatcac gtttccagca agcgcaacga cagggtctcc ttcagtgttc gctgtcagca 1920

ggcctgttgc caagttagag gcacccggtc ctgatgtgac taacacgact cccggttttc 1980

cagttaaacg gccgactgct tgggcatga atgctgcgtt ttgttcgtgc cgggcaacga 2040

taatttcagg tcctttatct tgtaaagcgt caaataccgc atcaattttt gcacctggaa 2100

tgccaaatac atgtgtgaca cttgtctcca ctaagcaatc aacaacaagc tccgccctc 2160

tgttttcac aagggttttt tgttcttttg ttgcttttgt caacatcctc agctctagat 2220

ttgaatatgt attacttggt tatggttata tatgacaaa gaaaaagaag aacagaagaa 2280

taacgcaagg aagaacaata actgaaattg atagagaagt attatgtctt tgtcttttta 2340

taataaatca agtgcagaaa tccgttagac aacatgaggg ataaaattta acgtgggcga 2400

agaagaagga aaaaagtttt tgtgaggcg taattgaagc gatctgttga ttgtagattt 2460

ttttttttt aggagtcaaa gtcagaagag aacagacaaa tggatttaac catccaatac 2520

ttttttggag caacgctaag ctcatgcttt tccattgggt acgtgctcag ttgttagata 2580

tggaaagaga ggatgtcac ggcagcgtga ctccaattga gcccgaaaga gaggatgcca 2640

cgttttcccg acggctgcta gaatggaaaa aggaaaaata gaagaatccc attcctatca 2700

ttatttacgt aatgaccac acatttttga gattttcaac tattacgtat tacgataatc 2760

ctgctgtcat tatcattatt atctatatcg acgtatgcaa cgtatgtgaa gccaaagtagg 2820

caattattta gtactgtcag tattgttatt catttcagat cttaaacctt accaggccaa	2880
ttcaacagac tgcgggcaac ttcttgtctg gtctttccat ggtaagtac agtcagtaa	2940
taatatgaac caatttatit ttctgtacat aaaaatgctt ataaaacttt aactaataat	3000
tagagattaa atcgcaaacg gccgactcta gaggatcccc caccttggtt aactcgttgt	3060
atcatcactg gataacttcg tatagcatac attatacgaa gttatctagg gattcataac	3120
cattttctca atcgaaftac acagaacaca ccgtacaaac ctctctatca taactactta	3180
atagtcacac acgtactcgt ctaaatacac atcatcgtcc tacaagttca tcaaagtgtt	3240
ggacagacaa ctataccagc atggatctct tgtatcggtt cttttctccc gctctctcgc	3300
aataacaatg aacactgggt caatcatagc ctacacaggt gaacagagta gcgtttatac	3360
agggtttata cggtgattcc tacggcaaaa atttttcatt tctaaaaaa aaaagaaaaa	3420
tttttctttc caacgctaga aggaaaagaa aaatctaatt aaattgattt ggtgattttc	3480
tgagagtacc ctttttcata tatcgaaatt tgaatataaa aggagatcga aaaaattttt	3540
ctattcaatc tgittttctgg ttttatttga tagttttttt gtgtattatt attatggatt	3600
agtactgggt tatatgggtt tttctgtata acttcttttt attttagttt gtttaattct	3660
attttgagtt acattatagt tccctaactg caagagaagt aacattaaaa ctcgagatgg	3720
gtaaggaaaa gactcacgtt tccaggccgc gattaaattc caacatggat gctgatttat	3780
atgggtataa atgggctcgc gataatgtcg ggcaatcagg tgcgacaatc tatcgattgt	3840
atgggaagcc cgatgcgccg gagttgtttc tgaacatgg caaaggtagc gttgccaatg	3900
atgttacaga tgagatggtc agactaaact ggctgacgga atttatgcct ctccgacca	3960
tcaagcattt tatccgtact cctgatgatg catggttact caccactgcg atccccggca	4020
aaacagcatt ccaggatata gaagaatatc ctgattcagg tgaaaatatt gttgatgcgc	4080
tggcagtgtt cctgcgccgg ttgcattcga ttctgtttg taattgtcct tttacacgcg	4140
atcgctgatt tctctcgtc caggcgcaat cacgaatgaa taacggtttg gttgatgcga	4200
gtgattttga tgacgagcgt aatggctggc ctgttgaaca agtctggaaa gaaatgcata	4260
agcttttgcc attctaccg gattcagtcg tcaactatgg tgattttctca cttgataacc	4320
ttatttttga cgaggggaaa ttaatagggt gtattgatgt tggacgagtc ggaatcgag	4380
accgatacca ggatcttgcc atcctatgga actgcctcgg tgagttttct ccttcattac	4440
agaaacggct ttttcaaaaa tatggtattg ataactctga tatgaataaa ttgcagtttc	4500
atttgatgct cgatgagttt ttctaagttt aacttgatac tactagattt tttctcttca	4560
tttataaaat ttttggttat aattgaagct ttagaagtat gaaaaaatcc tttttttca	4620
ttctttgcaa ccaaaataag aagcttcttt tattcattga aatgatgaat ataaacctaa	4680

caaaagaaaa agactcgaat atcaaacatt aaaaaaaaaat aaaagaggtt atctgttttc	4740
ccatttagtt ggagtttgca ttttctaata gatagaactc tcaattaatg tggatttagt	4800
ttctctgttc gataacttcg tatagcatatc attatacgaa gttatctgaa cattagaata	4860
cgtaatccgc aatgcggggc cgcttaatta atctagagtc gacctgcagg catgcaagct	4920
tggcgtaatc atggtcatac ctgtttcctg tgtgaaattg ttatccgctc acaattccac	4980
acaacatacg agccggaagc ataaagtgtg aagcctgggg tgcctaatga gtgagctaac	5040
tcacattaat tgcgttcgac tcaactgccc ctttcagtc gggaaacctg tcgtgccagc	5100
tgcattaatg aatcgcccaa gcgcggggga gaggcgggtt gcgtattggg cgctcttcgc	5160
cttctcgtct cactgactcg ctgcgctcgg tcgttcggct gcggcgagcg gtatcagctc	5220
actcaaaggc ggtaatacgg ttatccacag aatcagggga taacgcagga aagaacatgt	5280
gagcaaaagg ccagcaaaag gccaggaacc gtaaaaaggc cgcttgctg gcgtttttcc	5340
ataggctccg ccccccagc gagcatcaca aaaatcgacg ctcaagtcag aggtggcgaa	5400
acccgacagg actataaaga taccaggcgt tccccctgg aagctccctc gtgcgctctc	5460
ctgttccgac cctgcgctt accggatacc tgcgcgctt tctcccttcg ggaagcgtgg	5520
cgctttctca tagctcagc tgtaggtatc tcagttcggg taggtcgtt cgctccaagc	5580
tgggctgtgt gcacgaacct cccgttcagc ccgaccgtg cgcttatcc ggtaactatc	5640
gtcttgagtc caacccgta agacacgact tatcgccact ggagcagcc actggtaaca	5700
ggattagcag agcgaggtat gtaggcggtg ctacagagtt cttgaagtgg tggcctaact	5760
acggctacac tagaaggaca gtatttggtg tctgcgctct gctgaagcca gttaccttcg	5820
gaaaaagagt tggtagctct tgatccggca aacaaaccac cgctggtagc ggtgggtttt	5880
ttgtttgcaa gcagcagatt acgcgcagaa aaaaaggatc tcaagaagat cctttgatct	5940
tttctacggg gtctgacgt cagtggaaacg aaaactcacg ttaagggatt ttggtcatga	6000
gattatcaaa aaggatcttc acctagatcc ttttaatta aaaatgaagt tttaatcaa	6060
tctaaagtat atatgagtaa acttggtctg acagttacca atgcttaatc agtgaggcac	6120
ctatctcagc gatctgtcta tttcgttcat ccatagttgc ctgactcccc gtcgtgtaga	6180
taactacgat acgggagggc ttaccatctg gccccagtgc tgcaatgata ccgcgagacc	6240
cacgtcacc ggtccagat ttatcagcaa taaaccagcc agccggaagg gccgagcgca	6300
gaagtggctc tgcaacttta tccgcctcca tccagtctat taattgttgc cgggaagcta	6360
gagtaagtag ttcgccagtt aatagtttgc gcaacgttgt tgccattgct acaggcatcg	6420
tggtgtcacg ctctgtgtt ggtatggctt cattcagctc cggttcccaa cgatcaaggc	6480
gagttacatg atccccatg ttgtgcaaaa aagcggttag ctcttcggg cctccgatcg	6540

ttgtcagaag taagttggcc gcagtgttat cactcatggt tatggcagca ctgcataatt 6600

ctcttactgt catgccatcc gtaagatgct tttctgtgac tggtagtac tcaaccaagt 6660

cattctgaga atagtgtatg cggcgaccga gttgctcttg cccggcgta atacgggata 6720

ataccgcgcc acatagcaga actttaaaag tgctcatcat tggaaaacgt tcttcggggc 6780

gaaaactctc aaggatctta ccgctgttga gatccagttc gatgtaaccc actcgtgcac 6840

ccaactgate ttcagcatct tttactttca ccagcgtttc tgggtgagca aaaacaggaa 6900

ggcaaaatgc cgcaaaaaag ggaataaggg cgacacggaa atgttgaata ctcatactct 6960

tcctttttca atattattga agcatttattc agggttattg tctcatgagc ggatacatat 7020

ttgaatgtat ttagaaaaat aaacaaatag gggttccgcg cacatttccc cgaaaagtgc 7080

cacctgacgt ctaagaaacc attattatca tgacattaac ctataaaaat aggcgtatca 7140

cgaggccctt tcgtctcgcg cgtttcgggtg atgacggtga aaacctctga cacatgcagc 7200

tcccggagac ggtcacagct tgtctgtaag cggatgccgg gagcagacaa gcccgtcagg 7260

gcgcgtcagc ggggtgttggc ggggtgtcggg gctggcttaa ctatgcggca tcagagcaga 7320

ttgtactgag agtgcacat atgcggtgtg aaataccgca cagatgcgta aggagaaaaat 7380

accgcatcag gcgccattcg ccattcaggc tgcgcaactg ttgggaaggg cgatcggtgc 7440

gggcctcttc gctattacgc cagctggcga aagggggatg tgctgcaagg cgattaagtt 7500

gggtaacgcc agggttttcc cagtcacgac gttgtaaaac gacggccagt gaattcgagc 7560

tcggtaccgg gggatccggc gcgccgttt 7589

<210> 388

<211> 19

<212> DNA

<213> Artificial sequence

<220><223> N160SeqR5 primer

<400> 388

tgagcccgaa agagaggat 19

<210> 389

<211> 338

<212> PRT

<213> Artificial sequence

<220><223> k9_Grace_SH

<400> 389

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

245 250 255

Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile Thr Glu Asp Thr Lys

260 265 270

Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln Asp Gly Thr Phe Ala

275 280 285

Lys Asp Phe Leu Val Asp Met Ser Asp Asp Gly Ser Gln Val His Phe

290 295 300

Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro Ala Glu Val Val Gly

305 310 315 320

Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp Glu Asp Lys Leu Ile

325 330 335

Asn Asn

<210> 390

<211> 24

<212> DNA

<213> Artificial sequence

<220><223> primer K9G9_EZ_F1

<400> 390

aaacatggaa gaatgtaaga tggc 24

<210> 391

<211> 25

<212> DNA

<213> Artificial sequence

<220><223> Primer K9G9_EZ_R1

<400> 391

tcagttgtta atcaacttgt cttcg 25

<210> 392

<211> 9613

<212> DNA

<213> Artificial sequence

<220><223> pHR81-PI1v5-KARI-K9.G9

<400> 392

aaacagtatg gaagaatgta agatggctaa gatttactac caagaagact gtaacttgtc	60
cttgttggat ggtaagacta tcgccgttat cggttacggt tctcaaggtc acgctcatgc	120
cctgaatgct aaggaatccg gttgtaacgt tatcattggt ttatacgaag gtgcgaagga	180
gtggaaaaga gctgaagaac aaggtttcga agtctacacc gctgctgaag ctgctaagaa	240
ggctgacatc attatgatct tgatcaacga tgaagcag gctaccatgt acaaaaacga	300
catcgaacca aacttggaag ccggtaacat gttgatgttc gctcacggtt tcaacatcca	360
tttcggttgt attgttccac caaaggacgt tgatgtcact atgatcgctc caaagggtcc	420
aggctcacacc gttagatccg aatacgaaga aggtaaagggt gtcccatgct tgggtgctgt	480
cgaacaagac gctactggca aggctttgga tatggctttg gcctacgctt tagccatcgg	540
tgggtgctaga gccggtgtct tggaaactac cttcagaacc gaaactgaaa ccgacttggt	600
cgggtgaacaa gcgtttttat gtgggtggtgt ctgcgctttg atgcaggccg gttttgaaac	660
cttggttgaa gccggttacg acccaagaaa cgcttacttc gaatgtatcc acgaaatgaa	720
gttgatcggt gacttgatct accaatctgg tttctccggt atgcgttact ctatctccaa	780
cactgctgaa tacggtgact acattaccgg tccaaagatc attactgaag ataccaagaa	840
ggctatgaag aagattttgt ctgacattca agatggtacc tttgccaagg acttcttggt	900
tgacatgtct gatgctgggt cccagggtcca ctcaaggct atgagaaaagt tggcctccga	960
acaccagct gaagtgtcgt gtgaagaaat tagatccttg tactcctggt ccgacgaaga	1020
caagttgatt aacaactgag gccctgcagg ccagaggaaa ataatatcaa gtgctggaaa	1080
ctttttctct tggaaatttt gcaacatcaa gtcatagtca attgaattga cccaatttca	1140
catttaagat tttttttttt tcatccgaca tacatctgta cactaggaag ccctgttttt	1200
ctgaagcagc ttcaaataata tatatttttt acatatattat tatgattcaa tgaacaatct	1260
aattaaatcg aaaacaagaa ccgaaacgcg aataaataat ttatttagat ggtgacaagt	1320
gtataagtcc tcatcgggac agctacgatt tctctttcgg ttttggtgta gctactggtt	1380
gctgtgacgc agcggcatta gcgcggcggt atgagctacc ctctggcct gaaagatggc	1440
gggaataaag cggaactaaa aattactgac tgagccatat tgaggtcaat ttgtcaactc	1500
gtcaagtcaac gtttggtgga cggccctttt ccaacgaatc gtatatacta acatgcgcgc	1560
gcttctata tacacatata catatatata tatatatata tgtgtgcgtg tatgtgtaca	1620
cctgtattta atttcttac tcgcgggttt ttcttttttc tcaattcttg gcttctcttt	1680
tctcgagcgg accggtacct ccgcggtgcc ggcagatcta tttaaatggc gcgccgacgt	1740
caggtggcac ttttcgggga aatgtgcgcg gaaccctat ttgtttattt ttctaataac	1800

attcaaatat gtatccgctc atgagacaat aacctgata aatgcttcaa taatattgaa 1860
aaaggaagag tatgagtatt caacatttcc gtgtcgccct tattcccttt ttgcgccat 1920
tttgccttcc tgtttttgct caccagaaa cgctggtgaa agtaaaagat gctgaagatc 1980
agttgggtgc acgagtgggt tacatcgaac tggatctcaa cagcggtaa atccttgaga 2040

gttttcgccc cgaagaacgt tttccaatga tgagcacttt taaagttctg ctatgtggcg 2100
cggattatc ccgtattgac gccgggcaag agcaactcgg tcgccgcata cactattctc 2160
agaatgactt ggttagtac tcaccagtca cagaaaagca tcttacggat ggcatgacag 2220
taagagaatt atgcagtgt gccataacca tgagtataa cactgcggcc aacttacttc 2280
tgacaacgat cggaggaccg aaggagctaa ccgctttttt gcacaacatg ggggatcatg 2340
taactcgct tgatcgttg gaaccggagc tgaatgaagc cataccaaac gacgagcgtg 2400
acaccagat gcctgtagca atggcaacaa cgttgcgcaa actattaact ggcgaactac 2460

ttactctagc ttcccgcaa caattaatag actggatgga ggcggataaa gttgcaggac 2520
cacttctgcg ctcgccctt ccggttggt ggtttattgc tgataaatct ggagccggtg 2580
agcgtgggtc tcgcggtatc attgcagcac tggggccaga tggtaagccc tcccgtatcg 2640
tagttatcta cagacgggg agtcaggcaa ctatggatga acgaaataga cagatcgctg 2700
agataggtgc ctactgatt aagcattggt aactgtcaga ccaagtttac tcatatatac 2760
tttagattga tttaaaactt catttttaat ttaaaaggat ctaggtgaag atcctttttg 2820
ataatctcat gacaaaaac ccttaacgtg agttttcggt ccactgagcg tcagaccccg 2880

tagaaaagat caaaggatct tcttgagatc ctttttttct gcgcgtaac tgctgcttgc 2940
aaacaaaaa accaccgcta ccagcgggtg tttgtttgcc ggatcaagag ctaccaactc 3000
ttttccgaa ggtaactggc ttcagcagag cgcagatacc aaatactgtt cttctagtgt 3060
agccgtagtt aggccaccac ttcaagaact ctgtagcacc gcctacatac ctgctctgc 3120
taatcctgtt accagtggct gctgccagt gcgataagtc gtgtcttacc gggttggact 3180
caagacgata gttaccgat aaggcgagc ggtcgggtg aacggggggt tcgtgcacac 3240
agcccagctt ggagcgaacg acctacaccg aactgagata cctacagcgt gagctatgag 3300

aaagcgccac gcttcccgaa gggagaaagg cggacaggta tccgtaagc ggcagggtcg 3360
gaacaggaga gcgcacgagg gagcttccag ggggaaacgc ctggtatctt tatagtcctg 3420
tcgggtttcg ccacctctga cttgagcgtc gattttttgt atgctcgtca ggggggcgga 3480
gcctatggaa aaacgccagc aacgcggcct ttttacggtt cctggccttt tgctggcctt 3540
ttgctcacat gttcttttct cggttatecc ctgattctgt ggataaccgt attaccgct 3600
ttgagtgcgc tgataccgct cgccgcagcc gaacgaccga gcgcagcgag tcagtgcgcg 3660

aggaagcgga agagcgccca atacgcaaac cgcctctccc cgcgcggttg cgcattcatt	3720
aatgcagctg gcacgacagg tttcccgaact ggaaagcggg cagtgcgcgc aacgcaatta	3780
atgtgagtta gtcactcat taggcacccc aggcctttaca ctttatgctt ccggctcgta	3840
tgttgtgtgg aattgtgagc ggataacaat ttcacacagg aaacagctat gaccatgatt	3900
acgccaagct ttttctttcc aatttttttt ttttcgtcat tataaaaaatc attacgaccg	3960
agattcccgg gtaataactg atataattaa attgaagctc taatttgtga gtttagtata	4020
catgcattta cttataatac agtttttttag ttttgctggc cgcactctct caaatatgct	4080
tcccagcctg cttttctgta acgttcaccc tctaccttag catcccttcc ctttgcaaat	4140
agtcctcttc caacaataat aatgtcagat cctgtagaga ccacatcatc cacggttcta	4200
tactgttgac ccaatgcgtc tcccttgta tctaaacca caccgggtgt cataatcaac	4260
caatcgtaac cttcatctct tccacccatg tctctttgag caataaagcc gataacaaaa	4320
tctttgtgc tcttcgcaat gtcaacagta cccttagtat attctccagt agataggag	4380
cccttgcatg acaattctgc taacatcaaa aggcctctag gttcctttgt tacttcttct	4440
gccgcctgct tcaaaccgt aacaatacct gggcccacca caccgtgtgc attcgtaatg	4500
tctgcccatt ctgctattct gtatacacc gcagagtact gcaatttgac tgtattacca	4560
atgtcagcaa attttctgtc ttccaagagt aaaaaattgt acttgcgga taatgccttt	4620
agcggcttaa ctgtgccctc catggaaaaa tcagtcaaga tatccacatg tgtttttagt	4680
aaacaaattt tgggacctaa tgcttcaact aactccagta attccttggt ggtacgaaca	4740
tccaatgaag cacacaagtt tgtttgcttt tcgtgcatga tattaaatag ctgggcagca	4800
acaggactag gatgagtagc agcacgttcc ttatatgtag ctttcgacat gatttatctt	4860
cgtttctgc aggtttttgt tctgtgcagt tgggttaaga atactgggca atttcatgtt	4920
tcttcaacac tacatatgcg tatatatacc aatctaagtc tgtgtctctt ccttcgttct	4980
tccttctgtt cggagattac cgaatcaaaa aaatttcaag gaaaccgaaa tcaaaaaaaaa	5040
gaataaaaaa aaaaatgatg attgaaaagc ttgcatgcct gcaggtcgac tctagtatac	5100
tccgtctact gtacgatata cttccgctca ggtccttgct ctttaacgag gccttaccac	5160
tcttttgta ctctattgat ccagctcagc aaaggcagtg tgatctaaga ttctatcttc	5220
gcgatgtagt aaaactagct agaccgagaa agagactaga aatgcaaaag gcacttctac	5280
aatggctgcc atcattatta tccgatgtga cgctgcattt tttttttttt tttttttttt	5340
tttttttttt tttttttttt tttttttttg tacaaatata ataaaaaaag agaattcttt	5400

taagcaagga ttttcttaac ttcttcggcg acagcatcac cgacttcggt ggtactgttg	5460
gaaccaccta aatcaccagt tctgatacct gcatccaaaa cttttttaac tgcattctca	5520
atggctttac cttcttcagg caagttcaat gacaatttca acatcattgc agcagacaag	5580
atagtggcga taggggttgac cttattcttt ggcaaactcg gagcggaacc atggcatggt	5640
tcgtacaaac caaatgcggt gttcttgtct ggcaaagagg ccaaggacgc agatggcaac	5700
aaaccaagg agcctgggat aacggaggct tcatcggaga tgatatacc aaacatgttg	5760
ctggtgatta taataccatt taggtgggtt gggttcttaa ctaggatcat ggccgcagaa	5820
tcaatcaatt gatgttgaac tttcaatgta gggaattcgt tcttgatggt ttctccaca	5880
gtttttctcc ataacttga agaggccaaa acattagctt tatccaagga ccaaataaggc	5940
aatggtggct catgtttag ggccatgaaa gcggccattc ttgtgattct ttgcattct	6000
ggaacggtgt attgttcaat atcccaagcg acaccatcac catcgtcttc ctttctctta	6060
ccaaagtaaa tacctccac taattctcta acaacaacga agtcagtacc tttagcaaat	6120
tgtggcttga ttggagataa gtctaaaaga gagtcggatg caaagttaca tggctttaag	6180
ttggcgtaac attgaagttc ttacggatt tttagtaaac cttgttcagg tctaacta	6240
ccggtacccc atttaggacc acccacagca cctaacaaaa cggcatacgc cttcttgag	6300
gcttcacgc cctcatctgg aagtgaaca cctgtagcat cgatagcgc accaccaatt	6360
aatgatgttt cgaaatcgaa cttgacattg gaacgaacat cagaaatagc ttaagaacc	6420
ttaatggctt cggtgtgat ttcttgacca acgtgggtcac ctggcaaac gacgatcttc	6480
ttagggcgag acattacaat ggtatatact tgaaatata ataaaaaaaa aaaaaaaaa	6540
aaaaaaaaa aatgcagctt ctcaatgata ttgaatacg ctttgaggag atacagccta	6600
atatccgaca aactgtttta cagatttacg atcgtacttg ttacccatca ttgaatttg	6660
aacatccgaa cctgggagtt ttccctgaaa cagatagtat atttgaacct gtataataat	6720
atatagtcta gcgctttacg gaagacaatg tatgtatttc ggttcctgga gaaactattg	6780
catctattgc ataggtaatc ttgcacgtcg catccccggt tcattttctg cgtttccatc	6840
ttgcatttca atagcatatc ttgtttaacg aagcatctgt gcttcatttt gtagaacaaa	6900
aatgcaacgc gagagcgcta atttttcaaa caaagaatct gagctgcatt ttacagaac	6960
agaaatgcaa cgcgaaagcg ctattttacc aacgaagaat ctgtgcttca tttttgtaaa	7020
acaaaaatgc aacgcgagag cgctaatttt tcaacaaaag aatctgagct gcatttttac	7080
agaacagaaa tgcaacgcga gagcgtatt ttaccaacaa agaactata cttctttttt	7140
gttctacaaa aatgcatccc gagagcgcta tttttctaac aaagcatctt agattacttt	7200
ttttctctt tgtcgctct ataatacag ctcttgataa cttttgcac ttaggtccg	7260

ttaaggtag aagaaggcta ctttgggtgc tttttctct tccataaaaa aagcctgact	7320
ccacttcccc cgtttactga ttactagcga agctgcgggt gcatttttcc aagataaagg	7380
catccccgat tatattctat accgatgtgg attgcgcata ctttgtgaac agaaagtgat	7440
agcgttgatg attcttcatt ggtcagaaaa ttatgaacgg tttcttctat ttgtctctta	7500
tatactacgt ataggaaatg tttacatttt cgtattgttt tcgattcact ctatgaatag	7560
ttcttactac aatttttttg tctaaagagt aatactagag ataaacataa aaaatgtaga	7620
ggtcgagttt agatgcaagt tcaaggagcg aaaggtggat gggtaggtta tatagggata	7680
tagcacagag atatatagca aagagatact tttgagcaat gtttgtggaa gcggtattcg	7740
caatatitaa gtagctcgtt acagtccggt gcgtttttgg tttttgaaa gtgcgtcttc	7800
agagcgtttt tggttttcaa aagcgtcttg aagtccctat actttctaga gaataggaac	7860
ttcggaatag gaacttcaaa gcgtttccga aaacgagcgc ttccgaaaat gcaacgcgag	7920
ctgcgcacat acagctcact gttcacgtcg cacctatata tgcgtgttgc ctgtatatat	7980
atatacatga gaagaacggc atagtgcgtg tttatgctta aatgcgtact tatatgcgtc	8040
tatttatgta ggatgaaagg tagtctagta cctcctgtga tattatecca ttccatgcgg	8100
ggtatcgtat gcttccctca gcactaccct ttagctgttc tatatgctgc cactcctcaa	8160
ttggattagt ctcatccttc aatgctatca tttcctttga tattggatca tatgcatagt	8220
accgagaaac tagaggatct ccattaccg acatttgggc gctatacgtg catatgttca	8280
tgtatgtatc tgtattttaa acacttttgi attatttttc ctcatatatg tgtataggtt	8340
tatacggatg atttaattat tacttcacca ccttttatit caggetgata tcttagcctt	8400
gttactagtc accggtggcg gccgcacctg gtaaaacctc tagtggagta gtagatgtaa	8460
tcaatgaagc ggaagccaaa agaccagagt agaggcctat agaagaaact gcgatacctt	8520
ttgtgatggc taaacaaaca gacatctttt tatatgtttt tacttctgta tatcgtgaag	8580
tagtaagtga taagcgaatt tggctaagaa cgttgtaagt gaacaaggga cctcttttgc	8640
ctttcaaaaa aggattaaat ggagttaatc attgagattt agttttcgtt agattctgta	8700
tccctaaata actcccttac ccgacgggaa ggcacaaaag acttgaataa tagcaaacgg	8760
ccagtagcca agaccaaata atactagagt taactgatgg tcttaaacag gcattacgtg	8820
gtgaactcca agaccaatat acaaaatata gataagtat tcttgccac caatttaagg	8880
agcctacatc aggacagtag taccattcct cagagaagag gtatacataa caagaaaatc	8940
gcgtgaacac cttatataac ttagcccggt attgagctaa aaaaccttgc aaaatttcct	9000
atgaataaga atacttcaga cgtgataaaa atttactttc taactcttct cacgtgccc	9060
ctatctgttc ttccgctcta ccgtgagaaa taaagcatcg agtacggcag ttcgctgtca	9120

ctgaactaaa acaataaggc tagttcgaat gatgaacttg cttgctgtca aacttctgag 9180

ttgccgtga tgtgacactg tgacaataaa ttcaaaccgg ttatagcggg ctccctccggg 9240

accggttctg ccacctccaa tagagctcag taggagtcag aacctctgcg gtggctgtca 9300

gtgactcatc cgcgtttcgt aagttgtgcg cgtgcacatt tcgcccgttc ccgctcatct 9360

tgcagcaggc ggaaattttc atcacgctgt aggacgcaaa aaaaaataa ttaatcgtac 9420

aagaatcttg gaaaaaaaaat tgaaaaattt tgtataaaag ggatgaccta acttgactca 9480

atggctttta caccagtat tttccctttc cttgtttgtt acaattatag aagcaagaca 9540

aaaacatata gacaacctat tcctaggagt tatatttttt tacctacca gcaatataag 9600

taaaaaactg ttt 9613

<210> 393

<211> 20

<212> DNA

<213> Artificial sequence

<220><223> Primer pBAD-For

<400> 393

atgccatagc atttttatcc 20

<210> 394

<211> 22

<212> DNA

<213> Artificial sequence

<220><223> Primer pBAD-Rev

<400> 394

ctgatttaat ctgtatcagg ct 22

<210> 395

<211> 50

<212> DNA

<213> Artificial sequence

<220><223> Primer N87PC1

<400> 395

ctgacatcat tatgatcttg atcccagatg aaaagcaggc taccatgtac 50

<210> 396

<211> 50

<212> DNA
 <213> Artificial sequence
 <220><223> Primer N87PC1r
 <400> 396
 gtacatggta gcctgctttt catctgggat caagatcata atgatgtcag 50
 <210> 397
 <211> 37
 <212> DNA
 <213> Artificial sequence
 <220><223> Primer oK57E1
 <400> 397
 ggttttattcg aaggtgcgga ggagtggaaa agagctg 37
 <210> 398
 <211> 37

 <212> DNA
 <213> Artificial sequence
 <220><223> Primer oK57E1r
 <400> 398
 cagctctttt ccactcctcc gcaccttcga ataaacc 37
 <210> 399
 <211> 38
 <212> DNA
 <213> Artificial sequence
 <220><223> Primer oY53F1
 <400> 399
 gtaacgttat cattggttta tacgaagggtg cggaggag 38
 <210> 400
 <211> 38
 <212> DNA
 <213> Artificial sequence
 <220><223> Primer oY53F1r
 <400> 400
 ctctccgca ccttcgaata aaccaatgat aacgttac 38

<210> 401
 <211> 43
 <212> DNA
 <213> Artificial sequence
 <220><223> Primer oK90L1
 <400> 401
 gatcttgatc ccagatgaat tgcaggctac catgtacaaa aac 43
 <210> 402
 <211> 42
 <212> DNA
 <213> Artificial sequence
 <220><223> Primer oK90L1r
 <400> 402
 gtttttgatc atggtagcct gcaattcatc tgggatcaag at 42
 <210> 403
 <211> 40
 <212> DNA
 <213> Artificial sequence
 <220><223> Primer oK90M1
 <400> 403
 ctgatccca gatgaaatgc aggctaccat gtacaaaaac 40
 <210> 404
 <211> 40
 <212> DNA
 <213> Artificial sequence
 <220><223> Primer oK90M1r
 <400> 404
 gtttttgatc atggtagcct gcatttcac tgggatcaag 40
 <210> 405
 <211> 338
 <212> PRT
 <213> Artificial sequence
 <220><223> K9_Ingrid_SH
 <400> 405

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

245 250 255
Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile Thr Glu Asp Thr Lys
260 265 270
Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln Asp Gly Thr Phe Ala
275 280 285

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

<210> 406

<211> 719

<212> DNA

<213> Artificial sequence

<220><223> Promoter UAS(PGK1)-PFBA1

<400> 406

aattaccgtc gctcgtgatt tgtttgcaaa aagaacaaaa ctgaaaaaac ccagacacgc 60
tcgacttctt gtcttcctat tgattgcagc ttccaatttc gtcacacaac aaggctcctgc 120
ctacttggct tcacatacgt tgcatacgtc gatatagata ataatagataa tgacagcagg 180
attatcgtaa tacgtaatat ttgaaaatct caaaaatgtg tgggtcatta cgtaaataat 240
gataggaatg ggattcttct atttttcttt ttccattct agcagccgtc gggaaaacgt 300
ggcatcctct ctttcgggct caattggagt cacgtgccg tgagcatcct ctctttccat 360
atctaacaac tgagcacgta accaatggaa aagcatgagc ttagcgttgc tccaaaaaag 420

tattggatgg ttaataccat ttgtctgttc tcttctgact ttgactcctc aaaaaaaaaa 480
aatctacaat caacagatcg ctccaattac gccctcaca aaactttttt ccttcttctt 540
cgccacggtt aaattttatc cctcatgttg tctaacggat ttctgcactt gatttattat 600
aaaaagacaa agacataata cttctctatc aatttcagtt attgttcttc cttgcgttat 660
tcttctgttc ttctttttct ttgtcatat ataaccataa ccaagtaata catattcaa 719

<210> 407

<211> 36
 <212> DNA
 <213> Artificial sequence
 <220><223> Primer oBP821
 <400> 407
 aattgttttaa actaagcgaa tttcttatga tttatg 36

<210> 408
 <211> 34
 <212> DNA
 <213> Artificial sequence
 <220><223> Primer oBP484
 <400> 408
 aattggccgg ccaaaaaaag catgcacgta taca 34

<210> 409
 <211> 37
 <212> DNA
 <213> Artificial sequence
 <220><223> Primer oBP822
 <400> 409
 aattgcgccc gcgtgcctca tctatatttc tgaaatc 37

<210> 410
 <211> 37
 <212> DNA
 <213> Artificial sequence
 <220><223> Primer oBP823
 <400> 410
 aattttaatt aatcagctct tgttttgttc tgcaaat 37

<210> 411
 <211> 1647
 <212> DNA
 <213> Artificial sequence
 <220><223> Listeria grayi ketoisovalerate decarboxylase coding
 sequence--codon optimized for expression in Saccharomyces

cerevisiae

<400> 411

atgtataccg taggacagta cttggttagat agactagaag agattggtat cgataaggtt 60
 ttcggtgtgc caggggatta caatttgact ttcttagatt acattcaaaa tcacgaagga 120
 ctttcttggc aagggaaatc taatgaacta aacgcagcat atgcagcaga tggctacgcc 180
 cgtgaaagag gcgtatcagc tcttggttact acattcggag tgggtgaact gtcagccatt 240

aacggaacag ctggtagttt tgcagaacaa gtcctgtca tccacatcgt gggttctcca 300
 actatgaatg tgcaatccaa caaaaagctg gttcatcatt ccttaggaat gggttaacttt 360
 cataacttta gtgaaatggc taaggaagtc actgccgcta caaccatgct tactgaagag 420
 aatgcagctt cagagatcga cagagtatta gaaacagcct tgttggaaaa gaggccagta 480
 tacatcaatc ttccaattga tatagctcat aaagcaatag ttaaactgc aaaagcacta 540
 caaacagaga aatcatctgg tgagagagag gcacaacttg cagaaatcat actatcacac 600
 ttagaaaagg ccgtcaacc tatcgtaatc gccggtcatg agatcgcccg ttccagata 660

agagaaagat ttgaaaactg gataaaccaa acaaagtgc cagtaaccaa ttggcatat 720
 ggcaaaggct ctttcaatga agagaacgaa catttcattg gtacctatta cccagctttt 780
 tctgacaaaa acgttctgga ttacgttgac aatagtact tctgtttaca ttttggtggg 840
 aaaatcattg acaattctac ctctctattt tctcaaggct ttaagactga aaacacttta 900
 accgctgcaa atgacatcat tatgtgcca gatgggtcta cttactctgg gatttctctt 960
 aacggtcttt tggcagagct ggaaaaacta aactttactt ttgctgatac tgctgctaaa 1020
 caagctgaat tagctgtttt cgaaccacag gccgaaacac cactaaagca agacagattt 1080

caccaagctg ttatgaactt tttgcaagct gatgatgtgt tggctactga gcaggggaca 1140
 tcatctttcg gtttgatgtt ggcacctctg aaaaagggtg tgaatttgat cagtcaaaca 1200
 ttatggggct ccataggata cacattacct gctatgattg gttcacaat tgctgcccc 1260
 gaaaggagac acattctatc catcggtgat ggatcttttc aactgacagc acaggaaatg 1320
 tccaccatct tcagagagaa attgacacca gtgatattca ttatcaataa cgatggctat 1380
 acagtcgaaa gagccatcca tggagaggat gagagttaca atgatatacc aacttgaac 1440
 ttgaattag ttgctgaaac atttggtggt gatgccgaaa ctgtcgacac tcacaacgtt 1500

ttcacagaaa cagacttcgc taatacttta gctgctatcg atgctactcc tcaaaaagca 1560
 catgtcgttg aagttcatat ggaacaaatg gatatgccag aatcattgag acagattggc 1620
 ttagccttat ctaagcaaaa ctcttaa 1647

<210> 412

<211> 36
 <212> DNA
 <213> Artificial sequence
 <220><223> Primer oBP828
 <400> 412
 aattgtttaa acatgtatac cgtaggacag tacttg 36
 <210> 413
 <211> 35
 <212> DNA
 <213> Artificial sequence
 <220><223> oBP829
 <400> 413
 aattgtttaa acttaagagt ttgcttaga taagg 35
 <210> 414
 <211> 29
 <212> DNA
 <213> Artificial sequence
 <220><223> Primer FBAp-F
 <400> 414
 ctacaatcaa cagatcgctt caattacgc 29
 <210> 415
 <211> 29
 <212> DNA
 <213> Artificial sequence
 <220><223> Primer HY-50
 <400> 415
 gagagttcac tgtttaacaa tagcaatgg 29
 <210> 416
 <211> 26
 <212> DNA
 <213> Artificial sequence
 <220><223> Primer oBP834
 <400> 416

tgctgatact gctgctaaac aagctg

26

<210> 417

<211> 343

<212> PRT

<213> Artificial sequence

<220><223> Anaerostipes KARI variant-K9JB4

<400> 417

Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn

1 5 10 15

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

20 25 30

Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val

35 40 45

Ile Ile Gly Leu Tyr Glu Gly Ala Glu Glu Trp Lys Arg Ala Glu Glu

50 55 60

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

65 70 75 80

Ile Ile Met Ile Leu Ile Asn Asp Glu Lys Gln Ala Thr Met Tyr Lys

85 90 95

Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala

100 105 110

His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val

115 120 125

Asp Val Thr Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser

130 135 140

Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln

145 150 155 160

Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala

165 170 175

Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu

180 185 190

Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val

195 200 205
Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr
210 215 220
Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile
225 230 235 240
Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile
245 250 255
Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile
260 265 270
Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln
275 280 285
Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly

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

340

<210> 418

<211> 1032

<212> DNA

<213> Artificial sequence

<220><223> coding sequence for Anaerostipes KARI variant-K9JB4

<400> 418

atggaagaat gtaagatggc taagatttac taccaagaag actgtaactt gtccttggtg 60
gatggtaaga ctatcgccgt tatcggttac ggttctcaag gtcacgtca tgcctgaat 120
gctaaggaat ccggttgtaa cgttatcatt ggtttatacg aaggtgcgga ggagtggaaa 180
agagctgaag aacaagggtt cgaagtctac accgctgctg aagctgctaa gaaggctgac 240
atcattatga tcttgatcaa cgatgaaaag caggctacca tgtacaaaaa cgacatcgaa 300
ccaaacttgg aagccggtaa catgttgatg ttcgctcacg gtttcaacat ccatttcggt 360

tgtattgttc caccaaagga cgttgatgtc actatgatcg ctccaaaggg tccaggtcac 420

accgttagat ccgaatacga agaaggtaaa ggtgtcccat gcttggttgc tgcgaacaa 480

gacgctactg gcaaggettt ggatatggct ttggcctacg ctttagccat cgggtggtgct 540

agagccggtg tcttggaac taccttcaga accgaaactg aaaccgactt gttcggtgaa 600

caagctgttt tatgtggtgg tgtctgcgct ttgatgcagg ccggttttga aaccttggtt 660

gaagccggtt acgaccecaag aaacgcttac ttcgaatgta tccacgaaat gaagttgatc 720

gttgacttga tctaccaatc tggtttctcc ggtatgcgct actctatctc caacactgct 780

gaatacgggtg actacattac cggtccaaag atcattactg aagataccaa gaaggctatg 840

aagaagattt tgctgacat tcaagatggt acctttgcca aggacttctt ggttgacatg 900

tctgatgctg gttcccaggt ccacttcaag gctatgagaa agttggcctc cgaacaccca 960

gctgaagttg tcggtgaaga aattagatcc ttgtactcct ggtccgacga agacaagttg 1020

attaacaact ga 1032

<210> 419

<211> 343

<212> PRT

<213> Artificial sequence

<220><223> Anaerostipes KARI variant-K9JB4P

<400> 419

Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn

1 5 10 15

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

20 25 30

Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val

35 40 45

Ile Ile Gly Leu Tyr Glu Gly Ala Glu Glu Trp Lys Arg Ala Glu Glu

50 55 60

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

65 70 75 80

Ile Ile Met Ile Leu Ile Pro Asp Glu Lys Gln Ala Thr Met Tyr Lys

85 90 95

Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala

100	105	110	
His Gly Phe Asn Ile	His Phe Gly Cys Ile Val	Pro Pro Lys Asp Val	
115	120	125	
Asp Val Thr Met Ile	Ala Pro Lys Gly Pro Gly His Thr	Val Arg Ser	
130	135	140	
Glu Tyr Glu Glu Gly Lys Gly Val	Pro Cys Leu Val Ala Val	Glu Gln	
145	150	155	160
Asp Ala Thr Gly Lys Ala Leu	Asp Met Ala Leu Ala Tyr Ala Leu	Ala	
165	170	175	
Ile Gly Gly Ala Arg Ala Gly Val	Leu Glu Thr Thr Phe Arg Thr	Glu	
180	185	190	
Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val	Leu Cys Gly Gly Val		
195	200	205	
Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr			
210	215	220	
Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile			
225	230	235	240
Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile			
245	250	255	
Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile			
260	265	270	
Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln			
275	280	285	
Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly			
290	295	300	
Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro			
305	310	315	320
Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp			
325	330	335	
Glu Asp Lys Leu Ile Asn Asn			
340			

<210> 420

<211> 1032

<212> DNA

<213> Artificial sequence

<220><223> coding sequence for Anaerostipes KARI variant K9JB4P

<400> 420

```

atggaagaat gtaagatggc taagatttac taccaagaag actgtaactt gtccttggtg      60
gatggaaga ctatcgccgt tatcggttac ggttctcaag gtcacgtca tgcctgaat      120
gctaaggaat ccggttgtaa cgttatcatt ggtttatagc aaggtgcgga ggagtggaaa      180
agagctgaag aacaaggttt cgaagtctac accgctgctg aagctgctaa gaaggctgac      240

atcattatga tcttgatccc agatgaaaag caggctacca tgtacaaaaa cgacatcgaa      300
ccaaacttgg aagccggtaa catgttgatg ttcgctcacg gtttcaacat ccatttcggt      360
tgtattgttc caccaaagga cgttgatgtc actatgatcg ctccaaaggg tccaggtcac      420
accgttagat ccgaatacga agaaggtaaa ggtgtcccat gcttggttgc tgcgaacaa      480
gacgctactg gcaaggcttt ggatatggct ttggcctacg ctttagccat cgggtggtgct      540
agagccggtg tcttggaac taccttcaga accgaaactg aaaccgactt gttcggtgaa      600
caagctgttt tatgtgggtg tgtctgcgct ttgatgcagg ccggttttga aaccttggtt      660

gaagccggtt acgaccaag aaacgcttac ttcgaatgta tccacgaaat gaagttgatc      720
gttgacttga tctaccaatc tggtttctcc ggtatgcgtt actctatctc caacactgct      780
gaatacgggt actacattac cggtcctaaag atcattactg aagataccta gaaggetatg      840
aagaagattt tgtctgacat tcaagatggt acctttgcca aggacttctt ggttgacatg      900
tctgatgctg gttcccaggt ccacttcaag gctatgagaa agttggcctc cgaacaccca      960
gctgaagttg tcggtgaaga aattagatcc ttgtactcct ggtccgacga agacaagttg     1020
attaacaact ga                                     1032

```

<210> 421

<211> 343

<212> PRT

<213> Artificial sequence

<220><223> Anaerostipes caccae KARI variant-K9JG3

<400> 421

Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn

1 5 10 15

Leu Ser Leu Leu Asp Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser
 20 25 30
 Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val
 35 40 45

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

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

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

 Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile
 245 250 255
 Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile

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

Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro
 305 310 315 320
 Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
 325 330 335
 Glu Asp Lys Leu Ile Asn Asn
 340

<210> 422

<211> 1032

<212> DNA

<213> Artificial sequence

<220><223> coding sequence for Anaerostipes caccae KARI variant K9JG3

<400> 422

atggaagaat gtaagatggc taagatttac taccaagaag actgtaactt gtccttggtg 60

gatggaaga ctatgccgt tatcggttac ggttctcaag gtcacgtca tgcctgaat 120

gctaaggaat ccggttgtaa cgttatcatt ggtttattcg aaggtgcgaa ggagtggaaa 180

agagctgaag aacaagggtt cgaagtctac accgctgctg gagctgctaa gaaggctgac 240

atcattatga tcttgatcaa cgatgaaaag caggctacca tgtacaaaaa cgacatcgaa 300

ccaaacttgg aagccggtaa catgttgatg ttcgctcacg gtttcaacat ccatttcggt 360

tgtattgttc caccaaagga cgttgatgtc actatgatcg ctcaaaggg tccaggtcac 420

accgttagat ccgaatacga agaaggtaaa ggtgtcccat gcttggttgc tgtcgaacaa 480

gacgctactg gcaaggcttt ggatatggct ttggcctacg ctttagccat cgggtggtgct 540

agagccggtg tcttggaac taccttcaga accgaaactg aaaccgactt gttcggtgaa 600

caagctgttt taigtgggtg tgtctgcgct ttgatgcagg ccggttttga aaccttggtt 660

gaagccggtt acgaccaag aaacgcttac ttcgaatgta tccacgaaat gaagttgatc 720

gttgacttga tctaccaatc tggtttctcc ggtatgcgtt actctatctc caacactgct 780

gaatacgggtg actacattac cgggtccaaag atcattactg aagataccaa gaaggctatg 840

aagaagattt tgictgacat tcaagatggt acctttgcca aggacttctt ggttgacatg 900

tctgatgctg gttcccaggt ccacttcaag gctatgagaa agttggcctc cgaacaccca 960

gctgaagttg tcggtgaaga aattagatcc ttgtactcct ggtccgacga agacaagttg 1020

attaacaact ga 1032

<210> 423

<211> 343

<212> PRT

<213> Artificial sequence

<220><223> Anaerostipes caccae KARI variant-K9JG3P

<400> 423

Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn

1 5 10 15

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

20 25 30

Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val

35 40 45

Ile Ile Gly Leu Phe Glu Gly Ala Lys Glu Trp Lys Arg Ala Glu Glu

50 55 60

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

65 70 75 80

Ile Ile Met Ile Leu Ile Pro Asp Glu Lys Gln Ala Thr Met Tyr Lys

85 90 95

Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala

100 105 110

His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val

115 120 125

Asp Val Thr Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser

130 135 140

Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln

145 150 155 160

Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala

165 170 175
 Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu
 180 185 190
 Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val
 195 200 205
 Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr

210 215 220
 Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile
 225 230 235 240
 Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile
 245 250 255
 Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile
 260 265 270
 Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln

275 280 285
 Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly
 290 295 300
 Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro
 305 310 315 320
 Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
 325 330 335
 Glu Asp Lys Leu Ile Asn Asn
 340

<210> 424

<211> 1032

<212> DNA

<213> Artificial sequence

<220><223> coding sequence for Anaerostipes caccae KARI variant-K9JG3P

<400> 424

atggaagaat gtaagatggc taagatttac taccaagaag actgtaactt gtccttggtg 60
 gatggtgaaga ctatcgccgt tatcggttac ggttctcaag gtcacgtca tgccctgaat 120
 gctaaggaat ccggttgtaa cgttatcatt ggtttattcg aaggtgcaa ggagtggaaa 180

agagctgaag aacaaggttt cgaagtctac accgctgctg gagctgctaa gaaggctgac 240
atcattatga tcttgatccc agatgaaaag caggctacca tgtacaaaaa cgacatcgaa 300

ccaaacttgg aagccggtaa catgttgatg ttcgctcacg gtttcaacat ccatttcggt 360
tgtattgttc caccaaagga cgttgatgtc actatgatcg ctccaaaggg tccaggtcac 420
accgttagat ccgaatacga agaaggtaaa ggtgtcccat gcttggttgc tgcgaacaa 480
gacgctactg gcaaggcttt ggatatggct ttggcctacg ctttagccat cggtaggtgct 540
agagccgggt tcttgaaac taccttcaga accgaaactg aaaccgactt gttcggtgaa 600
caagctgttt tatgtgggtg tgtctgcgct ttgatgcagg ccggttttga aaccttggtt 660
gaagccggtt acgaccecaag aaacgcttac ttcgaatgta tccacgaaat gaagttgatc 720

gttgacttga tctaccaatc tggtttctcc ggtatgcgtt actctatctc caacactgct 780
gaatacgggt actacattac cggtcctaaag atcattactg aagataccaa gaaggtatg 840
aagaagattt tgctgacat tcaagatggg acctttgcca aggacttctt ggttgacatg 900
tctgatgctg gttcccaggt ccacttcaag gctatgagaa agttggcctc cgaacaccca 960
gctgaagttg tcggtgaaga aattagatcc ttgtactcct ggtccgacga agacaagttg 1020
attaacaact ga 1032

<210> 425
<211> 343
<212> PRT
<213> Artificial sequence

<220><223> Anaerostipes caccae KARI variant-K9JA1

<400> 425

Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn

1 5 10 15

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

20 25 30

Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val

35 40 45

Ile Ile Gly Leu Phe Glu Gly Ala Glu Glu Trp Lys Arg Ala Glu Glu

50 55 60

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

65 70 75 80

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

	325	330	335
Glu Asp Lys Leu Ile Asn Asn			
	340		
<210>	426		
<211>	1032		
<212>	DNA		
<213>	Artificial sequence		
<220><223>	coding sequence for Anaerostipes caccae KARI variant K9JA1		
<400>	426		
atggaagaat gtaagatggc taagatttac taccaagaag actgtaactt gtccttgttg			60
gatggtaaga ctatcgccgt tateggttac ggttctcaag gtcacgtca tgcctgaat			120
gctaaggaat ccggttgtaa cgttatcatt ggtttattcg aagggtcgga ggagtggaaa			180
agagctgaag aacaaggttt cgaagtctac accgctgtcg gagctgctaa gaaggctgac			240
atcattatga tcttgatccc agatgaaaag caggctacca tgtacaaaaa cgacatcgaa			300
ccaaacttgg aagccggtaa catgttgatg ttcgctcacg gtttcaacat ccatttcggg			360
tgtattgttc caccaaagga cgttgatgtc actatgatcg ctccaaaggg tccaggtcac			420
accgttagat ccgaatacga agaaggtaaa ggtgtcccat gcttggttgc tgtcgaacaa			480
gacgctactg gcaaggcttt ggatatggct ttggcctacg ctttagccat cgggtggtgct			540
agagccggtg tcttggaaac taccttcaga accgaaactg aaaccgactt gttcggtgaa			600
caagctgttt tatgtgggtg tgtctgcgct ttgatgcagg ccggttttga aaccttggtt			660
gaagccggtt acgaccaag aaacgcttac ttcgaatgta tccacgaaat gaagttgatc			720
gttgacttga tctaccaatc tggtttctcc ggtatgcgtt actctatctc caacactgct			780
gaatacgggt actacattac cggtccaaag atcattactg aagataccaa gaaggctatg			840
aagaagattt tgtctgacat tcaagatggg acctttgccaggacttctt ggttgacatg			900
tctgatgctg gttcccaggt ccacttcaag gctatgagaa agttggcctc cgaacacca			960
gctgaagttg tcggtgaaga aattagatcc ttgtactcct ggtccgacga agacaagttg			1020
attaacaact ga			1032
<210>	427		
<211>	343		
<212>	PRT		
<213>	Artificial sequence		
<220><223>	Anaerostipes caccae KARI variant- K9SB2		

<400> 427

Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn

1 5 10 15

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

20 25 30

Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val

35 40 45

Ile Ile Gly Leu Phe Glu Gly Ala Glu Glu Trp Lys Arg Ala Glu Glu

50 55 60

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

65 70 75 80

Ile Ile Met Ile Leu Ile Pro Asp Glu Lys Gln Ala Thr Met Tyr Lys

85 90 95

Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala

100 105 110

His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val

115 120 125

Asp Val Thr Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser

130 135 140

Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln

145 150 155 160

Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala

165 170 175

Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu

180 185 190

Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val

195 200 205

Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr

210 215 220

Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile

225 230 235 240

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

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

<210> 428

<211> 1032

<212> DNA

<213> Artificial sequence

<220><223> Anaerostipes caccae KARI variant-K9SB2

<400> 428

atggaagaat gtaagatggc taagatttac taccaagaag actgtaactt gtccttggtg	60
gatggtaaga ctatgccgt tatcggttac ggttctcaag gtcacgtca tgccctgaat	120
gctaaggaat ccggttgtaa cgttatcatt ggtttattcg aaggtgcgga ggagtggaaa	180
agagctgaag aacaaggttt cgaagtctac accgctgctg aagctgctaa gaaggtgac	240
atcattatga tcttgatccc agatgaaaag caggctacca tgtacaaaaa cgacatcgaa	300
ccaaacttgg aagccggtaa catgttgatg ttcgctcacg gtttcaacat ccatttcggt	360
tgtattgttc caccaaagga cgttgatgtc actatgatcg ctccaaaggg tccaggtcac	420
accgttagat ccgaatacga agaaggtaaa ggtgtcccat gcttggttgc tgtcgaacaa	480
gacgctactg gcaaggcttt ggatatggct ttggcctacg ctttagccat cgggtgtgct	540
agagccggtg tcttggaac taccttcaga accgaaactg aaaccgactt gttcggtgaa	600
caagctgttt tatgtgggtg tgctctgcgt ttgatgcagg ccggttttga aaccttggtt	660
gaagccggtt acgaccaag aaacgcttac ttcgaatgta tccacgaaat gaagttgatc	720

gttgacttga tctaccaatc tggtttctcc ggtatgcggtt actctatctc caacactgct 780

gaatacgggtg actacattac cggtccaaag atcattactg aagataccaa gaaggctatg 840

aagaagattt tgtctgacat tcaagatggt acctttgccca aggacttctt ggttgacatg 900

tctgatgctg gttcccaggt ccacttcaag gctatgagaa agttggcctc cgaacaccca 960

gctgaagttg tcggtgaaga aattagatcc ttgtactcct ggtccgacga agacaagttg 1020

attaacaact ga 1032

<210> 429

<211> 343

<212> PRT

<213> Artificial sequence

<220><223> Anaerostipes caccae KARI variant-K9SB2-K90L

<400

> 429

Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn

1 5 10 15

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

20 25 30

Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val

35 40 45

Ile Ile Gly Leu Phe Glu Gly Ala Glu Glu Trp Lys Arg Ala Glu Glu

50 55 60

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

65 70 75 80

Ile Ile Met Ile Leu Ile Pro Asp Glu Leu Gln Ala Thr Met Tyr Lys

85 90 95

Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala

100 105 110

His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val

115 120 125

Asp Val Thr Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser

130 135 140

Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln

145 150 155 160
 Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala
 165 170 175
 Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu
 180 185 190

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

Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile
 260 265 270
 Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln
 275 280 285
 Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly
 290 295 300
 Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro
 305 310 315 320

Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
 325 330 335
 Glu Asp Lys Leu Ile Asn Asn
 340

<210> 430

<211> 1032

<212> DNA

<213> Artificial sequence

<220><223> coding sequence for Anaerostipes KARI variant K9SB2-K90L

<400> 430

atggaagaat gtaagatggc taagatttac taccaagaag actgtaactt gtccttgttg 60

gatggaaga ctatgccgt tatcggttac ggttctcaag gtcacgtca tgccctgaat 120
gctaaggaat ccggttgtaa cgttatcatt ggtttattcg aaggtgcgga ggagtggaaa 180

agagctgaag aacaaggttt cgaagtctac accgtgctg aagctgctaa gaaggtgac 240
atcattatga tcttgatccc agatgaattg caggctacca tgtacaaaa cgacatcgaa 300
ccaaacttgg aagccggtaa catgttgatg ttcgctcacg gtttcaacat ccatttcggt 360
tgtattgttc caccaaagga cgttgatgtc actatgatcg ctccaaaggg tccaggtcac 420
accgttagat ccgaatacga agaaggtaaa ggtgtcccat gcttgggttc tgtcgaacaa 480
gacgctactg gcaaggcttt ggatatggct ttggcctacg ctttagccat cgggtggtgct 540
agagccggtg tcttggaac taccttcaga accgaaactg aaaccgactt gttcggtgaa 600

caagctgttt tatgtgggtg tgtctgcgt ttgatgcagg ccggttttga aaccttggtt 660
gaagccggtt acgaccaag aaacgcttac ttggaatgta tccacgaaat gaagttgatc 720
gttgacttga tctaccaatc tggtttctcc ggtatgcgtt actctatctc caacactgct 780
gaatacgggtg actacattac cggtccaaag atcattactg aagatacaca gaaggctatg 840
aagaagattt tgctgacat tcaagatggt accttgcca aggacttctt ggttgacatg 900
tctgatgctg gtccccaggt ccacttcaag gctatgagaa agttggcctc cgaacaccca 960
gctgaagttg tcggtgaaga aattagatcc ttgtactcct ggtccgacga agacaagttg 1020

attaacaact ga 1032

<210> 431
<211> 343
<212> PRT
<213> Artificial sequence
<220><223> Anaerostipes caccae KARI variant-K9SB2-K90M
<400> 431

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

Gln Gly Phe Glu Val Tyr Thr Ala Ala Glu Ala Ala Lys Lys Ala Asp
65 70 75 80
Ile Ile Met Ile Leu Ile Pro Asp Glu Met Gln Ala Thr Met Tyr Lys
85 90 95
Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala
100 105 110
His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val
115 120 125
Asp Val Thr Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser
130 135 140
Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln
145 150 155 160
Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala
165 170 175
Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu
180 185 190
Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val
195 200 205
Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr
210 215 220
Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile
225 230 235 240
Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile
245 250 255
Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile
260 265 270
Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln
275 280 285
Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly
290 295 300
Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro

305	310	315	320
Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp			
	325	330	335
Glu Asp Lys Leu Ile Asn Asn			
	340		
<210>	432		
<211>	1032		
<212>	DNA		
<213>	Artificial sequence		
<220><223>	coding sequence for Anaerostipes caccae KARI variant-K9SB2-K90M		
<400>	432		
atggaagaat gtaagatggc taagatttac taccaagaag actgtaactt gtccttgttg			60
gatggtaaga ctatcgccgt tatcggttac ggtttctcaag gtcacgctca tgccctgaat			120
gctaaggaat ccggttgtaa cgttatcatt ggtttatttcg aagggtcgga ggagtggaaa			180
agagctgaag aacaaggttt cgaagtctac accgctgctg aagctgctaa gaaggctgac			240
atcattatga tcttgatccc agatgaaatg caggctacca tgtacaaaaa cgacatcgaa			300
ccaaacttgg aagccggtaa catgttgatg ttgcgtcacg gtttcaacat ccatttcggt			360
tgtattgttc caccaaagga cgttgatgtc actatgatcg ctccaaaggg tccaggtcac			420
accgttagat ccgaatacga agaaggtaaa ggtgtcccat gcttggttgc tgtcgaacaa			480
gacgctactg gcaaggtcttt ggatatggct ttggcctacg ctttagccat cgggtggtgt			540
agagccgggtg tcttggaaac taccttcaga accgaaactg aaaccgactt gttcggtgaa			600
caagctgttt tatgtggtgg tgtctgcgct ttgatgcagg ccggttttga aaccttggtt			660
gaagccggtt acgacccaag aaacgcttac ttcgaatgta tccacgaaat gaagttgatc			720
gttgacttga tctaccaatc tggtttctcc ggtatgcggt actctatctc caacactgct			780
gaatacggtg actacattac cgttccaaag atcattactg aagataccaa gaaggctatg			840
aagaagattt tgtctgacat tcaagatggf acctttgcca aggacttctt ggttgacatg			900
tctgatgctg gttcccaggt ccacttcaag gctatgagaa agttggcctc cgaacacca			960
gctgaagttg tcggtgaaga aattagatcc ttgtactcct ggtccgacga agacaagttg			1020
attaacaact ga			1032
<210>	433		
<211>	343		
<212>	PRT		

<213> Artificial sequence

<220><223> Anaerostipes caccae KARI variant- 2011_A2

<400> 433

Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn

1 5 10 15

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

20 25 30

Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val

35 40 45

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

50 55 60

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

65 70 75 80

Ile Ile Met Ile Leu Ile Pro Asp Glu Lys Gln Ala Thr Met Tyr Lys

85 90 95

Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala

100 105 110

His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val

115 120 125

Asp Val Thr Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser

130 135 140

Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln

145 150 155 160

Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala

165 170 175

Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu

180 185 190

Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val

195 200 205

Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr

210 215 220

Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile
 225 230 235 240

Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile
 245 250 255

Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile
 260 265 270

Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln
 275 280 285

Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly
 290 295 300

Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro
 305 310 315 320

Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
 325 330 335

Glu Asp Lys Leu Ile Asn Asn
 340

<210> 434

<211> 343

<212> PRT

<213> Artificial sequence

<220><223> Anaerostipes caccae KARI variant-2011_A3

<400> 434

Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn
 1 5 10 15

Leu Ser Leu Leu Asp Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser
 20 25 30

Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val
 35 40 45

Ile Ile Gly Leu Phe Glu Gly Ala Glu Glu Trp Lys Arg Ala Glu Glu
 50 55 60

Gln Gly Phe Glu Val Tyr Thr Ala Ala Glu Ala Ala Lys Lys Ala Asp
 65 70 75 80

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

325 330 335

Glu Asp Lys Leu Ile Asn Asn

340

<210> 435

<211> 343

<212> PRT

<213> Artificial sequence

<220><223> Anaerostipes caccae KARI variant-2011_A5

<400> 435

Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn

1 5 10 15

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

20 25 30

Gln Gly Asn Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val

35 40 45

Ile Ile Gly Leu Phe Glu Gly Ala Glu Glu Trp Lys Arg Ala Glu Glu

50 55 60

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

65 70 75 80

Ile Ile Met Ile Leu Ile Pro Asp Glu Lys Gln Ala Thr Met Tyr Lys

85 90 95

Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala

100 105 110

His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val

115 120 125

Asp Val Thr Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser

130 135 140

Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln

145 150 155 160

Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala

165 170 175

Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu

180 185 190
 Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val
 195 200 205
 Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr
 210 215 220
 Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile
 225 230 235 240
 Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile
 245 250 255

 Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile
 260 265 270
 Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln
 275 280 285
 Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly
 290 295 300
 Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro
 305 310 315 320

 Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
 325 330 335
 Glu Asp Lys Leu Ile Asn Asn
 340
 <210> 436
 <211> 343
 <212> PRT
 <213> Artificial sequence
 <220><223> Anaerostipes caccae KARI variant 2011_A7
 <400> 436
 Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn
 1 5 10 15
 Leu Ser Leu Leu Asp Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser
 20 25 30

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

275 280 285
 Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly
 290 295 300
 Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro
 305 310 315 320
 Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
 325 330 335
 Glu Asp Lys Leu Ile Asn Asn
 340

<210> 437

<211> 343

<212> PRT

<213> Artificial sequence

<220><223> Anaerostipes caccae KARI variant-2011_A9

<400> 437

Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn
 1 5 10 15
 Leu Ser Leu Leu Asp Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser
 20 25 30
 Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val
 35 40 45

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

His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val
 115 120 125
 Asp Val Thr Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser

130 135 140
 Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln
 145 150 155 160
 Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala
 165 170 175

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

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

 Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro
 305 310 315 320
 Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
 325 330 335
 Glu Asp Lys Leu Ile Asn Asn
 340

 <210> 438
 <211> 343
 <212> PRT
 <213> Artificial sequence
 <220><223> Anaerostipes caccae KARI variant-2011_A11

<400> 438

Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn

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

65 70 75 80
Ile Ile Met Ile Leu Ile Pro Asp Glu Lys Gln Ala Thr Thr Tyr Lys
85 90 95
Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala
100 105 110
His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Glu Asp Val
115 120 125
Asp Val Thr Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser

130 135 140
Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln
145 150 155 160
Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala
165 170 175
Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu
180 185 190
Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val

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

Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile
245 250 255

Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile
260 265 270

Ala Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln
275 280 285

Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly
290 295 300

Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro
305 310 315 320

Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
325 330 335

Glu Asp Lys Leu Ile Asn Asn
340

<210> 439
<211> 343
<212> PRT
<213> Artificial sequence
<220><223> Anaerostipes caccae KARI variant-2011_B1
<400> 439

Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn
1 5 10 15

Leu Ser Leu Leu Asp Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser
20 25 30

Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val
35 40 45

Ile Ile Gly Leu Phe Glu Gly Ala Glu Glu Trp Lys Arg Ala Glu Glu
50 55 60

Gln Gly Phe Glu Val Tyr Thr Ala Ala Glu Ala Ala Lys Lys Ala Asp
65 70 75 80

Ile Ile Met Ile Leu Ile Pro Asp Glu Lys Gln Ala Thr Ile Tyr Lys

85

90

95

Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala

100

105

110

His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val

115

120

125

Asp Val Thr Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser

130

135

140

Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln

145

150

155

160

Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala

165

170

175

Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu

180

185

190

Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val

195

200

205

Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr

210

215

220

Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile

225

230

235

240

Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile

245

250

255

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

260

265

270

Thr Glu Asp Thr Lys Lys Thr Met Lys Lys Ile Leu Ser Asp Ile Gln

275

280

285

Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly

290

295

300

Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro

305

310

315

320

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

325

330

335

Glu Asp Lys Leu Ile Asn Asn

340

<210> 440

<211> 343

<212> PRT

<213> Artificial sequence

<220><223> Anaerostipes caccae KARI variant-2011_B2

<400> 440

Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn

1 5 10 15

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

20 25 30

Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val

35 40 45

Ile Ile Gly Leu Phe Glu Gly Ala Glu Glu Trp Lys Arg Ala Glu Glu

50 55 60

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

65 70 75 80

Ile Ile Met Ile Leu Ile Pro Asp Glu Lys Leu Ala Thr Met Tyr Lys

85 90 95

Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala

100 105 110

His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val

115 120 125

Asp Val Thr Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser

130 135 140

Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln

145 150 155 160

Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala

165 170 175

Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu

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

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

 305 310 315 320
 Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
 325 330 335
 Glu Asp Lys Leu Ile Asn Asn
 340
 <210> 441
 <211> 343
 <212> PRT
 <213> Artificial sequence
 <220><223> Anaerostipes caccae KARI variant-2011_B3
 <400> 441
 Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn
 1 5 10 15

 Leu Ser Leu Leu Asp Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser
 20 25 30
 Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val

35	40	45	
Ile Ile Gly Leu Leu Glu Gly Ala Glu Glu Trp Lys Arg Ala Glu Glu			
50	55	60	
Gln Gly Phe Glu Val Tyr Thr Ala Ala Glu Ala Ala Lys Lys Ala Asp			
65	70	75	80
Ile Ile Met Ile Leu Ile Pro Asp Glu Lys Gln Ala Thr Met Tyr Lys			
85	90	95	
Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala			
100	105	110	
His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val			
115	120	125	
Asp Val Thr Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser			
130	135	140	
Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln			
145	150	155	160
Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala			
165	170	175	
Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu			
180	185	190	
Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val			
195	200	205	
Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr			
210	215	220	
Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile			
225	230	235	240
Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile			
245	250	255	
Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile			
260	265	270	
Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln			
275	280	285	

Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly
 290 295 300

Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro
 305 310 315 320

Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
 325 330 335

Glu Asp Lys Leu Ile Asn Asn
 340

<210> 442

<211> 343

<212> PRT

<213> Artificial sequence

<220><223> Anaerostipes caccae KARI variant-2011_B4

<400> 442

Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn
 1 5 10 15

Leu Ser Leu Leu Asp Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser
 20 25 30

Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val
 35 40 45

Ile Ile Gly Leu Phe Glu Gly Ala Glu Glu Trp Lys Arg Ala Glu Glu
 50 55 60

Gln Gly Phe Glu Val Tyr Thr Ala Ala Glu Ala Ala Lys Lys Ala Asp
 65 70 75 80

Ile Ile Met Ile Leu Ile Pro Asp Glu Lys Gln Ala Thr Met Tyr Lys
 85 90 95

Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala
 100 105 110

His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val
 115 120 125

Asp Val Thr Met Ile Ala Thr Lys Gly Pro Gly His Thr Val Arg Ser
 130 135 140

Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln
 145 150 155 160
 Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala

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

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

 290 295 300
 Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro
 305 310 315 320
 Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
 325 330 335
 Glu Asp Lys Leu Ile Asn Asn
 340
 <210> 443
 <211> 343
 <212> PRT
 <213> Artificial sequence
 <220><223> Anaerostipes caccae KARI variant-2011_B7

<400> 443

Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn

1 5 10 15

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

20 25 30

Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val

35 40 45

Ile Ile Gly Leu Phe Glu Gly Ala Glu Glu Trp Lys Arg Ala Glu Glu

50 55 60

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

65 70 75 80

Ile Ile Met Ile Leu Ile Pro Asp Glu Lys Gln Ala Thr Met Tyr Lys

85 90 95

Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala

100 105 110

His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val

115 120 125

Asp Val Thr Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser

130 135 140

Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln

145 150 155 160

Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala

165 170 175

Ile Gly Gly Ala Arg Ala Gly Val Met Glu Thr Thr Phe Arg Thr Glu

180 185 190

Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val

195 200 205

Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr

210 215 220

Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile

225 230 235 240

Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile

245 250 255

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

260 265 270

Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln

275 280 285

Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly

290 295 300

Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro

305 310 315 320

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

325 330 335

Glu Asp Lys Leu Ile Asn Asn

340

<210> 444

<211> 343

<212> PRT

<213> Artificial sequence

<220><223> Anaerostipes caccae KARI variant-2011_B8

<400> 444

Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn

1 5 10 15

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

20 25 30

Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val

35 40 45

Ile Ile Gly Leu Phe Glu Gly Ala Glu Glu Trp Lys Arg Ala Glu Glu

50 55 60

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

65 70 75 80

Ile Ile Met Ile Leu Ile Pro Asp Glu Lys Gln Ala Thr Met Tyr Lys

85

90

95

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

 145 150 155 160
 Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala
 165 170 175
 Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu
 180 185 190
 Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val
 195 200 205
 Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr

 210 215 220
 Asp Pro Arg Asn Ala Tyr Phe Glu Ser Ile His Glu Met Lys Leu Ile
 225 230 235 240
 Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile
 245 250 255
 Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile
 260 265 270
 Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln

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

340

<210> 445

<211> 343

<212> PRT

<213> Artificial sequence

<220><223> Anaerostipes caccae KARI variant-2011_B10

<400> 445

Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn

1 5 10 15

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

20 25 30

Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val

35 40 45

Ile Ile Gly Leu Phe Glu Gly Ala Glu Glu Trp Lys Arg Ala Glu Glu

50 55 60

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

65 70 75 80

Ile Ile Met Ile Leu Ile Pro Asp Glu Lys Gln Ala Thr Met Tyr Lys

85 90 95

Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala

100 105 110

His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val

115 120 125

Asp Val Thr Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser

130 135 140

Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln

145 150 155 160

Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala

165 170 175

Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu

180 185 190

Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val
195 200 205

Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr
210 215 220

Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile
225 230 235 240

Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile
245 250 255

Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile
260 265 270

Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln
275 280 285

Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Asp Gly
290 295 300

Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro
305 310 315 320

Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
325 330 335

Glu Asp Lys Leu Ile Asn Asn
340

<210> 446

<211> 343

<212> PRT

<213> Artificial sequence

<220><223> Anaerostipes caccae KARI variant-2011_C1

<400> 446

Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn

1 5 10 15

Leu Ser Leu Leu Asp Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser
20 25 30

Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val
35 40 45

Ile Ile Gly Leu Phe Glu Gly Ala Glu Glu Trp Lys Arg Ala Lys Glu
 50 55 60
 Gln Gly Phe Glu Val Tyr Thr Ala Ala Glu Ala Ala Lys Lys Ala Asp

 65 70 75 80
 Ile Ile Met Ile Leu Ile Pro Asp Glu Lys Gln Ala Thr Met Tyr Lys
 85 90 95
 Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala
 100 105 110
 His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val
 115 120 125
 Asp Val Thr Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser

 130 135 140
 Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln
 145 150 155 160
 Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala
 165 170 175
 Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu
 180 185 190
 Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val

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

 260 265 270
 Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln
 275 280 285
 Asp Gly Thr Phe Ala Arg Asp Phe Leu Val Asp Met Ser Asp Ala Gly

290 295 300
 Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro
 305 310 315 320
 Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp

 325 330 335
 Glu Asp Lys Leu Ile Asn Asn
 340
 <210> 447
 <211> 343
 <212> PRT
 <213> Artificial sequence
 <220><223> Anaerostipes caccae KARI variant 2011_C9
 <400> 447
 Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn
 1 5 10 15
 Leu Ser Leu Leu Asp Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser
 20 25 30

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

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

145 150 155 160
 Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala
 165 170 175
 Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu
 180 185 190
 Thr Asp Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val
 195 200 205
 Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr
 210 215 220

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

 Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly
 290 295 300
 Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro
 305 310 315 320
 Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
 325 330 335
 Glu Asp Lys Leu Ile Asn Asn
 340
 <210> 448
 <211> 343
 <212> PRT
 <213> Artificial sequence

 <220><223> Anaerostipes caccae KARI variant-2011_C3
 <400> 448

Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn
1 5 10 15
Leu Ser Leu Leu Asp Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser
 20 25 30
Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val
 35 40 45
Ile Ile Gly Leu Phe Glu Gly Ala Glu Glu Trp Lys Arg Ala Glu Glu

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

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

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

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

 305 310 315 320
 Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
 325 330 335
 Glu Asp Lys Leu Ile Asn Asn
 340
 <210> 449
 <211> 343
 <212> PRT
 <213> Artificial sequence
 <220><223> Anaerostipes caccae KARI variant-2011_C6
 <400> 449
 Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn
 1 5 10 15

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

 Ile Ile Met Phe Leu Ile Pro Asp Glu Lys Gln Ala Thr Met Tyr Lys
 85 90 95
 Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala

100	105	110	
His Gly Phe Asn Ile	His Phe Gly Cys Ile Val	Pro Pro Lys Asp Val	
115	120	125	
Asp Val Thr Met Ile	Ala Pro Lys Gly Pro Gly His Thr	Val Arg Ser	
130	135	140	
Glu Tyr Glu Glu Gly Lys Gly Val	Pro Cys Leu Val Ala Val	Glu Gln	
145	150	155	160
Asp Ala Thr Gly Lys Ala Leu Asp Met	Ala Leu Ala Tyr Ala Leu Ala		
165	170	175	
Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr	Phe Arg Thr Glu		
180	185	190	
Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val			
195	200	205	
Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr			
210	215	220	
Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile			
225	230	235	240
Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile			
245	250	255	
Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile			
260	265	270	
Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln			
275	280	285	
Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly			
290	295	300	
Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro			
305	310	315	320
Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp			
325	330	335	
Glu Asp Lys Leu Ile Asn Asn			
340			

<210> 450

<211> 343

<212> PRT

<213> Artificial sequence

<220><223> Anaerostipes caccae KARI variant-2011_C7

<400> 450

Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Val Asp Cys Asn

1 5 10 15

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

20 25 30

Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val

35 40 45

Ile Ile Gly Leu Phe Glu Gly Ala Glu Glu Trp Lys Arg Ala Glu Glu

50 55 60

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

65 70 75 80

Ile Ile Met Ile Leu Ile Pro Asp Glu Lys Gln Ala Thr Ile Tyr Lys

85 90 95

Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala

100 105 110

His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val

115 120 125

Asp Val Thr Met Ile Ala Pro Lys Gly Pro Gly His Ile Val Arg Ser

130 135 140

Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln

145 150 155 160

Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala

165 170 175

Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu

180 185 190

Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val

195 200 205

Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr
 210 215 220
 Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile

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

 290 295 300
 Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro
 305 310 315 320
 Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
 325 330 335
 Glu Asp Lys Leu Ile Asn Asn
 340
 <210> 451
 <211> 343
 <212> PRT
 <213> Artificial sequence
 <220><223> Anaerostipes caccae KARI variant-2011_C8

 <400> 451
 Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn
 1 5 10 15
 Leu Ser Leu Leu Asp Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser
 20 25 30
 Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val
 35 40 45
 Ile Ile Gly Leu Phe Glu Gly Ala Glu Glu Trp Lys Arg Ala Glu Glu

50	55	60
Gln Gly Phe Glu Val Tyr Thr Val Ala Glu Ala Ala Lys Lys Ala Asp		
65	70	75
Ile Ile Met Ile Leu Ile Pro Asp Glu Lys Gln Ala Thr Met Tyr Lys		80
	85	90
Asn Asp Ile Glu Pro Tyr Leu Glu Ala Gly Asn Met Leu Met Phe Ala		95
	100	105
His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val		110
	115	120
		125
Asp Val Thr Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser		
130	135	140
Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln		
145	150	155
Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala		160
	165	170
Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Ile Arg Thr Glu		175
	180	185
		190
Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val		
195	200	205
Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr		
210	215	220
Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile		
225	230	235
Val Asp Leu Ile His Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile		240
	245	250
		255
Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile		
260	265	270
Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln		
275	280	285
Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly		
290	295	300

Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro
305 310 315 320

Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
325 330 335

Glu Asp Lys Leu Ile Asn Asn
340

<210> 452

<211> 343

<212> PRT

<213> Artificial sequence

<220><223> Anaerostipes caccae KARI variant-2011_C10

<400> 452

Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn
1 5 10 15

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

20 25 30

Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val

35 40 45

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

50 55 60

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

65 70 75 80

Ile Ile Met Ile Leu Ile Pro Asp Glu Lys Gln Ala Thr Met Tyr Lys

85 90 95

Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala

100 105 110

His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val

115 120 125

Asp Val Thr Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser

130 135 140

Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln

145 150 155 160
 Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala
 165 170 175
 Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu
 180 185 190
 Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val
 195 200 205
 Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr

 210 215 220
 Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile
 225 230 235 240
 Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile
 245 250 255
 Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile
 260 265 270
 Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln

 275 280 285
 Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly
 290 295 300
 Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro
 305 310 315 320
 Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
 325 330 335
 Glu Asp Lys Leu Ile Asn Asn
 340

<210> 453

<211> 343

<212> PRT

<213> Artificial sequence

<220><223> Anaerostipes caccae KARI variant-2011_C12

<400> 453

Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn

1 5 10 15
 Leu Ser Leu Leu Asp Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser
 20 25 30
 Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val
 35 40 45

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

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

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

 Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile
 245 250 255

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

Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro
305 310 315 320
Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
325 330 335
Glu Asp Lys Leu Ile Asn Asn
340

<210> 454

<211> 343

<212> PRT

<213> Artificial sequence

<220><223> Anaerostipes caccae KARI variant-2011_D1

<400> 454

Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn

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

65 70 75 80
Ile Ile Met Ile Leu Ile Pro Asp Glu Lys Gln Ala Thr Met Tyr Lys
85 90 95
Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala
100 105 110

His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val
115 120 125
Asp Val Thr Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser
130 135 140
Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln
145 150 155 160
Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala
165 170 175
Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu
180 185 190
Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val
195 200 205
Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr
210 215 220
Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile
225 230 235 240
Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile
245 250 255
Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile
260 265 270
Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln
275 280 285
Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly
290 295 300
Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro
305 310 315 320
Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
325 330 335
Glu Asp Lys Leu Ile Asn Asn
340
<210> 455

<211> 343

<212> PRT

<213> Artificial sequence

<220><223> Anaerostipes caccae KARI variant-2011_D2

<400> 455

Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn

1 5 10 15

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

20 25 30

Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val

35 40 45

Ile Ile Gly Leu Phe Glu Gly Ala Glu Glu Trp Lys Arg Ala Glu Glu

50 55 60

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

65 70 75 80

Ile Ile Met Ile Leu Ile Pro Asp Glu Lys Gln Ala Thr Ile Tyr Lys

85 90 95

Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala

100 105 110

His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val

115 120 125

Asp Val Thr Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser

130 135 140

Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln

145 150 155 160

Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala

165 170 175

Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu

180 185 190

Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val

195 200 205

Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr

210 215 220

Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile

225 230 235 240

Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile

245 250 255

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

260 265 270

Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln

275 280 285

Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly

290 295 300

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

305 310 315 320

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

325 330 335

Glu Asp Lys Leu Ile Asn Asn

340

<210> 456

<211> 343

<212> PRT

<213> Artificial sequence

<220><223> Anaerostipes caccae KARI variant-2011_D3

<400> 456

Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn

1 5 10 15

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

20 25 30

Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val

35 40 45

Ile Ile Gly Leu Phe Glu Gly Ala Glu Glu Trp Lys Arg Ala Glu Glu

50 55 60

Gln	Gly	Ile	Glu	Val	Tyr	Thr	Ala	Ala	Glu	Ala	Ala	Lys	Lys	Ala	Asp				
65					70					75					80				
Ile	Ile	Met	Ile	Leu	Ile	Pro	Asp	Glu	Lys	Gln	Ala	Thr	Met	Tyr	Lys				
85					90					95									
Asn	Asp	Ile	Glu	Pro	Asn	Leu	Glu	Ala	Gly	Asn	Met	Leu	Met	Phe	Ala				
100					105					110									
His	Gly	Phe	Asn	Ile	His	Phe	Gly	Cys	Ile	Val	Pro	Pro	Lys	Asp	Val				
115					120					125									
Asp	Val	Thr	Met	Ile	Ala	Pro	Lys	Gly	Pro	Gly	His	Thr	Val	Arg	Ser				
130					135					140									
Glu	Tyr	Glu	Glu	Gly	Lys	Gly	Val	Pro	Cys	Leu	Val	Ala	Val	Glu	Gln				
145					150					155					160				
Asp	Ala	Thr	Gly	Lys	Ala	Leu	Asp	Met	Ala	Leu	Ala	Tyr	Ala	Leu	Ala				
165					170					175									
Ile	Gly	Gly	Ala	Arg	Ala	Gly	Val	Leu	Glu	Thr	Thr	Phe	Arg	Thr	Glu				
180					185					190									
Thr	Glu	Thr	Asp	Leu	Phe	Gly	Glu	Gln	Ala	Val	Leu	Cys	Gly	Gly	Val				
195					200					205									
Cys	Ala	Leu	Met	Gln	Ala	Gly	Phe	Glu	Thr	Leu	Val	Glu	Ala	Gly	Tyr				
210					215					220									
Asp	Pro	Arg	Asn	Ala	Tyr	Phe	Glu	Cys	Ile	His	Glu	Met	Lys	Leu	Ile				
225					230					235					240				
Val	Asp	Leu	Ile	Tyr	Gln	Ser	Gly	Phe	Ser	Gly	Met	Arg	Tyr	Ser	Ile				
245					250					255									
Ser	Asn	Thr	Ala	Glu	Tyr	Gly	Asp	Tyr	Ile	Thr	Gly	Pro	Lys	Ile	Ile				
260					265					270									
Thr	Glu	Asp	Thr	Lys	Lys	Ala	Met	Lys	Lys	Ile	Leu	Ser	Asp	Ile	Gln				
275					280					285									
Asp	Gly	Thr	Phe	Ala	Lys	Asp	Phe	Leu	Val	Asp	Met	Ser	Asp	Ala	Gly				
290					295					300									
Ser	Gln	Val	His	Phe	Lys	Ala	Met	Arg	Lys	Leu	Ala	Ser	Glu	His	Pro				

305 310 315 320
 Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
 325 330 335
 Glu Asp Lys Leu Ile Asn Asn
 340
 <210> 457
 <211> 343
 <212> PRT
 <213> Artificial sequence
 <220><223> Anaerostipes caccae KARI variant-2011_D4
 <400> 457
 Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn
 1 5 10 15

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

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

 Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln
 145 150 155 160
 Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala

165 170 175
 Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu
 180 185 190
 Thr Glu Thr Asp Leu Phe Gly Glu Gln Val Val Leu Cys Gly Gly Val
 195 200 205

 Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr
 210 215 220
 Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile
 225 230 235 240
 Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile
 245 250 255
 Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile
 260 265 270

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

 Glu Asp Lys Leu Ile Asn Asn
 340

 <210> 458
 <211> 343
 <212> PRT
 <213> Artificial sequence
 <220><223> Anaerostipes caccae KARI variant-2011_D5
 <400> 458

 Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn
 1 5 10 15
 Leu Ser Leu Leu Asp Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser

20 25 30
 Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val

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

 100 105 110
 His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val
 115 120 125
 Asp Val Thr Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser
 130 135 140
 Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln
 145 150 155 160
 Asp Ala Thr Ser Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala

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

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

Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln
275 280 285

Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly
290 295 300

Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro
305 310 315 320

Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
325 330 335

Glu Asp Lys Leu Ile Asn Asn
340

<210> 459
<211> 343
<212> PRT
<213> Artificial sequence
<220><223> Anaerostipes caccae KARI variant-2011_D8

<400> 459

Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn
1 5 10 15

Leu Ser Leu Leu Asp Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser
20 25 30

Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val
35 40 45

Ile Ile Gly Leu Phe Glu Gly Ala Glu Glu Trp Lys Arg Ala Glu Glu
50 55 60

Gln Gly Phe Glu Val Tyr Thr Ala Ala Glu Ala Ala Lys Lys Ala Asp
65 70 75 80

Ile Ile Met Ile Leu Ile Pro Asp Glu Lys Gln Ala Thr Met Tyr Lys
85 90 95

Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala
100 105 110

His Gly Phe Thr Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val

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

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

 Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile
 260 265 270
 Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln
 275 280 285
 Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly
 290 295 300
 Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro
 305 310 315 320

 Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
 325 330 335
 Glu Asp Lys Leu Ile Asn Asn
 340
 <210> 460
 <211> 343

<212> PRT

<213> Artificial sequence

<220><223> Anaerostipes caccae KARI variant-2011_D9

<400> 460

Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn

1 5 10 15

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

20 25 30

Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val

35 40 45

Ile Ile Gly Leu Phe Glu Gly Ala Glu Glu Trp Lys Arg Ala Glu Glu

50 55 60

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

65 70 75 80

Ile Ile Met Ile Leu Ile Pro Asp Glu Met Gln Ala Thr Met Tyr Lys

85 90 95

Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala

100 105 110

His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val

115 120 125

Asp Val Thr Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser

130 135 140

Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln

145 150 155 160

Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala

165 170 175

Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu

180 185 190

Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val

195 200 205

Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr

210 215 220
 Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile
 225 230 235 240
 Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile
 245 250 255
 Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile
 260 265 270
 Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln

275 280 285
 Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly
 290 295 300
 Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro
 305 310 315 320
 Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
 325 330 335
 Glu Asp Lys Leu Ile Asn Asn
 340

<210> 461

<211> 343

<212> PRT

<213> Artificial sequence

<220><223> Anaerostipes caccae KARI variant-2011_D11

<400> 461

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

65	70	75	80
Ile Ile Met Ile Leu Ile Pro Asp Glu Lys Gln Ala Thr Leu Tyr Lys			
	85	90	95
Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala			
	100	105	110
His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val			
	115	120	125
Asp Val Thr Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser			
	130	135	140
Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln			
	145	150	155
Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala			
	165	170	175
Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu			
	180	185	190
Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val			
	195	200	205
Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr			
	210	215	220
Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile			
	225	230	235
			240
Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile			
	245	250	255
Ser Asn Ile Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile			
	260	265	270
Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln			
	275	280	285
Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly			
	290	295	300
Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro			
	305	310	315
			320

Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
325 330 335

Glu Asp Lys Leu Ile Asn Asn
340

<210> 462

<211> 343

<212> PRT

<213> Artificial sequence

<220><223> Anaerostipes caccae KARI variant-2011_D12

<400> 462

Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn

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

65 70 75 80
Ile Ile Lys Ile Leu Ile Pro Asp Glu Lys Gln Ala Thr Met Tyr Lys
85 90 95
Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala
100 105 110
His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val
115 120 125
Asp Val Thr Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser

130 135 140
Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln
145 150 155 160
Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala
165 170 175

Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu
180 185 190
Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val
195 200 205
Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr
210 215 220
Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile
225 230 235 240
Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile
245 250 255
Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile
260 265 270
Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln
275 280 285
Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly
290 295 300
Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro
305 310 315 320
Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
325 330 335
Glu Asp Lys Leu Ile Asn Asn
340
<210> 463
<211> 343
<212> PRT
<213> Artificial sequence
<220><223> Anaerostipes caccae KARI variant-2011_E3
<400> 463
Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn
1 5 10 15
Leu Ser Leu Leu Asp Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser

				20				25				30			
Gln	Gly	His	Ala	His	Ala	Leu	Asn	Ala	Lys	Glu	Ser	Gly	Cys	Asn	Val
35				40				45							
Ile	Ile	Gly	Leu	Phe	Glu	Gly	Ala	Glu	Glu	Trp	Lys	Arg	Ala	Glu	Glu
50				55				60							
Gln	Gly	Phe	Glu	Val	Tyr	Thr	Ala	Ala	Glu	Ala	Ala	Lys	Lys	Ala	Asp
65				70				75				80			
Ile	Ile	Met	Ile	Leu	Ile	Pro	Asp	Glu	Lys	Gln	Ala	Thr	Met	Tyr	Lys
85				90				95							
Asn	Asp	Ile	Glu	Pro	Asn	Leu	Glu	Ala	Gly	Asn	Met	Leu	Met	Phe	Ala
100				105				110							
His	Gly	Phe	Asn	Ile	His	Phe	Gly	Cys	Val	Val	Pro	Pro	Lys	Asp	Val
115				120				125							
Asp	Val	Thr	Met	Ile	Ala	Pro	Lys	Gly	Pro	Gly	His	Thr	Val	Arg	Ser
130				135				140							
Glu	Tyr	Glu	Glu	Gly	Lys	Gly	Val	Pro	Cys	Leu	Val	Ala	Val	Glu	Gln
145				150				155				160			
Asp	Ala	Thr	Gly	Lys	Ala	Leu	Asp	Met	Ala	Leu	Ala	Tyr	Ala	Leu	Ala
165				170				175							
Ile	Gly	Gly	Ala	Arg	Ala	Gly	Val	Leu	Glu	Thr	Thr	Phe	Arg	Thr	Glu
180				185				190							
Thr	Glu	Thr	Asp	Leu	Phe	Gly	Glu	Gln	Ala	Val	Leu	Cys	Gly	Gly	Val
195				200				205							
Cys	Ala	Leu	Met	Gln	Ala	Gly	Phe	Glu	Thr	Leu	Val	Glu	Ala	Gly	Tyr
210				215				220							
Asp	Pro	Arg	Asn	Ala	Tyr	Phe	Glu	Cys	Ile	His	Glu	Met	Lys	Leu	Ile
225				230				235				240			
Val	Asp	Leu	Ile	Tyr	Gln	Ser	Gly	Phe	Ser	Gly	Met	Arg	Tyr	Ser	Ile
245				250				255							
Ser	Asn	Thr	Ala	Glu	Tyr	Gly	Asp	Tyr	Ile	Thr	Gly	Pro	Lys	Ile	Ile
260				265				270							

Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln
275 280 285

Asp Gly Thr Phe Ala Lys Asp Phe Trp Val Asp Met Ser Asp Ala Gly
290 295 300

Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro
305 310 315 320

Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
325 330 335

Glu Asp Lys Leu Ile Asn Asn
340

<210> 464

<211> 343

<212> PRT

<213> Artificial sequence

<220><223> Anaerostipes caccae KARI variant-2011_E4

<400> 464

Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn
1 5 10 15

Leu Ser Leu Leu Asp Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser
20 25 30

Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val
35 40 45

Ile Ile Gly Leu Phe Glu Gly Ala Glu Glu Trp Lys Arg Ala Glu Glu

50 55 60

Gln Gly Phe Glu Val Tyr Thr Ala Ala Glu Ala Ala Lys Lys Ala Asp
65 70 75 80

Ile Ile Met Ile Leu Ile Pro Asp Glu Lys Gln Ala Thr Met Tyr Lys
85 90 95

Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ser
100 105 110

His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val

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

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

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

 305 310 315 320
 Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
 325 330 335
 Glu Asp Lys Leu Ile Asn Asn
 340
 <210> 465
 <211> 343
 <212> PRT

<213> Artificial sequence

<220><223> Anaerostipes caccae KARI variant-2011_E8

<400> 465

Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn

1 5 10 15

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

20 25 30

Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val

35 40 45

Ile Ile Gly Leu Phe Glu Gly Ala Glu Glu Trp Lys Arg Ala Glu Glu

50 55 60

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

65 70 75 80

Ile Ile Met Ile Leu Ile Pro Asp Glu Lys Gln Ala Thr Met Tyr Lys

85 90 95

Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala

100 105 110

His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val

115 120 125

Asp Val Thr Met Ile Ala Pro Lys Gly Pro Gly His Thr Ile Arg Ser

130 135 140

Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln

145 150 155 160

Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala

165 170 175

Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu

180 185 190

Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val

195 200 205

Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr

210 215 220

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

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

 Glu Asp Lys Leu Ile Asn Asn
 340

 <210> 466
 <211> 343
 <212> PRT
 <213> Artificial sequence
 <220><223> Anaerostipes caccae KARI variant-2011_E11
 <400> 466

 Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn
 1 5 10 15
 Leu Ser Leu Leu Asp Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser
 20 25 30
 Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val

 35 40 45
 Ile Ile Gly Leu Phe Glu Gly Ala Glu Glu Trp Lys Arg Ala Glu Glu
 50 55 60
 Gln Gly Phe Glu Val Tyr Thr Ala Ala Glu Ala Ala Lys Lys Ala Asp
 65 70 75 80

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

325 330 335

Glu Asp Lys Leu Ile Asn Asn

340

<210> 467

<211> 343

<212> PRT

<213> Artificial sequence

<220><223> Anaerostipes caccae KARI variant-2011_F2

<400> 467

Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn

1 5 10 15

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

20 25 30

Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val

35 40 45

Ile Ile Gly Leu Phe Glu Gly Val Glu Glu Trp Lys Arg Ala Glu Glu

50 55 60

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

65 70 75 80

Ile Ile Met Ile Leu Ile Pro Asp Glu Lys Gln Ala Thr Met Tyr Lys

85 90 95

Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala

100 105 110

His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val

115 120 125

Asp Val Thr Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser

130 135 140

Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln

145 150 155 160

Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala

165 170 175

Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu

180 185 190
 Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val
 195 200 205
 Cys Thr Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr
 210 215 220
 Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile
 225 230 235 240
 Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile
 245 250 255

 Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile
 260 265 270
 Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln
 275 280 285
 Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly
 290 295 300
 Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro
 305 310 315 320

 Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
 325 330 335
 Glu Asp Lys Leu Ile Asn Asn
 340
 <210> 468
 <211> 343
 <212> PRT
 <213> Artificial sequence
 <220><223> Anaerostipes caccae KARI variant-2011_F4
 <400> 468
 Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn
 1 5 10 15
 Leu Ser Leu Leu Asp Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser
 20 25 30

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

275 280 285
 Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly
 290 295 300
 Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro
 305 310 315 320
 Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
 325 330 335
 Glu Asp Lys Leu Ile Asn Asn
 340

<210> 469

<211> 343

<212> PRT

<213> Artificial sequence

<220><223> Anaerostipes caccae KARI variant-2011_F6

<400> 469

Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn
 1 5 10 15
 Leu Ser Leu Leu Asp Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser
 20 25 30
 Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val
 35 40 45

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

His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val
 115 120 125
 Asp Val Thr Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser

130 135 140
 Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln
 145 150 155 160
 Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala
 165 170 175

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

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

 Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro
 305 310 315 320
 Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
 325 330 335
 Glu Asp Lys Leu Ile Asn Asn
 340

 <210> 470
 <211> 343
 <212> PRT
 <213> Artificial sequence
 <220><223> Anaerostipes caccae KARI variant-2011_F9

<400> 470

Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn

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

65 70 75 80
Ile Ile Met Ile Leu Ile Pro Asp Glu Lys Gln Ala Thr Met Tyr Lys
85 90 95
Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala
100 105 110
His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val
115 120 125
Asp Val Thr Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser

130 135 140
Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln
145 150 155 160
Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala
165 170 175
Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu
180 185 190
Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val

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

Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile
245 250 255

Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile
260 265 270

Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln
275 280 285

Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly
290 295 300

Pro Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro
305 310 315 320

Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
325 330 335

Glu Asp Lys Leu Ile Asn Asn
340

<210> 471

<211> 343

<212> PRT

<213> Artificial sequence

<220><223> Anaerostipes caccae KARI variant-2011_F10

<400> 471

Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn
1 5 10 15

Leu Ser Leu Leu Asp Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser
20 25 30

Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val
35 40 45

Ile Ile Gly Leu Phe Glu Gly Ala Glu Glu Trp Lys Arg Ala Glu Glu
50 55 60

Gln Gly Phe Glu Val Tyr Thr Ala Ala Glu Ala Ala Lys Lys Ala Asp
65 70 75 80

Ile Ile Met Ile Met Ile Pro Asp Glu Lys Gln Ala Thr Met Tyr Lys

85

90

95

Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala

100

105

110

His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val

115

120

125

Asp Val Thr Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser

130

135

140

Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln

145

150

155

160

Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala

165

170

175

Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu

180

185

190

Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val

195

200

205

Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr

210

215

220

Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile

225

230

235

240

Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile

245

250

255

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

260

265

270

Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln

275

280

285

Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly

290

295

300

Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro

305

310

315

320

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

325

330

335

Glu Asp Lys Leu Ile Asn Asn

340

<210> 472

<211> 343

<212> PRT

<213> Artificial sequence

<220><223> Anaerostipes caccae KARI variant-2011_G1

<400> 472

Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Val Asp Cys Asn

1 5 10 15

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

20 25 30

Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val

35 40 45

Ile Ile Gly Leu Phe Glu Gly Ala Glu Glu Trp Lys Arg Ala Glu Glu

50 55 60

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

65 70 75 80

Ile Ile Met Ile Leu Ile Pro Asp Glu Lys Gln Ala Thr Ile Tyr Lys

85 90 95

Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala

100 105 110

His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val

115 120 125

Asp Val Thr Met Ile Ala Pro Lys Gly Pro Gly His Ile Val Arg Ser

130 135 140

Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln

145 150 155 160

Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala

165 170 175

Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu

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

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

 305 310 315 320
 Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
 325 330 335
 Glu Asp Lys Leu Ile Asn Asn
 340
 <210> 473
 <211> 343
 <212> PRT
 <213> Artificial sequence
 <220><223> Anaerostipes caccae KARI variant-2011_G3
 <400> 473
 Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn
 1 5 10 15

 Leu Ser Leu Leu Asp Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser
 20 25 30
 Gln Gly Gln Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val

35	40	45	
Ile Ile Gly Leu Phe Glu Gly Ala Glu Glu Trp Lys Arg Ala Glu Glu			
50	55	60	
Gln Gly Phe Glu Val Tyr Thr Ala Ala Glu Ala Ala Lys Lys Ala Asp			
65	70	75	80
Ile Ile Met Ile Leu Ile Pro Asp Glu Lys Gln Ala Thr Met Tyr Lys			
85	90	95	
Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala			
100	105	110	
His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val			
115	120	125	
Asp Val Thr Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser			
130	135	140	
Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln			
145	150	155	160
Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala			
165	170	175	
Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu			
180	185	190	
Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val			
195	200	205	
Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr			
210	215	220	
Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile			
225	230	235	240
Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile			
245	250	255	
Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile			
260	265	270	
Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln			
275	280	285	

Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly
290 295 300

Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro
305 310 315 320

Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
325 330 335

Glu Asp Lys Leu Ile Asn Asn
340

<210> 474

<211> 343

<212> PRT

<213> Artificial sequence

<220><223> Anaerostipes caccae KARI variant-2011_G4

<400> 474

Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn
1 5 10 15

Leu Ser Leu Leu Asp Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser
20 25 30

Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val
35 40 45

Ile Ile Gly Leu Phe Glu Gly Ala Glu Glu Trp Lys Arg Ala Glu Glu
50 55 60

Gln Gly Phe Glu Val Tyr Thr Ala Ala Glu Ala Ala Lys Lys Ala Asp
65 70 75 80

Ile Ile Met Phe Leu Ile Pro Asp Glu Lys Gln Ala Thr Met Tyr Lys
85 90 95

Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala
100 105 110

His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val
115 120 125

Asp Val Thr Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser
130 135 140

Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln
 145 150 155 160
 Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala

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

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

 290 295 300
 Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro
 305 310 315 320
 Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
 325 330 335
 Glu Asp Lys Leu Ile Asn Asn
 340
 <210> 475
 <211> 343
 <212> PRT
 <213> Artificial sequence
 <220><223> Anaerostipes caccae KARI variant-2011_G8

<400> 475

Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn

1 5 10 15

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

20 25 30

Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val

35 40 45

Ile Ile Gly Leu Phe Glu Gly Ala Glu Glu Trp Lys Arg Ala Glu Glu

50 55 60

Gln Gly Phe Glu Val Tyr Thr Ala Ala Glu Ala Ala Asn Lys Ala Asp

65 70 75 80

Ile Ile Met Ile Leu Ile Pro Asp Glu Lys Gln Val Thr Met Tyr Lys

85 90 95

Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala

100 105 110

His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val

115 120 125

Asp Val Thr Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser

130 135 140

Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln

145 150 155 160

Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala

165 170 175

Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu

180 185 190

Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val

195 200 205

Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr

210 215 220

Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile

225 230 235 240

Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile

245 250 255

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

260 265 270

Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln

275 280 285

Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly

290 295 300

Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro

305 310 315 320

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

325 330 335

Glu Asp Lys Leu Ile Asn Asn

340

<210> 476

<211> 343

<212> PRT

<213> Artificial sequence

<220><223> Anaerostipes caccae KARI variant-2011_G9

<400> 476

Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn

1 5 10 15

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

20 25 30

Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val

35 40 45

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

50 55 60

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

65 70 75 80

Ile Ile Met Ile Leu Ile Pro Asp Glu Lys Gln Ala Thr Met Tyr Lys

85

90

95

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

 145 150 155 160
 Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala
 165 170 175
 Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu
 180 185 190
 Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val
 195 200 205
 Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr

 210 215 220
 Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile
 225 230 235 240
 Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile
 245 250 255
 Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile
 260 265 270
 Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln

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

340

<210> 477

<211> 343

<212> PRT

<213> Artificial sequence

<220><223> Anaerostipes caccae KARI variant-2011_G10

<400> 477

Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn

1 5 10 15

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

20 25 30

Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val

35 40 45

Ile Ile Gly Leu Phe Glu Gly Ala Glu Glu Trp Lys Arg Ala Glu Glu

50 55 60

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

65 70 75 80

Ile Ile Met Phe Leu Ile Pro Asp Glu Lys Gln Ala Thr Met Tyr Lys

85 90 95

Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala

100 105 110

His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val

115 120 125

Asp Val Thr Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser

130 135 140

Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln

145 150 155 160

Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala

165 170 175

Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu

180 185 190

Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val
195 200 205

Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr
210 215 220

Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile
225 230 235 240

Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile
245 250 255

Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile
260 265 270

Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln
275 280 285

Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly
290 295 300

Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro
305 310 315 320

Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
325 330 335

Glu Asp Lys Leu Ile Asn Asn
340

<210> 478

<211> 343

<212> PRT

<213> Artificial sequence

<220><223> Anaerostipes caccae KARI variant-2011_H1

<400> 478

Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn

1 5 10 15

Leu Ser Leu Leu Asp Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser
20 25 30

Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val
35 40 45

Ile Ile Gly Leu Phe Glu Gly Ala Glu Glu Trp Lys Arg Ala Glu Glu
 50 55 60
 Gln Gly Phe Glu Val Tyr Thr Ala Ala Glu Ala Ala Lys Lys Ala Asp

 65 70 75 80
 Ile Ile Met Ile Leu Ile Pro Asp Glu Lys Gln Ala Thr Ile Tyr Lys
 85 90 95
 Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala
 100 105 110
 His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val
 115 120 125
 Asp Val Thr Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser

 130 135 140
 Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Leu Ala Val Glu Gln
 145 150 155 160
 Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala
 165 170 175
 Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu
 180 185 190
 Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val

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

 260 265 270
 Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln
 275 280 285
 Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly

290 295 300
 Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro
 305 310 315 320
 Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp

 325 330 335
 Glu Asp Lys Leu Ile Asn Asn
 340
 <210> 479
 <211> 343
 <212> PRT
 <213> Artificial sequence
 <220><223> Anaerostipes caccae KARI variant-2011_H5
 <400> 479
 Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn
 1 5 10 15
 Leu Ser Leu Leu Asp Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser
 20 25 30

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

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

145 150 155 160
 Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala
 165 170 175
 Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu
 180 185 190
 Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val
 195 200 205
 Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr
 210 215 220

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

 Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly
 290 295 300
 Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro
 305 310 315 320
 Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
 325 330 335
 Glu Asp Lys Leu Ile Asn Asn
 340
 <210> 480
 <211> 343
 <212> PRT
 <213> Artificial sequence

 <220><223> Anaerostipes caccae KARI variant-2011_H7
 <400> 480

Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn
1 5 10 15
Leu Ser Leu Leu Asp Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser
 20 25 30
Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val
 35 40 45
Ile Ile Gly Leu Phe Glu Gly Thr Glu Glu Trp Lys Arg Ala Glu Glu

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

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

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

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

 305 310 315 320
 Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
 325 330 335
 Glu Asp Lys Leu Ile Asn Asn
 340
 <210> 481
 <211> 343
 <212> PRT
 <213> Artificial sequence
 <220><223> Anaerostipes caccae KARI variant-2011_H9
 <400> 481
 Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn
 1 5 10 15

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

 Ile Ile Met Ile Leu Ile Pro Asp Glu Lys Gln Ala Thr Met Tyr Lys
 85 90 95
 Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala

100	105	110	
His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val			
115	120	125	
Asp Val Thr Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser			
130	135	140	
Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln			
145	150	155	160
Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala			
165	170	175	
Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu			
180	185	190	
Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val			
195	200	205	
Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr			
210	215	220	
Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile			
225	230	235	240
Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile			
245	250	255	
Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile			
260	265	270	
Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln			
275	280	285	
Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly			
290	295	300	
Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro			
305	310	315	320
Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp			
325	330	335	
Glu Asp Lys Leu Ile Asn Asn			
340			

<210> 482

<211> 343

<212> PRT

<213> Artificial sequence

<220><223> Anaerostipes caccae KARI variant-2012_A2

<400> 482

Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn

1 5 10 15

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

20 25 30

Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val

35 40 45

Ile Ile Gly Leu Phe Glu Gly Ala Glu Glu Trp Lys Gly Ala Glu Glu

50 55 60

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

65 70 75 80

Ile Ile Met Ile Leu Val Pro Asp Glu Lys Gln Ala Thr Met Tyr Lys

85 90 95

Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala

100 105 110

His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val

115 120 125

Asp Val Thr Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser

130 135 140

Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln

145 150 155 160

Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala

165 170 175

Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu

180 185 190

Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val

195 200 205

Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr
 210 215 220
 Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile

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

 290 295 300
 Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro
 305 310 315 320
 Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
 325 330 335
 Glu Asp Lys Leu Ile Asn Asn
 340
 <210> 483
 <211> 343
 <212> PRT
 <213> Artificial sequence
 <220><223> Anaerostipes caccae KARI variant-2012_A7

 <400> 483
 Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn
 1 5 10 15
 Leu Ser Leu Leu Asp Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser
 20 25 30
 Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val
 35 40 45
 Ile Ile Gly Leu Phe Glu Gly Ala Glu Glu Trp Lys Arg Ala Glu Glu

50	55	60
Gln Gly Phe Glu Val Tyr Ser Ala Ala Glu Ala Val Lys Lys Ala Asp		
65	70	75
Ile Ile Met Ile Leu Ile Pro Asp Glu Lys Gln Ala Thr Met Tyr Lys		
	85	90
Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala		
	100	105
His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val		
	115	120
		125
Asp Val Thr Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser		
	130	135
Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln		
145	150	155
Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala		
	165	170
Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu		
	180	185
		190
Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val		
	195	200
Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr		
	210	215
Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile		
225	230	235
Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile		
	245	250
		255
Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile		
	260	265
Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln		
	275	280
Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly		
	290	295
		300

Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro
305 310 315 320

Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
325 330 335

Glu Asp Lys Leu Ile Asn Asn
340

<210> 484

<211> 343

<212> PRT

<213> Artificial sequence

<220><223> Anaerostipes caccae KARI variant-2012_A8

<400> 484

Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn
1 5 10 15

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

20 25 30

Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val

35 40 45

Ile Ile Gly Leu Phe Glu Gly Ala Glu Glu Trp Lys Arg Ala Glu Glu

50 55 60

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

65 70 75 80

Ile Ile Met Ile Leu Ile Pro Asp Glu Lys Gln Ala Thr Met Tyr Lys

85 90 95

Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala

100 105 110

His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val

115 120 125

Asp Val Thr Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser

130 135 140

Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln

145 150 155 160
 Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala
 165 170 175
 Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu
 180 185 190
 Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val
 195 200 205
 Cys Ala Leu Thr Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr

 210 215 220
 Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile
 225 230 235 240
 Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile
 245 250 255
 Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile
 260 265 270
 Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln

 275 280 285
 Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly
 290 295 300
 Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro
 305 310 315 320
 Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
 325 330 335
 Glu Asp Lys Leu Ile Asn Asn
 340

<210> 485

<211> 343

<212> PRT

<213> Artificial sequence

<220><223> Anaerostipes caccae KARI variant-2012_A9

<400> 485

Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn

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

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

Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro
 305 310 315 320
 Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
 325 330 335
 Glu Asp Lys Leu Ile Asn Asn
 340

<210> 486

<211> 343

<212> PRT

<213> Artificial sequence

<220><223> Anaerostipes caccae KARI variant-2012_A10

<400> 486

Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn

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

65 70 75 80
 Ile Ile Met Ile Leu Ile Pro Asp Glu Lys Gln Ala Thr Met Tyr Lys
 85 90 95
 Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala
 100 105 110

His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val
115 120 125
Asp Val Thr Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser
130 135 140
Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln
145 150 155 160
Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala
165 170 175
Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu
180 185 190
Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val
195 200 205
Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr
210 215 220
Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile
225 230 235 240
Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile
245 250 255
Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile
260 265 270
Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln
275 280 285
Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly
290 295 300
Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro
305 310 315 320
Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
325 330 335
Glu Asp Lys Leu Ile Asn Asn
340
<210> 487

<211> 343

<212> PRT

<213> Artificial sequence

<220><223> Anaerostipes caccae KARI variant-2012_A11

<400> 487

Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn

1 5 10 15

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

20 25 30

Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val

35 40 45

Ile Ile Gly Leu Phe Glu Gly Ala Glu Glu Trp Lys Arg Ala Glu Glu

50 55 60

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

65 70 75 80

Ile Ile Met Ile Leu Ile Pro Asp Glu Lys Gln Ala Thr Met Tyr Lys

85 90 95

Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala

100 105 110

His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val

115 120 125

Asp Val Thr Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser

130 135 140

Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln

145 150 155 160

Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala

165 170 175

Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu

180 185 190

Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val

195 200 205

Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr

210 215 220

Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile

225 230 235 240

Val Asp Leu Ile Tyr Gln Thr Gly Phe Ser Gly Met Arg Tyr Ser Ile

245 250 255

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

260 265 270

Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln

275 280 285

Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly

290 295 300

Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro

305 310 315 320

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

325 330 335

Glu Asp Lys Leu Ile Asn Asn

340

<210> 488

<211> 343

<212> PRT

<213> Artificial sequence

<220><223> Anaerostipes caccae KARI variant-2012_B3

<400> 488

Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn

1 5 10 15

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

20 25 30

Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val

35 40 45

Ile Ile Gly Leu Phe Glu Gly Ala Glu Glu Trp Lys Arg Ala Glu Glu

50 55 60

Gln	Gly	Phe	Glu	Val	Tyr	Thr	Ala	Ala	Glu	Ala	Ala	Lys	Lys	Ala	Asp				
65					70					75					80				
Ile	Ile	Met	Ile	Leu	Ile	Pro	Asp	Glu	Lys	Gln	Ala	Thr	Met	Tyr	Lys				
85					90					95									
Asn	Asp	Ile	Glu	Pro	Asn	Leu	Glu	Ala	Gly	Asn	Met	Leu	Met	Phe	Ala				
100					105					110									
His	Gly	Phe	Asn	Ile	His	Phe	Gly	Cys	Ile	Val	Pro	Pro	Lys	Asp	Val				
115					120					125									
Asp	Val	Thr	Met	Ile	Ala	Ser	Lys	Gly	Pro	Gly	His	Thr	Val	Arg	Ser				
130					135					140									
Glu	Tyr	Glu	Glu	Gly	Lys	Gly	Val	Pro	Cys	Leu	Val	Ala	Val	Glu	Gln				
145					150					155					160				
Asp	Ala	Thr	Gly	Lys	Ala	Leu	Asp	Met	Ala	Leu	Ala	Tyr	Ala	Leu	Ala				
165					170					175									
Ile	Gly	Gly	Ala	Arg	Ala	Gly	Val	Leu	Glu	Thr	Thr	Phe	Arg	Thr	Glu				
180					185					190									
Thr	Glu	Thr	Asp	Leu	Phe	Gly	Glu	Gln	Ala	Val	Leu	Cys	Gly	Gly	Val				
195					200					205									
Cys	Ala	Leu	Met	Gln	Ala	Gly	Phe	Glu	Thr	Leu	Val	Glu	Ala	Gly	Tyr				
210					215					220									
Asp	Pro	Arg	Asn	Ala	Tyr	Phe	Glu	Cys	Ile	His	Glu	Met	Lys	Leu	Ile				
225					230					235					240				
Val	Asp	Leu	Ile	Tyr	Gln	Ser	Gly	Phe	Ser	Gly	Met	Arg	Tyr	Ser	Ile				
245					250					255									
Ser	Asn	Thr	Ala	Glu	Tyr	Gly	Asp	Tyr	Ile	Thr	Gly	Pro	Lys	Ile	Ile				
260					265					270									
Thr	Glu	Asp	Thr	Lys	Lys	Ala	Met	Lys	Lys	Ile	Leu	Ser	Asp	Ile	Gln				
275					280					285									
Asp	Gly	Thr	Phe	Ala	Lys	Asp	Phe	Leu	Val	Asp	Met	Ser	Asp	Ala	Gly				
290					295					300									
Ser	Gln	Val	His	Phe	Lys	Ala	Met	Arg	Lys	Leu	Ala	Ser	Glu	His	Pro				

305 310 315 320
 Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
 325 330 335
 Glu Asp Lys Leu Ile Asn Asn
 340
 <210> 489
 <211> 343
 <212> PRT
 <213> Artificial sequence
 <220><223> Anaerostipes caccae KARI variant-2012_B4
 <400> 489
 Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn
 1 5 10 15

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

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

 Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln
 145 150 155 160
 Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala

165 170 175
 Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu
 180 185 190
 Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val
 195 200 205

 Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr
 210 215 220
 Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile
 225 230 235 240
 Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile
 245 250 255
 Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile
 260 265 270

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

 Glu Asp Lys Leu Ile Asn Asn
 340

 <210> 490
 <211> 343
 <212> PRT
 <213> Artificial sequence
 <220><223> Anaerostipes caccae KARI variant-2012_B5
 <400> 490

 Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn
 1 5 10 15
 Leu Ser Leu Leu Asp Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser

20 25 30
 Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val

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

 100 105 110
 His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val
 115 120 125
 Asp Val Thr Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser
 130 135 140
 Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln
 145 150 155 160
 Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala

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

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

Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln
275 280 285

Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly
290 295 300

Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro
305 310 315 320

Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
325 330 335

Glu Asp Lys Leu Ile Asn Asn
340

<210> 491
<211> 343
<212> PRT
<213> Artificial sequence
<220><223> Anaerostipes caccae KARI variant-2012_B6

<400> 491

Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn
1 5 10 15

Leu Ser Leu Leu Asp Gly Lys Thr Ile Ala Val Ile Gly His Gly Ser
20 25 30

Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val
35 40 45

Ile Ile Gly Leu Phe Glu Gly Ala Glu Glu Trp Lys Arg Ala Glu Glu
50 55 60

Gln Gly Phe Glu Val Tyr Thr Ala Ala Glu Ala Ala Lys Lys Ala Asp
65 70 75 80

Ile Ile Met Ile Leu Ile Pro Asp Glu Lys Gln Ala Thr Met Tyr Lys
85 90 95

Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala
100 105 110

His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val

115	120	125	
Asp Val Thr Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser			
130	135	140	
Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln			
145	150	155	160
Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala			
165	170	175	
Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu			
180	185	190	
Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val			
195	200	205	
Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr			
210	215	220	
Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile			
225	230	235	240
Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile			
245	250	255	
Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile			
260	265	270	
Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln			
275	280	285	
Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly			
290	295	300	
Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro			
305	310	315	320
Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp			
325	330	335	
Glu Asp Lys Leu Ile Asn Asn			
340			
<210> 492			
<211> 343			

<212> PRT

<213> Artificial sequence

<220><223> Anaerostipes caccae KARI variant-2012_B9

<400> 492

Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn

1 5 10 15

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

20 25 30

Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val

35 40 45

Ile Met Gly Leu Phe Glu Gly Ala Glu Glu Trp Lys Arg Ala Glu Glu

50 55 60

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

65 70 75 80

Ile Ile Met Ile Leu Ile Pro Asp Glu Lys Gln Ala Thr Met Tyr Lys

85 90 95

Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala

100 105 110

His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val

115 120 125

Asp Val Thr Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser

130 135 140

Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln

145 150 155 160

Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala

165 170 175

Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu

180 185 190

Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val

195 200 205

Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr

210 215 220
 Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile
 225 230 235 240
 Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile
 245 250 255
 Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile
 260 265 270
 Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln

275 280 285
 Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly
 290 295 300
 Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro
 305 310 315 320
 Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
 325 330 335
 Glu Asp Lys Leu Ile Asn Asn
 340

<210> 493

<211> 343

<212> PRT

<213> Artificial sequence

<220><223> Anaerostipes caccae KARI variant-2012_C2

<400> 493

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

65	70	75	80
Ile Ile Met Ile Leu Ile Pro Asp Glu Lys Gln Ala Thr Met Tyr Lys			
	85	90	95
Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala			
	100	105	110
His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val			
	115	120	125
Asp Val Thr Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser			
	130	135	140
Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln			
145	150	155	160
Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala			
	165	170	175
Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu			
	180	185	190
Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val			
	195	200	205
Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr			
	210	215	220
Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile			
225	230	235	240
Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile			
	245	250	255
Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile			
	260	265	270
Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln			
	275	280	285
Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly			
	290	295	300
Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro			
305	310	315	320

Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
325 330 335

Glu Asp Lys Leu Ile Asn Asn
340

<210> 494

<211> 343

<212> PRT

<213> Artificial sequence

<220><223> Anaerostipes caccae KARI variant-2012_C3

<400> 494

Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn

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

65 70 75 80
Ile Ile Met Ile Leu Ile Pro Asp Glu Lys Gln Ala Thr Met Tyr Lys
85 90 95
Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala
100 105 110
His Gly Val Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val
115 120 125
Asp Val Thr Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser

130 135 140
Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln
145 150 155 160
Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala
165 170 175

Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Ser Glu
180 185 190
Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Ile
195 200 205
Trp Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr
210 215 220
Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile
225 230 235 240
Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile
245 250 255
Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile
260 265 270
Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln
275 280 285
Asp Gly Thr Ile Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly
290 295 300
Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro
305 310 315 320
Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
325 330 335
Glu Asp Lys Leu Ile Asn Asn
340
<210> 495
<211> 343
<212> PRT
<213> Artificial sequence
<220><223> Anaerostipes caccae KARI variant-2012_C5
<400> 495
Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn
1 5 10 15
Leu Ser Leu Leu Asp Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser

20	25	30	
Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val			
35	40	45	
Ile Ile Gly Leu Phe Glu Gly Ala Glu Glu Trp Lys Arg Ala Glu Glu			
50	55	60	
Gln Gly Phe Glu Val Tyr Thr Ala Ala Glu Ala Ala Lys Lys Ala Asp			
65	70	75	80
Ile Ile Met Ile Leu Ile Pro Asp Glu Lys Gln Ala Thr Ile Tyr Lys			
85	90	95	
Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala			
100	105	110	
His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val			
115	120	125	
Asp Val Thr Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser			
130	135	140	
Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln			
145	150	155	160
Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala			
165	170	175	
Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu			
180	185	190	
Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val			
195	200	205	
Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr			
210	215	220	
Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile			
225	230	235	240
Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile			
245	250	255	
Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile			
260	265	270	

Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln
275 280 285

Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly
290 295 300

Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro
305 310 315 320

Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
325 330 335

Glu Asp Lys Leu Ile Asn Asn
340

<210> 496

<211> 343

<212> PRT

<213> Artificial sequence

<220><223> Anaerostipes caccae KARI variant

<400> 496

Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn
1 5 10 15

Leu Ser Leu Leu Asp Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser
20 25 30

Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val
35 40 45

Ile Ile Gly Leu Phe Glu Gly Ala Glu Glu Trp Lys Arg Ala Glu Glu

50 55 60

Gln Gly Leu Glu Val Tyr Thr Ala Ala Glu Ala Ala Lys Lys Ala Asp
65 70 75 80

Ile Ile Met Ile Leu Ile Pro Asp Glu Lys Gln Ala Thr Met Tyr Lys
85 90 95

Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala
100 105 110

His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val

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

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

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

305 310 315 320
 Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
 325 330 335
 Glu Asp Lys Leu Ile Asn Asn

340

<210> 497

<211> 343

<212> PRT

<213> Artificial sequence

<220><223> Anaerostipes caccae KARI variant-2012_C8

<400> 497

Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn

1 5 10 15

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

20 25 30

Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val

35 40 45

Ile Ile Gly Leu Phe Glu Gly Ala Glu Glu Trp Lys Arg Ala Glu Glu

50 55 60

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

65 70 75 80

Ile Ile Met Ile Leu Ile Pro Asp Glu Lys Gln Ala Thr Ile Tyr Lys

85 90 95

Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala

100 105 110

His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val

115 120 125

Asp Val Thr Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser

130 135 140

Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln

145 150 155 160

Asp Ala Thr Gly Lys Ala Leu Asp Thr Ala Leu Ala Tyr Ala Leu Ala

165 170 175

Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu

180 185 190

Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val

195 200 205

Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr

210 215 220

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

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

 Glu Asp Lys Leu Ile Asn Asn
 340

 <210> 498
 <211> 343
 <212> PRT
 <213> Artificial sequence
 <220><223> Anaerostipes caccae KARI variant-2012_C10
 <400> 498

 Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn
 1 5 10 15
 Leu Ser Leu Leu Asp Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser
 20 25 30
 Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val

 35 40 45
 Ile Ile Gly Leu Phe Glu Gly Ala Glu Glu Trp Lys Arg Ala Glu Glu
 50 55 60
 Gln Gly Ile Glu Val Tyr Thr Ala Ala Glu Ala Ala Lys Lys Ala Asp
 65 70 75 80

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

325 330 335

Glu Asp Lys Leu Ile Asn Asn

340

<210> 499

<211> 343

<212> PRT

<213> Artificial sequence

<220><223> Anaerostipes caccae KARI variant-2012_D1

<400> 499

Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn

1 5 10 15

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

20 25 30

Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val

35 40 45

Ile Ile Gly Leu Phe Glu Gly Ala Glu Glu Trp Lys Arg Ala Glu Glu

50 55 60

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

65 70 75 80

Ile Ile Met Phe Leu Ile Pro Asp Glu Lys Gln Ala Thr Met Tyr Lys

85 90 95

Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala

100 105 110

His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val

115 120 125

Asp Val Thr Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser

130 135 140

Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln

145 150 155 160

Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala

165 170 175

Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu

180 185 190
 Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val
 195 200 205
 Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr
 210 215 220
 Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile
 225 230 235 240
 Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile
 245 250 255

 Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile
 260 265 270
 Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln
 275 280 285
 Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly
 290 295 300
 Ser Gln Val His Phe Lys Ala Thr Arg Lys Leu Ala Ser Glu His Pro
 305 310 315 320

 Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
 325 330 335
 Glu Asp Lys Leu Ile Asn Asn
 340
 <210> 500
 <211> 343
 <212> PRT
 <213> Artificial sequence
 <220><223> Anaerostipes caccae KARI variant D6_2011
 <400> 500
 Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn
 1 5 10 15
 Leu Ser Leu Leu Asp Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser
 20 25 30

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

275 280 285
 Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly
 290 295 300
 Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro
 305 310 315 320
 Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
 325 330 335
 Glu Asp Lys Leu Ile Asn Asn
 340

<210> 501

<211> 343

<212> PRT

<213> Artificial sequence

<220><223> Anaerostipes caccae KARI variant-2012_D8

<400> 501

Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn
 1 5 10 15
 Leu Ser Leu Leu Asp Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser
 20 25 30
 Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val
 35 40 45

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

His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val
 115 120 125
 Asp Val Thr Met Ile Ala Ser Lys Gly Pro Gly His Thr Val Arg Ser

130 135 140
 Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln
 145 150 155 160
 Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala
 165 170 175

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

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

 Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro
 305 310 315 320
 Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
 325 330 335
 Glu Asp Lys Leu Ile Asn Asn
 340

 <210> 502
 <211> 343
 <212> PRT
 <213> Artificial sequence
 <220><223> Anaerostipes caccae KARI variant-2012_D11

<400> 502

Met Glu Glu Cys Lys Met Ala Asn Ile Tyr Tyr Gln Glu Asp Cys Asn

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

65 70 75 80
Ile Ile Met Ile Leu Ile Pro Asp Glu Met Gln Ala Thr Met Tyr Lys
85 90 95
Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala
100 105 110
His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val
115 120 125
Asp Val Thr Met Ile Ala Pro Lys Gly Pro Gly His Ile Val Arg Ser

130 135 140
Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln
145 150 155 160
Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala
165 170 175
Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu
180 185 190
Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val

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

Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile
245 250 255

Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile
260 265 270

Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln
275 280 285

Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly
290 295 300

Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro
305 310 315 320

Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
325 330 335

Glu Asp Lys Leu Ile Asn Asn
340

<210> 503
<211> 343
<212> PRT
<213> Artificial sequence
<220><223> Anaerostipes caccae KARI variant-2012_E5
<400> 503

Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn
1 5 10 15

Leu Ser Leu Leu Asp Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser
20 25 30

Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val
35 40 45

Ile Ile Gly Leu Phe Glu Gly Ala Glu Glu Trp Lys Arg Ala Glu Glu
50 55 60

Gln Gly Phe Glu Val Tyr Thr Ala Ala Glu Ala Ala Lys Lys Ala Asp
65 70 75 80

Ile Ile Met Leu Leu Ile Pro Asp Glu Lys Gln Ala Thr Met Tyr Lys

85

90

95

Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala

100

105

110

His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val

115

120

125

Asp Val Thr Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser

130

135

140

Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln

145

150

155

160

Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala

165

170

175

Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu

180

185

190

Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val

195

200

205

Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr

210

215

220

Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile

225

230

235

240

Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile

245

250

255

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

260

265

270

Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln

275

280

285

Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly

290

295

300

Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro

305

310

315

320

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

325

330

335

Glu Asp Lys Leu Ile Asn Asn

340

<210> 504

<211> 343

<212> PRT

<213> Artificial sequence

<220><223> Anaerostipes caccae KARI variant-2012_E9

<400> 504

Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn

1 5 10 15

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

20 25 30

Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val

35 40 45

Ile Ile Gly Leu Phe Glu Gly Ala Glu Glu Trp Lys Arg Ala Glu Glu

50 55 60

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

65 70 75 80

Ile Ile Ile Ile Leu Ile Pro Asp Glu Lys Gln Ala Thr Met Tyr Lys

85 90 95

Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala

100 105 110

His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val

115 120 125

Asp Val Thr Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser

130 135 140

Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln

145 150 155 160

Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala

165 170 175

Ile Gly Gly Ala Arg Ala Gly Val Met Glu Thr Thr Phe Arg Thr Glu

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

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

 305 310 315 320
 Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
 325 330 335
 Glu Asp Lys Leu Ile Asn Asn
 340
 <210> 505
 <211> 343
 <212> PRT
 <213> Artificial sequence
 <220><223> Anaerostipes caccae KARI variant-2012_F1
 <400> 505
 Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn
 1 5 10 15

 Leu Ser Leu Leu Asp Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser
 20 25 30
 Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val

35	40	45	
Ile Ile Gly Leu Phe Glu Gly Ala Glu Glu Trp Lys Arg Ala Glu Glu			
50	55	60	
Gln Gly Phe Glu Val Tyr Thr Ala Ala Glu Ala Ala Lys Lys Ala Asp			
65	70	75	80
Ile Ile Met Ile Leu Ile Pro Asp Glu Lys Gln Ala Thr Met Tyr Lys			
85	90	95	
Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala			
100	105	110	
His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val			
115	120	125	
Asp Val Thr Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser			
130	135	140	
Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln			
145	150	155	160
Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala			
165	170	175	
Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu			
180	185	190	
Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val			
195	200	205	
Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr			
210	215	220	
Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile Tyr Glu Met Lys Leu Ile			
225	230	235	240
Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile			
245	250	255	
Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile			
260	265	270	
Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln			
275	280	285	

Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly
 290 295 300

Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro
 305 310 315 320

Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
 325 330 335

Glu Asp Lys Leu Ile Asn Asn
 340

<210> 506
 <211> 343
 <212> PRT
 <213> Artificial sequence
 <220><223> Anaerostipes caccae KARI variant-2012_F2
 <400> 506

Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn
 1 5 10 15

Leu Ser Leu Leu Asp Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser
 20 25 30

Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val
 35 40 45

Ile Ile Gly Leu Phe Glu Gly Val Glu Glu Trp Lys Arg Ala Glu Glu
 50 55 60

Gln Gly Phe Glu Val Tyr Thr Ala Ala Glu Ala Ala Lys Lys Ala Asp
 65 70 75 80

Ile Ile Met Ile Leu Ile Pro Asp Glu Lys Gln Ala Thr Met Tyr Lys
 85 90 95

Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala
 100 105 110

His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val
 115 120 125

Asp Val Thr Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser
 130 135 140

Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln
 145 150 155 160
 Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala

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

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

 290 295 300
 Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro
 305 310 315 320
 Ala Glu Asp Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
 325 330 335
 Glu Asp Lys Leu Ile Asn Asn
 340
 <210> 507
 <211> 343
 <212> PRT
 <213> Artificial sequence
 <220><223> Anaerostipes caccae KARI variant-2012_F3

<400> 507

Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn

1 5 10 15

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

20 25 30

Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val

35 40 45

Ile Ile Gly Leu Phe Glu Gly Ala Glu Glu Trp Lys Arg Ala Glu Glu

50 55 60

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

65 70 75 80

Ile Ile Met Ile Leu Ile Pro Asp Glu Met Gln Ala Thr Met Tyr Lys

85 90 95

Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala

100 105 110

His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val

115 120 125

Asp Val Thr Met Ile Ala Pro Lys Gly Pro Gly His Thr Ile Arg Ser

130 135 140

Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln

145 150 155 160

Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala

165 170 175

Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Ser Thr Phe Arg Thr Glu

180 185 190

Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val

195 200 205

Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr

210 215 220

Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile

225 230 235 240

Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile

245 250 255

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

260 265 270

Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln

275 280 285

Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly

290 295 300

Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro

305 310 315 320

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

325 330 335

Glu Asp Lys Leu Ile Asn Asn

340

<210> 508

<211> 343

<212> PRT

<213> Artificial sequence

<220><223> Anaerostipes caccae KARI variant-2012_F4

<400> 508

Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn

1 5 10 15

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

20 25 30

Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val

35 40 45

Ile Ile Gly Leu Phe Glu Gly Ala Glu Glu Trp Lys Arg Ala Glu Glu

50 55 60

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

65 70 75 80

Ile Ile Met Ile Leu Ile Pro Asp Glu Lys Gln Ala Thr Ile Tyr Lys

85

90

95

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

 145 150 155 160
 Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala
 165 170 175
 Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu
 180 185 190
 Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val
 195 200 205
 Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr

 210 215 220
 Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile
 225 230 235 240
 Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile
 245 250 255
 Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile
 260 265 270
 Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln

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

340

<210> 509

<211> 343

<212> PRT

<213> Artificial sequence

<220><223> Anaerostipes caccae KARI variant-2012_F7

<400> 509

Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Val Asp Cys Asn

1 5 10 15

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

20 25 30

Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val

35 40 45

Ile Ile Gly Leu Phe Glu Gly Ala Glu Glu Trp Lys Arg Ala Glu Glu

50 55 60

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

65 70 75 80

Ile Ile Met Ile Leu Ile Pro Asp Glu Lys Gln Ala Thr Ile Tyr Lys

85 90 95

Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala

100 105 110

His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val

115 120 125

Asp Val Thr Met Ile Ala Pro Lys Gly Pro Gly His Ile Val Arg Ser

130 135 140

Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln

145 150 155 160

Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala

165 170 175

Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu

180 185 190

Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val
195 200 205

Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr
210 215 220

Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile
225 230 235 240

Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile
245 250 255

Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile
260 265 270

Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln
275 280 285

Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly
290 295 300

Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro
305 310 315 320

Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
325 330 335

Glu Asp Lys Leu Ile Asn Asn
340

<210> 510

<211> 343

<212> PRT

<213> Artificial sequence

<220><223> Anaerostipes caccae KARI variant-2012_F10

<400> 510

Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn

1 5 10 15

Leu Ser Leu Leu Asp Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser
20 25 30

Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val
35 40 45

Ile Ile Gly Leu Phe Glu Gly Ala Glu Glu Trp Lys Arg Ala Glu Glu
 50 55 60
 His Gly Phe Glu Val Tyr Thr Ala Ala Glu Ala Ala Lys Lys Ala Asp

 65 70 75 80
 Ile Ile Met Ile Leu Ile Pro Asp Glu Lys Gln Ala Thr Met Tyr Lys
 85 90 95
 Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala
 100 105 110
 His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val
 115 120 125
 Asp Val Thr Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser

 130 135 140
 Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln
 145 150 155 160
 Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala
 165 170 175
 Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Ile Arg Thr Glu
 180 185 190
 Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val

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

 260 265 270
 Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln
 275 280 285
 Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly

290 295 300
 Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro
 305 310 315 320
 Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp

 325 330 335
 Glu Asp Lys Leu Ile Asn Asn
 340
 <210> 511
 <211> 343
 <212> PRT
 <213> Artificial sequence
 <220><223> Anaerostipes caccae KARI variant-2012_F12
 <400> 511
 Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn
 1 5 10 15
 Leu Ser Leu Leu Asp Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser
 20 25 30

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

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

145 150 155 160

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

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

 Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly
 290 295 300
 Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro
 305 310 315 320
 Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
 325 330 335
 Glu Asp Lys Leu Ile Asn Asn
 340
 <210> 512
 <211> 343
 <212> PRT
 <213> Artificial sequence

 <220><223> Anaerostipes caccae KARI variant-2012_G3
 <400> 512

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

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

 305 310 315 320
 Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
 325 330 335
 Glu Asp Lys Leu Ile Asn Asn
 340
 <210> 513
 <211> 343
 <212> PRT
 <213> Artificial sequence
 <220><223> Anaerostipes caccae KARI variant-2012_G4
 <400> 513
 Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn
 1 5 10 15

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

 Ile Ile Met Ile Leu Ile Pro Asp Glu Lys Gln Ala Thr Met Tyr Lys
 85 90 95
 Asn Asp Ile Glu Pro Ser Leu Glu Ala Gly Asn Met Leu Met Phe Ala

100	105	110	
His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val			
115	120	125	
Asp Val Thr Met Ile Ala Pro Lys Gly Pro Gly His Thr Ile Arg Ser			
130	135	140	
Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln			
145	150	155	160
Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala			
165	170	175	
Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu			
180	185	190	
Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val			
195	200	205	
Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr			
210	215	220	
Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile			
225	230	235	240
Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile			
245	250	255	
Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile			
260	265	270	
Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln			
275	280	285	
Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly			
290	295	300	
Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro			
305	310	315	320
Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp			
325	330	335	
Glu Asp Lys Leu Ile Asn Asn			
340			

<210> 514

<211> 343

<212> PRT

<213> Artificial sequence

<220><223> Anaerostipes caccae KARI variant-2012_G5

<400> 514

Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn

1 5 10 15

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

20 25 30

Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val

35 40 45

Ile Ile Gly Leu Phe Glu Gly Ala Glu Glu Trp Lys Arg Ala Glu Glu

50 55 60

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

65 70 75 80

Ile Ile Met Leu Leu Ile Pro Asp Glu Lys Gln Ala Thr Met Tyr Lys

85 90 95

Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala

100 105 110

His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val

115 120 125

Asp Val Thr Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser

130 135 140

Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln

145 150 155 160

Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala

165 170 175

Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu

180 185 190

Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val

195 200 205

Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr
 210 215 220
 Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile

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

 290 295 300
 Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro
 305 310 315 320
 Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
 325 330 335
 Glu Asp Lys Leu Ile Asn Asn
 340
 <210> 515
 <211> 343
 <212> PRT
 <213> Artificial sequence
 <220><223> Anaerostipes caccae KARI variant-2012_G8

 <400> 515
 Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn
 1 5 10 15
 Leu Ser Leu Leu Asp Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser
 20 25 30
 Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val
 35 40 45
 Ile Ile Gly Leu Phe Glu Gly Ala Glu Glu Trp Lys Arg Ala Glu Glu

50 55 60
 Gln Gly Phe Glu Val Tyr Thr Thr Ala Glu Ala Ala Asn Lys Ala Asp
 65 70 75 80
 Ile Ile Met Ile Leu Ile Pro Asp Glu Lys Gln Val Thr Met Tyr Lys
 85 90 95
 Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala
 100 105 110
 His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val
 115 120 125

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

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

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

Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro
305 310 315 320

Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
325 330 335

Glu Asp Lys Leu Ile Asn Asn
340

<210> 516

<211> 343

<212> PRT

<213> Artificial sequence

<220><223> Anaerostipes caccae KARI variant-2012_G9

<400> 516

Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn
1 5 10 15

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

20 25 30

Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val

35 40 45

Ile Ile Gly Leu Phe Glu Gly Ala Glu Glu Trp Lys Arg Ala Glu Glu

50 55 60

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

65 70 75 80

Ile Ile Met Ile Leu Ile Pro Asp Glu Lys Gln Ala Thr Ile Tyr Lys

85 90 95

Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala

100 105 110

His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val

115 120 125

Asp Val Thr Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser

130 135 140

Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln

145 150 155 160
 Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala
 165 170 175
 Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu
 180 185 190
 Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val
 195 200 205
 Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr

 210 215 220
 Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile
 225 230 235 240
 Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile
 245 250 255
 Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile
 260 265 270
 Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln

 275 280 285
 Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly
 290 295 300
 Ser Gln Ile His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro
 305 310 315 320
 Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
 325 330 335
 Glu Asp Lys Leu Ile Asn Asn
 340

<210> 517

<211> 343

<212> PRT

<213> Artificial sequence

<220><223> Anaerostipes caccae KARI variant-2012_G10

<400> 517

Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn

1 5 10 15
 Leu Ser Leu Leu Asp Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser
 20 25 30
 Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val
 35 40 45

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

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

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

 Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile
 245 250 255

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

Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro
305 310 315 320
Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
325 330 335
Glu Asp Lys Leu Ile Asn Asn
340

<210> 518

<211> 343

<212> PRT

<213> Artificial sequence

<220><223> Anaerostipes caccae KARI variant-2012_G12

<400> 518

Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn

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

65 70 75 80
Ile Ile Met Ile Leu Ile Pro Asp Glu Lys Gln Ala Thr Met Tyr Lys
85 90 95
Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala
100 105 110

His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val
115 120 125
Asp Val Thr Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser
130 135 140
Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln
145 150 155 160
Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala
165 170 175
Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu
180 185 190
Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val
195 200 205
Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr
210 215 220
Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile
225 230 235 240
Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile
245 250 255
Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile
260 265 270
Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln
275 280 285
Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly
290 295 300
Ser Gln Val His Ile Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro
305 310 315 320
Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
325 330 335
Glu Asp Lys Leu Ile Asn Asn
340
<210> 519

<211> 343

<212> PRT

<213> Artificial sequence

<220><223> Anaerostipes caccae KARI variant-2012_H1

<400> 519

Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn

1 5 10 15

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

20 25 30

Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val

35 40 45

Ile Ile Gly Leu Phe Glu Gly Ala Glu Glu Trp Lys Arg Ala Glu Glu

50 55 60

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

65 70 75 80

Ile Ile Met Ile Leu Ile Pro Asp Glu Thr Gln Ala Thr Met Tyr Lys

85 90 95

Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala

100 105 110

His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val

115 120 125

Asp Val Thr Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser

130 135 140

Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln

145 150 155 160

Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala

165 170 175

Ile Gly Gly Ser Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu

180 185 190

Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val

195 200 205

Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr

210 215 220

Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile

225 230 235 240

Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile

245 250 255

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

260 265 270

Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln

275 280 285

Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly

290 295 300

Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro

305 310 315 320

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

325 330 335

Glu Asp Lys Leu Ile Asn Asn

340

<210> 520

<211> 343

<212> PRT

<213> Artificial sequence

<220><223> Anaerostipes caccae KARI variant-2012_H3

<400> 520

Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn

1 5 10 15

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

20 25 30

Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val

35 40 45

Ile Ile Gly Leu Phe Glu Gly Ala Glu Glu Cys Lys Arg Ala Glu Glu

50 55 60

Gln	Gly	Phe	Glu	Val	Tyr	Thr	Ala	Ala	Glu	Ala	Ala	Lys	Lys	Ala	Asp				
65					70					75					80				
Ile	Ile	Met	Ile	Leu	Ile	Pro	Asp	Glu	Lys	Gln	Ala	Thr	Met	Tyr	Lys				
85					90					95									
Asn	Asp	Ile	Glu	Pro	Asn	Leu	Glu	Ala	Gly	Asn	Met	Leu	Met	Phe	Ala				
100					105					110									
His	Gly	Phe	Asn	Ile	His	Phe	Gly	Cys	Ile	Val	Pro	Pro	Lys	Asp	Val				
115					120					125									
Asp	Val	Thr	Met	Ile	Ala	Pro	Lys	Gly	Pro	Gly	His	Thr	Val	Arg	Ser				
130					135					140									
Glu	Tyr	Glu	Glu	Gly	Lys	Gly	Val	Pro	Cys	Leu	Val	Ala	Val	Glu	Gln				
145					150					155					160				
Asp	Ala	Thr	Gly	Lys	Ala	Leu	Asp	Met	Ala	Leu	Ala	Tyr	Ala	Leu	Ala				
165					170					175									
Ile	Gly	Gly	Ala	Arg	Ala	Gly	Val	Leu	Glu	Thr	Thr	Phe	Arg	Thr	Glu				
180					185					190									
Thr	Glu	Thr	Asp	Leu	Phe	Gly	Glu	Gln	Ala	Val	Leu	Cys	Gly	Gly	Val				
195					200					205									
Cys	Ala	Leu	Met	Gln	Ala	Gly	Phe	Glu	Thr	Leu	Val	Glu	Ala	Gly	Tyr				
210					215					220									
Asp	Pro	Arg	Asn	Ala	Tyr	Phe	Glu	Cys	Ile	His	Glu	Met	Lys	Leu	Ile				
225					230					235					240				
Val	Asp	Leu	Ile	Tyr	Gln	Ser	Gly	Phe	Ser	Gly	Met	Arg	Tyr	Ser	Ile				
245					250					255									
Ser	Asn	Thr	Ala	Glu	Tyr	Gly	Asp	Tyr	Ile	Thr	Gly	Pro	Lys	Ile	Ile				
260					265					270									
Thr	Glu	Asp	Thr	Lys	Lys	Ala	Met	Lys	Lys	Ile	Leu	Ser	Asp	Ile	Gln				
275					280					285									
Asp	Gly	Thr	Phe	Ala	Lys	Asp	Phe	Leu	Val	Asp	Met	Ser	Asp	Ala	Gly				
290					295					300									
Ser	Gln	Val	His	Phe	Lys	Ala	Met	Arg	Lys	Leu	Ala	Ser	Glu	His	Pro				

305 310 315 320
Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
 325 330 335
Glu Asp Lys Leu Ile Asn Asn
 340
<210> 521
<211> 343
<212> PRT
<213> Artificial sequence
<220><223> Anaerostipes caccae KARI variant-2012_H7
<400> 521
Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn
1 5 10 15

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

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

Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln
145 150 155 160
Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala

165 170 175
 Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu
 180 185 190
 Thr Glu Thr Asp Leu Phe Gly Glu Gln Thr Val Leu Cys Gly Gly Val
 195 200 205

 Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr
 210 215 220
 Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile
 225 230 235 240
 Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile
 245 250 255
 Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile
 260 265 270

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

 Glu Asp Lys Leu Ile Asn Asn
 340

 <210> 522
 <211> 343
 <212> PRT
 <213> Artificial sequence
 <220><223> Anaerostipes caccae KARI variant
 <400> 522

 Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn
 1 5 10 15
 Leu Ser Leu Leu Asp Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser

20 25 30
 Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val

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

 100 105 110
 His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val
 115 120 125
 Asp Val Thr Met Ile Ala Pro Lys Gly Pro Gly His Thr Ile Arg Ser
 130 135 140
 Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln
 145 150 155 160
 Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala

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

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

Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln
275 280 285

Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly
290 295 300

Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro
305 310 315 320

Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
325 330 335

Glu Asp Lys Leu Ile Asn Asn
340

<210> 523
<211> 343
<212> PRT
<213> Artificial sequence
<220><223> Anaerostipes caccae KARI variant-2012_H11

<400> 523

Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn
1 5 10 15

Leu Ser Leu Leu Asp Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser
20 25 30

Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val
35 40 45

Ile Ile Gly Leu Phe Glu Gly Ala Glu Glu Trp Lys Arg Ala Glu Glu
50 55 60

Gln Gly Phe Glu Val Tyr Thr Val Ala Glu Ala Ala Lys Lys Ala Asp
65 70 75 80

Ile Ile Met Ile Leu Ile Pro Asp Glu Lys Gln Ala Thr Met Tyr Lys
85 90 95

Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala
100 105 110

His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val

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

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

 Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile
 260 265 270
 Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln
 275 280 285
 Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly
 290 295 300
 Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro
 305 310 315 320

 Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
 325 330 335
 Glu Asp Lys Leu Ile Asn Asn
 340
 <210> 524
 <211> 343

<212> PRT

<213> Artificial sequence

<220><223> Anaerostipes caccae KARI variant-2013_A2

<400> 524

Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn

1 5 10 15

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

20 25 30

Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val

35 40 45

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

50 55 60

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

65 70 75 80

Ile Ile Met Ile Leu Ile Pro Asp Glu Lys Gln Ala Thr Met Tyr Lys

85 90 95

Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala

100 105 110

His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val

115 120 125

Asp Val Thr Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser

130 135 140

Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln

145 150 155 160

Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala

165 170 175

Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu

180 185 190

Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val

195 200 205

Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr

210 215 220
 Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile
 225 230 235 240
 Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile
 245 250 255
 Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile
 260 265 270
 Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile His

275 280 285
 Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly
 290 295 300
 Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro
 305 310 315 320
 Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
 325 330 335
 Glu Asp Lys Leu Ile Asn Asn
 340

<210> 525

<211> 343

<212> PRT

<213> Artificial sequence

<220><223> Anaerostipes caccae KARI variant 2013_A4

<400> 525

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

65	70	75	80
Ile Ile Met Ile Met Ile Pro Asp Glu Lys Gln Ala Thr Met Tyr Lys			
	85	90	95
Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala			
	100	105	110
His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val			
	115	120	125
Asp Val Thr Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser			
	130	135	140
Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln			
	145	150	155
Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala			
	165	170	175
Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu			
	180	185	190
Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val			
	195	200	205
Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr			
	210	215	220
Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile			
	225	230	235
			240
Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile			
	245	250	255
Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile			
	260	265	270
Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln			
	275	280	285
Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly			
	290	295	300
Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro			
	305	310	315
			320

Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
325 330 335

Glu Asp Lys Leu Ile Asn Asn
340

<210> 526

<211> 343

<212> PRT

<213> Artificial sequence

<220><223> Anaerostipes caccae KARI variant-2013_A5

<400> 526

Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn

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

65 70 75 80
Ile Ile Met Ile Leu Ile Pro Asp Glu Lys Gln Ala Thr Met Tyr Lys
85 90 95
Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala
100 105 110
His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val
115 120 125
Asp Val Thr Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser

130 135 140
Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln
145 150 155 160
Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala
165 170 175

Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu
180 185 190
Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val
195 200 205
Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr
210 215 220
Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile
225 230 235 240
Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile
245 250 255
Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile
260 265 270
Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln
275 280 285
Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly
290 295 300
Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro
305 310 315 320
Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
325 330 335
Glu Asp Lys Leu Ile Asn Asn
340
<210> 527
<211> 343
<212> PRT
<213> Artificial sequence
<220><223> Anaerostipes caccae KARI variant-2013_B2
<400> 527
Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn
1 5 10 15
Leu Ser Leu Leu Asp Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser

20 25 30
 Gln Gly His Thr His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val
 35 40 45
 Ile Ile Gly Leu Phe Glu Gly Ala Glu Glu Trp Lys Arg Ala Glu Glu
 50 55 60
 Gln Gly Phe Glu Val Tyr Thr Ala Ala Glu Ala Ala Lys Lys Ala Asp
 65 70 75 80
 Ile Ile Met Ile Leu Ile Pro Asp Glu Lys Gln Ala Thr Met Tyr Lys
 85 90 95
 Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala
 100 105 110
 His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val
 115 120 125
 Asp Val Thr Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser
 130 135 140
 Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln
 145 150 155 160
 Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala
 165 170 175
 Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu
 180 185 190
 Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Ile Leu Cys Gly Gly Val
 195 200 205
 Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr
 210 215 220
 Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile
 225 230 235 240
 Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile
 245 250 255
 Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile
 260 265 270

Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln
275 280 285

Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly
290 295 300

Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro
305 310 315 320

Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
325 330 335

Glu Asp Lys Leu Ile Asn Asn
340

<210> 528

<211> 343

<212> PRT

<213> Artificial sequence

<220><223> Anaerostipes caccae KARI variant-2013_B5

<400> 528

Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn
1 5 10 15

Leu Ser Leu Leu Asp Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser
20 25 30

Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val
35 40 45

Ile Ile Gly Leu Phe Glu Gly Ala Glu Glu Trp Lys Arg Ala Glu Glu

50 55 60

Gln Gly Phe Glu Val Tyr Thr Ala Ala Glu Ala Ala Lys Lys Ala Asp
65 70 75 80

Ile Ile Met Ile Leu Ile Pro Asp Glu Lys Gln Ala Thr Met Tyr Lys
85 90 95

Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala
100 105 110

His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val

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

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

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

 305 310 315 320
 Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
 325 330 335
 Glu Asp Lys Leu Ile Asn Asn
 340
 <210> 529
 <211> 343
 <212> PRT

<213> Artificial sequence

<220><223> Anaerostipes caccae KARI variant-2013_B7

<400> 529

Met Glu Glu Cys Lys Met Ala Lys Met Tyr Tyr Gln Glu Asp Cys Asn

1 5 10 15

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

20 25 30

Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val

35 40 45

Ile Ile Gly Leu Phe Glu Gly Ala Glu Glu Trp Lys Arg Ala Glu Glu

50 55 60

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

65 70 75 80

Ile Ile Met Ile Leu Ile Pro Asp Glu Glu Gln Ala Thr Met Tyr Lys

85 90 95

Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala

100 105 110

His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val

115 120 125

Asp Val Thr Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser

130 135 140

Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln

145 150 155 160

Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala

165 170 175

Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu

180 185 190

Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val

195 200 205

Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr

210 215 220

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

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

 Glu Asp Lys Leu Ile Asn Asn
 340

 <210> 530
 <211> 343
 <212> PRT
 <213> Artificial sequence
 <220><223> Anaerostipes caccae KARI variant-2013_B8
 <400> 530

 Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn
 1 5 10 15
 Leu Ser Leu Leu Asp Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser
 20 25 30
 Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val

 35 40 45
 Ile Ile Gly Leu Phe Glu Cys Ala Glu Glu Trp Lys Arg Ala Glu Glu
 50 55 60
 Gln Gly Phe Glu Val Tyr Thr Ala Ala Glu Ala Ala Lys Lys Ala Asp
 65 70 75 80

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

325 330 335

Glu Asp Lys Leu Ile Asn Asn

340

<210> 531

<211> 343

<212> PRT

<213> Artificial sequence

<220><223> Anaerostipes caccae KARI variant-2013_B9

<400> 531

Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn

1 5 10 15

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

20 25 30

Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val

35 40 45

Ile Ile Gly Leu Phe Glu Gly Val Glu Glu Trp Lys Arg Ala Glu Glu

50 55 60

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

65 70 75 80

Ile Ile Met Ile Leu Ile Pro Asp Glu Lys Gln Ala Thr Met Tyr Lys

85 90 95

Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala

100 105 110

His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val

115 120 125

Asp Val Thr Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser

130 135 140

Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln

145 150 155 160

Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala

165 170 175

Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu

180 185 190
 Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val
 195 200 205
 Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr
 210 215 220
 Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile
 225 230 235 240
 Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile
 245 250 255

 Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile
 260 265 270
 Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln
 275 280 285
 Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly
 290 295 300
 Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro
 305 310 315 320

 Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
 325 330 335
 Glu Asp Lys Leu Ile Asn Asn
 340
 <210> 532
 <211> 343
 <212> PRT
 <213> Artificial sequence
 <220><223> Anaerostipes caccae KARI variant-2013_B11
 <400> 532
 Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn
 1 5 10 15
 Leu Ser Leu Leu Asp Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser

20 25 30

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

275 280 285
 Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly
 290 295 300
 Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro
 305 310 315 320
 Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
 325 330 335
 Glu Asp Lys Leu Ile Asn Asn
 340

<210> 533

<211> 343

<212> PRT

<213> Artificial sequence

<220><223> Anaerostipes caccae KARI variant-2013_C1

<400> 533

Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn
 1 5 10 15
 Leu Ser Leu Leu Asp Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser
 20 25 30
 Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val
 35 40 45

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

His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val
 115 120 125
 Asp Val Thr Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser

130 135 140
 Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln
 145 150 155 160
 Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala
 165 170 175

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

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

 Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro
 305 310 315 320
 Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
 325 330 335
 Glu Asp Lys Leu Ile Asn Asn
 340

 <210> 534
 <211> 343
 <212> PRT
 <213> Artificial sequence
 <220><223> Anaerostipes caccae KARI variant-2013_C6

<400> 534

Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn

1 5 10 15

Leu Ser Leu Leu Asp Gly Met Thr Ile Ala Val Ile Gly Tyr Gly Ser

20 25 30

Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val

35 40 45

Ile Ile Gly Leu Phe Glu Gly Ala Glu Glu Trp Lys Arg Ala Glu Glu

50 55 60

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

65 70 75 80

Ile Ile Met Ile Leu Ile Pro Asp Glu Lys Gln Ala Thr Met Tyr Lys

85 90 95

Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala

100 105 110

His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val

115 120 125

Asp Val Thr Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser

130 135 140

Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln

145 150 155 160

Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala

165 170 175

Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu

180 185 190

Thr Asp Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val

195 200 205

Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr

210 215 220

Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile

225 230 235 240

Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile
245 250 255

Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile
260 265 270

Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln
275 280 285

Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly
290 295 300

Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro
305 310 315 320

Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
325 330 335

Glu Asp Lys Leu Ile Asn Asn
340

<210> 535
<211> 343
<212> PRT
<213> Artificial sequence
<220><223> Anaerostipes caccae KARI variant-2013_C8
<400> 535

Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn
1 5 10 15

Leu Ser Leu Leu Asp Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser
20 25 30

Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val
35 40 45

Ile Ile Gly Leu Phe Glu Gly Ala Glu Glu Trp Lys Arg Ala Glu Glu
50 55 60

Gln Gly Phe Glu Val Tyr Thr Ala Ala Glu Ala Ala Lys Lys Ala Asp
65 70 75 80

Ile Ile Met Ile Leu Ile Pro Asp Glu Lys Gln Ala Thr Met Tyr Lys

85

90

95

Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala

100

105

110

His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val

115

120

125

Asp Val Thr Met Ile Ala Thr Lys Gly Pro Gly His Thr Val Arg Ser

130

135

140

Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln

145

150

155

160

Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala

165

170

175

Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu

180

185

190

Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val

195

200

205

Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr

210

215

220

Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile

225

230

235

240

Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile

245

250

255

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

260

265

270

Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln

275

280

285

Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly

290

295

300

Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro

305

310

315

320

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

325

330

335

Glu Asp Lys Leu Ile Asn Asn

340

<210> 536

<211> 343

<212> PRT

<213> Artificial sequence

<220><223> Anaerostipes caccae KARI variant-2013_C12

<400> 536

Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn

1 5 10 15

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

20 25 30

Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val

35 40 45

Ile Ile Gly Leu Phe Glu Gly Ala Glu Glu Trp Lys Arg Ala Glu Glu

50 55 60

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

65 70 75 80

Ile Ile Met Ile Leu Ile Pro Asp Glu Lys Gln Ala Ser Met Tyr Lys

85 90 95

Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala

100 105 110

His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val

115 120 125

Asp Val Thr Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser

130 135 140

Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln

145 150 155 160

Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Val

165 170 175

Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu

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

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

 305 310 315 320
 Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
 325 330 335
 Glu Asp Lys Leu Ile Asn Asn
 340
 <210> 537
 <211> 343
 <212> PRT
 <213> Artificial sequence
 <220><223> Anaerostipes caccae KARI variant-2013_D1
 <400> 537
 Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn
 1 5 10 15

 Leu Ser Leu Leu Asp Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser
 20 25 30
 Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val

35	40	45	
Ile Ile Gly Leu Phe Glu Gly Ala Glu Glu Trp Lys Arg Ala Glu Glu			
50	55	60	
Gln Gly Phe Glu Val Tyr Thr Ala Ala Glu Ala Ala Lys Lys Ala Asp			
65	70	75	80
Ile Ile Met Ile Leu Ile Pro Asp Glu Lys Gln Ala Thr Arg Tyr Lys			
85	90	95	
Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala			
100	105	110	
His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val			
115	120	125	
Asp Val Thr Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser			
130	135	140	
Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln			
145	150	155	160
Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala			
165	170	175	
Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu			
180	185	190	
Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val			
195	200	205	
Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr			
210	215	220	
Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile			
225	230	235	240
Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile			
245	250	255	
Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile			
260	265	270	
Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln			
275	280	285	

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

Glu Asp Lys Leu Ile Asn Asn
 340

<210> 538

<211> 636

<212> DNA

<213> Artificial sequence

<220><223> Promoter UAS(PGK1)-ENO2p

<400> 538

aattaccgtc gctcgtgatt tgtttgcaaa aagaacaaaa ctgaaaaaac ccagacacgc 60
 tcgacttctt gtcttcttat tgattgcagc ttccaatttc gtcacacaac aaggtcctgc 120
 ggcgttatgt cactaacgac gtgcaccaac ttgcggaaag tggaaatccg ttccaaaact 180
 ggcatccact aattgataca tctacacacc gcacgccttt tttctgaagc ccactttcgt 240
 ggactttgcc atatgcaaaa ttcatagaat gtgataccaa gtcagcatal acctcactag 300

ggtagtttct ttggttgtat tgatcatttg gttcatcgtg gttcattaat tttttttctc 360
 cattgctttc tggctttgat cttactatca ttgggatttt tgcgaaggt ttagaattg 420
 tatgtgacaa gtggcaccaa gcatatataa aaaaaaaaaag cattatcttc ctaccagagt 480
 tgattgttaa aaacgtatct atagcaaagc caattgtaat taattcttat ttgtatctt 540
 ttcttccctt gtctcaatct tttattttta tttatttttt cttttcttag tttctttcat 600
 aacaccaagc aactaatact ataacataca ataata 636

<210> 539

<211> 979

<212> DNA

<213> Artificial sequence

<220><223> UAS(FBA1)-PDC1p

<400> 539

tgacagcagg attatcgtaa tacgtaatag ttgaaaatct caaaaatgtg tgggtcatta 60

cgtaaataat gataggaatg ggattcttct atttttcctt tttccattct cgcacgccga 120
aatgcatgca agtaacctat tcaaagtaat atctcataca tgtttcatga gggtacaac 180
atgcgactgg gtgagcatat gttccgctga tgtgatgtgc aagataaaca agcaaggcag 240
aaactaactt ctcttcatg taataaacac accccgcgtt tatttaccta tctctaaact 300
tcaacacctt atatcataac taatatttct tgagataagc acactgcacc cataccttcc 360
ttaaaaacgt agcttccagt ttttgggtgt tccggcttcc ttcctcgattc cgcccgttaa 420

acgcatattt ttgttgcctg gtggcatttg caaaatgcat aacctatgca tttaaaagat 480
tatgtatgct ctctgactt ttcgtgtgat gaggtcgtg gaaaaaatga ataatttatg 540
aatttgagaa caattttgtg ttgttacggt attttactat ggaataatca atcaattgag 600
gattttatgc aaatatcgtt tgaatatttt tccgaccctt tgagtacttt tcttcataat 660
tgcataatat tgcgctgc ccttttttct gttagacggt gtcttgatct acttgctatc 720
gttcaacacc acctatttt ctaactattt tttttttagc tcatttgaat cagcttatgg 780
tgatggcaca tttttgcata aacctagctg tcctcgttga acataggaaa aaaaaatata 840

taaacaaggc tctttcactc tccttgcaat cagatttggg tttgttcctt ttattttcat 900
atttctgtc atattccttt ctcaattatt attttctact cataacctca cgcaaaataa 960
cacagtcaaa tcaatcaaa 979

<210> 540

<211> 994

<212> DNA

<213> Artificial sequence

<220><223> Promoter UAS(PGK1)-PDC1p

<400> 540

aattaccgtc gctcgtgatt tgtttgcaaa aagaacaaaa ctgaaaaaac ccagacacgc 60
tcgacttctt gtcttcttat tgattgcagc ttccaatttc gtcacacaac aaggctcctgt 120
cgaccgcgac gccgaaatgc atgcaagtaa cctattcaaa gtaatatctc atacatgttt 180

catgagggtg acaacatgcg actgggtgag catatgttcc gctgatgtga tgtgcaagat 240
aaacaagcaa ggcagaaact aacttcttct tcatgtaata aacacacccc gcgtttattt 300
acctatctct aaacttcaac acctatatac ataactaata tttcttgaga taagcacact 360
gcaccatac cttccttaaa aacgtagctt ccagtttttg gtggttccgg cttccttccc 420
gattccgccc gctaaacgca tatttttgtt gcctgggtggc atttgcaaaa tgcataacct 480
atgcatttaa aagattatgt atgctcttct gacttttctg gtgatgaggc tcgtggaaaa 540

aatgaataat ttatgaattt gagaacaatt ttgtgttggt acggtatttt actatggaat 600

aatcaatcaa ttgaggattt tatgcaaata tcgtttgaat atttttccga ccctttgagt 660

actttttctt ataattgcat aatattgtcc gctgcccctt tttctgttag acggtgtctt 720

gatctacttg ctatcgttca acaccacctt attttctaac ttttttttt ttagctcatt 780

tgaatcagct tatggatgat gcacattttt gcataaacct agctgtcctc gttgaacata 840

ggaaaaaaaa atatataaac aaggtctttt cactctcctt gcaatcagat ttgggtttgt 900

tccctttatt ttcatatttc ttgtcatatt cctttctcaa ttattatttt ctactcataa 960

cctcacgcaa aataacacag tcaaatcaat caaa 994

<210> 541

<211> 731

<212> DNA

<213> Artificial sequence

<220><223> Promoter UAS(PGK)-OLE1p

<400> 541

aattaccgtc gctcgtgatt tgtttgcaaa aagaacaaaa ctgaaaaaac ccagacacgc 60

tcgacttcct gtcttcctat tgattgcagc ttccaatttc gtcacacaac aaggtcctgt 120

cgaccggggc gcccggtcaa aaactaactc cgagcccggt catgtcccgg gttagcgggc 180

ccaacaaagg cgcttatctg gtgggcttcc gtagaagaaa aaaagctgtt gagcgagcta 240

tttcgggtat ccagccttc tctgcagacc gccccagtgt gcttggtctt ggtgctgttc 300

gtagcatca catcgctgt gacaggcaga ggtaataacg gcttaagggt ctcttcgcat 360

agtcggcagc tttctttcgg acgttgaaca ctcaacaaac cttatctagt gccaaccag 420

gtgtgcttct acgagtcttg ctactcaga cacacctatc cctatgtta cggctatggg 480

gatggcacac aaaggtggaa ataatagtag ttaacaatat atgcagcaaa tcatcggtc 540

ctggctcctc gagtcttgca aatcagcata tacatatata tatgggggca gatcttgatt 600

catttattgt tctatttcca tctttcctac ttctgtttcc gtttatattt tgtattacgt 660

agaatagaac atcatagtaa tagatagttg tggatgatcat attataaaca gcactaaaac 720

attacaacaa a 731

<210> 542

<211> 546

<212> PRT

<213> *Macrococcus caseolyticus*

<400> 542

```

Met Lys Gln Arg Ile Gly Gln Tyr Leu Ile Asp Ala Leu His Val Asn
1           5           10           15
Gly Val Asp Lys Ile Phe Gly Val Pro Gly Asp Phe Thr Leu Ala Phe
           20           25           30
Leu Asp Asp Ile Ile Arg His Asp Asn Val Glu Trp Val Gly Asn Thr
           35           40           45
Asn Glu Leu Asn Ala Ala Tyr Ala Ala Asp Gly Tyr Ala Arg Val Asn

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

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

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

```

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

245 250 255

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

260 265 270

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

275 280 285

Gly Phe Ser Tyr Lys Phe Asp Thr Asn Asn Ile Ile Tyr Ile Asn His

290 295 300

Asn Asp Phe Lys Ala Glu Asp Val Ile Ser Asp Asn Val Ser Leu Ile

305 310 315 320

Asp Leu Val Asn Gly Leu Asn Ser Ile Asp Tyr Arg Asn Glu Thr His

325 330 335

Tyr Pro Ser Tyr Gln Arg Ser Asp Met Lys Tyr Glu Leu Asn Asp Ala

340 345 350

Pro Leu Thr Gln Ser Asn Tyr Phe Lys Met Met Asn Ala Phe Leu Glu

355 360 365

Lys Asp Asp Ile Leu Leu Ala Glu Gln Gly Thr Ser Phe Phe Gly Ala

370 375 380

Tyr Asp Leu Ser Leu Tyr Lys Gly Asn Gln Phe Ile Gly Gln Pro Leu

385 390 395 400

Trp Gly Ser Ile Gly Tyr Thr Phe Pro Ser Leu Leu Gly Ser Gln Leu

405 410 415

Ala Asp Met His Arg Arg Asn Ile Leu Leu Ile Gly Asp Gly Ser Leu

420 425 430

Gln Leu Thr Val Gln Ala Leu Ser Thr Met Ile Arg Lys Asp Ile Lys

435 440 445

Pro Ile Ile Phe Val Ile Asn Asn Asp Gly Tyr Thr Val Glu Arg Leu

450 455 460

Ile His Gly Met Glu Glu Pro Tyr Asn Asp Ile Gln Met Trp Asn Tyr

465 470 475 480

Lys Gln Leu Pro Glu Val Phe Gly Gly Lys Asp Thr Val Lys Val His

485 490 495
Asp Ala Lys Thr Ser Asn Glu Leu Lys Thr Val Met Asp Ser Val Lys

500 505 510
Ala Asp Lys Asp His Met His Phe Ile Glu Val His Met Ala Val Glu

515 520 525
Asp Ala Pro Lys Lys Leu Ile Asp Ile Ala Lys Ala Phe Ser Asp Ala

530 535 540
Asn Lys

545

<210> 543

<211> 548

<212> PRT

<213> *Listeria grayi*

<400> 543

Met Tyr Thr Val Gly Gln Tyr Leu Val Asp Arg Leu Glu Glu Ile Gly
1 5 10 15

Ile Asp Lys Val Phe Gly Val Pro Gly Asp Tyr Asn Leu Thr Phe Leu
20 25 30

Asp Tyr Ile Gln Asn His Glu Gly Leu Ser Trp Gln Gly Asn Thr Asn
35 40 45

Glu Leu Asn Ala Ala Tyr Ala Ala Asp Gly Tyr Ala Arg Glu Arg Gly
50 55 60

Val Ser Ala Leu Val Thr Thr Phe Gly Val Gly Glu Leu Ser Ala Ile
65 70 75 80

Asn Gly Thr Ala Gly Ser Phe Ala Glu Gln Val Pro Val Ile His Ile
85 90 95

Val Gly Ser Pro Thr Met Asn Val Gln Ser Asn Lys Lys Leu Val His
100 105 110

His Ser Leu Gly Met Gly Asn Phe His Asn Phe Ser Glu Met Ala Lys
115 120 125

Glu Val Thr Ala Ala Thr Thr Met Leu Thr Glu Glu Asn Ala Ala Ser

130 135 140
 Glu Ile Asp Arg Val Leu Glu Thr Ala Leu Leu Glu Lys Arg Pro Val
 145 150 155 160
 Tyr Ile Asn Leu Pro Ile Asp Ile Ala His Lys Ala Ile Val Lys Pro
 165 170 175
 Ala Lys Ala Leu Gln Thr Glu Lys Ser Ser Gly Glu Arg Glu Ala Gln
 180 185 190
 Leu Ala Glu Ile Ile Leu Ser His Leu Glu Lys Ala Ala Gln Pro Ile
 195 200 205

 Val Ile Ala Gly His Glu Ile Ala Arg Phe Gln Ile Arg Glu Arg Phe
 210 215 220
 Glu Asn Trp Ile Asn Gln Thr Lys Leu Pro Val Thr Asn Leu Ala Tyr
 225 230 235 240
 Gly Lys Gly Ser Phe Asn Glu Glu Asn Glu His Phe Ile Gly Thr Tyr
 245 250 255
 Tyr Pro Ala Phe Ser Asp Lys Asn Val Leu Asp Tyr Val Asp Asn Ser
 260 265 270

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

 Thr Ala Ala Lys Gln Ala Glu Leu Ala Val Phe Glu Pro Gln Ala Glu
 340 345 350
 Thr Pro Leu Lys Gln Asp Arg Phe His Gln Ala Val Met Asn Phe Leu
 355 360 365
 Gln Ala Asp Asp Val Leu Val Thr Glu Gln Gly Thr Ser Ser Phe Gly
 370 375 380

Leu Met Leu Ala Pro Leu Lys Lys Gly Met Asn Leu Ile Ser Gln Thr
385 390 395 400

Leu Trp Gly Ser Ile Gly Tyr Thr Leu Pro Ala Met Ile Gly Ser Gln
405 410 415

Ile Ala Ala Pro Glu Arg Arg His Ile Leu Ser Ile Gly Asp Gly Ser
420 425 430

Phe Gln Leu Thr Ala Gln Glu Met Ser Thr Ile Phe Arg Glu Lys Leu
435 440 445

Thr Pro Val Ile Phe Ile Ile Asn Asn Asp Gly Tyr Thr Val Glu Arg
450 455 460

Ala Ile His Gly Glu Asp Glu Ser Tyr Asn Asp Ile Pro Thr Trp Asn
465 470 475 480

Leu Gln Leu Val Ala Glu Thr Phe Gly Gly Asp Ala Glu Thr Val Asp
485 490 495

Thr His Asn Val Phe Thr Glu Thr Asp Phe Ala Asn Thr Leu Ala Ala
500 505 510

Ile Asp Ala Thr Pro Gln Lys Ala His Val Val Glu Val His Met Glu
515 520 525

Gln Met Asp Met Pro Glu Ser Leu Arg Gln Ile Gly Leu Ala Leu Ser
530 535 540

Lys Gln Asn Ser

545

<210> 544

<211> 24

<212> DNA

<213> Artificial sequence

<220><223> Primer K9_53F_031411f

<400> 544

gttatcatcg gattattcga agga

24

<210> 545

<211> 24

<212> DNA

<213> Artificial sequence

<220><223> Primer K9_53L_031411f

<400> 545

gttatcatcg gattattgga agga

24

<210> 546

<211> 24

<212> DNA

<213> Artificial sequence

<220><223> Primer K9_53Y_031411f

<400> 546

gttatcatcg gattatatga agga

24

<210> 547

<211> 24

<212> DNA

<213> Artificial sequence

<220><223> Primer K9_53C_031411f

<400> 547

gttatcatcg gattatgtga agga

24

<210> 548

<211> 24

<212> DNA

<213> Artificial sequence

<220><223> Primer K9_53W_031411f

<400> 548

gttatcatcg gattatggga agga

24

<210> 549

<211> 24

<212> DNA

<213> Artificial sequence

<220><223> Primer K9_53P_031411f

<400> 549

gttatcatcg gattaccaga agga

24

<210> 550

<211> 24
 <212> DNA
 <213> Artificial sequence
 <220><223> Primer K9_53H_031411f
 <400> 550
 gttatcatcg gattacatga agga 24
 <210> 551
 <211> 24
 <212> DNA
 <213> Artificial sequence
 <220><223> Primer K9_53Q_031411f
 <400> 551
 gttatcatcg gattacaaga agga 24

 <210> 552
 <211> 24
 <212> DNA
 <213> Artificial sequence
 <220><223> Primer K9_53I_031411f
 <400> 552
 gttatcatcg gattaattga agga 24
 <210> 553
 <211> 24
 <212> DNA
 <213> Artificial sequence
 <220><223> Primer K9_53M_031411f
 <400> 553
 gttatcatcg gattaatgga agga 24
 <210> 554
 <211> 24
 <212> DNA
 <213> Artificial sequence
 <220><223> Primer K9_53N_031411f
 <400> 554

gttatcatcg gattaaatga agga	24
<210> 555	
<211> 24	
<212> DNA	
<213> Artificial sequence	
<220><223> Primer K9_53V_031411f	
<400> 555	
gttatcatcg gattagtga agga	24
<210> 556	
<211> 24	
<212> DNA	
<213> Artificial sequence	
<220><223> Primer K9_53A_031411f	
<400> 556	
gttatcatcg gattagctga agga	24
<210> 557	
<211> 24	
<212> DNA	
<213> Artificial sequence	
<220><223> Primer K9_53D_031411f	
<400> 557	
gttatcatcg gattagatga agga	24
<210> 558	
<211> 24	
<212> DNA	
<213> Artificial sequence	
<220><223> Primer K9_53E_031411f	
<400> 558	
gttatcatcg gattagaaga agga	24
<210> 559	
<211> 24	
<212> DNA	
<213> Artificial sequence	

<220><223> Primer K9_53G_031411f

<400> 559

gttatcatcg gattaggtga agga

24

<210> 560

<211> 21

<212> DNA

<213> Artificial sequence

<220><223> Primer K9_56F_031411f

<400> 560

ggattacctg aaggattcaa a

21

<210> 561

<211> 21

<212> DNA

<213> Artificial sequence

<220><223> Primer K9_56L_031411f

<400> 561

ggattacctg aaggattgaa a

21

<210> 562

<211> 21

<212> DNA

<213> Artificial sequence

<220><223> Primer K9_56Y_031411f

<400> 562

ggattacctg aaggatataa a

21

<210> 563

<211> 21

<212> DNA

<213> Artificial sequence

<220><223> Primer K9_56C_031411f

<400> 563

ggattacctg aaggatgtaa a

21

<210> 564

<211> 21

<212> DNA
 <213> Artificial sequence
 <220><223> Primer K9_56W_031411f
 <400> 564
 ggattacctg aaggatggaa a 21
 <210> 565
 <211> 21
 <212> DNA
 <213> Artificial sequence
 <220><223> Primer K9_56P_031411f
 <400> 565
 ggattacctg aaggacaaaa a 21
 <210> 566
 <211> 21
 <212> DNA
 <213> Artificial sequence
 <220><223> Primer K9_56H_031411f
 <400> 566
 ggattacctg aaggacataa a 21

 <210> 567
 <211> 21
 <212> DNA
 <213> Artificial sequence
 <220><223> Primer K9_56Q_031411f
 <400> 567
 ggattacctg aaggacaaaa a 21
 <210> 568
 <211> 21
 <212> DNA
 <213> Artificial sequence
 <220><223> Primer K9_56I_031411f
 <400> 568
 ggattacctg aaggaattaa a 21
 <210> 569

<211> 21
 <212> DNA
 <213> Artificial sequence
 <220><223> Primer K9_56M_031411f
 <400> 569
 ggattacctg aaggaatgaa a 21

<210> 570
 <211> 21
 <212> DNA
 <213> Artificial sequence
 <220><223> Primer K9_56N_031411f
 <400> 570
 ggattacctg aaggaaataa a 21

<210> 571
 <211> 21
 <212> DNA
 <213> Artificial sequence
 <220><223> Primer K9_56V_031411f
 <400> 571
 ggattacctg aaggagttaa a 21

<210> 572
 <211> 21
 <212> DNA
 <213> Artificial sequence
 <220><223> Primer K9_56A_031411f
 <400> 572
 ggattacctg aaggagctaa a 21

<210> 573
 <211> 21
 <212> DNA
 <213> Artificial sequence
 <220><223> Primer K9_56D_031411f
 <400> 573

ggattacctg aaggagataa a 21

<210> 574

<211> 21

<212> DNA

<213> Artificial sequence

<220><223> Primer K9_56E_031411f

<400> 574

ggattacctg aaggagaaaa a 21

<210> 575

<211> 21

<212> DNA

<213> Artificial sequence

<220><223> Primer K9_56G_031411f

<400> 575

ggattacctg aaggaggtaa a 21

<210> 576

<211> 34

<212> DNA

<213> Artificial sequence

<220><223> Primer K9_58F_051611f

<400> 576

gattacctga aggatttaaa ttctggaaga gaggc 34

<210> 577

<211> 34

<212> DNA

<213> Artificial sequence

<220><223> Primer K9_58L_051611f

<400> 577

gattacctga aggatttaaa ttgtggaaga gaggc 34

<210> 578

<211> 34

<212> DNA

<213> Artificial sequence

<220><223> Primer K9_58Y_051611f

<400> 578
gattacctga aggatttaaa tattggaaga gagc 34

<210> 579
<211> 34
<212> DNA
<213> Artificial sequence
<220><223> Primer K9_58C_051611f
<400> 579
gattacctga aggatttaaa tgttggaaga gagc 34

<210> 580
<211> 34
<212> DNA
<213> Artificial sequence
<220><223> Primer K9_58W_051611f
<400> 580
gattacctga aggatttaaa tgggtggaaga gagc 34

<210> 581
<211> 34
<212> DNA
<213> Artificial sequence
<220><223> Primer K9_58P_051611f
<400> 581
gattacctga aggatttaaa ccatggaaga gagc 34

<210> 582
<211> 34
<212> DNA
<213> Artificial sequence
<220><223> Primer K9_58H_051611f
<400> 582
gattacctga aggatttaaa cattggaaga gagc 34

<210> 583
<211> 34
<212> DNA

<213> Artificial sequence

<220><223> Primer K9_58Q_051611f

<400> 583

gattacctga aggatttaaa caatggaaga gagc

34

<210> 584

<211> 34

<212> DNA

<213> Artificial sequence

<220><223> Primer K9_58I_051611f

<400> 584

gattacctga aggatttaaa atttggaaga gagc

34

<210> 585

<211> 34

<212> DNA

<213> Artificial sequence

<220><223> Primer K9_58M_051611f

<400> 585

gattacctga aggatttaaa atgtggaaga gagc

34

<210> 586

<211> 34

<212> DNA

<213> Artificial sequence

<220><223> Primer K9_58N_051611f

<400> 586

gattacctga aggatttaaa aattggaaga gagc

34

<210> 587

<211> 34

<212> DNA

<213> Artificial sequence

<220><223> Primer K9_58V_051611f

<400> 587

gattacctga aggatttaaa gtttggaaga gagc

34

<210> 588

<211> 34
 <212> DNA
 <213> Artificial sequence
 <220><223> Primer K9_58A_051611f
 <400> 588
 gattacctga aggatttaaa gcttggaaga gagc 34
 <210> 589
 <211> 34
 <212> DNA
 <213> Artificial sequence
 <220><223> Primer K9_58D_051611f
 <400> 589
 gattacctga aggatttaaa gattggaaga gagc 34
 <210> 590
 <211> 34
 <212> DNA
 <213> Artificial sequence
 <220><223> Primer K9_58E_051611f
 <400> 590
 gattacctga aggatttaaa gaatggaaga gagc 34
 <210> 591
 <211> 34
 <212> DNA
 <213> Artificial sequence
 <220><223> Primer K9_58G_051611f
 <400> 591
 gattacctga aggatttaaa ggttggaaga gagc 34
 <210> 592
 <211> 26
 <212> DNA
 <213> Artificial sequence
 <220><223> Primer K9_56F_071211r
 <400> 592
 gctctcttcc atggtttgaa tccttc 26

<210> 593
 <211> 26
 <212> DNA
 <213> Artificial sequence
 <220><223> Primer K9_56L_071211r
 <400> 593
 gctctcttcc atggtttcaa tccttc 26

<210> 594
 <211> 26
 <212> DNA
 <213> Artificial sequence
 <220><223> Primer K9_56Y_071211r
 <400> 594
 gctctcttcc atggtttata tccttc 26

<210> 595
 <211> 26
 <212> DNA
 <213> Artificial sequence
 <220><223> Primer K9_56C_071211r
 <400> 595
 gctctcttcc atggtttaca tccttc 26

<210> 596
 <211> 26
 <212> DNA
 <213> Artificial sequence
 <220><223> Primer K9_56W_071211r
 <400> 596
 gctctcttcc atggtttcca tccttc 26

<210> 597
 <211> 26
 <212> DNA
 <213> Artificial sequence
 <220><223> Primer K9_56P_071211r

<400> 597
gctctcttcc atggttttgg tccttc 26

<210> 598
<211> 26
<212> DNA
<213> Artificial sequence
<220><223> Primer K9_56H_071211r
<400> 598
gctctcttcc atggttttatg tccttc 26

<210> 599
<211> 26
<212> DNA
<213> Artificial sequence
<220><223> Primer K9_56Q_071211r
<400> 599
gctctcttcc atggtttttg tccttc 26

<210> 600
<211> 26
<212> DNA
<213> Artificial sequence
<220><223> Primer K9_56I_071211r
<400> 600
gctctcttcc atggttttaat tccttc 26

<210> 601
<211> 26
<212> DNA
<213> Artificial sequence
<220><223> Primer K9_56M_071211r
<400> 601
gctctcttcc atggtttcat tccttc 26

<210> 602
<211> 26
<212> DNA
<213> Artificial sequence

<220><223> Primer K9_56N_071211r

<400> 602

gctctcttcc atggtttatt tccttc

26

<210> 603

<211> 26

<212> DNA

<213> Artificial sequence

<220><223> Primer K9_56V_071211r

<400> 603

gctctcttcc atggtttaac tccttc

26

<210> 604

<211> 26

<212> DNA

<213> Artificial sequence

<220><223> Primer K9_56A_071211r

<400> 604

gctctcttcc atggttttagc tccttc

26

<210> 605

<211> 26

<212> DNA

<213> Artificial sequence

<220><223> Primer K9_56D_071211r

<400> 605

gctctcttcc atggtttatc tccttc

26

<210> 606

<211> 26

<212> DNA

<213> Artificial sequence

<220><223> Primer K9_56E_071211r

<400> 606

gctctcttcc atggtttttc tccttc

26

<210> 607

<211> 26

<212> DNA
 <213> Artificial sequence
 <220><223> Primer K9_56G_071211r
 <400> 607
 gctctcttcc atggtttacc tccttc 26
 <210> 608
 <211> 25
 <212> DNA
 <213> Artificial sequence
 <220><223> Primer K9_58F_071211r
 <400> 608
 gttcttctgc tctcttccag aattt 25
 <210> 609
 <211> 25
 <212> DNA
 <213> Artificial sequence
 <220><223> Primer K9_58L_071211r
 <400> 609
 gttcttctgc tctcttccac aattt 25
 <210> 610
 <211> 25
 <212> DNA
 <213> Artificial sequence
 <220><223> Primer K9_58Y_071211r
 <400> 610
 gttcttctgc tctcttccaa tatatt 25
 <210> 611
 <211> 25
 <212> DNA
 <213> Artificial sequence
 <220><223> Primer K9_58C_071211r
 <400> 611
 gttcttctgc tctcttccaa cattt 25

<210> 612
 <211> 25
 <212> DNA
 <213> Artificial sequence
 <220><223> Primer K9_58W_071211r
 <400> 612
 gttctttctgc tctcttccac cattt 25
 <210> 613
 <211> 25
 <212> DNA
 <213> Artificial sequence
 <220><223> Primer K9_58P_071211r
 <400> 613
 gttctttctgc tctcttccat ggttt 25
 <210> 614
 <211> 25
 <212> DNA
 <213> Artificial sequence
 <220><223> Primer K9_58H_071211r
 <400> 614
 gttctttctgc tctcttccaa tgttt 25
 <210> 615
 <211> 25
 <212> DNA
 <213> Artificial sequence
 <220><223> Primer K9_58Q_071211r
 <400> 615
 gttctttctgc tctcttccat tgttt 25
 <210> 616
 <211> 25
 <212> DNA
 <213> Artificial sequence
 <220><223> Primer K9_58I_071211r
 <400> 616

gtttcttctgc tctcttccaa atttt 25

<210> 617

<211> 25

<212> DNA

<213> Artificial sequence

<220><223> Primer K9_58M_071211r

<400> 617

gtttcttctgc tctcttccac atttt 25

<210> 618

<211> 25

<212> DNA

<213> Artificial sequence

<220><223> Primer K9_58N_071211r

<400> 618

gtttcttctgc tctcttccaa ttttt 25

<210> 619

<211> 25

<212> DNA

<213> Artificial sequence

<220><223> Primer K9_58V_071211r

<400> 619

gtttcttctgc tctcttccaa acttt 25

<210> 620

<211> 25

<212> DNA

<213> Artificial sequence

<220><223> Primer K9_58A_071211r

<400> 620

gtttcttctgc tctcttccaa gcttt 25

<210> 621

<211> 25

<212> DNA

<213> Artificial sequence

<220><223> Primer K9_58D_071211r

<400> 621

gttctttctgc tctcttccaa tcttt

25

<210> 622

<211> 25

<212> DNA

<213> Artificial sequence

<220><223> Primer K9_58E_071211r

<400> 622

gttctttctgc tctcttccat tcttt

25

<210> 623

<211> 25

<212> DNA

<213> Artificial sequence

<220><223> Primer K9_58G_071211r

<400> 623

gttctttctgc tctcttccaa ccttt

25

<210> 624

<211> 343

<212> PRT

<213> Artificial sequence

<220><223> Anaerostipes caccae KARI variant-BI7D12

<400> 624

Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn

1 5 10 15

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

20 25 30

Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val

35 40 45

Ile Ile Gly Leu Gln Glu Gly Val Lys Asp Trp Lys Arg Ala Glu Glu

50 55 60

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

65 70 75 80

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

	325	330	335
Glu Asp Lys Leu Ile Asn Asn			
	340		
<210>	625		
<211>	1029		
<212>	DNA		
<213>	Artificial sequence		
<220><223>	coding sequence for Anaerostipes caccae KARI variant- BI7D12		
<400>	625		
atggaagaat gtaagatggc taagatttac taccaagaag actgtaactt gtccttgttg 60			
gatggttaaga ctatcgccgt tatcggttac ggttctcaag gtcacgctca tgcctgaat 120			
gctaaggaat cgggttgtaa cgttatcatt ggtttacaag aaggtgtaa ggactggaaa 180			
agagctgaag aacaaggttt cgaagtctac accgctgtcg aagctgctaa gaaggctgac 240			
atcattatga tcttggtgcc ggatgaaaag caggctacca tgtacaaaa cgacatcgaa 300			
ccaaacttgg aagccggtaa catgttgatg ttcgctcacg gtttcaacat ccatttcggt 360			
tgtattgttc caccaaagga cgttgatgtc atgatgatcg ctccaaaggg tccaggtcac 420			
accgttagat ccgaatacga agaaggtaaa ggtgtcccat gcttggttgc tgtcgaacaa 480			
gacgctactg gcaaggcttt ggatatggct ttggcctacg ctttagccat cggtggtgct 540			
agagccggtg tcttggaac taccttcaga ggggaaactg aaaccgactt gttcggtgaa 600			
caagctgttt tatgtggtgg tgtctgcgct ttgatgcagg ccggttttga aaccttggtt 660			
gaagccggtt acgaccaag aaacgcttac ttcgaatgta tccacgaaat gaagttgatc 720			
gttgacttga tctaccaatc tggtttctcc ggtatgcgtt actctatctc caacactgct 780			
gaatacggtg actacattac cggtccaaag atcattactg aagataccea gaaggctatg 840			
aagaagattt tgtctgacat tcaagatggg acctttgccaggacttctt ggttgacatg 900			
tctgatgctg gtgccagggt ccacttcaag gctatgagaa agttggcctc cgaacacca 960			
gctgaagttg tcggtgaaga aattagatcc ttgtactcct ggtccgacga agacaagttg 1020			
attaacaac 1029			
<210>	626		
<211>	343		
<212>	PRT		
<213>	Artificial sequence		
<220><223>	Anaerostipes caccae KARI variant-BI10F1		

<400> 626

Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn

1 5 10 15

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

20 25 30

Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val

35 40 45

Ile Ile Gly Leu Glu Glu Gly Val Lys Asp Trp Lys Arg Ala Glu Glu

50 55 60

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

65 70 75 80

Ile Ile Met Ile Leu Val Pro Asp Glu Lys Gln Ala Thr Met Tyr Lys

85 90 95

Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala

100 105 110

His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val

115 120 125

Asp Val Met Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser

130 135 140

Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln

145 150 155 160

Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala

165 170 175

Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Gly Glu

180 185 190

Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val

195 200 205

Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr

210 215 220

Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile

225 230 235 240

Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile
245 250 255
Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile
260 265 270
Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln
275 280 285
Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly
290 295 300
Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro
305 310 315 320
Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
325 330 335
Glu Asp Lys Leu Ile Asn Asn
340

<210> 627

<211> 1029

<212> DNA

<213> Artificial sequence

<220><223> coding sequence for Anaerostipes caccae KARI variant-BI10F1

<400> 627

atggaagaat gtaagatggc taagatttac taccaagaag actgtaactt gtccttgttg	60
gatggtaaga ctatcgccgt tatcggttac ggtttctcaag gtcacgtca tgccctgaat	120
gctaaggaat ccggttgtaa cgttatcatt ggtttagaag aaggtgttaa ggactggaaa	180
agagctgaag aacaaggttt cgaagtctac accgtgctg aagctgctaa gaaggctgac	240
atcattatga tcttggtgcc ggatgaaaag caggctacca tgtacaaaaa cgacatcgaa	300
ccaaacttgg aagccggtaa catgttgatg ttcgctcacg gtttcaacat ccatttcggt	360
tgtattgttc caccaaagga cgttgatgac atgatgatcg ctccaaaggg tccaggtcac	420
accgttagat ccgaatacga agaaggtaaa ggtgtcccat gcttggttgc tgcgaacaa	480
gacgctactg gcaaggcttt ggatatggct ttggcctacg ctttagccat cgggtgtgct	540
agagccggtg tcttggaac taccttcaga ggggaaactg aaaccgactt gttcggtgaa	600
caagctgttt tatgtgggtg tgtctgcgct ttgatgcagg ccggttttga aaccttggtt	660

gaagccggtt acgaccaag aaacgcttac ttcgaatgta tccacgaaat gaagttgatc 720

gttgacttga tctaccaatc tggtttctcc ggtatgcgtt actctatctc caacactgct 780

gaatacgggtg actacattac cggtcctaaag atcattactg aagataccaa gaaggtatg 840

aagaagattt tgtctgacat tcaagatggt acctttgcca aggacttctt ggttgacatg 900

tctgatgctg gttcccaggt ccacttcaag gctatgagaa agttggcctc cgaacaccca 960

gctgaagttg tcggtgaaga aattagatcc ttgtactcct ggtccgacga agacaagttg 1020

attaacaac 1029

<210> 628

<211> 343

<212> PRT

<213> Artificial sequence

<220><223> Anaerostipes caccae KARI variant-BJ6G6

<400> 628

Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn

1 5 10 15

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

20 25 30

Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val

35 40 45

Ile Ile Gly Leu Pro Glu Gly Asp Lys Gln Trp Lys Arg Ala Glu Glu

50 55 60

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

65 70 75 80

Ile Ile Met Ile Leu Val Pro Asp Glu Lys Gln Ala Thr Met Tyr Lys

85 90 95

Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala

100 105 110

His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val

115 120 125

Asp Val Met Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser

130 135 140

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

<210> 629

<211> 1029

<212> DNA

<213> Artificial sequence

<220><223> Anaerostipes caccae KARI variant-BJ6G6

<400> 629

atggaagaat gtaagatggc taagatttac taccaagaag actgtaactt gtccttggtg 60
gatggtaaga ctatcgccgt tategggttac ggttctcaag gtcacgtca tgcctgaat 120

gctaaggaat ccggttgtaa cgttatcatt ggtttacctg aagtgataa gcactggaaa 180
agagctgaag aacaagggtt cgaagtctac accgctgctg aagctgctaa gaaggctgac 240
atcattatga tcttggtgcc ggatgaaaag caggctacca tgtacaaaaa cgacatcgaa 300
ccaaacttgg aagccggtaa catgttgatg ttcgctcacg gtttcaacat ccatttcggt 360
tgtattgttc caccaaagga cgttgatgtc atgatgatcg ctcaaaggg tccaggtcac 420
accgttagat ccgaatacga agaaggtaaa ggtgtcccat gcttggttgc tgcgaacaa 480
gacgctactg gcaaggcttt ggatatggct ttggcctacg ctttagccat cgggtgtgct 540

agagccggtg tcttggaac taccttcaga ggggaaactg aaaccgactt gttcggtgaa 600
caagctgttt tatgtgggtg tgtctgcgtt ttgatgcagg ccggttttga aaccttggtt 660
gaagccggtt acgaccaag aaacgcttac ttcgaatgta tccacgaaat gaagttgatc 720
gttgacttga tctaccaatc tggtttctcc ggtatgcgtt actctatctc caacactgct 780
gaatacgggt actacattac cggtccaaag atcattactg aagatacaca gaaggtatg 840
aagaagattt tgctgacat tcaagatggt acctttgcca aggacttctt ggttgacatg 900
tctgatgctg gttcccaggt ccacttcaag gctatgagaa agttggcctc cgaacacca 960

gctgaagttg tcggtgaaga aattagatcc ttgtactcct ggtccgacga agacaagttg 1020
attaacaac 1029

<210> 630

<211> 343

<212> PRT

<213> Artificial sequence

<220><223> Anaerostipes caccae KARI variant-BJ7D6

<400> 630

Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn

1 5 10 15

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

20 25 30

Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val

35 40 45

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

50	55	60	
Gln Gly Phe Glu Val Tyr Thr Ala Ala Glu Ala Ala Lys Lys Ala Asp			
65	70	75	80
Ile Ile Met Ile Leu Val Pro Asp Glu Lys Gln Ala Thr Met Tyr Lys			
	85	90	95
Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala			
	100	105	110
His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val			
	115	120	125
Asp Val Met Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser			
	130	135	140
Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln			
145	150	155	160
Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala			
	165	170	175
Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Gly Glu			
	180	185	190
Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val			
	195	200	205
Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr			
	210	215	220
Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile			
225	230	235	240
Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile			
	245	250	255
Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile			
	260	265	270
Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln			
	275	280	285
Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly			
	290	295	300

<212> PRT

<213> Artificial sequence

<220><223> Anaerostipes caccae KARI variant-BJ7F7

<400> 632

Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn

1 5 10 15

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

20 25 30

Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val

35 40 45

Ile Ile Gly Leu Ala Glu Gly Asp Lys Gln Trp Lys Arg Ala Glu Glu

50 55 60

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

65 70 75 80

Ile Ile Met Ile Leu Val Pro Asp Glu Lys Gln Ala Thr Met Tyr Lys

85 90 95

Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala

100 105 110

His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val

115 120 125

Asp Val Met Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser

130 135 140

Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln

145 150 155 160

Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala

165 170 175

Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Gly Glu

180 185 190

Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val

195 200 205

Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr

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

 Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile
 260 265 270
 Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln
 275 280 285
 Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly
 290 295 300
 Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro
 305 310 315 320

Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
 325 330 335
 Glu Asp Lys Leu Ile Asn Asn
 340

<210> 633

<211> 1029

<212> DNA

<213> Artificial sequence

<220><223> coding sequence for Anaerostipes caccae KARI variant-BJ7F7

<400> 633

atggaagaat gtaagatggc taagatttac taccaagaag actgtaactt gtccttggtg	60
gatggtaaga ctatgccgt tatcggttac ggttctcaag gtcacgtca tgccctgaat	120
gctaaggaat ccggttgtaa cgttatcatt ggttagctg aagtgataa gcactgaaa	180
agagctgaag aacaaggttt cgaagtctac accgctgctg aagctgctaa gaagctgac	240
atcattatga tcttggtgcc ggatgaaaag caggctacca tgtacaaaa cgacatcgaa	300
ccaaacttgg aagccggtaa catgttgatg ttcgctcacg gtttcaacat ccatttcggt	360
tgtattgttc caccaaagga cgttgatgtc atgatgatcg ctcaaaggg tccaggtcac	420
accgttagat ccgaatacga agaaggtaaa ggtgtcccat gcttggttgc tgcgaacaa	480
gacgctactg gcaaggcttt ggatatggct ttggcctacg ctttagccat cgggtgtgct	540

agagccggtg tcttggaac taccttcaga ggggaaactg aaaccgactt gttcggtgaa 600

caagctgttt tatgtgggtg tgtctgcgct ttgatgcagg ccggttttga aaccttggtt 660

gaagccggtt acgaccecaag aaacgcttac ttcgaatgta tccacgaaat gaagttgatc 720

gttgacttga tctaccaatc tggtttctcc ggtatgcgctt actctatctc caacactgct 780

gaatacgggtg actacattac cggtccaaag atcattactg aagataccaa gaaggctatg 840

aagaagattt tgtctgacat tcaagatggt acctttgcca aggacttctt ggttgacatg 900

tctgatgctg gttcccaggt ccacttcaag gctatgagaa agttggcctc cgaacaccca 960

gctgaagttg tcggtgaaga aattagatcc ttgtactcct ggtccgacga agacaagttg 1020

attaacaac 1029

<210> 634

<211> 24

<212> DNA

<213> Artificial sequence

<220><223> Primer pBAD_266f

<400> 634

ctctctactg tttctccata cccg 24

<210> 635

<211> 338

<212> PRT

<213> Artificial sequence

<220><223> N-terminal truncated Anaerostipes caccae KARI variant

<400> 635

Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn Leu Ser Leu Leu Asp

1 5 10 15

Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser Gln Gly His Ala His

20 25 30

Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val Ile Ile Gly Leu Tyr

35 40 45

Glu Gly Ser Lys Ser Trp Lys Arg Ala Glu Glu Gln Gly Phe Glu Val

50 55 60

Tyr Thr Ala Ala Glu Ala Ala Lys Lys Ala Asp Ile Ile Met Ile Leu

65					70								75								80										
Ile	Asn	Asp	Glu	Lys	Gln	Ala	Thr	Met	Tyr	Lys	Asn	Asp	Ile	Glu	Pro																
				85								90								95											
Asn	Leu	Glu	Ala	Gly	Asn	Met	Leu	Met	Phe	Ala	His	Gly	Phe	Asn	Ile																
				100								105								110											
His	Phe	Gly	Cys	Ile	Val	Pro	Pro	Lys	Asp	Val	Asp	Val	Thr	Met	Ile																
				115								120								125											
Ala	Pro	Lys	Gly	Pro	Gly	His	Thr	Val	Arg	Ser	Glu	Tyr	Glu	Glu	Gly																
				130								135								140											
Lys	Gly	Val	Pro	Cys	Leu	Val	Ala	Val	Glu	Gln	Asp	Ala	Thr	Gly	Lys																
				145								150								155								160			
Ala	Leu	Asp	Met	Ala	Leu	Ala	Tyr	Ala	Leu	Ala	Ile	Gly	Gly	Ala	Arg																
				165								170								175											
Ala	Gly	Val	Leu	Glu	Thr	Thr	Phe	Arg	Thr	Glu	Thr	Glu	Thr	Asp	Leu																
				180								185								190											
Phe	Gly	Glu	Gln	Ala	Val	Leu	Cys	Gly	Gly	Val	Cys	Ala	Leu	Met	Gln																
				195								200								205											
Ala	Gly	Phe	Glu	Thr	Leu	Val	Glu	Ala	Gly	Tyr	Asp	Pro	Arg	Asn	Ala																
				210								215								220											
Tyr	Phe	Glu	Cys	Ile	His	Glu	Met	Lys	Leu	Ile	Val	Asp	Leu	Ile	Tyr																
				225								230								235								240			
Gln	Ser	Gly	Phe	Ser	Gly	Met	Arg	Tyr	Ser	Ile	Ser	Asn	Thr	Ala	Glu																
				245								250								255											
Tyr	Gly	Asp	Tyr	Ile	Thr	Gly	Pro	Lys	Ile	Ile	Thr	Glu	Asp	Thr	Lys																
				260								265								270											
Lys	Ala	Met	Lys	Lys	Ile	Leu	Ser	Asp	Ile	Gln	Asp	Gly	Thr	Phe	Ala																
				275								280								285											
Lys	Asp	Phe	Leu	Val	Asp	Met	Ser	Asp	Ala	Gly	Ser	Gln	Val	His	Phe																
				290								295								300											
Lys	Ala	Met	Arg	Lys	Leu	Ala	Ser	Glu	His	Pro	Ala	Glu	Val	Val	Gly																
				305								310								315								320			

Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp Glu Asp Lys Leu Ile

325

330

335

Asn Asn

<210> 636

<211> 1017

<212> DNA

<213> Artificial sequence

<220><223> coding sequence for N-terminal truncated Anaerostipes caccae KARI
variant

<400> 636

atggctaaga ttactacca agaagactgt aacttgctct tgttgatgg taagactatc 60
gccgttatcg gttacgggtc tcaaggcac gctcatgccc tgaatgctaa ggaatccggt 120
tgtaacgtta tcattgggtt atacgaaggt tctaagtcct ggaaaagagc tgaagaacaa 180
ggtttcgaag tctacaccgc tgctgaagct gctaagaagg ctgacatcat tatgatcttg 240
atcaacgatg aaaagcaggc taccatgtac aaaaacgaca tcgaaccaa cttggaagcc 300

ggtaacatgt tgatgttcgc tcacggtttc aacatccatt tcggttgat tgttcacca 360
aaggacgttg atgtcactat gatcgctcca aagggtccag gtcacaccgt tagatccgaa 420
tacgaagaag gtaaagggtg cccatgcttg gttgctgtcg aacaagacgc tactggcaag 480
gctttggata tggctttggc ctacgcttta gccatcgggt gtgctagagc cgggtgcttg 540
gaaactacct tcagaaccga aactgaaacc gacttggtcg gtgaacaagc tgttttatgt 600
ggtggtgtct gcgctttgat gcaggccggt ttgaaacct tgggtgaagc cggttacgac 660
ccaagaaacg cttacttcga atgtatccac gaaatgaagt tgatcgttga cttgatctac 720

caatctgggt tctccggtat gcgttactct atctccaaca ctgctgaata cggtgactac 780
attaccggtc caaagatcat tactgaagat accaagaagg ctatgaagaa gatattgtct 840
gacattcaag atggtacctt tgccaaggac ttcttggttg acatgtctga tgctggttcc 900
cagggtccact tcaaggctat gagaaagttg gcctccgaac acccagctga agttgtcggt 960
gaagaaatta gatccttgta ctcttggtcc gacgaagaca agttgattaa caactga 1017

<210> 637

<211> 338

<212> PRT

<213> Artificial sequence

<220><223> N-terminal truncated *Anaerostipes caccae* KARI K9SB2 variant

<400> 637

Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn Leu Ser Leu Leu Asp

1 5 10 15

Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser Gln Gly His Ala His

20 25 30

Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val Ile Ile Gly Leu Phe

35 40 45

Glu Gly Ala Glu Glu Trp Lys Arg Ala Glu Glu Gln Gly Phe Glu Val

50 55 60

Tyr Thr Ala Ala Glu Ala Ala Lys Lys Ala Asp Ile Ile Met Ile Leu

65 70 75 80

Ile Pro Asp Glu Lys Gln Ala Thr Met Tyr Lys Asn Asp Ile Glu Pro

85 90 95

Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala His Gly Phe Asn Ile

100 105 110

His Phe Gly Cys Ile Val Pro Pro Lys Asp Val Asp Val Thr Met Ile

115 120 125

Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser Glu Tyr Glu Glu Gly

130 135 140

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

145 150 155 160

Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala Ile Gly Gly Ala Arg

165 170 175

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

180 185 190

Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val Cys Ala Leu Met Gln

195 200 205

Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr Asp Pro Arg Asn Ala

210 215 220

Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile Val Asp Leu Ile Tyr

225 230 235 240
 Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile Ser Asn Thr Ala Glu
 245 250 255

 Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile Thr Glu Asp Thr Lys
 260 265 270
 Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln Asp Gly Thr Phe Ala
 275 280 285
 Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly Ser Gln Val His Phe
 290 295 300
 Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro Ala Glu Val Val Gly
 305 310 315 320

Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp Glu Asp Lys Leu Ile
 325 330 335
 Asn Asn

<210> 638

<211> 1017

<212> DNA

<213> Artificial sequence

<220><223> coding sequence for N-terminal truncated Anaerostipes caccae
 K9SB2 KARI variant

<400> 638

atggctaaga ttactacca agaagactgt aactgtcct tgttgatgg taagactatc 60
 gccgttatcg gttacggttc tcaaggtcac gtcgatgcc tgaatgctaa ggaatccggt 120
 tgtaacgtta tcattggttt attcgaaggt gcggaggagt ggaaaagagc tgaagaacaa 180

 ggtttcgaag tctacaccgc tgctgaagct gctaagaagg ctgacatcat tatgatcttg 240
 atcccagatg aaaagcaggc taccatgtac aaaaacgaca tcgaaccaa cttggaagcc 300
 ggtaacatgt tgatgttcgc tcacggtttc aacatccatt tcggttgat tgttcacca 360
 aaggacgttg atgtcactat gatcgtcca aagggtccag gtcacaccgt tagatccgaa 420
 tacgaagaag gtaaaggtgt cccatgcttg gttgctgtcg aacaagacgc tactggcaag 480
 gctttggata tggctttggc ctacgcttta gccatcggtg gtgctagagc cgggtgtcttg 540

gaaactacct tcagaaccga aactgaaacc gacttggtcg gtgaacaagc tgttttatgt 600

ggtgggtgtct gcgctttgat gcaggccggt tttgaaacct tggttgaagc cggttacgac 660

ccaagaaacg cttacttcga atgtatccac gaaatgaagt tgatcggtga cttgatctac 720

caatctgggtt tctccgggat gcgttactct atctccaaca ctgctgaata cggtgactac 780

attaccggtc caaagatcat tactgaagat accaagaagg ctatgaagaa gattttgtct 840

gacattcaag atggtacctt tgccaaggac ttcttggttg acatgtctga tgctggttcc 900

caggtcact tcaaggctat gagaaagttg gcctccgaac acccagctga agttgtcggc 960

gaagaaatta gatccttgta ctcttggtcc gacgaagaca agttgattaa caactga 1017

<210> 639

<211> 343

<212> PRT

<213> Artificial sequence

<220><223> Anaerostipes caccae KARI variant-K9SB2-T93I

<400> 639

Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn

1 5 10 15

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

20 25 30

Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val

35 40 45

Ile Ile Gly Leu Phe Glu Gly Ala Glu Glu Trp Lys Arg Ala Glu Glu

50 55 60

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

65 70 75 80

Ile Ile Met Ile Leu Ile Pro Asp Glu Lys Gln Ala Ile Met Tyr Lys

85 90 95

Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala

100 105 110

His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val

115 120 125

Asp Val Thr Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser

130 135 140
 Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln
 145 150 155 160
 Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala
 165 170 175

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

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

Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro
 305 310 315 320
 Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
 325 330 335
 Glu Asp Lys Leu Ile Asn Asn
 340

<210> 640

<211> 1032

<212> DNA

<213> Artificial sequence

<220><223> coding sequence for Anaerostipes caccae KARI variant-K9SB2-T93I

<400> 640

atggaagaat gtaagatggc taagatttac taccaagaag actgtaactt gtccttggtg	60
gatggtgaaga ctatcgccgt tatcggttac ggtttctcaag gtcacgtca tgccctgaat	120
gctaaggaat ccggttgtaa cgttatcatt ggtttattcg aaggtgcgga ggagtggaaa	180
agagctgaag aacaagggtt cgaagtctac accgctgctg aagctgctaa gaaggtgac	240
atcattatga tcttgatccc agatgaaaag caggctatca tgtacaaaaa cgacatcgaa	300
ccaaacttgg aagccggtaa catgttgatg ttgctcacg gtttcaacat ccatttcggt	360
tgtattgttc caccaaagga cgttgatgtc actatgatcg ctccaaaggg tccaggtcac	420
accgttagat ccgaatacga agaaggtaaa ggtgtcccat gcttggttgc tgtcgaacaa	480
gacgctactg gcaaggcttt ggatatggct ttggcctacg ctttagccat cgggtgtgct	540
agagccggtg tcttggaac taccttcaga accgaaactg aaaccgactt gttcggtgaa	600
caagctgttt tatgtgggtg tgtctgcgt ttgatgcagg ccggttttga aaccttggtt	660
gaagccggtt acgaccaag aaacgcttac ttcgaatgta tccacgaaat gaagttgatc	720
gttgacttga tctaccaatc tggtttctcc ggtatgcgtt actctatctc caacactgct	780
gaatacgggtg actacattac cggtccaaag atcattactg aagataccaa gaaggctatg	840
aagaagattt tgtctgacat tcaagatggt acctttgcca aggacttctt ggttgacatg	900
tctgatgctg gtcccagggt ccacttcaag gctatgagga agttggcctc cgaacaccca	960
gctgaagttg tcggtgaaga aattagatcc ttgtactcct ggtccgacga agacaagttg	1020
attaacaact ga	1032

<210> 641

<211> 343

<212> PRT

<213> Artificial sequence

<220><223> Anaerostipes caccae KARI variant-K9SB2-T93A

<400> 641

Met	Glu	Glu	Cys	Lys	Met	Ala	Lys	Ile	Tyr	Tyr	Gln	Glu	Asp	Cys	Asn
1				5				10						15	
Leu	Ser	Leu	Leu	Asp	Gly	Lys	Thr	Ile	Ala	Val	Ile	Gly	Tyr	Gly	Ser
				20				25						30	
Gln	Gly	His	Ala	His	Ala	Leu	Asn	Ala	Lys	Glu	Ser	Gly	Cys	Asn	Val
				35				40						45	

Ile Ile Gly Leu Phe Glu Gly Ala Glu Glu Trp Lys Arg Ala Glu Glu
 50 55 60
 Gln Gly Phe Glu Val Tyr Thr Ala Ala Glu Ala Ala Lys Lys Ala Asp
 65 70 75 80
 Ile Ile Met Ile Leu Ile Pro Asp Glu Lys Gln Ala Ala Met Tyr Lys
 85 90 95
 Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala
 100 105 110
 His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val
 115 120 125
 Asp Val Thr Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser
 130 135 140
 Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln
 145 150 155 160
 Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala
 165 170 175
 Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu
 180 185 190
 Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val
 195 200 205
 Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr
 210 215 220
 Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile
 225 230 235 240
 Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile
 245 250 255
 Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile
 260 265 270
 Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln
 275 280 285
 Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly

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

<210> 642

<211> 1032

<212> DNA

<213> Artificial sequence

<220><223> coding sequence for Anaerostipes caccae KARI variant K9SB2-T93A

<400> 642

atggaagaat gtaagatggc taagatttac taccaagaag actgtaactt gtccttggtg 60
 gatggtaaga ctatcgccgt tatcggttac ggttctcaag gtcacgtca tgccctgaat 120
 gctaaggaat ccggttgtaa cgttatcatt ggtttattcg aagggtcgga ggagtggaaa 180
 agagctgaag aacaagggtt cgaagtctac accgtgctg aagctgctaa gaaggtgac 240
 atcattatga tcttgatccc agatgaaaag caggctgcca tgtacaaaa cgacatcgaa 300

ccaaacttgg aagccggtaa catgttgatg ttgcgtcacg gtttcaacat ccatttcggt 360
 tgtattgttc caccaaagga cgttgatgtc actatgatcg ctccaaaggg tccaggtcac 420
 accgttagat ccgaatacga agaaggtaaa ggtgtcccat gcttggttgc tgcgaacaa 480
 gacgtactg gcaaggcttt ggatatggct ttggcctacg ctttagccat cgggtgtgct 540
 agagccggtg tcttggaac taccttcaga accgaaactg aaaccgactt gttcggtgaa 600
 caagctgttt tatgtgggtg tgtctgcgct ttgatgcagg ccggttttga aaccttggtt 660
 gaagccgggt acgaccaag aaacgcttac ttcgaatgta tccacgaaat gaagttgatc 720

gttgacttga tctaccaatc tggtttctcc ggtatgcgtt actctatctc caacactgct 780
 gaatacgggt actacattac cgggtccaaag atcattactg aagatacga gaaggtatg 840
 aagaagattt tgctgacat tcaagatggt acctttgcca aggacttctt ggttgacatg 900
 tctgatgctg gttcccaggt ccacttcaag gctatgagaa agttggcctc cgaacacca 960
 gctgaagttg tcggtgaaga aattagatcc ttgtactcct ggtccgacga agacaagttg 1020
 attaacaact ga 1032

<210> 643

<211> 343

<212> PRT

<213> Artificial Sequence

<220><223> Anaerostipes caccae DSM 14662

<400> 643

Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn

1 5 10 15

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

20 25 30

Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val

35 40 45

Ile Ile Gly Leu Tyr Glu Gly Ser Lys Ser Trp Lys Arg Ala Glu Glu

50 55 60

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

65 70 75 80

Ile Ile Met Ile Leu Ile Asn Asp Glu Lys Gln Ala Thr Met Tyr Lys

85 90 95

Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala

100 105 110

His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val

115 120 125

Asp Val Thr Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser

130 135 140

Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln

145 150 155 160

Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala

165 170 175

Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu

180 185 190

Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val

195 200 205

Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr
 210 215 220
 Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile
 225 230 235 240
 Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile
 245 250 255
 Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile
 260 265 270
 Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln
 275 280 285
 Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly
 290 295 300
 Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro
 305 310 315 320
 Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
 325 330 335
 Glu Asp Lys Leu Ile Asn Asn
 340
 <210> 644
 <211> 343
 <212> PRT
 <213> Artificial Sequence
 <220><223> Anaerostipes caccae KARI AB1G9 variant
 <400> 644
 Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn
 1 5 10 15
 Leu Ser Leu Leu Asp Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser
 20 25 30
 Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val
 35 40 45
 Ile Ile Gly Leu Tyr Glu Gly Ala Lys Glu Trp Lys Arg Ala Glu Glu
 50 55 60

Gln Gly Phe Glu Val Tyr Thr Ala Ala Glu Ala Ala Lys Lys Ala Asp
65 70 75 80

Ile Ile Met Ile Leu Ile Asn Asp Glu Lys Gln Ala Thr Met Tyr Lys
85 90 95

Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala
100 105 110

His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val
115 120 125

Asp Val Thr Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser
130 135 140

Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln
145 150 155 160

Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala
165 170 175

Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu
180 185 190

Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val
195 200 205

Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr
210 215 220

Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile
225 230 235 240

Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile
245 250 255

Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile
260 265 270

Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln
275 280 285

Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly
290 295 300

Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro

305	310										315					320				
Ala	Glu	Val	Val	Gly	Glu	Glu	Ile	Arg	Ser	Leu	Tyr	Ser	Trp	Ser	Asp					
				325			330					335								
Glu Asp Lys Leu Ile Asn Asn																				
340																				
<210> 645																				
<211> 343																				
<212> PRT																				
<213> Artificial Sequence																				
<220><223> Anaerostipes caccae KARI-AB1D3 variant																				
<400> 645																				
Met	Glu	Glu	Cys	Lys	Met	Ala	Lys	Ile	Tyr	Tyr	Gln	Glu	Asp	Cys	Asn					
1				5					10				15							
Leu	Ser	Leu	Leu	Asp	Gly	Lys	Thr	Ile	Ala	Val	Ile	Gly	Tyr	Gly	Ser					
				20			25				30									
Gln	Gly	His	Ala	His	Ala	Leu	Asn	Ala	Lys	Glu	Ser	Gly	Cys	Asn	Val					
				35			40				45									
Ile	Ile	Gly	Leu	Tyr	Glu	Gly	Ala	Lys	Asp	Trp	Lys	Arg	Ala	Glu	Glu					
50						55			60											
Gln	Gly	Phe	Glu	Val	Tyr	Thr	Ala	Ala	Glu	Ala	Ala	Lys	Lys	Ala	Asp					
65		70				75				80										
Ile	Ile	Met	Ile	Leu	Ile	Asn	Asp	Glu	Lys	Gln	Ala	Thr	Met	Tyr	Lys					
				85			90				95									
Asn	Asp	Ile	Glu	Pro	Asn	Leu	Glu	Ala	Gly	Asn	Met	Leu	Met	Phe	Ala					
				100			105				110									
His	Gly	Phe	Asn	Ile	His	Phe	Gly	Cys	Ile	Val	Pro	Pro	Lys	Asp	Val					
				115			120				125									
Asp	Val	Thr	Met	Ile	Ala	Pro	Lys	Gly	Pro	Gly	His	Thr	Val	Arg	Ser					
130						135			140											
Glu	Tyr	Glu	Glu	Gly	Lys	Gly	Val	Pro	Cys	Leu	Val	Ala	Val	Glu	Gln					
145		150				155				160										

Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala

165 170 175

Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu

180 185 190

Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val

195 200 205

Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr

210 215 220

Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile

225 230 235 240

Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile

245 250 255

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

260 265 270

Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln

275 280 285

Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly

290 295 300

Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro

305 310 315 320

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

325 330 335

Glu Asp Lys Leu Ile Asn Asn

340

<210> 646

<211> 343

<212> PRT

<213> Artificial Sequence

<220><223> Anaerostipes caccae KARI- K9D3 variant

<400> 646

Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn

1 5 10 15
 Leu Ser Leu Leu Asp Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser
 20 25 30
 Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val
 35 40 45
 Ile Ile Gly Leu Tyr Glu Gly Ala Lys Asp Trp Lys Arg Ala Glu Glu
 50 55 60

 Gln Gly Phe Glu Val Tyr Thr Ala Ala Glu Ala Ala Lys Lys Ala Asp
 65 70 75 80
 Ile Ile Met Ile Leu Ile Asn Asp Glu Lys Gln Ala Thr Met Tyr Lys
 85 90 95
 Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala
 100 105 110
 His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val
 115 120 125

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

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

Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile
 260 265 270
 Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln
 275 280 285
 Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly
 290 295 300
 Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro
 305 310 315 320

Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
 325 330 335
 Glu Asp Lys Leu Ile Asn Asn
 340

<210> 647

<211> 343

<212> PRT

<213> Artificial Sequence

<220><223> Anaerostipes caccae KARI-K9G9 variant

<400> 647

Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn
 1 5 10 15
 Leu Ser Leu Leu Asp Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser

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

85 90 95
 Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala
 100 105 110

His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val
115 120 125
Asp Val Thr Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser
130 135 140
Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln
145 150 155 160
Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala
165 170 175
Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu
180 185 190
Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val
195 200 205
Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr
210 215 220
Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile
225 230 235 240
Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile
245 250 255
Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile
260 265 270
Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln
275 280 285
Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly
290 295 300
Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro
305 310 315 320
Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
325 330 335
Glu Asp Lys Leu Ile Asn Asn
340

<210> 648

<211> 1689

<212> DNA

<213> Artificial Sequence

<220><223> PDC1 pyruvate decarboxylase from *Saccharomyces cerevisiae*

<400> 648

atgtctgaaa ttactttggg taaatatttg ttcgaaagat taaagcaagt caacgttaac	60
accgtttttcg gtttgccagg tgacttcaac ttgtccttgt tggacaagat ctacgaagtt	120
gaaggtatga gatgggctgg taacgccaac gaattgaacg ctgcttacgc cgctgatggt	180
tacgctcgta tcaagggtat gtctttgtatc atcaccacct tcggtgtcgg tgaattgtct	240
gctttgaacg gtattgccgg ttcttacgct gaacacgtcg gtgttttgca cgttgtttgt	300
gtcccatcca tctctgctca agctaagcaa ttgtttgtgc accacacctt gggtaacggt	360
gacttcactg ttttccacag aatgtctgcc aacatttctg aaaccactgc tatgatcact	420
gacattgcta cgcggccagc tgaaattgac agatgtatca gaaccactta cgtcacccaa	480
agaccagtct acttaggttt gccagctaac ttggctgact tgaacgtccc agctaagttg	540
ttgcaaactc caattgacat gtctttgaag ccaaacgatg ctgaatccga aaaggaagtc	600
attgacacca tcttggcttt ggtcaaggat gctaagaacc cagttatctt ggctgatgct	660
tgtttgtcca gacacgacgt caaggctgaa actaagaagt tgattgactt gactcaattc	720
ccagctttcg tcacccaat gggttaagggt tccattgacg aacaacaccc aagatacgg	780
ggtgtttacg tcggtacctt gtccaagcca gaagttaagg aagccgttga atctgctgac	840
ttgattttgt ctgtcgggtc ttgtttgtct gatttcaaca ccggttcttt ctcttactct	900
tacaagacca agaacattgt cgaattccac tccgaccaca tgaagatcag aaacgccact	960
ttcccagggtg tccaaatgaa attcgttttg caaaagtgtg tgaccactat tgctgacgcc	1020
gctaagggtt acaagccagt tgctgtccca gctagaactc cagctaacgc tgctgtccca	1080
gcttctaccc cattgaagca agaattggatg tggaaccaat tgggttaactt cttgcaagaa	1140
ggatgatgtg tcattgctga aaccggtacc tccgctttcg gtatcaacca aaccactttc	1200
ccaaacaaca cctacgggtat ctctcaagtc ttatgggggt ccatgggttt caccactggt	1260
gctaccttgg gtgctgcttt cgctgctgaa gaaattgac caaagaagag agttatctta	1320
ttcattgggtg acggttcttt gcaattgact gttcaagaaa tctccacat gatcagatgg	1380
ggcttgaagc catacttggt cgtcttgaac aacgatggtt acaccattga aaagttgatt	1440
cacgggtcca aggtcaata caacgaaatt caaggttggg accacctatc cttgttgcca	1500
actttcgggtg ctaaggacta tgaaaccac agatgcgcta ccaccggtga atgggacaag	1560

ttgacccaag acaagtcttt caacgacaac tctaagatca gaatgattga aatcatgttg 1620
ccagtcttcg atgtccaca aaacttggtt gaacaagcta agttgactgc tgctaccaac 1680
gctaagcaa 1689

<210> 649

<211> 563

<212> PRT

<213> Artificial Sequence

<220><223> PDC1 pyruvate decarboxylase from *Saccharomyces cerevisiae*

<400> 649

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	320
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	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 Lys Leu Ile
465 470 475 480

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> 650

<211> 1689

<212> DNA

<213> Artificial Sequence

<220><223> PDC5 pyruvate decarboxylase from *Saccharomyces cerevisiae*

<400> 650

atgtctgaaa taaccttagg taaatatatta ttgaaagat tgagccaagt caactgtaac 60

accgtcttcg gtttgccagg tgactttaac ttgtctcttt tggataagct ttatgaagtc 120

aaaggtatga gatgggctgg taacgctaac gaattgaacg ctgcctatgc tgctgatggt 180

tacgctcgta tcaagggtat gtcctgtatt attaccacct tcggtgttgg tgaattgtct 240

gctttgaatg gtattgccgg ttcttacgct gaacatgtcg gtgttttgca cgttgttggt 300

gttccatcca tctcttctca agctaagcaa ttgttgttgc atcatacctt gggtaacggt 360

gacttcactg ttttccacag aatgtctgcc aacatttctg aaaccactgc catgatcact 420

gatattgcta acgctccagc tgaaattgac agatgtatca gaaccaccta cactacccaa 480

agaccagtct acttgggttt gccagctaac ttggttgact tgaacgtccc agccaagtta 540

ttggaaactc caattgactt gtctttgaag ccaaacgacg ctgaagctga agctgaagtt 600

gttagaactg ttgttgaatt gatcaaggat gctaagaacc cagttatctt ggctgatgct 660

tgtgcttcta gacatgatgt caaggctgaa actaagaagt tgatggactt gactcaattc 720

ccagtttacg tcacccaat gggttaagggt gctattgacg aacaacaccc aagatacggg 780

gggtgtttacg ttggtacctt gtctagacca gaagttaaga aggctgtaga atctgctgat 840

ttgatattgt ctatcggtgc tttgttgtct gatttcaata ccggttcttt ctcttactcc 900

tacaagacca aaaatatcgt tgaattccac tctgaccaca tcaagatcag aaacgccacc 960

ttcccagggtg ttcaaatgaa atttgccttg caaaaattgt tggatgctat tccagaagtc 1020

gtcaaggact acaaacctgt tgctgtccca gctagagttc caattaccaa gtctactcca 1080

gctaacactc caatgaagca agaattggatg tggaaccatt tgggtlaactt cttgagagaa 1140

gggtgatattg ttattgctga aaccgggtact tccgccttcg gtattaacca aactactttc 1200

ccaacagatg tatagctat cgtecaagtc ttgtgggggtt ccattgggtt cacagtcggc 1260

gctctattgg gtgctactat ggccgctgaa gaacttgatc caaagaagag agttatttta 1320

ttcattgggtg acggttctct acaattgact gtcaagaaa tctctacat gattagatgg 1380

ggtttgaagc catacathtt tgtcttgaat aacaacgggt acaccattga aaaattgatt 1440

cacggtcctc atgccgaata taatgaaatt caaggttggg accacttggc cttattgcca 1500

acttttgggtg ctagaaacta cgaaaccac agagttgcta ccaactggtga atgggaaaag 1560

ttgactcaag acaaggactt ccaagacaac tctaagatta gaatgattga agttatgttg 1620

ccagtctttg atgctccaca aaacttgggt aaacaagctc aattgactgc cgctactaac 1680

gctaaacaa 1689

<210> 651

<211> 563

<212> PRT

<213> Artificial Sequence

<220><223> PDC5 pyruvate decarboxylase from *Saccharomyces cerevisiae*

<400> 651

Met Ser Glu Ile Thr Leu Gly Lys Tyr Leu Phe Glu Arg Leu Ser Gln

1 5 10 15

Val Asn Cys Asn Thr Val Phe Gly Leu Pro Gly Asp Phe Asn Leu Ser

20 25 30
 Leu Leu Asp Lys Leu Tyr Glu Val Lys 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 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 Ala Asn Ile Ser Glu Thr Thr Ala Met Ile Thr Asp Ile Ala Asn
 130 135 140
 Ala Pro Ala Glu Ile Asp Arg Cys Ile Arg Thr Thr Tyr Thr 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 Glu Thr Pro Ile Asp Leu Ser Leu Lys Pro Asn
 180 185 190
 Asp Ala Glu Ala Glu Ala Glu Val Val Arg Thr Val Val Glu Leu Ile
 195 200 205
 Lys Asp Ala Lys Asn Pro Val Ile Leu Ala Asp Ala Cys Ala Ser Arg
 210 215 220
 His Asp Val Lys Ala Glu Thr Lys Lys Leu Met Asp Leu Thr Gln Phe
 225 230 235 240

 Pro Val Tyr Val Thr Pro Met Gly Lys Gly Ala Ile Asp Glu Gln His
 245 250 255
 Pro Arg Tyr Gly Gly Val Tyr Val Gly Thr Leu Ser Arg Pro Glu Val
 260 265 270

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

 Trp Met Trp Asn His Leu Gly Asn Phe Leu Arg Glu Gly Asp Ile Val
 370 375 380
 Ile Ala Glu Thr Gly Thr Ser Ala Phe Gly Ile Asn Gln Thr Thr Phe
 385 390 395 400
 Pro Thr Asp Val Tyr Ala Ile Val Gln Val Leu Trp Gly Ser Ile Gly
 405 410 415
 Phe Thr Val Gly Ala Leu Leu Gly Ala Thr Met Ala Ala Glu Glu Leu
 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 Ile Phe Val Leu Asn Asn Asn Gly Tyr Thr Ile Glu Lys Leu Ile
 465 470 475 480
 His Gly Pro His Ala Glu Tyr Asn Glu Ile Gln Gly Trp Asp His Leu
 485 490 495

 Ala Leu Leu Pro Thr Phe Gly Ala Arg Asn Tyr Glu Thr His Arg Val
 500 505 510
 Ala Thr Thr Gly Glu Trp Glu Lys Leu Thr Gln Asp Lys Asp Phe Gln

gctagcacgc ccttgaacaa agagtgggtg tggaacgaat tgtccaaatt ctgcaagaa 1140

ggtgatgtta tcatttccga gaccggcacg tctgccttcg gtatcaatca aactatcttt 1200

cctaaggacg cctacgggat ctgcaggtg ttgtgggggt ccatcggttt tacaacagga 1260

gcaactttag gtgctgcctt tgccgctgag gagattgacc ccaacaagag agtcatttta 1320

ttcatagggtg acgggtcttt gcagttaacc gtccaagaaa tctccacat gatcagatgg 1380

gggttaaagc cgtatctttt tgctcttaac aacgacggct acactatcga aaagctgatt 1440

catgggcctc acgcagagta caacgaaatc cagacctggg atcacctcgc cctgttgccc 1500

gcatttgggtg cgaaaaagta cgaaaatcac aagatcgcca ctacgggtga gtgggatgcc 1560

ttaaccactg attcagagtt ccagaaaaac tcggtgatc 1599

<210> 653

<211> 533

<212> PRT

<213> Artificial Sequence

<220><223> PDC6 pyruvate decarboxylase from *Saccharomyces cerevisiae*

<400> 653

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 Ile Phe Gly Leu Pro Gly Asp Phe Asn Leu Ser

20 25 30

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

370 375 380
 Ile Ser Glu Thr Gly Thr Ser Ala Phe Gly Ile Asn Gln Thr Ile Phe
 385 390 395 400
 Pro Lys Asp Ala 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 Asn 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
 465 470 475 480
 His Gly Pro His Ala Glu Tyr Asn Glu Ile Gln Thr Trp Asp His Leu

 485 490 495
 Ala Leu Leu Pro Ala Phe Gly Ala Lys Lys Tyr Glu Asn His Lys Ile
 500 505 510
 Ala Thr Thr Gly Glu Trp Asp Ala Leu Thr Thr Asp Ser Glu Phe Gln
 515 520 525
 Lys Asn Ser Val Ile
 530

<210> 654

<211> 1692

<212> DNA

<213> Artificial Sequence

<220><223> pyruvate decarboxylase from *Candida glabrata*

<400> 654

atgtctgaga ttactttggg tagatacttg ttcgagagat tgaaccaagt cgacgttaag	60
accatcttcg gtttgccagg tgacttcaac ttgtccctat tggacaagat ctacgaagtt	120
gaaggtatga gatgggctgg taacgctaac gaattgaacg ctgcttacgc tgctgacggt	180
tacgctagaa tcaagggtat gtcctgtatc atcaccacct tcggtgtcgg tgaattgtct	240
gccttgaacg gtattgccgg ttcttacgct gaacacgtcg gtgtcttgca cgtcgtcgg	300

gtcccatcca tctcctctca agctaagcaa ttgtttgtgc accacacctt gggtaacggt 360
gacttcactg tcttccacag aatgtccgct aacatctctg agaccaccgc tatggtcact 420

gacatcgcta ccgtccagc tgagatcgac agatgtatca gaaccaccta catcacccaa 480
agaccagtct acttgggtct accagctaac ttggtcgacc taaagggtccc agccaagctt 540
ttggaaaccc caattgactt gtccttgaag ccaaaccgacc cagaagccga aactgaagtc 600
gttgacaccg tcttggaaatt gatcaaggct gctaagaacc cagttatctt ggctgatgct 660
tgtgcttcca gacacgacgt caaggctgaa accaagaagt tgattgacgc cactcaattc 720
ccatccttcg ttaccccaat gggtaagggt tccatcgacg aacaacaccc aagattcggt 780
gggtgtctacg tcggtacctt gtccagacca gaagttaagg aagctgttga atccgctgac 840

ttgatcttgt ctgtcgggtgc ttgttgtcc gatttcaaca ctggttcttt ctcttactct 900
tacaagacca agaacatcgt cgaattccac tctgactaca tcaagatcag aaacgctacc 960
ttcccagggtg tccaaatgaa gtctcgtttg caaaagtgtg tgaacgccgt ccagaagct 1020
atcaagggtt acaagccagt cctgttccca gctagagtcc cagaaaacaa gtctctgtgac 1080
ccagctaccc cattgaagca agaattggatg tggaaccaag tttccaagtt ctgcaagaa 1140
ggtagatgtt ttatcactga aaccggtacc tccgcttttg gtatcaacca aacccattc 1200
ccaaacaacg cttacggtat ctccaagtt ctatgggggtt ccatcggttt caccaccggt 1260

gcttgtttgg gtgccgcttt cgctgctgaa gaaatcgacc caaagaagag agttatcttg 1320
ttcattgggtg acggttcttt gcaattgact gtccaagaaa tctccaccat gatcagatgg 1380
ggcttgaagc catacttggt cgtcttgaac aacgacggtt acaccatcga aagattgatt 1440
cacggtgaaa aggtctggta caacgacatc caaaactggg accacttggc tctattgcca 1500
accttcggtg ctaaggacta cgaaaaccac agagtcgcca ccaccggtga atgggacaag 1560
ttgaccaag acaaggaatt caacaagaac tccaagatca gaatgatcga agttatgttg 1620
ccagttatgg acgtccaac ttccttgatt gaacaagcta agttgaccgc ttccatcaac 1680

gctaagcaag aa 1692

<210> 655
<211> 564
<212> PRT
<213> Artificial Sequence
<220><223> pyruvate decarboxylase from Candida glabrata
<400> 655
Met Ser Glu Ile Thr Leu Gly Arg Tyr Leu Phe Glu Arg Leu Asn Gln

1 5 10 15
 Val Asp Val Lys Thr Ile 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 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 Ala Asn Ile Ser Glu Thr Thr Ala Met Val Thr Asp Ile Ala Thr
 130 135 140
 Ala Pro Ala Glu Ile Asp Arg Cys Ile Arg Thr Thr Tyr Ile Thr Gln
 145 150 155 160
 Arg Pro Val Tyr Leu Gly Leu Pro Ala Asn Leu Val Asp Leu Lys Val

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

 225 230 235 240
 Pro Ser Phe Val Thr Pro Met Gly Lys Gly Ser Ile Asp Glu Gln His
 245 250 255

Pro Arg Phe Gly Gly Val Tyr Val Gly Thr Leu Ser Arg 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 Tyr Ile Lys Ile Arg Asn Ala Thr
305 310 315 320

Phe Pro Gly Val Gln Met Lys Phe Ala Leu Gln Lys Leu Leu Asn Ala
325 330 335

Val Pro Glu Ala Ile Lys Gly Tyr Lys Pro Val Pro Val Pro Ala Arg
340 345 350

Val Pro Glu Asn Lys Ser Cys Asp Pro Ala Thr Pro Leu Lys Gln Glu
355 360 365

Trp Met Trp Asn Gln Val Ser Lys Phe Leu Gln Glu Gly Asp Val Val
370 375 380

Ile Thr Glu Thr Gly Thr Ser Ala Phe Gly Ile Asn Gln Thr Pro Phe
385 390 395 400

Pro Asn Asn Ala Tyr Gly Ile Ser Gln Val Leu Trp Gly Ser Ile Gly
405 410 415

Phe Thr Thr Gly Ala Cys 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 Lys Ala Gly Tyr Asn Asp Ile Gln Asn Trp Asp His Leu
485 490 495

Ala Leu Leu Pro Thr Phe Gly Ala Lys Asp Tyr Glu Asn His Arg Val

500 505 510
Ala Thr Thr Gly Glu Trp Asp Lys Leu Thr Gln Asp Lys Glu Phe Asn
515 520 525
Lys Asn Ser Lys Ile Arg Met Ile Glu Val Met Leu Pro Val Met Asp
530 535 540
Ala Pro Thr Ser Leu Ile Glu Gln Ala Lys Leu Thr Ala Ser Ile Asn

545 550 555 560
Ala Lys Gln Glu

<210> 656

<211> 1788

<212> DNA

<213> Artificial Sequence

<220><223> PDC1 pyruvate decarboxylase from Pichia stipitis

<400> 656

atggctgaag tctcattagg aagatatctc ttcgagagat tgtaccaatt gcaagtgcag	60
accatcttcg gtgtccctgg tgatttcaac ttgtcgcttt tggacaagat ctacgaagtg	120
gaagatgccc atggcaagaa ttcgtttaga tgggctggta atgccaacga attgaatgca	180
tcgtacgctg ctgacgggta ctgagagtc aagcgtttag ggtgtttggt cactaccttt	240
gggtgcggtg aattgtctgc ttgtaattgt attgccggtt cttatgccga acatgttggt	300
ttgcttcattg tcgtagggtg tccatcgatt tctcgcaag ctaagcaatt gttacttcac	360
cacactttgg gtaattggta ttctactgtt ttccatagaa tgtccaacaa cttttctcag	420
accacagcct ttatctccga tatcaactcg gctccagctg aaattgatag atgtatcaga	480
gaggcctacg tcaaacaaag accagtttat atcgggttac cagctaactt agttgatttg	540
aatgttccgg cctctttgct tgagtctcca atcaacttgt cgttggaata gaacgaccca	600
gaggctcaag atgaagtcac tgactctgtc ttagacttga tcaaaaagtc gctgaaccca	660
atcatcttgg tcgatgcctg tgccctcgaga catgactgta aggctgaagt tactcagttg	720
attgaacaaa cccaattccc agtatttctc actccaatgg gtaaaggtag cgttgatgag	780
ggtaggttag acggagaatt gttagaagat gatcctcatt tgattgcaa ggtcgtgct	840
aggttgtctg ctggcaagaa cgctgcctct agattcggag gtgtttatgt cggaaccttg	900
tcgaagcccg aagtcaagga cgctgtagag agtgcagatt tgattttgtc tgtcgggtgcc	960
cttttgtctg atttcaacac tgggtcattt tctactcct acagaaccaa gaacatcgtc	1020

gaattccatt ctgattacac taagattaga caagccactt tcccaggtgt gcagatgaag 1080

gaagccttgc aagaattgaa caagaaagt tcatctgctg ctatgacta tgaagtcaag 1140

cctgtgccca agatcaagtt ggccaataca ccagccacca gagaagtcaa gtttaactcag 1200

gaatggttgt ggaccagagt gtcttcgtgg ttcagagaag gtgatattat taccacgaa 1260

accggtacat cctccttcgg tatagttcaa tccagattcc caaacaacac catcggtatc 1320

tcccaagtat tgtggggttc tattggtttc tctgttggtg ccactttggg tgctgccatg 1380

gtgcccgaag aactcgaccc taacaagaga accatcttgt ttgttgaga tggttctttg 1440

caattgaccg ttcaggaaat ctccaccata atcagatggg gtaccacacc ttaccttttc 1500

gtgttgaaca atgacggtta caccatcgag cgtttgatcc acggtgtaaa tgcctcatat 1560

aatgacatcc aaccatggca aaacttggaa atcttgccca ctttctcggc caagaactac 1620

gacgtgtga gaatctccaa catcgagaaa gcagaagata tcttgaaaga caaggaattc 1680

ggaaagaact ccaagattag attgatagaa gtcattgtac caagattgga tgcaccatct 1740

aaccttgcca aacaagctgc cattacagct gccaccaacg ccgaagct 1788

<210> 657

<211> 596

<212> PRT

<213> Artificial Sequence

<220><223> PDC1 pyruvate decarboxylase from Pichia stipitis

<400> 657

Met Ala Glu Val Ser Leu Gly Arg Tyr Leu Phe Glu Arg Leu Tyr Gln

1 5 10 15

Leu Gln Val Gln Thr Ile Phe Gly Val Pro Gly Asp Phe Asn Leu Ser

20 25 30

Leu Leu Asp Lys Ile Tyr Glu Val Glu Asp Ala His Gly Lys Asn Ser

35 40 45

Phe Arg Trp Ala Gly Asn Ala Asn Glu Leu Asn Ala Ser Tyr Ala Ala

50 55 60

Asp Gly Tyr Ser Arg Val Lys Arg Leu Gly Cys Leu Val Thr Thr Phe

65 70 75 80

Gly Val Gly Glu Leu Ser Ala Leu Asn Gly Ile Ala Gly Ser Tyr Ala

85 90 95

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

340 345 350
 Thr Phe Pro Gly Val Gln Met Lys Glu Ala Leu Gln Glu Leu Asn Lys
 355 360 365
 Lys Val Ser Ser Ala Ala Ser His Tyr Glu Val Lys Pro Val Pro Lys
 370 375 380

 Ile Lys Leu Ala Asn Thr Pro Ala Thr Arg Glu Val Lys Leu Thr Gln
 385 390 395 400
 Glu Trp Leu Trp Thr Arg Val Ser Ser Trp Phe Arg Glu Gly Asp Ile
 405 410 415
 Ile Ile Thr Glu Thr Gly Thr Ser Ser Phe Gly Ile Val Gln Ser Arg
 420 425 430
 Phe Pro Asn Asn Thr Ile Gly Ile Ser Gln Val Leu Trp Gly Ser Ile
 435 440 445

 Gly Phe Ser Val Gly Ala Thr Leu Gly Ala Ala Met Ala Ala Gln Glu
 450 455 460
 Leu Asp Pro Asn Lys Arg Thr Ile Leu Phe Val Gly Asp Gly Ser Leu
 465 470 475 480
 Gln Leu Thr Val Gln Glu Ile Ser Thr Ile Ile Arg Trp Gly Thr Thr
 485 490 495
 Pro Tyr Leu Phe Val Leu Asn Asn Asp Gly Tyr Thr Ile Glu Arg Leu
 500 505 510

 Ile His Gly Val Asn Ala Ser Tyr Asn Asp Ile Gln Pro Trp Gln Asn
 515 520 525
 Leu Glu Ile Leu Pro Thr Phe Ser Ala Lys Asn Tyr Asp Ala Val Arg
 530 535 540
 Ile Ser Asn Ile Gly Glu Ala Glu Asp Ile Leu Lys Asp Lys Glu Phe
 545 550 555 560
 Gly Lys Asn Ser Lys Ile Arg Leu Ile Glu Val Met Leu Pro Arg Leu
 565 570 575

 Asp Ala Pro Ser Asn Leu Ala Lys Gln Ala Ala Ile Thr Ala Ala Thr
 580 585 590

Asn Ala Glu Ala

595

<210> 658

<211> 1707

<212> DNA

<213> Artificial Sequence

<220><223> PDC2 pyruvate decarboxylase from Pichia stipitis

<400> 658

atggtatcaa cctaccaga atcagagggt actctaggaa ggtacctctt tgagcgactc	60
caccaattga aagtggacac cattttcggc ttgccgggtg acttcaacct ttccttattg	120
gacaaagtgt atgaagtcc ggatatgagg tgggctggaa atgccaacga attgaatgct	180
gcctatgctg ccgatgggta ctccagaata aagggtattgt cttgcttggt cacaactttt	240
ggtgttggtg aattgtctgc tttaacgga gttggtgggtg cctatgctga acacgtagga	300
cttctacatg tcgttggagt tccatccata tcgtcacagg ctaaacagtt gttgctccac	360
cataccttgg gtaatgggtg cttcactgtt tttcacagaa tgtccaatag catttctcaa	420
actacagcat ttctctcaga tatctctatt gcaccaggtc aaatagatag atgcatcaga	480
gaagcatatg ttcatcagag accagtttat gttgggtttac cggcaaatat ggttgatctc	540
aaggttcctt ctagtctctt agaaactcca attgatttga aattgaaaca aaatgatcct	600
gaagctcaag aagtgttga aacagtcttg aagtgtgtgt cccaagctac aaacccatt	660
atcttggtag acgttgtgc cctcagacac aattgcaaag aggaagtcaa acaattggtt	720
gatgccacta attttcaagt ctttacaact ccaatgggta aatctggtat ctccgaatct	780
catccaagat tgggcgggtg ctatgtcggg acaatgtcga gtcttcaagt caaaaaagcc	840
gttgaaaatg ccgatcttat actatctgtt ggttcgttgt tatcggactt caatacaggt	900
tcattttcat actcctacaa gacgaagaat gttgttgaat tccactctga ctatatgaaa	960
atcagacagg ccaccttccc aggagttaa atgaaagaag ccttgcaaca gttgataaaa	1020
agggtctctt cttacatcaa tccaagctac attcctactc gagttcctaa aaggaaacag	1080
ccattgaaag ctccatcaga agctcctttg acccaagaat atttgtgtgc taaagtatcc	1140
ggctggttta gagagggtga tattatcgta accgaaactg gtacatctgc tttcggaatt	1200
attcaatccc attttccag caacactatc ggtatatccc aagtcttgtg gggctcaatt	1260
ggtttcacag taggtgcaac agttgggtgct gccatggcag cccaggaaat cgaccctagc	1320
aggagagtaa ttttgttcgt cggtgatggt tcattgcagt tgacggttca ggaaatctct	1380

acgttggtga aatgggattg taacaatact tatctttacg tgttgaacaa tgatggttac 1440

actatagaaa ggttgatcca cggcaaaagt gccagctaca acgatataca gccttggaac 1500

catttatect tgccttcgtt attcaatgct aagaaatacc aaaatgtcag agtatcgact 1560

gctggagaat tggactcttt gttctctgat aagaaatttg cttctccaga taggataaga 1620

atgattgagg tgatgttacc gagattggat gcaccagcaa atcttggtgc tcaagcaaag 1680

ttgtctgaac gggtaaacct tgaaaaat 1707

<210> 659

<211> 569

<212> PRT

<213> Artificial Sequence

<220><223> PDC2 pyruvate decarboxylase from Pichia stipitis

<400> 659

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

1 5 10 15

Phe Glu Arg Leu His Gln Leu Lys Val Asp Thr Ile Phe Gly Leu Pro

20 25 30

Gly Asp Phe Asn Leu Ser Leu Leu Asp Lys Val Tyr Glu Val Pro Asp

35 40 45

Met Arg Trp Ala Gly Asn Ala Asn Glu Leu Asn Ala Ala Tyr Ala Ala

50 55 60

Asp Gly Tyr Ser Arg Ile Lys Gly Leu Ser Cys Leu Val Thr Thr Phe

65 70 75 80

Gly Val Gly Glu Leu Ser Ala Leu Asn Gly Val Gly Gly Ala Tyr Ala

85 90 95

Glu His Val Gly Leu Leu His Val Val Gly Val Pro Ser Ile Ser Ser

100 105 110

Gln Ala Lys Gln Leu Leu Leu His His Thr Leu Gly Asn Gly Asp Phe

115 120 125

Thr Val Phe His Arg Met Ser Asn Ser Ile Ser Gln Thr Thr Ala Phe

130 135 140

Leu Ser Asp Ile Ser Ile Ala Pro Gly Gln Ile Asp Arg Cys Ile Arg

145 150 155 160
 Glu Ala Tyr Val His Gln Arg Pro Val Tyr Val Gly Leu Pro Ala Asn
 165 170 175
 Met Val Asp Leu Lys Val Pro Ser Ser Leu Leu Glu Thr Pro Ile Asp
 180 185 190

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

 Ile Ser Glu Ser His Pro Arg Leu Gly Gly Val Tyr Val Gly Thr Met
 260 265 270
 Ser Ser Pro Gln Val Lys Lys Ala Val Glu Asn Ala Asp Leu Ile Leu
 275 280 285
 Ser Val Gly Ser Leu Leu Ser Asp Phe Asn Thr Gly Ser Phe Ser Tyr
 290 295 300
 Ser Tyr Lys Thr Lys Asn Val Val Glu Phe His Ser Asp Tyr Met Lys
 305 310 315 320

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

 Glu Gly Asp Ile Ile Val Thr Glu Thr Gly Thr Ser Ala Phe Gly Ile
 385 390 395 400

Ile Gln Ser His Phe Pro Ser Asn Thr Ile Gly Ile Ser Gln Val Leu
405 410 415
Trp Gly Ser Ile Gly Phe Thr Val Gly Ala Thr Val Gly Ala Ala Met
420 425 430
Ala Ala Gln Glu Ile Asp Pro Ser Arg Arg Val Ile Leu Phe Val Gly
435 440 445
Asp Gly Ser Leu Gln Leu Thr Val Gln Glu Ile Ser Thr Leu Cys Lys
450 455 460
Trp Asp Cys Asn Asn Thr Tyr Leu Tyr Val Leu Asn Asn Asp Gly Tyr
465 470 475 480
Thr Ile Glu Arg Leu Ile His Gly Lys Ser Ala Ser Tyr Asn Asp Ile
485 490 495
Gln Pro Trp Asn His Leu Ser Leu Leu Arg Leu Phe Asn Ala Lys Lys
500 505 510
Tyr Gln Asn Val Arg Val Ser Thr Ala Gly Glu Leu Asp Ser Leu Phe
515 520 525
Ser Asp Lys Lys Phe Ala Ser Pro Asp Arg Ile Arg Met Ile Glu Val
530 535 540
Met Leu Ser Arg Leu Asp Ala Pro Ala Asn Leu Val Ala Gln Ala Lys
545 550 555 560
Leu Ser Glu Arg Val Asn Leu Glu Asn
565

<210> 660

<211> 1692

<212> DNA

<213> Artificial Sequence

<220><223> pyruvate decarboxylase from Kluyveromyces lactis

<400> 660

atgtctgaaa ttacattagg tcgttacttg ttcgaaagat taaagcaagt cgaagttcaa	60
accatctttg gtctaccagg tgatttcaac ttgtccctat tggacaatat ctacgaagtc	120
ccaggtatga gatgggctgg taatgccaac gaattgaacg ctgcttacgc tgctgatggt	180
tacgccagat taaagggtat gtcctgtatc atcaccacct tcggtgtcgg tgaattgtct	240

gctttgaacg gtattgccgg ttcttacgct gaacacgttg gtgtcttgca cgttgtcggt 300
gttccatcgg tctctttctca agctaagcaa ttgtttgtgc accacacctt gggtaacggt 360

gacttcactg ttttccacag aatgtgctcc aacatttctg aaaccactgc tatgatcacc 420
gatatcaaca ctgccccagc tgaaatcgac agatgtatca gaaccactta cgtttcccaa 480
agaccagtct acttgggttt gccagctaac ttggctgact tgactgtccc agcttctttg 540
ttggacactc caattgattt gagcttgaag ccaaatgacc cagaagccga agaagaagtc 600
atcgaaaacg tcttgcaact gatcaaggaa gctaagaacc cagttatctt ggctgatgct 660
tgtttgtcca gacacgatgc caagctgag accaagaagt tgatcgactt gactcaattc 720
ccagccttcg ttacccaat gggttaagggt tccattgacg aaaagcacc aagattcggt 780

ggtgtctacg tcggtaccct atctttctca gctgtcaagg aagccgttga atctgtcac 840
ttggttctat cggtcggtgc tctattgtcc gatttcaaca ctggttcttt ctcttactct 900
tacaagacca agaacattgt cgaattccac tctgactaca ccaagatcag aaggcctacc 960
ttcccagggtg tccaaatgaa gtctgcttta caaaaattgt tgactaaggt tgccgatgct 1020
gctaagggtt acaagccagt tccagttcca tctgaaccag aacacaacga agatgtcgct 1080
gactccactc cattgaagca agaatgggtc tggactcaag tcggtgaatt cttgagagaa 1140
ggtgatgttg ttatcactga aaccggtacc tctgccttcg gtatcaacca aactcatttc 1200

ccaaacaaca catacgggat ctctcaagtt ttatgggggtt ccattgggtt caccactggt 1260
gctaccttgg gtgtgcctt cgctgccgaa gaaattgatc caaagaagag agttatctta 1320
ttcattggtg acggttcttt gcaattgact gttaagaaa tctccaccat gatcagatgg 1380
ggcttgaagc catacttggt cgtattgaac aacgacggtt acaccattga aagattgatt 1440
cacggtgaaa cgctcaata caactgtatc caaaactggc aacacttgga attattgcca 1500
actttcggtg ccaaggacta cgaagctgtc agagtttcca ccaactggtga atggaacaag 1560
ttgaccactg acgaaaagtt ccaagacaac accagaatca gattgatcga agttatgttg 1620

ccaactatgg atgtccatc taacttggtt aagcaagctc aattgactgc tgcaccaac 1680
gctaagaact aa 1692

<210> 661
<211> 563
<212> PRT
<213> Artificial Sequence
<220><223> pyruvate decarboxylase from Kluyveromyces lactis
<400> 661

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
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
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
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> 662

<211> 1716

<212> DNA

<213> Artificial Sequence

<220><223> pyruvate decarboxylase from *Yarrowia lipolytica*

<400> 662

atgagcgcact ccgaacccca aatggctcgac ctgggcgcact atctctttgc cgcattcaag	60
cagctaggcg tggactccgt ctttggagtg cccggcgact tcaacctcac cctgttggac	120
cacgtgtaca atgtcgacat gcggtgggtt gggaacacaa acgagctgaa tgccggtac	180
tcggccgacg gctactcccg ggtcaagcgg ctggcatgtc ttgtcaccac ctttggcgtg	240
ggagagctgt ctgccgtggc tgctgtggca ggctcgtacg ccgagcatgt gggcgtgggtg	300
catgttgtgg gcgttccag cacctctgct gagaacaagc atctgctgct gcaccacaca	360
ctcggtaacg gcgacttccg ggtctttgcc cagatgtcca aatcatctc cgagtacacc	420
caccatattg aggaccccgag cgagctgcc gacgtaatcg acaccgcat ccgaatcgcc	480
tacaccacac agcggcccgt ttacattgct gtgcctcca acttctccga ggtcgatatt	540
gccgaccagg ctagactgga taccctctg gacctttcgc tgcagcccaa cgaccccgag	600
agccagtacg aggtgattga ggagatttgc tcgcgtatca aggccgcaa gaagcccgtg	660
attctcgtcg acgcctgcgc ttgcgatac agatgtgtgg acgagaccaa ggagctggcc	720
aagatcacca actttgccta ctttgcact cccatgggta agggttctgt ggacgaggat	780
actgaccggt acggaggaac atacgtcgga tcgctgactg ctctgctac tgccgaggtg	840
gttgagacag ctgatctcat catctccgta ggagctcttc tgtcggactt caacaccggt	900
tccttctcgt actcctactc cacaaaaaac gtggtggaat tgcattcgga ccacgtcaaa	960

atcaagtcgc ccacctacaa caacgtcggc atgaaaatgc tgttcccgcc cctgctcgaa 1020

gccgtcaaga aactggttgc cgagaccct gactttgcat ccaaggctct ggctgttccc 1080

gacaccactc ccaagatccc cgaggtaccc gatgatcaca ttacgacca ggcattggctg 1140

tggcagcgtc tcagttactt tctgaggccc accgacatcg tggtcaccga gaccggaacc 1200

tcgtcctttg gaatcatcca gaccaagtgc ccccaaacg tccgaggtat ctgcaggtg 1260

ctgtggggct ctattggata ctgggtggga gcagcctgtg gagcctccat tgctgcacag 1320

gagattgacc ccagcagcg agtgattctg tttgtggcg acggctctct tcagctgacg 1380

gtgaccgaga tctcgtgcat gatccgaac aacgtcaagc cgtacatttt tgtgctcaac 1440

aacgacggct acaccatcga gaggtcatt cacggcgaac acgcctcgta caacgatgtg 1500

cacatgtgga agtactccaa gattctcgac acgttcaacg ccaaggccca cgagtcgatt 1560

tggtcaaca ccaaggcgca gatggacgt ctgttcgaca acgaagagtt tgccaagccc 1620

gacaagatcc gggtcattga ggtcatgtgc gacaagatgg acgcgcctgc ctgctgac 1680

aagcaggctg agctctctgc caagaccaac gtttag 1716

<210> 663

<211> 571

<212> PRT

<213> Artificial Sequence

<220><223> pyruvate decarboxylase from Yarrowia lipolytica

<400> 663

Met Ser Asp Ser Glu Pro Gln Met Val Asp Leu Gly Asp Tyr Leu Phe

1 5 10 15

Ala Arg Phe Lys Gln Leu Gly Val Asp Ser Val Phe Gly Val Pro Gly

20 25 30

Asp Phe Asn Leu Thr Leu Leu Asp His Val Tyr Asn Val Asp Met Arg

35 40 45

Trp Val Gly Asn Thr Asn Glu Leu Asn Ala Gly Tyr Ser Ala Asp Gly

50 55 60

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

65 70 75 80

Gly Glu Leu Ser Ala Val Ala Ala Val Ala Gly Ser Tyr Ala Glu His

85 90 95

Val Gly Val Val His Val Val Gly Val Pro Ser Thr Ser Ala Glu Asn
100 105 110

Lys His Leu Leu Leu His His Thr Leu Gly Asn Gly Asp Phe Arg Val
115 120 125

Phe Ala Gln Met Ser Lys Leu Ile Ser Glu Tyr Thr His His Ile Glu
130 135 140

Asp Pro Ser Glu Ala Ala Asp Val Ile Asp Thr Ala Ile Arg Ile Ala
145 150 155 160

Tyr Thr His Gln Arg Pro Val Tyr Ile Ala Val Pro Ser Asn Phe Ser
165 170 175

Glu Val Asp Ile Ala Asp Gln Ala Arg Leu Asp Thr Pro Leu Asp Leu
180 185 190

Ser Leu Gln Pro Asn Asp Pro Glu Ser Gln Tyr Glu Val Ile Glu Glu
195 200 205

Ile Cys Ser Arg Ile Lys Ala Ala Lys Lys Pro Val Ile Leu Val Asp
210 215 220

Ala Cys Ala Ser Arg Tyr Arg Cys Val Asp Glu Thr Lys Glu Leu Ala
225 230 235 240

Lys Ile Thr Asn Phe Ala Tyr Phe Val Thr Pro Met Gly Lys Gly Ser
245 250 255

Val Asp Glu Asp Thr Asp Arg Tyr Gly Gly Thr Tyr Val Gly Ser Leu
260 265 270

Thr Ala Pro Ala Thr Ala Glu Val Val Glu Thr Ala Asp Leu Ile Ile
275 280 285

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

Ser Tyr Ser Thr Lys Asn Val Val Glu Leu His Ser Asp His Val Lys
305 310 315 320

Ile Lys Ser Ala Thr Tyr Asn Asn Val Gly Met Lys Met Leu Phe Pro
325 330 335

Pro Leu Leu Glu Ala Val Lys Lys Leu Val Ala Glu Thr Pro Asp Phe

340 345 350
 Ala Ser Lys Ala Leu Ala Val Pro Asp Thr Thr Pro Lys Ile Pro Glu
 355 360 365
 Val Pro Asp Asp His Ile Thr Thr Gln Ala Trp Leu Trp Gln Arg Leu
 370 375 380

 Ser Tyr Phe Leu Arg Pro Thr Asp Ile Val Val Thr Glu Thr Gly Thr
 385 390 395 400
 Ser Ser Phe Gly Ile Ile Gln Thr Lys Phe Pro His Asn Val Arg Gly
 405 410 415
 Ile Ser Gln Val Leu Trp Gly Ser Ile Gly Tyr Ser Val Gly Ala Ala
 420 425 430
 Cys Gly Ala Ser Ile Ala Ala Gln Glu Ile Asp Pro Gln Gln Arg Val
 435 440 445

 Ile Leu Phe Val Gly Asp Gly Ser Leu Gln Leu Thr Val Thr Glu Ile
 450 455 460
 Ser Cys Met Ile Arg Asn Asn Val Lys Pro Tyr Ile Phe Val Leu Asn
 465 470 475 480
 Asn Asp Gly Tyr Thr Ile Glu Arg Leu Ile His Gly Glu Asn Ala Ser
 485 490 495
 Tyr Asn Asp Val His Met Trp Lys Tyr Ser Lys Ile Leu Asp Thr Phe
 500 505 510

 Asn Ala Lys Ala His Glu Ser Ile Val Val Asn Thr Lys Gly Glu Met
 515 520 525
 Asp Ala Leu Phe Asp Asn Glu Glu Phe Ala Lys Pro Asp Lys Ile Arg
 530 535 540
 Leu Ile Glu Val Met Cys Asp Lys Met Asp Ala Pro Ala Ser Leu Ile
 545 550 555 560
 Lys Gln Ala Glu Leu Ser Ala Lys Thr Asn Val
 565 570

<210> 664

<211> 1716

<212> DNA

<213> Artificial Sequence

<220><223> pyruvate decarboxylase from Schizosaccharomyces pombe

<400> 664

atgagtgggg atattttagt cggatgaatat ctattcaaaa ggcttgaaca attaggggtc	60
aagtccattc ttggtgttcc aggagatttc aatttagctc tacttgactt aattgagaaa	120
gttgagatg agaaatttcg ttgggttggc aataccaatg agttgaatgg tgcttatgcc	180
gctgatggtt atgctcgtgt taatgggtctt tcagccattg ttacaacgtt cggcgtggga	240
gagctttccg ctattaatgg agtggcaggt tcttatgcgg agcatgtccc agtagttcat	300
attgttggaa tgccttcac aaaggtgcaa gatactggag ctttgcttca tcatacttta	360
ggagatggag actttcgac tttcatggat atgtttaaga aagtttctgc ctacagtata	420
atgatcgata acggaacga tgcagctgaa aagatcgatg aagccttgct gatttgttat	480
aaaaaggcta ggctgttta catttggtatt ccttctgatg ctggctactt caaagcatct	540
tcatcaaate ttgggaaaag actaaagctc gaggaggata ctaacgatcc agcagttgag	600
caagaagtca tcaatcatat ctcggaatg gttgtcaatg caaagaaacc agtgatttta	660
attgacgctt gtgctgtaag acatcgtgtc gttccagaag tacatgagct gattaaattg	720
accttttcc ctacatatgt aactcccatg ggtaaactcg caattgacga aacttcgcaa	780
ttttttgacg gcgtttatgt tggttcaatt tcagatcctg aagttaaaga cagaattgaa	840
tccactgatc tgttgcatac catcggtgct ctcaaatcag actttaacac gggttccttc	900
tcttaccacc tcagccaaaa gaatgccgtt gagtttcatt cagaccacat gcgcattcga	960
tatgctcttt atccaaatgt agccatgaag tatattcttc gcaaactgtt gaaagtactt	1020
gatgcttcta tgtgtcattc caagctgct cctaccattg gctacaacat caagcctaag	1080
catgcggaag gatattcttc caacgagatt actcattgct ggttttggcc taaatttagt	1140
gaatttttga agccccgaga tgttttgatc accgagactg gaactgcaa ctttggtgtc	1200
cttgattgca ggtttccaaa ggatgtaaca gccatttccc aggtattatg gggatctatt	1260
ggatactccg ttggtgcaat gtttgggtgct gttttggccg tccacgattc taaagagccc	1320
gatcgtcgta ccattcttgt agtaggtgat ggatccttac aactgacgat tacagagatt	1380
tcaacctgca ttcgccataa cctcaaacca attattttca taattaacaa cgacggttac	1440
accattgagc gttaatttca tggtttgcgt gctagctata acgaaattaa cactaaatgg	1500
ggctaccaac agattcccaa gtttttcgga gctgctgaaa accacttccg cacttactgt	1560
gttaaaactc ctactgacgt tgaaaagttg tttagcgaca aggagtttgc aaatgcagat	1620

gtcattcaag tagttgagct tgtaatgcct atgttggatg cacctcgtgt cctagttgag 1680
caagccaagt tgacgtctaa gatcaataag caatga 1716

<210> 665

<211> 571

<212> PRT

<213> Artificial Sequence

<220><223> pyruvate decarboxylase from Schizosaccharomyces pombe

<400> 665

Met Ser Gly Asp Ile Leu Val Gly Glu Tyr Leu Phe Lys Arg Leu Glu

1 5 10 15

Gln Leu Gly Val Lys Ser Ile Leu Gly Val Pro Gly Asp Phe Asn Leu

20 25 30

Ala Leu Leu Asp Leu Ile Glu Lys Val Gly Asp Glu Lys Phe Arg Trp

35 40 45

Val Gly Asn Thr Asn Glu Leu Asn Gly Ala Tyr Ala Ala Asp Gly Tyr

50 55 60

Ala Arg Val Asn Gly Leu Ser Ala Ile Val Thr Thr Phe Gly Val Gly

65 70 75 80

Glu Leu Ser Ala Ile Asn Gly Val Ala Gly Ser Tyr Ala Glu His Val

85 90 95

Pro Val Val His Ile Val Gly Met Pro Ser Thr Lys Val Gln Asp Thr

100 105 110

Gly Ala Leu Leu His His Thr Leu Gly Asp Gly Asp Phe Arg Thr Phe

115 120 125

Met Asp Met Phe Lys Lys Val Ser Ala Tyr Ser Ile Met Ile Asp Asn

130 135 140

Gly Asn Asp Ala Ala Glu Lys Ile Asp Glu Ala Leu Ser Ile Cys Tyr

145 150 155 160

Lys Lys Ala Arg Pro Val Tyr Ile Gly Ile Pro Ser Asp Ala Gly Tyr

165 170 175

Phe Lys Ala Ser Ser Ser Asn Leu Gly Lys Arg Leu Lys Leu Glu Glu

180 185 190

Asp Thr Asn Asp Pro Ala Val Glu Gln Glu Val Ile Asn His Ile Ser
 195 200 205
 Glu Met Val Val Asn Ala Lys Lys Pro Val Ile Leu Ile Asp Ala Cys
 210 215 220

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

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

 Ile Gly Tyr Asn Ile Lys Pro Lys His Ala Glu Gly Tyr Ser Ser Asn
 355 360 365
 Glu Ile Thr His Cys Trp Phe Trp Pro Lys Phe Ser Glu Phe Leu Lys
 370 375 380
 Pro Arg Asp Val Leu Ile Thr Glu Thr Gly Thr Ala Asn Phe Gly Val
 385 390 395 400
 Leu Asp Cys Arg Phe Pro Lys Asp Val Thr Ala Ile Ser Gln Val Leu
 405 410 415

 Trp Gly Ser Ile Gly Tyr Ser Val Gly Ala Met Phe Gly Ala Val Leu
 420 425 430
 Ala Val His Asp Ser Lys Glu Pro Asp Arg Arg Thr Ile Leu Val Val

435 440 445
 Gly Asp Gly Ser Leu Gln Leu Thr Ile Thr Glu Ile Ser Thr Cys Ile
 450 455 460
 Arg His Asn Leu Lys Pro Ile Ile Phe Ile Ile Asn Asn Asp Gly Tyr
 465 470 475 480

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

 Val Glu Leu Val Met Pro Met Leu Asp Ala Pro Arg Val Leu Val Glu
 545 550 555 560
 Gln Ala Lys Leu Thr Ser Lys Ile Asn Lys Gln
 565 570

 <210> 666
 <211> 1689
 <212> DNA
 <213> Artificial Sequence
 <220><223> pyruvate decarboxylase from *Zygosaccharomyces rouxii*
 <400> 666
 atgtctgaaa ttactctagg tcgttacttg ttcgaaagat taaagcaagt tgacactaac 60
 accatcttcg gtgttccagg tgacttcaac ttgtccttgt tggacaaggt ctacgaagtg 120

 caaggtctaa gatgggctgg taacgctaac gaattgaacg ctgcctacgc tgctgacggt 180
 tacgccagag ttaaggggtt ggctgctttg atcaccacct tcggtgtcgg tgaattgtct 240
 gctttgaacg gtattgcagg ttcttacgct gaacacgttg gtgttttgca cattgttggt 300
 gttccatctg tctcttctca agctaagcaa ttgttgttgc accacacctt gggtaacggt 360
 gacttcactg ttttccacag aatgtccgcc aacatctctg aaaccaccgc tatgttgacc 420
 gacatcactg ctgctccagc tgaattgac cgttgcatca gagttgctta cgtaaccaa 480

agaccagtct acttgggtct accagctaac ttggttgacc aaaaggtccc agcttctttg 540

ttgaacactc caattgatct atctctaaag gagaacgacc cagaagctga aaccgaagtt 600

gttgacaccg ttttgaatt gatcaaggaa gctaagaacc cagttatctt ggctgatgct 660

tgctgtcca gacacgacgt caaggctgaa accaagaagt tgatcgactt gactcaattc 720

ccatctttcg ttactcttat gggtaagggt tccatcgacg aacaaaaccc aagattcgggt 780

ggtgtctacg tcggtactct atccagccca gaagttaagg aagctgttga atctgctgac 840

ttggttctat ctgtcgggtc tctattgtcc gatttcaaca ctggttcttt ctcttactct 900

tacaagacca agaacgttgt tgaattccac tctgaccaca tcaagatcag aaacgtacc 960

ttcccagggtg ttcaaatgaa attcgttttg aagaaactat tgcaagctgt cccagaagct 1020

gtcaagaact acaagccagg tccagtccca gctccgccat ctccaaacgc tgaagttgct 1080

gactctacca ccttgaagca agaatgggta tggagacaag tcggtagctt cttgagagaa 1140

ggtgatgttg ttattaccga aactggtacc tctgctttcg gtatcaacca aactcacttc 1200

cctaaccaaa cttacgggtat ctctcaagtc ttgtgggggt ctattgggta caccactggt 1260

tccactttgg gtgtgcctt cgctgctgaa gaaattgacc ctaagaagag agttatcttg 1320

ttcattgggtg acggttctct acaattgacc gttcaagaaa tctccaccat gatcagatgg 1380

ggtctaaagc catacttggt cgttttgaac aacgatgggt acaccattga aagattgatt 1440

cacggtgaaa ccgtgaata caactgtatc caaccatgga agcacttgga attgttgaac 1500

accttcgggt ccaaggacta cgaaaaccac agagtctcca ctgtcgggtga atggaacaag 1560

ttgactcaag atccaaaatt caacgaaaac tctagaatta gaatgatcga agttatgctt 1620

gaagtcatgg acgtccatc ttctttgggt gctcaagctc aattgaccgc tgctactaac 1680

gctaagcaa 1689

<210> 667

<211> 563

<212> PRT

<213> Artificial Sequence

<220><223> pyruvate decarboxylase from *Zygosaccharomyces rouxii*

<400> 667

Met Ser Glu Ile Thr Leu Gly Arg Tyr Leu Phe Glu Arg Leu Lys Gln

1 5 10 15

Val Asp Thr Asn Thr Ile Phe Gly Val Pro Gly Asp Phe Asn Leu Ser

20 25 30

Leu Leu Asp Lys Val Tyr Glu Val Gln Gly Leu Arg Trp Ala Gly Asn
 35 40 45
 Ala Asn Glu Leu Asn Ala Ala Tyr Ala Ala Asp Gly Tyr Ala Arg Val
 50 55 60
 Lys Gly Leu Ala Ala Leu 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 Ile 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 Ala Asn Ile Ser Glu Thr Thr Ala Met Leu Thr Asp Ile Thr Ala
 130 135 140
 Ala Pro Ala Glu Ile Asp Arg Cys Ile Arg Val Ala Tyr Val Asn Gln
 145 150 155 160
 Arg Pro Val Tyr Leu Gly Leu Pro Ala Asn Leu Val Asp Gln Lys Val
 165 170 175
 Pro Ala Ser Leu Leu Asn Thr Pro Ile Asp Leu Ser Leu Lys Glu Asn
 180 185 190
 Asp Pro Glu Ala Glu Thr Glu Val Val Asp Thr Val Leu Glu 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 Val Lys Ala Glu Thr Lys Lys Leu Ile Asp Leu Thr Gln Phe
 225 230 235 240
 Pro Ser Phe Val Thr Pro Met Gly Lys Gly Ser Ile Asp Glu Gln Asn
 245 250 255
 Pro Arg Phe Gly Gly Val Tyr Val Gly Thr Leu Ser Ser Pro Glu 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 Val Val Glu Phe His Ser Asp His Ile Lys Ile Arg Asn Ala Thr

 305 310 315 320
 Phe Pro Gly Val Gln Met Lys Phe Val Leu Lys Lys Leu Leu Gln Ala
 325 330 335
 Val Pro Glu Ala Val Lys Asn Tyr Lys Pro Gly Pro Val Pro Ala Pro
 340 345 350
 Pro Ser Pro Asn Ala Glu Val Ala Asp Ser Thr Thr Leu Lys Gln Glu
 355 360 365
 Trp Leu Trp Arg Gln Val Gly Ser 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 Gln Thr Tyr Gly Ile Ser Gln Val Leu Trp Gly Ser Ile Gly
 405 410 415
 Tyr Thr Thr Gly Ser 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 Glu Tyr Asn Cys Ile Gln Pro Trp Lys His Leu
 485 490 495
 Glu Leu Leu Asn Thr Phe Gly Ala Lys Asp Tyr Glu Asn His Arg Val

 500 505 510
 Ser Thr Val Gly Glu Trp Asn Lys Leu Thr Gln Asp Pro Lys Phe Asn
 515 520 525

Glu Asn Ser Arg Ile Arg Met Ile Glu Val Met Leu Glu Val Met Asp
 530 535 540
 Ala Pro Ser Ser Leu Val Ala Gln Ala Gln Leu Thr Ala Ala Thr Asn
 545 550 555 560
 Ala Lys Gln

<210> 668

<211> 1461

<212> DNA

<213> Artificial Sequence

<220><223> HXK2 from *S. cerevisiae*

<400> 668

atggttcatt taggtccaaa aaaaccacaa gccagaaagg gttccatggc cgatgtgcc	60
aaggaattga tgcaacaaat tgagaatttt gaaaaaattt tcactgttcc aactgaaact	120
ttacaagccg ttaccaagca cttcatttcc gaattggaaa agggtttgtc caagaagggt	180
ggtaacattc caatgattcc aggttgggtt atggatttcc caactggtaa ggaatccggt	240
gatttcttgg ccattgattt ggggtgtacc aacttgagag ttgtcttagt caagttgggc	300
ggtgaccgta ctttgacac cactcaatct aagtacagat taccagatgc tatgagaact	360
actcaaaatc cagacgaatt gtgggaattt attgccgact ctttgaaagc ttttattgat	420
gagcaattcc cacaaggtat ctctgagcca attccattgg gtttcacctt ttccttccca	480
gcttctcaaa acaaaatcaa tgaaggtatc ttgcaaagat ggactaaagg ttttgatatt	540
ccaaacattg aaaaccacga tgttgttcca atgttgcaaa agcaaatcac taagaggaat	600
atcccaattg aagtgtgtgc ttgataaac gacactaccg gtacttttgt tgcttcttac	660
tacactgacc cagaaactaa gatgggtgtt atcttcggta ctggtgtcaa tgggtgcttac	720
tacgatgttt gttccgatat cgaaaagcta caaggaaaac tatctgatga cattccacca	780
tctgctccaa tggccatcaa ctgtgaatac gggttccttcg ataatgaaca tgtcgttttg	840
ccaagaacta aatacgaat caccattgat gaagaatctc caagaccagg ccaacaaacc	900
tttgaaaaaa tgccttctgg ttactactta ggtgaaattt tgcgtttggc cttgatggac	960
atgtacaaac aaggtttcat cttcaagaac caagacttgt ctaagttcga caagcctttc	1020
gtcatggaca cttcttacc agccagaatc gaggaagatc cattcgagaa cctagaagat	1080
accgatgact tgttccaaaa tgagttcgggt atcaacacta ctgttcaaga acgtaaattg	1140

atcacagcgtt taictgaatt gattgggtgct agagctgcta gattgtccgt ttgtgggtatt 1200

gctgctatct gtcaaaagag aggttacaag accggtcaca tcgctgcaga cggttccgtt 1260

tacaacagat acccaggttt caaagaaaag gctgccaatg ctttgaagga catttacggc 1320

tggactcaaa cctcactaga cgactaccca atcaagattg ttcctgctga agatgggttcc 1380

ggtgctgggtg ccgctgttat tgctgctttg gcccaaaaaa gaattgctga aggtaagtcc 1440

gttggtatca tcggtgctta a 1461

<210> 669

<211> 486

<212> PRT

<213> Artificial Sequence

<220><223> HXK2 from *S. cerevisiae*

<400> 669

Met Val His Leu Gly Pro Lys Lys Pro Gln Ala Arg Lys Gly Ser Met

1 5 10 15

Ala Asp Val Pro Lys Glu Leu Met Gln Gln Ile Glu Asn Phe Glu Lys

20 25 30

Ile Phe Thr Val Pro Thr Glu Thr Leu Gln Ala Val Thr Lys His Phe

35 40 45

Ile Ser Glu Leu Glu Lys Gly Leu Ser Lys Lys Gly Gly Asn Ile Pro

50 55 60

Met Ile Pro Gly Trp Val Met Asp Phe Pro Thr Gly Lys Glu Ser Gly

65 70 75 80

Asp Phe Leu Ala Ile Asp Leu Gly Gly Thr Asn Leu Arg Val Val Leu

85 90 95

Val Lys Leu Gly Gly Asp Arg Thr Phe Asp Thr Thr Gln Ser Lys Tyr

100 105 110

Arg Leu Pro Asp Ala Met Arg Thr Thr Gln Asn Pro Asp Glu Leu Trp

115 120 125

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

130 135 140

Gln Gly Ile Ser Glu Pro Ile Pro Leu Gly Phe Thr Phe Ser Phe Pro

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

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

 260 265 270
 Phe Asp Asn Glu His Val Val Leu Pro Arg Thr Lys Tyr Asp Ile Thr
 275 280 285
 Ile Asp Glu Glu Ser Pro Arg Pro Gly Gln Gln Thr Phe Glu Lys Met
 290 295 300
 Ser Ser Gly Tyr Tyr Leu Gly Glu Ile Leu Arg Leu Ala Leu Met Asp
 305 310 315 320
 Met Tyr Lys Gln Gly Phe Ile Phe Lys Asn Gln Asp Leu Ser Lys Phe

 325 330 335
 Asp Lys Pro Phe Val Met Asp Thr Ser Tyr Pro Ala Arg Ile Glu Glu
 340 345 350
 Asp Pro Phe Glu Asn Leu Glu Asp Thr Asp Asp Leu Phe Gln Asn Glu
 355 360 365
 Phe Gly Ile Asn Thr Thr Val Gln Glu Arg Lys Leu Ile Arg Arg Leu
 370 375 380
 Ser Glu Leu Ile Gly Ala Arg Ala Ala Arg Leu Ser Val Cys Gly Ile

 385 390 395 400

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

450 455 460
 Ala Val Ile Ala Ala Leu Ala Gln Lys Arg Ile Ala Glu Gly Lys Ser
 465 470 475 480
 Val Gly Ile Ile Gly Ala
 485

<210> 670

<211> 1458

<212> DNA

<213> Artificial Sequence

<220><223> RAG5 from K. lactis

<400> 670

atggttcgtt taggtccaaa gaagcctcca gccagaaagg ggtccatggc agatgtgcca 60
 gctaatttga tggaacaaat ccacggtttg gaaactttgt tcaccgtctc ttcagaaaaa 120

atgagaagca ttgtcaagca tttcatcagt gaattggaca aaggtttgtc caaaaagggt 180
 ggtaacattc ctatgattcc aggttgggtt gttgagtatc caactggtaa ggaaactggt 240
 gatttcttag ctcttgattt ggggtgtacc aacttgagag ttgtgttggt taaattgggt 300
 ggtaatcatg atttcgacac cactcaaac aagtacagat taccagacca ttgagaact 360
 ggtacttctg aacaatttg gtcatttatt gcaaagtgtt tgaaggaatt cgtcgaatgaa 420
 tggtagccag atggtgtttc tgaaccattg ccattgggtt tcactttctc ataccctgca 480
 tctcaaaaga agatcaattc cgggtgtgtt caacgttga ccaaggggtt cgatattgaa 540

gggtgtgaag gtcacgatgt tgttccaatg ctacaagaac agattgaaaa gctgaatata 600
 ccaatcaatg tcgttcgatt gatcaacgat accactggta ccttgggtgc ctctttgtac 660
 actgatcctc aaactaagat ggggtatcatt atcgggtactg gtgtcaacgg tgcttactac 720
 gatgttgttt ctggtattga gaaattggaa ggtttgttgc cagaagatat cggtagcat 780
 tctccaatgg caatcaactg tgaatatggt tccttcgata acgaacattt ggtgttgcca 840

agaaccaa at acgatgttat aatcgatgaa gaatctccaa gaccaggtca acaagctttc 900
gaaaagatga ctcttggtta ctatctaggt gaaatcatgc gtctagtact attggacttg 960

tagcacagtg gtttcatctt taaggaccaa gatatctcca agttgaaaga ggcttacgtc 1020
atggacacca gttatccatc taagatcgaa gatgatccat tcgaaaactt ggaagacact 1080
gacgatctgt tcaagactaa ctggaacatc gaaactaccg ttgttgagag aaagttgatt 1140
agaaaattag ccgaattggt cggaacaaga gctgcaagat tgactgtttg tgggtgtttc 1200
gctatctgtg acaagagagg ctacaagact gctcacattg cagctgatgg ttctgtcttc 1260
aacagatacc caggtttaca ggaaaaggcc gctcaagcct tgaaggatat ctacaactgg 1320
gatgtcgaaa agatggaaga ccaccaatc caattggtgg ctgctgaaga tggttccggt 1380

gttgggtgctg ctatcattgc ttgtttgact caaaagagat tggctgccgg taagtctgtt 1440
ggtattaaag gcgaatag 1458

<210> 671

<211> 485

<212> PRT

<213> Artificial Sequence

<220><223> RAG5 from K. lactis

<400> 671

Met Val Arg Leu Gly Pro Lys Lys Pro Pro Ala Arg Lys Gly Ser Met

1 5 10 15

Ala Asp Val Pro Ala Asn Leu Met Glu Gln Ile His Gly Leu Glu Thr

20 25 30

Leu Phe Thr Val Ser Ser Glu Lys Met Arg Ser Ile Val Lys His Phe

35 40 45

Ile Ser Glu Leu Asp Lys Gly Leu Ser Lys Lys Gly Gly Asn Ile Pro

50 55 60

Met Ile Pro Gly Trp Val Val Glu Tyr Pro Thr Gly Lys Glu Thr Gly

65 70 75 80

Asp Phe Leu Ala Leu Asp Leu Gly Gly Thr Asn Leu Arg Val Val Leu

85 90 95

Val Lys Leu Gly Gly Asn His Asp Phe Asp Thr Thr Gln Asn Lys Tyr

100 105 110

Arg Leu Pro Asp His Leu Arg Thr Gly Thr Ser Glu Gln Leu Trp Ser
115 120 125

Phe Ile Ala Lys Cys Leu Lys Glu Phe Val Asp Glu Trp Tyr Pro Asp
130 135 140

Gly Val Ser Glu Pro Leu Pro Leu Gly Phe Thr Phe Ser Tyr Pro Ala
145 150 155 160

Ser Gln Lys Lys Ile Asn Ser Gly Val Leu Gln Arg Trp Thr Lys Gly
165 170 175

Phe Asp Ile Glu Gly Val Glu Gly His Asp Val Val Pro Met Leu Gln
180 185 190

Glu Gln Ile Glu Lys Leu Asn Ile Pro Ile Asn Val Val Arg Leu Ile
195 200 205

Asn Asp Thr Thr Gly Thr Leu Val Ala Ser Leu Tyr Thr Asp Pro Gln
210 215 220

Thr Lys Met Gly Ile Ile Ile Gly Thr Gly Val Asn Gly Ala Tyr Tyr
225 230 235 240

Asp Val Val Ser Gly Ile Glu Lys Leu Glu Gly Leu Leu Pro Glu Asp
245 250 255

Ile Gly Pro Asp Ser Pro Met Ala Ile Asn Cys Glu Tyr Gly Ser Phe
260 265 270

Asp Asn Glu His Leu Val Leu Pro Arg Thr Lys Tyr Asp Val Ile Ile
275 280 285

Asp Glu Glu Ser Pro Arg Pro Gly Gln Gln Ala Phe Glu Lys Met Thr
290 295 300

Ser Gly Tyr Tyr Leu Gly Glu Ile Met Arg Leu Val Leu Leu Asp Leu
305 310 315 320

Tyr Asp Ser Gly Phe Ile Phe Lys Asp Gln Asp Ile Ser Lys Leu Lys
325 330 335

Glu Ala Tyr Val Met Asp Thr Ser Tyr Pro Ser Lys Ile Glu Asp Asp
340 345 350

Pro Phe Glu Asn Leu Glu Asp Thr Asp Asp Leu Phe Lys Thr Asn Leu

355 360 365
 Asn Ile Glu Thr Thr Val Val Glu Arg Lys Leu Ile Arg Lys Leu Ala
 370 375 380
 Glu Leu Val Gly Thr Arg Ala Ala Arg Leu Thr Val Cys Gly Val Ser
 385 390 395 400
 Ala Ile Cys Asp Lys Arg Gly Tyr Lys Thr Ala His Ile Ala Ala Asp
 405 410 415

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

Gly Ile Lys Gly Glu
 485

<210> 672

<211> 1416

<212> DNA

<213> Artificial Sequence

<220><223> HPGLK1 from H. polymorpha

<400> 672

atgagtttgg atactgaagt cgataagatt gtgtcggagt ttgccgtcac ccaggagaca	60
ctccaaaagg gtgtggagcg tttcattgag cttgcaactg ccggactgaa tagtgatgag	120
gacaagtatg gtctgccaat gatcccaact ttgtttacct ccatcccaac cggtaaagag	180
aagggcattc tttttgccgc agacttggga ggaaccaatt tcagagtttg ctctgttgcc	240
ttgaacggag atcacacttt caaactgata cagcagaagt cacatattcc tgccgaactg	300
atgacctcca cctcggacga attgttttcg tatcttgcaa gcaaggtaa gaatttctta	360
gagactatc atgaagggcg tgttacttct acaggaagcc agaaattcaa gatgggtttc	420
actttcagtt tcctgtctc gcagaccgcc ttaaagccg gtactttgct aagatggacc	480
aagggtatca atattccgga tactgttggt caagaggtg tttctctatt ccaaatgcat	540

ttagacgccc aggaaattcc tgttactgtg tctgccctgt ccaacgatac tgtgggaacc 600
cttcttgcaa gatcctacac ggggtccaat aaggagggca ctactgttct aggatgcatc 660
ttcggaacgg gaacaaacgg tgcttacaac gagaagctcg agaatatcaa gaagcttccg 720

gccgaggtga gagagaagct gaaggtcaa ggtgtcacc acatgggtcat taatactgaa 780
tggggttctc tcgataacca gctcaaggtt ttgccaataa cgaagtatga cgctcaagtt 840
gacgaactta cgggcaataa gggcttccac atgtttgaaa agcgtgttcc cggaatgttc 900
ttgggtgaga ttctgagaca tattttgttc gaccttcact ctaaggaggt gctatttact 960
cagtaccca gctacgaatc cctgccccac agattgagga cgccgtggga tctggactct 1020
gaggttctct cactgattga gatcgacgaa tccaccaatt tgcaggccac tgagctgtct 1080
ttgaaacagg cattgagact gccaaactact actgaggaga gacttgctat tcaaaaactt 1140

actcgtgctg tggccaagag atctgcctat cttgtgcta ttctattgc tgctattcta 1200
cacatgaccg agtcttttaa gggccacaac gttgaggtgg acgttggagc agacgggtct 1260
gtggttgagt tctaccctgg attcagaact atgatgagag acgccattgc gcagacgcag 1320
ataggtgcca aaggagagag aagactgcac attaacattg ccaaagacgg ctcactctgtg 1380
ggcgtgcat tgtgcgcatt aagcgagaaa gactaa 1416

<210> 673

<211> 471

<212> PRT

<213> Artificial Sequence

<220><223> HPGLK1 from H. polymorpha

<400> 673

Met Ser Leu Asp Thr Glu Val Asp Lys Ile Val Ser Glu Phe Ala Val

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

65 70 75 80

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

325 330 335
 Asp Leu Asp Ser Glu Val Leu Ser Leu Ile Glu Ile Asp Glu Ser Thr
 340 345 350
 Asn Leu Gln Ala Thr Glu Leu Ser Leu Lys Gln Ala Leu Arg Leu Pro
 355 360 365
 Thr Thr Thr Glu Glu Arg Leu Ala Ile Gln Lys Leu Thr Arg Ala Val
 370 375 380
 Ala Lys Arg Ser Ala Tyr Leu Ala Ala Ile Pro Ile Ala Ala Ile Leu

385 390 395 400
 His Met Thr Glu Ser Phe Lys Gly His Asn Val Glu Val Asp Val Gly
 405 410 415
 Ala Asp Gly Ser Val Val Glu Phe Tyr Pro Gly Phe Arg Thr Met Met
 420 425 430
 Arg Asp Ala Ile Ala Gln Thr Gln Ile Gly Ala Lys Gly Glu Arg Arg
 435 440 445
 Leu His Ile Asn Ile Ala Lys Asp Gly Ser Ser Val Gly Ala Ala Leu

450 455 460
 Cys Ala Leu Ser Glu Lys Asp

465 470

<210> 674

<211> 1634

<212> DNA

<213> Artificial Sequence

<220><223> HXK2 from S. pombe

<400> 674

aacacttttc gcctcacttg cgaatctacg aaaggaatat ataggtgggt caccctttt 60
 cttttcattt cgtgttttta atagtatttt acatcaacag agataactat ttctgttaac 120
 gatttttttt cccacttggt ttcttccttt ttgtgtgaat tttaattaat ttataataag 180
 caatggaggc taattttcaa caagctgtta aaaagttagt caatgacttt gaatacccta 240

ccgagtcctt gagagaggcc gttaaggagt ttgacgaatt acgtcaaaag ggtttacaaa 300
 agaatggatga ggtgcttgct atggctcctg cctttatctc tacccttccc accggcgctg 360
 aaactggatga cttcttgccc cttgactttg gtggtaccaa cttgcgtgtt tgttgatcc 420

aacttctcgg tgacggcaag tatgagatga agcacagcaa gtccgtcttg ccccgatgaat 480
gcgttcgtaa cgagtctgtt aagcccatca ttgactttat gaggaccat gttgagcttt 540
tcatcaagga gcaattccct tccaagtttg gctgccctga ggaggaatac ctctctatgg 600
gtttcacctt ttcttatccc gccaaaccaag ttcccatcac cgagagctac ttgcttcgtt 660

ggaccaaggg tcttaacatt cctgaggcca tcaacaagga ctttgcccaa tttttgactg 720
aaggtttcaa ggctcgtaac ctctctatta gaatcgagge tgtcatcaac gataccgtcg 780
gtactctcgt taccctgtct tatacttcaa aggagagcga cacctttatg ggtatcattt 840
tcggaaccgg taccaacggt gcttacgtcg agcaaatgaa ccaaattccc aagcttgctg 900
gcaagtgtac tggatgatcat atgcttatca acatgggaatg gggagcaact gattttcttt 960
gccttcactc cactcgttat gatttacttc ttgatcatga tactcccaat gctggtcgtc 1020
aaatctttga gaagcgcgtt ggtggtatgt atctcgttga gcttttccgc cgtgccttat 1080

tccacttgat caaggtttac aacttcaacg aaggtatttt ccttcttcc attactgatg 1140
cttggctttt ggaaacttct gttctttcca gaatgatggt tgaacgttct gctgagaatg 1200
ttcgtaacgt tcttagtaca ttcaagttcc gtttccgcag cgacgaagag gctttgtacc 1260
tttgggatgc tgcctatgca attggccgtc gtgctgctcg tatgtctgcc gttcccattg 1320
cttctttgta tctttctacc ggccgcgtg gtaagaagag tgatgttggt gttgatggtt 1380
ctttagtcca acactatcct cactttgttg acatgctccg tgaagccttg cgtgagctta 1440
tcggtgataa cgaaaaattg atttccattg gtattgcca ggaaggcagt ggtattggtg 1500

ccgctctttg cgcctccaa gctgttaagg aaaagaaagg cttggcctaa atcatgttag 1560
atgtctgtta gctttttttg aattgtacgt agaaatgagc atgtaaatat gaaattgctt 1620
tttaacagct tttta 1634

<210> 675

<211> 455

<212> PRT

<213> Artificial Sequence

<220><223> HXK2 from S. pombe

<400> 675

Met Glu Ala Asn Phe Gln Gln Ala Val Lys Lys Leu Val Asn Asp Phe

1 5 10 15

Glu Tyr Pro Thr Glu Ser Leu Arg Glu Ala Val Lys Glu Phe Asp Glu

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

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

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

405 410 415
Arg Glu Leu Ile Gly Asp Asn Glu Lys Leu Ile Ser Ile Gly Ile Ala
420 425 430
Lys Asp Gly Ser Gly Ile Gly Ala Ala Leu Cys Ala Leu Gln Ala Val
435 440 445
Lys Glu Lys Lys Gly Leu Ala
450 455

<210> 676

<211> 804

<212> DNA

<213> Artificial Sequence

<220><223> YMR226C

<400> 676

atgtcccaag gtagaaaagc tgcagaaaga ttggctaaga agactgtcct cattacaggt 60

gcattctgctg gtattggtaa ggccaccgca ttagagtact tggaggcatc caatggtgat 120

atgaaactga tcttggctgc tagaagatta gaaaagctcg aggaattgaa gaagaccatt 180

gatcaagagt ttccaaacgc aaaagttcat gtggcccagc tggatatcac tcaagcagaa 240
 aaaatcaagc ccttcattga aaacttgcca caagagttca aggatattga cattctggtg 300
 aacaatgccg gaaaggctct tggcagtgac cgtgtgggcc agatcgcaac ggaggatatac 360
 caggacgtgt ttgacaccaa cgtcacggct ttaatcaata tcacacaagc tgtactgccc 420
 atattccaag ccaagaattc aggagatatt gtaaatttgg gttcaatcgc tggcagagac 480

gcatacccaa caggttctat ctattgtgcc tctaagttg ccgtgggggc gttcactgat 540
 agtttgagaa aggagctcat caacactaaa attagagtca ttctaattgc accagggcta 600
 gtcgagactg aattttcact agttagatac agaggtaacg aggaacaagc caagaatgtt 660
 tacaaggata ctacccatt gatggctgat gacgtggctg atctgatcgt ctatgcaact 720
 tccagaaaac aaaatactgt aattgcagac actttaatct ttccaacaaa ccaagcgtca 780
 cctcatcata tcttcctggtg ataa 804

<210> 677

<211> 267

<212> PRT

<213> Artificial Sequence

<220><223> YMR226C

<400> 677

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> 678

<211> 1410

<212> DNA

<213> Artificial Sequence

<220><223> YIL074C (Chr 9)

<400> 678

atgtcttatt cagctgccga taatttaca gattcattcc aacgtgccat gaacttttct	60
ggctctctgt gtgcagtctc aacctcacca actcagtcac ttatgaacac actacctcgt	120
cgtgtaagca ttacaaagca accaaagget ttaaaacctt tttctactgg tgacatgaat	180
attctactgt tggaaaatgt caatgcaact gcaatcaaaa tcttcaagga tcagggttac	240
caagtagagt tccacaagtc ttctctacct gaggatgaat tgattgaaaa aatcaaagac	300
gtacacgcta tcggtataag atccaaaact agattgactg aaaaaatact acagcatgcc	360

aggaatctag ttigtattgg ttgtttttgc ataggtacca atcaagtaga cctaaaatat 420
gcccctagta aaggtattgc tgttttcaat tcgccattct ccaattcaag atccgtagca 480
gaattggtaa ttggtgagat cattagtta gcaagacaat taggtgatag atccattgaa 540
ctgcatacag gtacatggaa taaagtcgct gctagggtgt gggaagtaag aggaaaaact 600
ctcggtatta ttgggtatgg tcacattggt tcgcaattat cagttcttgc agaagctatg 660

ggcctgcatg tgctatacta tgatatctg acaattatgg ccttaggtac tgccagacaa 720
gtttctacat tagatgaatt gttgaataaa tctgattttg taactactaca tgtaccagct 780
actccagaaa ctgaaaaaat gttatctgct ccacaattcg ctgctatgaa ggacggggct 840
tatgttatta atgcctcaag aggtactgtc gtggacattc catctctgat ccaagccgtc 900
aaggccaaca aaattgcagg tgctgcttta gatgtttatc cacatgaacc agctaagaac 960
ggtgaagggt catttaacga tgaacttaac agctggactt ctgagttggt ttcattacca 1020
aatataatcc tgacaccaca tattggtggc tctacagaag aagctcaaag ttcaatcggt 1080

attgaggtgg ctactgcatt gtccaaatac atcaatgaag gtaactctgt cggttctgtg 1140
aacttcccag aagtcagttt gaagtctttg gactacgate aagagaacac agtacgtgtc 1200
ttgtatatc atcgtaacgt tcctgggtgt ttgaagaccg ttaatgatat cttatccgat 1260
cataatatcg agaacagtt ttctgattct cacggcgaga tcgcttatct aatggcagac 1320
atctcttctg ttaatcaag tgaaatcaag gatatatatg aaaagttgaa ccaaacttct 1380
gccaaagttt ccatcaggtt attatactaa 1410

<210> 679

<211> 469

<212> PRT

<213> Artificial Sequence

<220><223> YIL074C (Chr 9)

<400> 679

Met Ser Tyr Ser Ala Ala Asp Asn Leu Gln Asp Ser Phe Gln Arg Ala

1 5 10 15

Met Asn Phe Ser Gly Ser Pro Gly Ala Val Ser Thr Ser Pro Thr Gln

20 25 30

Ser Phe Met Asn Thr Leu Pro Arg Arg Val Ser Ile Thr Lys Gln Pro

35 40 45

Lys Ala Leu Lys Pro Phe Ser Thr Gly Asp Met Asn Ile Leu Leu Leu

50			55			60									
Glu	Asn	Val	Asn	Ala	Thr	Ala	Ile	Lys	Ile	Phe	Lys	Asp	Gln	Gly	Tyr
65			70			75			80						
Gln	Val	Glu	Phe	His	Lys	Ser	Ser	Leu	Pro	Glu	Asp	Glu	Leu	Ile	Glu
85			90			95									
Lys	Ile	Lys	Asp	Val	His	Ala	Ile	Gly	Ile	Arg	Ser	Lys	Thr	Arg	Leu
100			105			110									
Thr	Glu	Lys	Ile	Leu	Gln	His	Ala	Arg	Asn	Leu	Val	Cys	Ile	Gly	Cys
115			120			125									
Phe	Cys	Ile	Gly	Thr	Asn	Gln	Val	Asp	Leu	Lys	Tyr	Ala	Ala	Ser	Lys
130			135			140									
Gly	Ile	Ala	Val	Phe	Asn	Ser	Pro	Phe	Ser	Asn	Ser	Arg	Ser	Val	Ala
145			150			155			160						
Glu	Leu	Val	Ile	Gly	Glu	Ile	Ile	Ser	Leu	Ala	Arg	Gln	Leu	Gly	Asp
165			170			175									
Arg	Ser	Ile	Glu	Leu	His	Thr	Gly	Thr	Trp	Asn	Lys	Val	Ala	Ala	Arg
180			185			190									
Cys	Trp	Glu	Val	Arg	Gly	Lys	Thr	Leu	Gly	Ile	Ile	Gly	Tyr	Gly	His
195			200			205									
Ile	Gly	Ser	Gln	Leu	Ser	Val	Leu	Ala	Glu	Ala	Met	Gly	Leu	His	Val
210			215			220									
Leu	Tyr	Tyr	Asp	Ile	Val	Thr	Ile	Met	Ala	Leu	Gly	Thr	Ala	Arg	Gln
225			230			235			240						
Val	Ser	Thr	Leu	Asp	Glu	Leu	Leu	Asn	Lys	Ser	Asp	Phe	Val	Thr	Leu
245			250			255									
His	Val	Pro	Ala	Thr	Pro	Glu	Thr	Glu	Lys	Met	Leu	Ser	Ala	Pro	Gln
260			265			270									
Phe	Ala	Ala	Met	Lys	Asp	Gly	Ala	Tyr	Val	Ile	Asn	Ala	Ser	Arg	Gly
275			280			285									
Thr	Val	Val	Asp	Ile	Pro	Ser	Leu	Ile	Gln	Ala	Val	Lys	Ala	Asn	Lys
290			295			300									

Ile Ala Gly Ala Ala Leu Asp Val Tyr Pro His Glu Pro Ala Lys Asn
305 310 315 320

Gly Glu Gly Ser Phe Asn Asp Glu Leu Asn Ser Trp Thr Ser Glu Leu
325 330 335

Val Ser Leu Pro Asn Ile Ile Leu Thr Pro His Ile Gly Gly Ser Thr
340 345 350

Glu Glu Ala Gln Ser Ser Ile Gly Ile Glu Val Ala Thr Ala Leu Ser
355 360 365

Lys Tyr Ile Asn Glu Gly Asn Ser Val Gly Ser Val Asn Phe Pro Glu
370 375 380

Val Ser Leu Lys Ser Leu Asp Tyr Asp Gln Glu Asn Thr Val Arg Val
385 390 395 400

Leu Tyr Ile His Arg Asn Val Pro Gly Val Leu Lys Thr Val Asn Asp
405 410 415

Ile Leu Ser Asp His Asn Ile Glu Lys Gln Phe Ser Asp Ser His Gly
420 425 430

Glu Ile Ala Tyr Leu Met Ala Asp Ile Ser Ser Val Asn Gln Ser Glu
435 440 445

Ile Lys Asp Ile Tyr Glu Lys Leu Asn Gln Thr Ser Ala Lys Val Ser
450 455 460

Ile Arg Leu Leu Tyr

465

<210> 680

<211> 792

<212> DNA

<213> Artificial Sequence

<220><223> YIR036C (Chr 9)

<400> 680

atgggcaagg ttatTTTgat tacaggtgcc tcccgTggga ttggcctgca attggtgaaa 60

actgttatcg aagaggacga tgaatgcac gtctacggcg tagcaagaac ggaagctggt 120

ctgcagtctt tgcaaagaga atacggtgca gacaaatttg tctatcgtgt cctcgacatc 180

acggacaggt ctgcaatgga agcgttgggtg gaggaatcc ggcaaaagca tggaaaactg 240
gacggtattg tcgcaaatgc ggggatgcta gaaccggtga agtccatctc ccagtccaac 300
tccgaacacg acatcaagca gtgggaacgg ctgttcgatg tgaacttttt cagcattgtc 360
tctttgggtg cactgtgttt acccctcttg aagagctcgc catttgtagg caacattgtc 420
ttcgtcagct ctggagccag tgtgaaacca tataacggat ggtcggcgta cggctgctcg 480
aaagccgcat taaaccactt tgccatggac attgccagtg aagagcccag tgataaagtg 540
cgtgccgtgt gtattgcacc gggcgtcgtt gacacgcaga tgcagaaaga tattagggaa 600

acattgggtc ctcagggcag gacaccaag gctctcgaga ggtttactca attgtacaag 660
acttcgtcac tgctggacc aaaggtgcct gcggcggtac tagcgcaact cgtcctgaaa 720
ggtattcccg actctttgaa cggatcaatat ctccgtaca acgatgagcg actggggccg 780
gtgcagggt ag 792

<210> 681

<211> 263

<212> PRT

<213> Artificial Sequence

<220><223> YIR036C (Chr 9)

<400> 681

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

1 5 10 15

Gln Leu Val Lys Thr Val Ile Glu Glu Asp Asp Glu Cys Ile Val Tyr

20 25 30

Gly Val Ala Arg Thr Glu Ala Gly Leu Gln Ser Leu Gln Arg Glu Tyr

35 40 45

Gly Ala Asp Lys Phe Val Tyr Arg Val Leu Asp Ile Thr Asp Arg Ser

50 55 60

Arg Met Glu Ala Leu Val Glu Glu Ile Arg Gln Lys His Gly Lys Leu

65 70 75 80

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

85 90 95

Ser Gln Ser Asn Ser Glu His Asp Ile Lys Gln Trp Glu Arg Leu Phe

100 105 110

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

115 120 125
Leu Leu Lys Ser Ser Pro Phe Val Gly Asn Ile Val Phe Val Ser Ser
130 135 140

Gly Ala Ser Val Lys Pro Tyr Asn Gly Trp Ser Ala Tyr Gly Cys Ser
145 150 155 160
Lys Ala Ala Leu Asn His Phe Ala Met Asp Ile Ala Ser Glu Glu Pro
165 170 175
Ser Asp Lys Val Arg Ala Val Cys Ile Ala Pro Gly Val Val Asp Thr
180 185 190
Gln Met Gln Lys Asp Ile Arg Glu Thr Leu Gly Pro Gln Gly Met Thr
195 200 205

Pro Lys Ala Leu Glu Arg Phe Thr Gln Leu Tyr Lys Thr Ser Ser Leu
210 215 220
Leu Asp Pro Lys Val Pro Ala Ala Val Leu Ala Gln Leu Val Leu Lys
225 230 235 240
Gly Ile Pro Asp Ser Leu Asn Gly Gln Tyr Leu Arg Tyr Asn Asp Glu
245 250 255
Arg Leu Gly Pro Val Gln Gly
260

<210> 682

<211> 1503

<212> DNA

<213> Artificial Sequence

<220><223> YPL061W (ALD6)(Chr 16)

<400> 682

atgactaagc tacactttga cactgctgaa ccagtcaaga tcacacttcc aaatggtttg	60
acatacgagc aaccaaccgg tctattcatt aacaacaagt ttatgaaagc tcaagacggt	120
aagacctatc ccgtcgaaga tccttcact gaaaacaccg tttgtgaggt ctcttctgcc	180
accactgaag atgttgaata tgctatcgaa tgtgccgacc gtgctttcca cgacactgaa	240
tgggctacc aagaccaag agaaagaggc cgtctactaa gtaagttggc tgacgaattg	300
gaaagccaaa ttgacttggg ttcttcatt gaagctttgg acaatggtaa aactttggcc	360

ttagcccggtg gggatgttac cattgcaatc aactgtctaa gagatgctgc tgcctatgcc 420

gacaaagtca acggtagaac aatcaacacc ggtgacggct acatgaactt caccacctta 480

gagccaatcg gtgtctgtgg tcaaattatt ccatggaact tccaataat gatgttggt 540

tggaagatcg cccagcatt ggccatgggt aacgtctgta tcttgaaacc cgctgctgtc 600

acacctttaa atgccctata ctttgcttct ttatgtaaga aggttggtat tccagctggt 660

gtcgtcaaca tcgttccagg tcctggtaga actgttggtg ctgctttgac caacgaccca 720

agaatcagaa agctggcttt taccggttct acagaagtcg gtaagagtgt tgctgtcgac 780

tcttctgaat ctaacttgaa gaaaatcact ttggaactag gtggttaagtc cgcccatgtg 840

gtctttgacg atgctaacat taagaagact ttaccaaacc tagtaaaccg tattttcaag 900

aacgtctgtc aaatttggtc ctctggttct agaatttacg ttcaagaagg tatttacgac 960

gaactattgg ctgctttcaa ggcttacttg gaaaccgaaa tcaaagttgg taatccattt 1020

gacaaggcta acttccaagg tgctatcact aaccgtcaac aattcgacac aattatgaac 1080

tacatcgata tcggttaagaa agaaggcgcc aagatcttaa ctggtggcga aaaagttggt 1140

gacaagggtt acttcatcag accaaccgtt ttctacgatg ttaatgaaga catgagaatt 1200

gttaaggaag aaatttttgg accagttgtc actgtcgcaa agttcaagac tttagaagaa 1260

gggtgtcgaaa tggetaacag ctctgaattc ggtctagggt ctggtatcga aacagaatct 1320

ttgagcacag gtttgaagggt ggccaagatg ttgaaggccg gtaccgtctg gatcaacaca 1380

tacaacgatt ttgactccag agttccattc ggtggtgtta agcaatctgg ttacggtaga 1440

gaaatgggtg aagaagtcta ccatgcatac actgaagtaa aagctgtcag aattaagttg 1500

taa 1503

<210> 683

<211> 500

<212> PRT

<213> Artificial Sequence

<220><223> YPL061W (ALD6)(Chr 16)

<400> 683

Met Thr Lys Leu His Phe Asp Thr Ala Glu Pro Val Lys Ile Thr Leu

1	5	10	15
Pro	Asn	Gly	Leu
Thr	Tyr	Glu	Gln
Pro	Thr	Gly	Leu
Phe	Ile	Asn	Asn
20	25	30	
Lys	Phe	Met	Lys
Ala	Gln	Asp	Gly
Lys	Thr	Tyr	Pro
Val	Glu	Asp	Pro

35 40 45
 Ser Thr Glu Asn Thr Val Cys Glu Val Ser Ser Ala Thr Thr Glu Asp
 50 55 60
 Val Glu Tyr Ala Ile Glu Cys Ala Asp Arg Ala Phe His Asp Thr Glu

 65 70 75 80
 Trp Ala Thr Gln Asp Pro Arg Glu Arg Gly Arg Leu Leu Ser Lys Leu
 85 90 95
 Ala Asp Glu Leu Glu Ser Gln Ile Asp Leu Val Ser Ser Ile Glu Ala
 100 105 110
 Leu Asp Asn Gly Lys Thr Leu Ala Leu Ala Arg Gly Asp Val Thr Ile
 115 120 125
 Ala Ile Asn Cys Leu Arg Asp Ala Ala Ala Tyr Ala Asp Lys Val Asn

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

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

 260 265 270
 Leu Gly Gly Lys Ser Ala His Leu Val Phe Asp Asp Ala Asn Ile Lys
 275 280 285

Lys Thr Leu Pro Asn Leu Val Asn Gly Ile Phe Lys Asn Ala Gly Gln
 290 295 300
 Ile Cys Ser Ser Gly Ser Arg Ile Tyr Val Gln Glu Gly Ile Tyr Asp
 305 310 315 320
 Glu Leu Leu Ala Ala Phe Lys Ala Tyr Leu Glu Thr Glu Ile Lys Val
 325 330 335
 Gly Asn Pro Phe Asp Lys Ala Asn Phe Gln Gly Ala Ile Thr Asn Arg
 340 345 350
 Gln Gln Phe Asp Thr Ile Met Asn Tyr Ile Asp Ile Gly Lys Lys Glu
 355 360 365
 Gly Ala Lys Ile Leu Thr Gly Gly Glu Lys Val Gly Asp Lys Gly Tyr
 370 375 380
 Phe Ile Arg Pro Thr Val Phe Tyr Asp Val Asn Glu Asp Met Arg Ile
 385 390 395 400
 Val Lys Glu Glu Ile Phe Gly Pro Val Val Thr Val Ala Lys Phe Lys
 405 410 415
 Thr Leu Glu Glu Gly Val Glu Met Ala Asn Ser Ser Glu Phe Gly Leu
 420 425 430
 Gly Ser Gly Ile Glu Thr Glu Ser Leu Ser Thr Gly Leu Lys Val Ala
 435 440 445
 Lys Met Leu Lys Ala Gly Thr Val Trp Ile Asn Thr Tyr Asn Asp Phe
 450 455 460
 Asp Ser Arg Val Pro Phe Gly Gly Val Lys Gln Ser Gly Tyr Gly Arg
 465 470 475 480
 Glu Met Gly Glu Glu Val Tyr His Ala Tyr Thr Glu Val Lys Ala Val
 485 490 495
 Arg Ile Lys Leu
 500
 <210> 684
 <211> 1029
 <212> DNA
 <213> Artificial Sequence

<220><223> YPL088W (Chr 16)

<400> 684

atggtttttag ttaagcaggt aagactcggg aactcaggtc ttaagatatc accgatagtg	60
ataggatgta tgcatacgg gtccaagaaa tgggcggact gggtcataga ggacaagacc	120
caaattttca agattatgaa gcattgttac gataaaggtc ttcgtacttt tgacacagca	180
gatttttatt ctaatgggtt gagtgaaaga ataattaagg agtttctgga gtactacagt	240
ataaagagag aaacgggtgt gattatgacc aaaatttact tcccagtga tgaacgctt	300
gatttgcatc ataacttcac tttaaatgaa ttgaagaat tggacttgtc caaccagcgg	360
ggtttatcca gaaagcatat aattgctggt gtcgagaact ctgtgaaaag actgggcaca	420
tatatagacc ttitacaaat tcacagatta gatcatgaaa cgccaatgaa agagatcatg	480
aaggcattga atgatgttgt tgaagcgggc cacgttagat acattggggc ttcgagtatg	540
ttggcaactg aatttgcaga actgcagttc acagccgata aatatggctg gtttcagttc	600
atttcttcgc agtcttacta caatttgctc tatcgtgaag atgaacgcga attgattcct	660
tttgccaaaa gacacaatat tggtttactt ccatggcttc ctaacgcacg aggcattgtg	720
actcgtctc tgaaccaaag cacggacagg attaagagt atccaacttt caagtcgtta	780
catttgata atctcgaaga agaacaaaag gaaattataa atcgtgtgga aaaggtgtcg	840
aaggacaaaa aagtcctgat ggctatgctc tccattgcat gggttttgca taaaggatgt	900
caccctattg tgggattgaa cactacagca agagtagacg aagcgattgc cgcactacaa	960
gtaactctaa cagaagaaga gataaagtac ctgaggagc cctacaaacc ccagaggcaa	1020
agatgttaa	1029

<210> 685

<211> 342

<212> PRT

<213> Artificial Sequence

<220><223> YPL088W(Chr 16)

<400> 685

Met	Val	Leu	Val	Lys	Gln	Val	Arg	Leu	Gly	Asn	Ser	Gly	Leu	Lys	Ile
1				5					10					15	
Ser	Pro	Ile	Val	Ile	Gly	Cys	Met	Ser	Tyr	Gly	Ser	Lys	Lys	Trp	Ala
				20					25					30	
Asp	Trp	Val	Ile	Glu	Asp	Lys	Thr	Gln	Ile	Phe	Lys	Ile	Met	Lys	His

35 40 45
 Cys Tyr Asp Lys Gly Leu Arg Thr Phe Asp Thr Ala Asp Phe Tyr Ser
 50 55 60
 Asn Gly Leu Ser Glu Arg Ile Ile Lys Glu Phe Leu Glu Tyr Tyr Ser
 65 70 75 80
 Ile Lys Arg Glu Thr Val Val Ile Met Thr Lys Ile Tyr Phe Pro Val

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

 145 150 155 160
 Lys Ala Leu Asn Asp Val Val Glu Ala Gly His Val Arg Tyr Ile Gly
 165 170 175
 Ala Ser Ser Met Leu Ala Thr Glu Phe Ala Glu Leu Gln Phe Thr Ala
 180 185 190
 Asp Lys Tyr Gly Trp Phe Gln Phe Ile Ser Ser Gln Ser Tyr Tyr Asn
 195 200 205
 Leu Leu Tyr Arg Glu Asp Glu Arg Glu Leu Ile Pro Phe Ala Lys Arg

 210 215 220
 His Asn Ile Gly Leu Leu Pro Trp Ser Pro Asn Ala Arg Gly Met Leu
 225 230 235 240
 Thr Arg Pro Leu Asn Gln Ser Thr Asp Arg Ile Lys Ser Asp Pro Thr
 245 250 255
 Phe Lys Ser Leu His Leu Asp Asn Leu Glu Glu Glu Gln Lys Glu Ile
 260 265 270
 Ile Asn Arg Val Glu Lys Val Ser Lys Asp Lys Lys Val Ser Met Ala

 275 280 285

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

<210> 686

<211> 1086

<212> DNA

<213> Artificial Sequence

<220><223> YCR105W (ADH7)(Chr 3)

<400> 686

atgctttacc cagaaaaatt tcagggcacg ggtatttcca acgcaaagga ttggaagcat	60
cctaaattag tgagttttga cccaaaaccc ttggcgatc atgacgttga tgttgaaatt	120
gaagcctgtg gtatctgcgg atctgatttt catatagccg ttggttaattg ggggccagtc	180
ccagaaaatc aaatccttgg acatgaaata attggccgcg tgggaaggt tggatccaag	240
tgccacactg gggtaaaaat cgggtaccgt gttggtgttg gtgccaagc cttggcgtgt	300
tttgagtgtg aacgttgcaa aagtgacaac gagcaatact gtaccaatga ccacgttttg	360
actatgtgga ctccttaca ggacggctac atttcacaag gaggctttgc ctcccacgtg	420
aggcttcacg aacactttgc tattcaaata ccagaaaata ttccaagtcc gctagccgct	480
ccattattgt gtggtggat tacagttttc tctccactac taagaaatgg ctgtggtcca	540
ggtaagagg taggtattgt tggcatcggt ggtattgggc atatgggat tctgttggt	600
aaagctatgg gagccgaggt ttatgcgttt tcgcgaggcc actccaagcg ggaggattct	660
atgaaactcg gtgctgatca ctatattgct atgttgaggg ataaaggctg gacagaacaa	720
tactctaacg ctttggacct tcttgcgtt tgcctcatc cttgtcgaa agttaat	780
gacagtatcg ttaagattat gaagattgga ggtccatcg tttcaattgc tgctcctgaa	840
gttaatgaaa agcttgtttt aaaaccgttg ggcctaattg gagtatcaat ctcaagcagt	900
gctatcggt ctaggaagga aatgaacaa ctattgaaat tagtttccga aaagaatgct	960
aaaatatggg tggaaaaact tccgatcagc gaagaaggcg tcagccatgc ctttacaagg	1020
atggaaagcg gagacgtcaa atacagattt actttggtcg attatgataa gaaattccat	1080

aaatag

1086

<210> 687

<211> 361

<212> PRT

<213> Artificial Sequence

<220><223> YCR105W (ADH7)(Chr 3)

<400> 687

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

1 5 10 15

Asp Trp Lys His Pro Lys Leu Val Ser Phe Asp Pro Lys Pro Phe Gly

20 25 30

Asp His Asp Val Asp Val Glu Ile Glu Ala Cys Gly Ile Cys Gly Ser

35 40 45

Asp Phe His Ile Ala Val Gly Asn Trp Gly Pro Val Pro Glu Asn Gln

50 55 60

Ile Leu Gly His Glu Ile Ile Gly Arg Val Val Lys Val Gly Ser Lys

65 70 75 80

Cys His Thr Gly Val Lys Ile Gly Asp Arg Val Gly Val Gly Ala Gln

85 90 95

Ala Leu Ala Cys Phe Glu Cys Glu Arg Cys Lys Ser Asp Asn Glu Gln

100 105 110

Tyr Cys Thr Asn Asp His Val Leu Thr Met Trp Thr Pro Tyr Lys Asp

115 120 125

Gly Tyr Ile Ser Gln Gly Gly Phe Ala Ser His Val Arg Leu His Glu

130 135 140

His Phe Ala Ile Gln Ile Pro Glu Asn Ile Pro Ser Pro Leu Ala Ala

145 150 155 160

Pro Leu Leu Cys Gly Gly Ile Thr Val Phe Ser Pro Leu Leu Arg Asn

165 170 175

Gly Cys Gly Pro Gly Lys Arg Val Gly Ile Val Gly Ile Gly Gly Ile

180 185 190

Gly His Met Gly Ile Leu Leu Ala Lys Ala Met Gly Ala Glu Val Tyr
195 200 205

Ala Phe Ser Arg Gly His Ser Lys Arg Glu Asp Ser Met Lys Leu Gly
210 215 220

Ala Asp His Tyr Ile Ala Met Leu Glu Asp Lys Gly Trp Thr Glu Gln
225 230 235 240

Tyr Ser Asn Ala Leu Asp Leu Leu Val Val Cys Ser Ser Ser Leu Ser
245 250 255

Lys Val Asn Phe Asp Ser Ile Val Lys Ile Met Lys Ile Gly Gly Ser
260 265 270

Ile Val Ser Ile Ala Ala Pro Glu Val Asn Glu Lys Leu Val Leu Lys
275 280 285

Pro Leu Gly Leu Met Gly Val Ser Ile Ser Ser Ser Ala Ile Gly Ser
290 295 300

Arg Lys Glu Ile Glu Gln Leu Leu Lys Leu Val Ser Glu Lys Asn Val
305 310 315 320

Lys Ile Trp Val Glu Lys Leu Pro Ile Ser Glu Glu Gly Val Ser His
325 330 335

Ala Phe Thr Arg Met Glu Ser Gly Asp Val Lys Tyr Arg Phe Thr Leu
340 345 350

Val Asp Tyr Asp Lys Lys Phe His Lys
355 360

<210> 688

<211> 1035

<212> DNA

<213> Artificial Sequence

<220><223> YDR541C (Chr 4)

<400>

688

atgtctaata cagttctagt ttctggcgct tcaggtttta ttgccttgca taccctgtca	60
caattgttaa aacaagatta taaggttatt ggaactgtga gatcccatga aaaagaagca	120
aaattgctaa gacaatttca acataaccct aatttaactt tagaaattgt tccggacatt	180
tctcatccaa atgctttcga taaggttctg cagaaacgtg gacgtgagat taggtatgtt	240

ctacacacgg ccctcctttt tcattatgat actaccgaat atgaaaaaga cttattgatt 300
 cccgcgttag aaggtaaaaa aaacatccta aattctatca agaaatatgc agcagacact 360
 gtagagcgtg ttgttgtgac ttcttcttgt actgctatta taacccttgc aaagatggac 420

gatcccagtg tggttttttac agaagagagt tggaacgaag caacctggga aagctgtcaa 480
 attgatggga taaatgctta ctttgcattc aagaagtttg ctgaaaagc tgcctgggag 540
 ttcacaaaag agaataaga tcacatcaaa ttcaactaa caacagtcaa cccttctctt 600
 ctttttggtc ctcaactttt cgatgaagat gtgcatggcc atttgaatac ttcttgcgaa 660
 atgatcaatg gcctaattca taccacagta aatgccagtg ttcttgattt tcattccatt 720
 tttattgatg taagggatgt ggccctagct catctgtatg ctttccagaa ggaaaatacc 780
 gcgggtaaaa gattagtggg aactaacggt aaatttggaa accaagatat cctggatatt 840

ttgaacgaag attttccaca attaagaggt ctatttcctt tgggtaagcc tggcacaggt 900
 gatcaagtca ttgaccgagg ttcaactaca gataatagtg caacgaggaa aatacttggc 960
 tttgagttca gaagtttaca cgaaagtgtc catgatactg ctgcccacaa tttgaagaag 1020
 cagaacagat tatga 1035

<210> 689

<211> 344

<212> PRT

<213> Artificial Sequence

<220><223> YDR541C (Chr 4)

<400> 689

Met Ser Asn Thr Val Leu Val Ser Gly Ala Ser Gly Phe Ile Ala Leu

1 5 10 15

His Ile Leu Ser Gln Leu Leu Lys Gln Asp Tyr Lys Val Ile Gly Thr

20 25 30

Val Arg Ser His Glu Lys Glu Ala Lys Leu Leu Arg Gln Phe Gln His

35 40 45

Asn Pro Asn Leu Thr Leu Glu Ile Val Pro Asp Ile Ser His Pro Asn

50 55 60

Ala Phe Asp Lys Val Leu Gln Lys Arg Gly Arg Glu Ile Arg Tyr Val

65 70 75 80

Leu His Thr Ala Ser Pro Phe His Tyr Asp Thr Thr Glu Tyr Glu Lys

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

Ile Leu Lys Lys Gln Asn Arg Leu

340

<210> 690

<211> 1410

<212> DNA

<213> Artificial Sequence

<220><223> YER081 (SER3)(Chr 5)

<400> 690

atgacaagca ttgacattaa caacttacia aatacctttc aacaagctat gaatatgagc	60
ggctccccag gcgctgtttg tacttcacct acgcaatctt tcatgaatac cgttccacag	120
cgcttgaatg ctgtaaagca cccaaaaatt ttgaagcctt tctcaacggg tgatatgaag	180
atcttactat tagaaaacgt taatcaaact gctattacia tcttcgaaga gcaaggttac	240
caagtcgaat tctataaatc ttcatgtccc gaggaagagt tgatcgaaaa gatcaaggac	300
gttcatgcta ttggtatcag atcaaagact agattaactt caaatgtctt acaacatgcg	360
aagaatctgg ttgtattgg ttgtttctgt atcggtacca accaagttga cttagactac	420
gctaccagca gaggtattgc tgttttcaac tcgcctttct ccaactcaag atcagtagca	480
gaattggtca tcgctgaaat cattagttta gcaagacaac taggtgatag atctatcgaa	540
ttacataccg gtacatggaa taaggttgct gctagatgtt gggaggtaag aggaaaaact	600
cttggattta ttgggtacgg tcacattggt tcccaattat cagttcttgc agaagctatg	660
ggtttgcatt tgttgtacta cgatatgtta actatcatgg ccttgggtac tgccagacaa	720
gtttctacat tagatgaatt gttgaataaa tctgattttg tgacactaca tgtaccagct	780
actcctgaaa ctgaaaaaat gttatctgcc ccacaatttg ctgctatgaa ggatggcgct	840
tatgttatta atgcttcaag aggtactgtc gtggacattc catctttgat ccaagccgtg	900
aaagccaaca aaattgcagg tgctgctttg gatgtttatc cacatgaacc agctaagaac	960
ggtgaagggt catttaacga tgagctaaat agctggactt ctgaattagt ttcattacca	1020
aatatcatct tgacaccaca cattggtggc tctaccgaag aagcccaaag ctcaatcggt	1080
attgaagtgg ctaccgcatt gtccaaatac atcaatgaag gtaactctgt cggttcagtc	1140
aacttcccag aagtggcatt gaaatcattg tcttacgacc aagagaacac tgtgcgtgtg	1200
ttatacattc accaaaatgt accaggtgtt ttgaagaccg tcaatgatat ttatcgaac	1260
cataacatcg aaaagcaatt ttccgattca aatggtgaaa ttgcttattt aatggctgat	1320
atctcttctg ttgaccaaag cgatatataa gatatttatg aacaactaaa tcaaacctct	1380
gctaagatct caattagatt gctatattaa	1410

<210> 691

<211> 469

<212> PRT

<213> Artificial Sequence

<220><223> YER081 (SER3)(Chr 5)

<400> 691

Met Thr Ser Ile Asp Ile Asn Asn Leu Gln Asn Thr Phe Gln Gln Ala

1 5 10 15
Met Asn Met Ser Gly Ser Pro Gly Ala Val Cys Thr Ser Pro Thr Gln

20 25 30
Ser Phe Met Asn Thr Val Pro Gln Arg Leu Asn Ala Val Lys His Pro

35 40 45
Lys Ile Leu Lys Pro Phe Ser Thr Gly Asp Met Lys Ile Leu Leu Leu

50 55 60
Glu Asn Val Asn Gln Thr Ala Ile Thr Ile Phe Glu Glu Gln Gly Tyr

65 70 75 80
Gln Val Glu Phe Tyr Lys Ser Ser Leu Pro Glu Glu Glu Leu Ile Glu

85 90 95
Lys Ile Lys Asp Val His Ala Ile Gly Ile Arg Ser Lys Thr Arg Leu

100 105 110
Thr Ser Asn Val Leu Gln His Ala Lys Asn Leu Val Cys Ile Gly Cys

115 120 125
Phe Cys Ile Gly Thr Asn Gln Val Asp Leu Asp Tyr Ala Thr Ser Arg

130 135 140
Gly Ile Ala Val Phe Asn Ser Pro Phe Ser Asn Ser Arg Ser Val Ala

145 150 155 160
Glu Leu Val Ile Ala Glu Ile Ile Ser Leu Ala Arg Gln Leu Gly Asp

165 170 175
Arg Ser Ile Glu Leu His Thr Gly Thr Trp Asn Lys Val Ala Ala Arg

180 185 190
Cys Trp Glu Val Arg Gly Lys Thr Leu Gly Ile Ile Gly Tyr Gly His

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

 260 265 270
 Phe Ala Ala Met Lys Asp Gly Ala Tyr Val Ile Asn Ala Ser Arg Gly
 275 280 285
 Thr Val Val Asp Ile Pro Ser Leu Ile Gln Ala Val Lys Ala Asn Lys
 290 295 300
 Ile Ala Gly Ala Ala Leu Asp Val Tyr Pro His Glu Pro Ala Lys Asn
 305 310 315 320
 Gly Glu Gly Ser Phe Asn Asp Glu Leu Asn Ser Trp Thr Ser Glu Leu

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

 385 390 395 400
 Leu Tyr Ile His Gln Asn Val Pro Gly Val Leu Lys Thr Val Asn Asp
 405 410 415
 Ile Leu Ser Asn His Asn Ile Glu Lys Gln Phe Ser Asp Ser Asn Gly
 420 425 430
 Glu Ile Ala Tyr Leu Met Ala Asp Ile Ser Ser Val Asp Gln Ser Asp
 435 440 445

Ile Lys Asp Ile Tyr Glu Gln Leu Asn Gln Thr Ser Ala Lys Ile Ser

450

455

460

Ile Arg Leu Leu Tyr

465

<210> 692

<211> 711

<212> DNA

<213> Artificial Sequence

<220><223> YPL275W (FDH2)(Chr 16)

<400> 692

atggtggtca tcaataagca attaatggtg agtgggatat tgccggcgtg gctaaaaaat 60

gagtatgata tggaagacaa aataatttca acggtaggtg ccggtagaat tggatatagg 120

gttctggaaa gattggtcgc atttaatccg aagaagttac tgtactacga ctaccaggaa 180

ctacctgcgg aagcaatcaa tagattgaac gaggccagca agcttttcaa tggcagaggt 240

gatattgttc agagagtaga gaaattggag gatatggttg ctcagtcaga tgttgttacc 300

atcaactgtc cattgcacaa ggactcaagg ggtttattca ataaaaagct tatttccac 360

atgaaagatg gtgcatactt ggtgaatacc gctagagggt ctatttgtgt cgcagaagat 420

gttgccgagg cagtcaagtc tggtaaattg gctggctatg gtggtgatgt ctgggataag 480

caaccagcac caaaagacca tccctggagg actatggaca ataaggacca cgtgggaaac 540

gcaatgactg ttcatatcag tggcacatct ctgcatgctc aaaagaggta cgctcaggga 600

gtaaagaaca tcctaaatag ttacttttcc aaaaagtgtg attaccgtcc acaggatatt 660

attgtgcaga atggtttctta tgccaccaga gcttatggac agaagaaata a 711

<210> 693

<211> 236

<212> PRT

<213> Artificial Sequence

<220><223> YPL275W (FDH2)(Chr 16)

<400> 693

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

1

5

10

15

Trp Leu Lys Asn Glu Tyr Asp Leu Glu Asp Lys Ile Ile Ser Thr Val

20

25

30

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

35 40 45

Asn Pro Lys Lys Leu Leu Tyr Tyr Asp Tyr Gln Glu Leu Pro Ala Glu

50 55 60

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

65 70 75 80

Asp Ile Val Gln Arg Val Glu Lys Leu Glu Asp Met Val Ala Gln Ser

85 90 95

Asp Val Val Thr Ile Asn Cys Pro Leu His Lys Asp Ser Arg Gly Leu

100 105 110

Phe Asn Lys Lys Leu Ile Ser His Met Lys Asp Gly Ala Tyr Leu Val

115 120 125

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

130 135 140

Val Lys Ser Gly Lys Leu Ala Gly Tyr Gly Gly Asp Val Trp Asp Lys

145 150 155 160

Gln Pro Ala Pro Lys Asp His Pro Trp Arg Thr Met Asp Asn Lys Asp

165 170 175

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

180 185 190

Ala Gln Lys Arg Tyr Ala Gln Gly Val Lys Asn Ile Leu Asn Ser Tyr

195 200 205

Phe Ser Lys Lys Phe Asp Tyr Arg Pro Gln Asp Ile Ile Val Gln Asn

210 215 220

Gly Ser Tyr Ala Thr Arg Ala Tyr Gly Gln Lys Lys

225 230 235

<210> 694

<211> 1494

<212> DNA

<213> Artificial Sequence

<220><223> YBR006W (UGA5)(Chr2)

<400> 694

atgactttga gtaagtattc taaaccaact ctaaacgacc ctaattttatt cagagaatct	60
ggttatattg acggaaaatg ggttaagggc actgacgaag tttttgaggt ggtagaccct	120
gttccggcg aaatcatagc aagagttccc gaacaaccag tctccgtggt tgaggaagcg	180
attgatgttg cctatgaac tttcaagacg tacaagaata caacaccaag agagagggca	240
aagtggctca gaaacatgta taacttaatg cttgaaaatt tggatgatct ggcaaccatc	300

attacttttag aaaatggtaa agctctaggg gaagctaaag gagaaatcaa atacgcggct	360
tcgtattttg agtggtagcg cgaggaagca ccccgtttat atggtgctac tattcaacct	420
ttgaacctc acaacagagt attcacaatt aggcaacctg ttggtgtatg cggataaatt	480
tgtccatgga atttccgag cgccatgatc acgagaaagg ccgccgtgc tttagctgtg	540
ggctgcacag tagtcatcaa gccagactct caaacgccgc tatctgcttt agcaatggca	600
tatttggtg aaaaggcagg ctttccaag ggttcgttta atgttattct ttcacatgcc	660
aacacaccaa agcttggtaa aacattatgt gaatcaccaa aagtcaagaa agttactttt	720

actggttcta caaacgtcgg taaaatcttg atgaaacaat cttcttctac ttgaagaaa	780
ctgtcttttg agctgggtgg taacgccctt ttcatagtct ttgaggatgc cgatttggat	840
caagccttgg aacaagccat ggcttgtaaa tttaggggtt tgggtcaaac atgtgtgtgc	900
gcaaatagac ttacgttca ctcatcata attgataaat ttgcgaaatt actcgcggag	960
agggtcaaaa aattcgtaat tggccatggt ttggacccaa aaactacaca tggttgtgtc	1020
attaactcca gcgtattga aaaagttgaa agacataaac aggatgccat tgataaggga	1080
gcaaaagtgt tgettgaagg tggacgttta actgagttag gtcctaactt ttatgtcca	1140

gtaattttgt cacacgttcc ctcaacagct attgtttcca aggaggagac ttttggcca	1200
ttatgtccaa tcttttcttt tgatactatg gaagaagttg tcggatatgc taatgatact	1260
gagtttgggt tagcagcata tgtcttttct aaaaatgtca acactttata cactgtgtct	1320
gaagctttgg aaactggatg ggtttcatgt aatacaggtg ttttttcgga ttgttctata	1380
ccatttgggt gtgttaaaga gtcaggattt ggaagagaag gttcgctata tggatttgaa	1440
gattacactg ttttgaagac catcacaatt gggaatttgc caaacagcat ttaa	1494

<210> 695

<211> 497

<212> PRT

<213> Artificial Sequence

<220><223> YBR006W (UGA5)(Chr2)

<400> 695

Met Thr Leu Ser Lys Tyr Ser Lys Pro Thr Leu Asn Asp Pro Asn Leu
1 5 10 15

Phe Arg Glu Ser Gly Tyr Ile Asp Gly Lys Trp Val Lys Gly Thr Asp
20 25 30

Glu Val Phe Glu Val Val Asp Pro Ala Ser Gly Glu Ile Ile Ala Arg
35 40 45

Val Pro Glu Gln Pro Val Ser Val Val Glu Glu Ala Ile Asp Val Ala

50 55 60

Tyr Glu Thr Phe Lys Thr Tyr Lys Asn Thr Thr Pro Arg Glu Arg Ala
65 70 75 80

Lys Trp Leu Arg Asn Met Tyr Asn Leu Met Leu Glu Asn Leu Asp Asp
85 90 95

Leu Ala Thr Ile Ile Thr Leu Glu Asn Gly Lys Ala Leu Gly Glu Ala
100 105 110

Lys Gly Glu Ile Lys Tyr Ala Ala Ser Tyr Phe Glu Trp Tyr Ala Glu

115 120 125

Glu Ala Pro Arg Leu Tyr Gly Ala Thr Ile Gln Pro Leu Asn Pro His
130 135 140

Asn Arg Val Phe Thr Ile Arg Gln Pro Val Gly Val Cys Gly Ile Ile
145 150 155 160

Cys Pro Trp Asn Phe Pro Ser Ala Met Ile Thr Arg Lys Ala Ala Ala
165 170 175

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

180 185 190

Pro Leu Ser Ala Leu Ala Met Ala Tyr Leu Ala Glu Lys Ala Gly Phe
195 200 205

Pro Lys Gly Ser Phe Asn Val Ile Leu Ser His Ala Asn Thr Pro Lys
210 215 220

Leu Gly Lys Thr Leu Cys Glu Ser Pro Lys Val Lys Lys Val Thr Phe
225 230 235 240

Thr Gly Ser Thr Asn Val Gly Lys Ile Leu Met Lys Gln Ser Ser Ser

245 250 255

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

260 265 270

Val Phe Glu Asp Ala Asp Leu Asp Gln Ala Leu Glu Gln Ala Met Ala

275 280 285

Cys Lys Phe Arg Gly Leu Gly Gln Thr Cys Val Cys Ala Asn Arg Leu

290 295 300

Tyr Val His Ser Ser Ile Ile Asp Lys Phe Ala Lys Leu Leu Ala Glu

305 310 315 320

Arg Val Lys Lys Phe Val Ile Gly His Gly Leu Asp Pro Lys Thr Thr

325 330 335

His Gly Cys Val Ile Asn Ser Ser Ala Ile Glu Lys Val Glu Arg His

340 345 350

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

355 360 365

Arg Leu Thr Glu Leu Gly Pro Asn Phe Tyr Ala Pro Val Ile Leu Ser

370 375 380

His Val Pro Ser Thr Ala Ile Val Ser Lys Glu Glu Thr Phe Gly Pro

385 390 395 400

Leu Cys Pro Ile Phe Ser Phe Asp Thr Met Glu Glu Val Val Gly Tyr

405 410 415

Ala Asn Asp Thr Glu Phe Gly Leu Ala Ala Tyr Val Phe Ser Lys Asn

420 425 430

Val Asn Thr Leu Tyr Thr Val Ser Glu Ala Leu Glu Thr Gly Met Val

435 440 445

Ser Cys Asn Thr Gly Val Phe Ser Asp Cys Ser Ile Pro Phe Gly Gly

450 455 460

Val Lys Glu Ser Gly Phe Gly Arg Glu Gly Ser Leu Tyr Gly Ile Glu

465 470 475 480

Asp Tyr Thr Val Leu Lys Thr Ile Thr Ile Gly Asn Leu Pro Asn Ser

485

490

495

Ile

<210> 696

<211> 1323

<212> DNA

<213> Artificial Sequence

<220><223> YOL059W (Chr 15)

<400> 696

atgcttgctg tcagaagatt aacaagatac acattcctta agcgaacgca tccgggtgta	60
tatactgtc gtgcatataa aattttgcct tcaagatcta ctttcctaag aagatcatta	120
ttacaaacac aactgcactc aaagatgact gtcatacta atatcaaaca gcacaaacac	180
tgtcatgagg accatcctat cagaagatcg gactctgccg tgtcaattgt acatttgaaa	240
cgtgcgcctt tcaaggttac agtgattggt tctggtaact gggggaccac catcgccaaa	300
gtcattgcgg aaaacacaga attgcattcc catatcttcg agccagaggt gagaatgtgg	360
gtttttgatg aaaagatcgg cgacgaaaat ctgacggata tcataaatac aagacaccag	420
aacgttaaat atctaccaa tattgacctg ccccataate tagtggccga tcctgatctt	480
ttacactcca tcaagggtgc tgacatcctt gttttcaaca tccctcatca atttttacca	540
aacatagtca aacaattgca aggccacgtg gcccctcatg taagggccat ctctgtctta	600
aaagggttcg agttgggtc caagggtgtg caattgctat cctcctatgt tactgatgag	660
ttaggaatcc aatgtggcgc actatctggt gcaaacttgg caccggaagt ggccaaggag	720
cattggtccg aaaccaccgt ggcttaccaa ctaccaaagg attatcaagg tgatggcaag	780
gatgtagatc ataagathtt gaaattgctg ttccacagac cttacttcca cgtcaatgtc	840
atcgatgatg ttgctggtat atccattgcc ggtgccttga agaacgtcgt ggcacttgca	900
tgtggtttcg tagaaggatg gggatggggt aacaatgcct ccgcagccat tcaaaggctg	960
ggtttaggtg aaattatcaa gttcggtaga atgtttttcc cagaatcaa agtcgagacc	1020
tactatcaag aatccgtcgg tgttcagat ctgatcacca cctgctcagg cggtagaaac	1080
gtcaagggtg ccacatacat ggccaagacc ggtaagtcag ccttgaagc agaaaaggaa	1140
ttgcttaacg gtcaatccgc ccaagggata atcacatgca gagaagtcca cgagtggcta	1200
caaacatgtg agttgaccca agaattccca ttattcgagg cagtctacca gatagtctac	1260
aacaacgtcc gcatggaaga cctaccggag atgattgaag agctagacat cgatgacgaa	1320

tag 1323

<210> 697

<211> 440

<212> PRT

<213> Artificial Sequence

<220><223> YOL059W (Chr 15)

<400> 697

Met Leu Ala Val Arg Arg Leu Thr Arg Tyr Thr Phe Leu Lys Arg Thr

1 5 10 15

His Pro Val Leu Tyr Thr Arg Arg Ala Tyr Lys Ile Leu Pro Ser Arg

20 25 30

Ser Thr Phe Leu Arg Arg Ser Leu Leu Gln Thr Gln Leu His Ser Lys

35 40 45

Met Thr Ala His Thr Asn Ile Lys Gln His Lys His Cys His Glu Asp

50 55 60

His Pro Ile Arg Arg Ser Asp Ser Ala Val Ser Ile Val His Leu Lys

65 70 75 80

Arg Ala Pro Phe Lys Val Thr Val Ile Gly Ser Gly Asn Trp Gly Thr

85 90 95

Thr Ile Ala Lys Val Ile Ala Glu Asn Thr Glu Leu His Ser His Ile

100 105 110

Phe Glu Pro Glu Val Arg Met Trp Val Phe Asp Glu Lys Ile Gly Asp

115 120 125

Glu Asn Leu Thr Asp Ile Ile Asn Thr Arg His Gln Asn Val Lys Tyr

130 135 140

Leu Pro Asn Ile Asp Leu Pro His Asn Leu Val Ala Asp Pro Asp Leu

145 150 155 160

Leu His Ser Ile Lys Gly Ala Asp Ile Leu Val Phe Asn Ile Pro His

165 170 175

Gln Phe Leu Pro Asn Ile Val Lys Gln Leu Gln Gly His Val Ala Pro

180 185 190

His Val Arg Ala Ile Ser Cys Leu Lys Gly Phe Glu Leu Gly Ser Lys

195 200 205
Gly Val Gln Leu Leu Ser Ser Tyr Val Thr Asp Glu Leu Gly Ile Gln

210 215 220
Cys Gly Ala Leu Ser Gly Ala Asn Leu Ala Pro Glu Val Ala Lys Glu
225 230 235 240

His Trp Ser Glu Thr Thr Val Ala Tyr Gln Leu Pro Lys Asp Tyr Gln
245 250 255

Gly Asp Gly Lys Asp Val Asp His Lys Ile Leu Lys Leu Leu Phe His
260 265 270

Arg Pro Tyr Phe His Val Asn Val Ile Asp Asp Val Ala Gly Ile Ser

275 280 285
Ile Ala Gly Ala Leu Lys Asn Val Val Ala Leu Ala Cys Gly Phe Val
290 295 300

Glu Gly Met Gly Trp Gly Asn Asn Ala Ser Ala Ala Ile Gln Arg Leu
305 310 315 320

Gly Leu Gly Glu Ile Ile Lys Phe Gly Arg Met Phe Phe Pro Glu Ser
325 330 335

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

340 345 350
Thr Thr Cys Ser Gly Gly Arg Asn Val Lys Val Ala Thr Tyr Met Ala
355 360 365

Lys Thr Gly Lys Ser Ala Leu Glu Ala Glu Lys Glu Leu Leu Asn Gly
370 375 380

Gln Ser Ala Gln Gly Ile Ile Thr Cys Arg Glu Val His Glu Trp Leu
385 390 395 400

Gln Thr Cys Glu Leu Thr Gln Glu Phe Pro Leu Phe Glu Ala Val Tyr

405 410 415
Gln Ile Val Tyr Asn Asn Val Arg Met Glu Asp Leu Pro Glu Met Ile
420 425 430

Glu Glu Leu Asp Ile Asp Asp Glu

435 440

<210> 698

<211> 804

<212> DNA

<213> Artificial Sequence

<220><223> Saccharomyces paradoxus AABY01000127

<400> 698

atgtcccaag gtagaaaagc tgcagaaaga ttggctaaca agaccgtgct cattacgggt 60
gcctctgctg gtattggtaa ggccaccgca ttagagtatt tggaggcatc caatggtgat 120

atgaaactgg tcttagctgc tagaagatta gaaaagctcg aggaattaaa gaaaactatt 180
gatcaggagt ttccaaacgc caaagttcat gtggccaac tggatatcac tcaagcagaa 240
aagatcaagc cctttattga gaatttgcca aaggagtcca aagacattga ctttttggtg 300
aacaacgctg ggaaggccct tggtagcgac cgtgtggggg agattgcaac acaagatata 360
caggatgtgt ttgacaccaa cgtcacagct ttaattaata tcaactcaagc tgtgctgccc 420
atttttcaag ccaagaactc aggggatatt gtgaacttgg gttcgggtggc tggcagggat 480
gcatacccaa cgggttccat ctattgtgcc tccaagtttg ccgtgggggc gttcactgat 540

agtttaagaa aggagcttat caacaccaag atcagagtca tcctaatacgc accagggcta 600
gtcgaaactg aattttcact ggtagatac agaggcaacg aggagcaagc caagaatgtc 660
tacaaggaca ccacccattt aatggctgat gacgtggctg attgatcgt gtacgcaact 720
tccaggaaac aaaacactgt aattgcagac acgctaatac ttccaaccaa ccaagcatcg 780
cctcaccaca tcttccgtgg atga 804

<210> 699

<211> 267

<212> PRT

<213> Artificial Sequence

<220><223> Saccharomyces paradoxus AABY01000127

<400> 699

Met Ser Gln Gly Arg Lys Ala Ala Glu Arg Leu Ala Asn 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 Val 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 Lys Glu Phe Lys Asp Ile
85 90 95

Asp Ile Leu Val Asn Asn Ala Gly Lys Ala Leu Gly Thr Asp Arg Val
100 105 110

Gly Glu Ile Ala Thr Gln 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 Val 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> 700

<211> 805

<212> DNA

<213> Artificial Sequence

<220><223> Saccharomyces bayanus AACA01000631

<400> 700

atgtccaag gtagaaaagc tgcagaaaga ttggccaaca agacggtgct cattacaggc	60
gcttctgctg gtattggtaa ggccaccgca ttggagtatt tggaagcatc caatggaaac	120
atgaaactga tcttggctgc gaggagattg gagaagctag aggagctgaa gaagaccatc	180
gacgaggagt ttccaatgc aaaggttcac gttggccaac tggatatcac acaggccgag	240
aagatcaagc ccttcattga aaacttgccg gaggcattca aggatattga catcctgata	300
aacaatgccg gcaaagccct gggctccgaa cgtgtcgggg aaattgccac acaggacatc	360
caggacgtgt tcgacaccaa cgtcacggcg ttgatcaacg tcacgcaagc agtgctgcca	420
attttccaag ccaagaactc aggggacatc gtcaacttgg ggctcgggtg ccggcagaga	480
cgcatacccc acaggctcca tctactgtgc ttccaagttt gccgtcgggtg cgttactga	540
cagtttgaga aaggaactga tcaacacgaa gatcagagtt atcttgatcg cgccggggct	600
ggttgagacc gagtctcac tggtcagata cagaggtaat gaggaacaag ctaaaaacgt	660
ctacaaggac actacgccgt tgatggccga cgacgtggct gacttaatcg tatattccac	720
ttccagaaag cagaacaccg tggttgccga caccctgac ttccccacca accaagcctc	780
gccctaccac atctttcgcg gttaa	805

<210> 701

<211> 262

<212> PRT

<213> Artificial Sequence

<220><223> Saccharomyces bayanus AACA01000631

<400> 701

Met	Ser	Gln	Gly	Arg	Lys	Ala	Ala	Glu	Arg	Leu	Ala	Asn	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	Asn	Met	Lys	Leu	Ile	Leu	Ala	Ala	Arg
			35							40				45	
Arg	Leu	Glu	Lys	Leu	Glu	Glu	Leu	Lys	Lys	Thr	Ile	Asp	Glu	Glu	Phe
			50							55				60	
Pro	Asn	Ala	Lys	Val	His	Val	Gly	Gln	Leu	Asp	Ile	Thr	Gln	Ala	Glu

65 70 75 80
 Lys Ile Lys Pro Phe Ile Glu Asn Leu Pro Glu Ala Phe Lys Asp Ile

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

 145 150 155 160
 Arg Ile Pro His Arg Leu His Leu Leu Cys Phe Gln Val Cys Arg Arg
 165 170 175
 Cys Val His Gln Phe Glu Lys Gly Thr Asp Gln His Glu Asp Gln Ser
 180 185 190
 Tyr Leu Asp Arg Ala Gly Ala Gly Asp Arg Val Leu Thr Gly Gln Ile
 195 200 205
 Gln Arg Gly Thr Ser Lys Arg Leu Gln Gly His Tyr Ala Val Asp Gly

 210 215 220
 Arg Arg Arg Gly Leu Asn Arg Ile Phe His Phe Gln Lys Ala Glu His
 225 230 235 240
 Arg Gly Cys Arg His Pro Asp Leu Pro His Gln Pro Ser Leu Ala Leu
 245 250 255
 Pro His Leu Ser Arg Leu
 260
 <210> 702
 <211> 804
 <212> DNA
 <213> Artificial Sequence
 <220><223> Saccharomyces castellii AACF01000116
 <400> 702

atgtctcaag gtcctaaagc tgccgaaaga ttgaatgaga agattgtgtt tatcactggt 60
gcttcagctg gtattgggca agccaccgct ttggaatata tggatgcgtc gaacggtact 120
gtgaaattgg ttctagtgc cagaagattg gagaaattac aacaattgaa ggaagtcatt 180
gaggcaaaat accctaagag taaagtctat attgggaagt tggatgtgac agagcttgag 240
accattcaac cattcttggc taatcttcct gaggaattta aggatattga tatcttgatt 300
aataatgccg ggaaggcatt aggttccgat cgtgtaggtg atattgatat aaaagatgtg 360
aagggaatga tggataccaa tgtcttgggg ttgatcaatg tgacgcaagc tgtgttgac 420

atattccaaa agaagaactc cggatgatt gtgaacttag gttcagttgc tggagagat 480
gcatacccaa cagggtccat ttactgtgct tctaaatttg ccgtgagggc ctttactgaa 540
agtttgagaa ggaattaat taataccaag attaggtga tattgatagc cccgggtatc 600
gtcgaaactg aattctcagt tgttagatac aagggtgata atgagcgtgc taaatctgtc 660
tacgatggag ttcaccctt ggaagcagac gacgtagcag atttaattgt atacaccact 720
tcaagaaaac agaacacagt aattgctgac actttgatat tccaacctc tcaaggttcc 780
gcattccacg tccatcgca ttaa 804

<210> 703

<211> 267

<212> PRT

<213> Artificial Sequence

<220><223> Saccharomyces castellii AACF01000116

<400> 703

Met Ser Gln Gly Pro Lys Ala Ala Glu Arg Leu Asn Glu Lys Ile Val

1 5 10 15

Phe Ile Thr Gly Ala Ser Ala Gly Ile Gly Gln Ala Thr Ala Leu Glu

20 25 30

Tyr Met Asp Ala Ser Asn Gly Thr Val Lys Leu Val Leu Val Ala Arg

35 40 45

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

50 55 60

Pro Lys Ser Lys Val Tyr Ile Gly Lys Leu Asp Val Thr Glu Leu Glu

65 70 75 80

Thr Ile Gln Pro Phe Leu Asp Asn Leu Pro Glu Glu Phe Lys Asp Ile

85 90 95

Asp Ile Leu Ile Asn Asn Ala Gly Lys Ala Leu Gly Ser Asp Arg Val
100 105 110

Gly Asp Ile Asp Ile Lys Asp Val Lys Gly Met Met Asp Thr Asn Val
115 120 125

Leu Gly Leu Ile Asn Val Thr Gln Ala Val Leu His Ile Phe Gln Lys
130 135 140

Lys Asn Ser Gly Asp Ile Val Asn Leu Gly Ser Val Ala Gly Arg Asp
145 150 155 160

Ala Tyr Pro Thr Gly Ser Ile Tyr Cys Ala Ser Lys Phe Ala Val Arg
165 170 175

Ala Phe Thr Glu Ser Leu Arg Arg Glu Leu Ile Asn Thr Lys Ile Arg
180 185 190

Val Ile Leu Ile Ala Pro Gly Ile Val Glu Thr Glu Phe Ser Val Val
195 200 205

Arg Tyr Lys Gly Asp Asn Glu Arg Ala Lys Ser Val Tyr Asp Gly Val
210 215 220

His Pro Leu Glu Ala Asp Asp Val Ala Asp Leu Ile Val Tyr Thr Thr
225 230 235 240

Ser Arg Lys Gln Asn Thr Val Ile Ala Asp Thr Leu Ile Phe Pro Thr
245 250 255

Ser Gln Gly Ser Ala Phe His Val His Arg Asp
260 265

<210> 704

<211> 804

<212> DNA

<213> Artificial Sequence

<220><223> Saccharomyces mikatae AACH01000019

<400> 704

atgtctcaag gtagaaaagc tgcagaaaga ttggctggca aaaccgttct catcacgggt 60

gcctctgctg gtattggcaa agccactgca ttagagtatt tggaggcatc caatggcgat 120

atgaaattaa tcttagccgc tagaagatta gaaaagctcg aggaattgaa gaagactatc 180

gatgaagagt ttccaaacgc aaaggtccat gtgaccaaac tggacatcac acagacagaa 240
aagatcaagc cctttattga aaacttgcca gaggagtca aagacattga tattctggtg 300
aacaacgctg gtaaggctct tggtagggac cgtgttgggg agattgatac acaggacgtc 360
caggacgtgt tcgacaccaa cgtctcggtt ttgattaatg tcacacaggc tgttctgccc 420
atcttccaag ctaagaactc aggggatatt gtgaacttgg gctcggtagc tggcagagat 480
gcatacccaa cggggtccat ctattgtgca tctaagttg ccgtcggggc tttcactgag 540
agtttgagaa tggaacttat aaacactaag attagagtca ttctaattgc accagggtta 600

gtcgaaactg agttttccct ggtagatac agaggtaacg aagaacaagc caagaatgtt 660
tacaaggaca ccactccgtt gatggccgat gacgtggctg atttgattgt gtatgcgact 720
tcaaggaagc agaacactgt aattgcagac acactaatct ttcctaccaa ccaagcgtca 780
ccttaccata tctttcgagg gtga 804

<210> 705

<211> 267

<212> PRT

<213> Artificial Sequence

<220><223> Saccharomyces mikatae AAC01000019

<400> 705

Met Ser Gln Gly Arg Lys Ala Ala Glu Arg Leu Ala Gly 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 Glu Glu Phe

50 55 60

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

65 70 75 80

Lys Ile Lys Pro Phe Ile Glu Asn Leu Pro Glu Glu Phe Lys Asp Ile

85 90 95

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

100 105 110

Gly Glu Ile Asp Thr Gln Asp Val Gln Asp Val Phe Asp Thr Asn Val

115 120 125
Ser Ala Leu Ile Asn Val Thr Gln Ala Val Leu Pro Ile Phe Gln Ala
130 135 140

Lys Asn Ser Gly Asp Ile Val Asn Leu Gly Ser Val 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 Glu Ser Leu Arg Met 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 Tyr His Ile Phe Arg Gly
260 265

<210> 706

<211> 804

<212> DNA

<213> Artificial Sequence

<220><223> Ashbya gossypii AE016819

<400> 706

atgtccctag gaagaaaagc agctgaaaga ttagccaaca aaattgtgct tgtgactggt	60
gcctctgcgg gcattggccg tgctacagcc attaatatg cagacgcgac ggacggggca	120
atcaagttga ttttggtggc aagacgcgca gaaaagctca ccagcttgaa acaggagatc	180
gaaagcaagt atcccaacgc caagatccat gtcggacaat tggatgtgac ccaactggac	240
cagatccgcc catttttggg gggactacct gaggagtcc gagacattga tattttaatt	300
aacaacgcag gtaaggccct cggcactgag aggggtgggg aaatctcgat ggacgatatc	360

caggaggttt tcaacactaa tgttatcggc ttggtgcact tgactcagga ggttctacct 420
 attatgaaag ccaagaattc cggggacatt gtcaatgttg ggtcgattgc cggccgcgaa 480
 gcctaccctg gtggctctat ttactgtgcc acgaaacatg cggtaaggc tttcaccagg 540
 gccatgcgga aggagctcat tagcaccaag atccgggtct tcgaaattgc gccgggctct 600
 gtagaaacgg aattctccat ggttcgtatg cgcggtaacg aagagaatgc caagaaagtg 660
 taccagggat ttgaaccct agatggatgat gatatcgctg atacaattgt ctatgccaca 720
 tccagaagat ccaacaccgt agttgcagag atggtcgttt acccatccgc gcaaggttct 780

ctgtacgata ctcaccgcaa ctaa 804

<210> 707

<211> 267

<212> PRT

<213> Artificial Sequence

<220><223> Ashbya gossypii AE016819

<400> 707

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

1 5 10 15

Leu Val Thr Gly Ala Ser Ala Gly Ile Gly Arg Ala Thr Ala Ile Asn

20 25 30

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

35 40 45

Arg Ala Glu Lys Leu Thr Ser Leu Lys Gln Glu Ile Glu Ser Lys Tyr

50 55 60

Pro Asn Ala Lys Ile His Val Gly Gln Leu Asp Val Thr Gln Leu Asp

65 70 75 80

Gln Ile Arg Pro Phe Leu Glu Gly Leu Pro Glu Glu Phe Arg Asp Ile

85 90 95

Asp Ile Leu Ile Asn Asn Ala Gly Lys Ala Leu Gly Thr Glu Arg Val

100 105 110

Gly Glu Ile Ser Met Asp Asp Ile Gln Glu Val Phe Asn Thr Asn Val

115 120 125

Ile Gly Leu Val His Leu Thr Gln Glu Val Leu Pro Ile Met Lys Ala

130 135 140

Lys Asn Ser Gly Asp Ile Val Asn Val Gly Ser Ile Ala Gly Arg Glu
 145 150 155 160
 Ala Tyr Pro Gly Gly Ser Ile Tyr Cys Ala Thr Lys His Ala Val Lys

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

225 230 235 240
 Ser Arg Arg Ser Asn Thr Val Val Ala Glu Met Val Val Tyr Pro Ser
 245 250 255
 Ala Gln Gly Ser Leu Tyr Asp Thr His Arg Asn
 260 265

<210> 708

<211> 807

<212> DNA

<213> Artificial Sequence

<220><223> Candida glabrata CR380959

<400> 708

atgtctcaag gaagaaaagc tgctgagagg ttacaaggga agattgcctt tattacgggt 60

gcctctgcgg gcatcggtaa agctacagcc attgagtatt tggatgcttc caatggtagt 120

gtgaagctag ttcttgggtc acgtagaatg gagaaattgg aggagttgaa gaaggaattg 180

ctggctcaat atcctgatgc aaagattcat ataggtaaac tggatgttac agactttgaa 240

aacgtcaagc agtttttggc tgacttgcca gaagagtcca aggacatcga catcctgatc 300

aataacgctg gtaaagcgtt ggggtctgac aaagttggag acattgaccc tgaggatata 360

gcaggaatgg ttaacaccaa cgtccttgca ttgatcaatt taacacaatt gttgttgcca 420

ttattcaaga agaagaacag tggtagatgc gtcaacttgg gatcgattgc tggtagagac 480

gcatacccaa cgggtgctat atactgtgca acaaaacatg ctgtcagggc attcacacaa 540

tccttaagga aggaattgat caacaccgac attagagtaa ttgaaattgc tcctggtatg 600
gtcgaaccg agttttctgt ggtcaggtac aaagtgaca agtccaaagc agacgacgtc 660
tacagaggta caacaccact atatgccgat gatatcgagg atttgattgt gtactctacc 720
agcagaaagc caaacatggt ggtagcagat gtcttggtct tcccaacaca ccaggcatcg 780
gcttcgcaca tctacagggg cgactaa 807

<210> 709

<211> 268

<212> PRT

<213> Artificial Sequence

<220><223> Candida glabrata CR380959

<400> 709

Met Ser Gln Gly Arg Lys Ala Ala Glu Arg Leu Gln Gly Lys Ile Ala

1 5 10 15
Phe Ile Thr Gly Ala Ser Ala Gly Ile Gly Lys Ala Thr Ala Ile Glu

20 25 30
Tyr Leu Asp Ala Ser Asn Gly Ser Val Lys Leu Val Leu Gly Ala Arg

35 40 45
Arg Met Glu Lys Leu Glu Glu Leu Lys Lys Glu Leu Leu Ala Gln Tyr

50 55 60
Pro Asp Ala Lys Ile His Ile Gly Lys Leu Asp Val Thr Asp Phe Glu

65 70 75 80
Asn Val Lys Gln Phe Leu Ala Asp Leu Pro Glu Glu Phe Lys Asp Ile

85 90 95
Asp Ile Leu Ile Asn Asn Ala Gly Lys Ala Leu Gly Ser Asp Lys Val

100 105 110
Gly Asp Ile Asp Pro Glu Asp Ile Ala Gly Met Val Asn Thr Asn Val

115 120 125
Leu Ala Leu Ile Asn Leu Thr Gln Leu Leu Leu Pro Leu Phe Lys Lys

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 Ala Ile Tyr Cys Ala Thr Lys His Ala Val Arg

165 170 175
Ala Phe Thr Gln Ser Leu Arg Lys Glu Leu Ile Asn Thr Asp Ile Arg
180 185 190
Val Ile Glu Ile Ala Pro Gly Met Val Glu Thr Glu Phe Ser Val Val

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

260 265
<210> 710
<211> 813
<212> DNA
<213> Artificial Sequence
<220><223> Debaryomyces hansenii CR382139
<400> 710
atgtcgtacg gatctaaagc tgctgaacgt gttgccaata agattgtctt aatcactggt 60
gcttcacatcg gaattgggtga agcaactgcc aaagaaattg catcagccgc taatggcaat 120
ttaaatttag tgttgtgtgc tagacgaaaa gaaaagtggg ataattttatc taaagaattg 180
actgacaaat attcatccat caaggttcat gttgctcaac tagatgtatc taagctcgag 240
actatcaagc catttatcaa tgattaccg aaagaattct ctgacgtgga tgtattagtc 300

aacaatgcag gcttggcttt gggccgtgat gaagttggaa ccattgacac agatgatatg 360
ttatcgatgt ttcaaactaa tgttttaggg ttaattacca tcacacaggc tgttttgcca 420
atcatgaaaa agaagaacag cggagatgtt gttaatatag gttcaattgc tggaagagac 480
tcttacctg gaggtggaat ttactgtcca actaaggcaa gtgtcaagtc gttttcgcaa 540
gttttaagaa aggaattgat tagcaccaag attagagttc ttgaggttga ccctggtaat 600
gttgaaactg aattttcaaa gtgcagattc aagggcgata tggaaaaggc aaagctggtt 660
tacgcgggta ctgaaccatt attatccgaa gacgtagctg aggttgtcgt attcggactt 720

acaagaaagc aaaataccgt tattgctgag acattagtct tttcaaccaa tcaagccagc 780
tcattctcact tataccgtga aagcgataaa taa 813

<210> 711
<211> 270
<212> PRT
<213> Artificial Sequence
<220><223> Debaryomyces hansenii CR382139
<400> 711

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

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

Asp Val Leu Val Asn Asn Ala Gly Leu Ala Leu Gly Arg Asp Glu Val
100 105 110
Gly Thr Ile Asp Thr Asp Asp Met Leu Ser Met Phe Gln Thr Asn Val
115 120 125
Leu Gly Leu Ile Thr Ile Thr Gln Ala Val Leu Pro Ile Met Lys Lys
130 135 140
Lys Asn Ser Gly Asp Val Val Asn Ile Gly Ser Ile Ala Gly Arg Asp
145 150 155 160

Ser Tyr Pro Gly Gly Gly Ile Tyr Cys Pro Thr Lys Ala Ser Val Lys
165 170 175
Ser Phe Ser Gln Val Leu Arg Lys Glu Leu Ile Ser Thr Lys Ile Arg
180 185 190

Val Leu Glu Val Asp Pro Gly Asn Val Glu Thr Glu Phe Ser Asn Val

195 200 205

Arg Phe Lys Gly Asp Met Glu Lys Ala Lys Leu Val Tyr Ala Gly Thr

210 215 220

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

225 230 235 240

Thr Arg Lys Gln Asn Thr Val Ile Ala Glu Thr Leu Val Phe Ser Thr

245 250 255

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

260 265 270

<210> 712

<211> 810

<212> DNA

<213> Artificial Sequence

<220><223> Scheffersomyces stipitis XM_001387479

<400> 712

atgtcgtttg gaaaaaagc tgctgaaaga cttgccaaca aaatcattct taccacggg 60
gcttcgtctg gtatttgtga agctacagct agagagtttg catctgctgc caatgggaat 120
atcagattga ttttgacagc cagaagaaaa gaaaagtgg ctcaattgtc agactcattg 180
accaaggaat ttccaactat caaaatccat tctgccaat tggatgtgac cgaacatgat 240
ggcatcaagc ctttcatttc tggtttacc aaggatttcg ccgacatcga tgtgttgatc 300
aacaatgctg gaaaagctct tggaaaagca tctgttggtg aaatcagtga cagtgatatc 360
caaggcatga tgcaaagaa tgtcttggga ctcacaaaca tgactcaggc tgtgattccc 420

atttttaagg ctaaaaattc tggagatata gtcaacatcg gttcgattgc tggaagagac 480
ccttacctg gtggatgat ctactgtgcc tccaaggctg ctgttaagtt cttctcgcat 540
tctttgagaa aggaactcat taacaccaga atcagagttt tggaagtga tccaggtgct 600
gtgttgaccg agttctcttt ggttcgtttc cagggtgatc agggagctgc tgatgctgtt 660
tatgaaggta cccaaccttt ggatgcctct gatatcgag aagttatcgt gtttggtatc 720
accagaaagc agaacaccgt catagccgaa accttggtat tccaagtca ccaggcttct 780
gcctctcatg tttaacaaggc tctaagtag 810

<210> 713

<211> 269

<212> PRT

<213> Artificial Sequence

<220><223> Scheffersomyces stipitis XM_001387479

<400> 713

Met Ser Phe Gly Lys Lys Ala Ala Glu Arg Leu Ala Asn Lys Ile Ile

1 5 10 15

Leu Ile Thr Gly Ala Ser Ser Gly Ile Gly Glu Ala Thr Ala Arg Glu

20 25 30

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

35 40 45

Arg Lys Glu Lys Leu Ala Gln Leu Ser Asp Ser Leu Thr Lys Glu Phe

50 55 60

Pro Thr Ile Lys Ile His Ser Ala Lys Leu Asp Val Thr Glu His Asp

65 70 75 80

Gly Ile Lys Pro Phe Ile Ser Gly Leu Pro Lys Asp Phe Ala Asp Ile

85 90 95

Asp Val Leu Ile Asn Asn Ala Gly Lys Ala Leu Gly Lys Ala Ser Val

100 105 110

Gly Glu Ile Ser Asp Ser Asp Ile Gln Gly Met Met Gln Thr Asn Val

115 120 125

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

130 135 140

Lys Asn Ser Gly Asp Ile Val Asn Ile Gly Ser Ile Ala Gly Arg Asp

145 150 155 160

Pro Tyr Pro Gly Gly Ser Ile Tyr Cys Ala Ser Lys Ala Ala Val Lys

165 170 175

Phe Phe Ser His Ser Leu Arg Lys Glu Leu Ile Asn Thr Arg Ile Arg

180 185 190

Val Leu Glu Val Asp Pro Gly Ala Val Leu Thr Glu Phe Ser Leu Val

195 200 205

Arg Phe His Gly Asp Gln Gly Ala Ala Asp Ala Val Tyr Glu Gly Thr

210	215	220	
Gln Pro Leu Asp Ala Ser Asp Ile Ala Glu Val Ile Val Phe Gly Ile			
225	230	235	240
Thr Arg Lys Gln Asn Thr Val Ile Ala Glu Thr Leu Val Phe Pro Ser			
245	250	255	
His Gln Ala Ser Ala Ser His Val Tyr Lys Ala Pro Lys			
260	265		
<210> 714			
<211> 891			
<212> DNA			
<213> Artificial Sequence			
<220><223> Meyerozyma guilliermondii XM_001482184			
<400> 714			
atgtgectct taccagccgg tagcactgta ttatgtcatc acccagtagt gagtgtggag	60		
attaaatcct caatcttcat gtctttcggg gccaaagccg ctgaacgcct tgccaacaag	120		
atcatattga tcaactggggc atcgtctggg ataggcgagg ctaccgccag agaattcgct	180		
gtgtgtgcca atggaaaaat tctgttgatt ttgaccgctc ggagagaaga caaactcaag	240		
tctctctcgc aacaattgag cctcatttac ccgcaaatta aaatccattc tgctcgtctt	300		
gatgtctctg agttttcgtc acttaagccg ttcattactg ggttgccaaa ggattttgct	360		
agcatcgacg ttttggtgaa taatgcgggg aaagcattgg gaagagccaa tgttggtgaa	420		
atttcccaag aggaaatcaa tggcatgttc cataccaatg ttcttgggtt gataaactta	480		
actcaggagg tgttaccat cttcaaaaag aaaaatgctg gagatattgt gaacattggc	540		
tcagtggccg gtagagaacc ttacctgga ggtgcagtat actgtgcttc aaaggcagca	600		
gttaactact tttctcattc tttagaaaag gaaactatca attccaaaat cagggtcatg	660		
gaggtggatc ctggggcagt agagacagag ttctcgttgg ttctgtttgg cggatgatgcc	720		
gaggctgcga aaaaggtgta tgagggaacc gagcctttgg gccagaaga tattgcggaa	780		
atcattgtgt ttgctgtgtc gaaaaagcc aaaactgtca ttgcggaaac tttggtgttt	840		
cctaccatc aggtcggagc agttcatgtt catagagggc cgcttgagtg a	891		
<210> 715			
<211> 296			
<212> PRT			

<213> Artificial Sequence

<220><223> Meyerozyma guilliermondii XM_001482184

<400> 715

```

Met Cys Leu Leu Pro Ala Gly Ser Thr Val Leu Cys His His Pro Val
1           5           10           15
Val Ser Val Glu Ile Lys Ser Ser Ile Phe Met Ser Phe Gly Ala Lys
           20           25           30
Ala Ala Glu Arg Leu Ala Asn Lys Ile Ile Leu Ile Thr Gly Ala Ser
           35           40           45
Ser Gly Ile Gly Glu Ala Thr Ala Arg Glu Phe Ala Ala Ala Ala Asn

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

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

           180          185          190
Val Tyr Cys Ala Ser Lys Ala Ala Val Asn Tyr Phe Ser His Ser Leu
           195          200          205
Arg Lys Glu Thr Ile Asn Ser Lys Ile Arg Val Met Glu Val Asp Pro
210          215          220

```

Gly Ala Val Glu Thr Glu Phe Ser Leu Val Arg Phe Gly Gly Asp Ala
 225 230 235 240
 Glu Ala Ala Lys Lys Val Tyr Glu Gly Thr Glu Pro Leu Gly Pro Glu

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

<210> 716

<211> 804

<212> DNA

<213> Artificial Sequence

<220><223> Vanderwaltozyma polyspora XM_001645671

<400> 716

atgtcacagg gtagaaaggc ttcagaaagg ttggctggta aaactgtatt aattacaggt	60
gcttcacag ggattgggaa agccactgca ttagaatata tagatgcctc caatggatcat	120
atgaagttaa ttttagttgc aagaagatta gaaaaattgc aagagttgaa ggaaacaatt	180
tgtaaagaat atccagaatc taaggttcat gttgaagaat tagatatttc tgatattaat	240
agaatcccag aatttatgac aaaattacct gaagaattca aagatattga tatattgatt	300
aacaatgcag gtaaagcatt aggaagtgat actattggta atatcgagaa tgaggatatt	360
aaaggtatgt ttgagactaa cgtttttggg ttaatctgtt taacacaagc tgtacttcca	420
atattcaagg ctaaaaatgg tggatgatatt gtcaatttag ggtcaattgc aggcatagaa	480
gcttacccaa caggatctat atattgtgca actaaatttg cagttaaagc attcactgaa	540
agttaaagaa aggaattgat taatacaaag atcagagtta ttgaaattgc accaggtatg	600
gttaacactg aattttctgt aattagatat aaaggtgacc aagaaaagc agataaagtt	660
tatgaaaaca ctactccttt atatgcagat gacatcgctg atttgatagt ttacaccact	720
tctagaaaat cgaataccgt tatcgctgat gttttggat tccaacatg ccaagcttct	780
gcattccata tctatcgtgg ataa	804

<210> 717

<211> 267

<212> PRT

<213> Artificial Sequence

<220><223> Vanderwaltozyma polyspora XM_001645671

<400> 717

Met Ser Gln Gly Arg Lys Ala Ser Glu Arg Leu Ala Gly Lys Thr Val

1 5 10 15

Leu Ile Thr Gly Ala Ser Ser Gly Ile Gly Lys Ala Thr Ala Leu Glu

20 25 30

Tyr Leu Asp Ala Ser Asn Gly His Met Lys Leu Ile Leu Val Ala Arg

35 40 45

Arg Leu Glu Lys Leu Gln Glu Leu Lys Glu Thr Ile Cys Lys Glu Tyr

50 55 60

Pro Glu Ser Lys Val His Val Glu Glu Leu Asp Ile Ser Asp Ile Asn

65 70 75 80

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

85 90 95

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

100 105 110

Gly Asn Ile Glu Asn Glu Asp Ile Lys Gly Met Phe Glu Thr Asn Val

115 120 125

Phe Gly Leu Ile Cys Leu Thr Gln Ala Val Leu Pro Ile Phe Lys Ala

130 135 140

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

145 150 155 160

Ala Tyr Pro Thr Gly Ser Ile Tyr Cys Ala Thr Lys Phe Ala Val Lys

165 170 175

Ala Phe Thr Glu Ser Leu Arg Lys Glu Leu Ile Asn Thr Lys Ile Arg

180 185 190

Val Ile Glu Ile Ala Pro Gly Met Val Asn Thr Glu Phe Ser Val Ile

195 200 205

Arg Tyr Lys Gly Asp Gln Glu Lys Ala Asp Lys Val Tyr Glu Asn Thr

210	215	220	
Thr Pro Leu Tyr Ala Asp Asp Ile Ala Asp Leu Ile Val Tyr Thr Thr			
225	230	235	240
Ser Arg Lys Ser Asn Thr Val Ile Ala Asp Val Leu Val Phe Pro Thr			
	245	250	255
Cys Gln Ala Ser Ala Ser His Ile Tyr Arg Gly			
	260	265	
<210> 718			
<211> 807			
<212> DNA			
<213> Artificial Sequence			
<220><223> Candida dubliniensis XM_002419771			
<400> 718			
atgtcatttg gtagaaaagc tgctgaaaga ttagccaata gatccattct taccactggt		60	
gcttcacatcg ggattgggtga agcatgtgct aaagttttcg ctgaagcatc taatgggtcaa		120	
gttaaattag ttttaggagc aagaagaaaa gaacgattag ttaaattatc tgatacttta		180	
attaaacaat atcctaatat taaaattcat catgattttt tggatgttac tattaagat		240	
tcaatttcaa aattcattgc tggaattcct catgaatttg aacctgatgt attaattaat		300	
aatagtggta aagccttggg gaaagaagaa gttggagaat tgaaagatga agatattacg		360	
gaaatgtttg atactaatgt cattggagtc attcgtatga ctcaagcagt ttacattta		420	
cttaaaaaaa aaccttatgc tgatgtgggt ttcatggaa gtattgctgg acgtgttcct		480	
tataaaaatg gaggtgggta ttgtgcatct aaagctgctg ttcgtagttt caccgatata		540	
tttagaaaag aaactattaa tactgggtatt agagtcattg aagttgatcc aggtgcagta		600	
cttactgaat ttagtgttgt tcgttataaa ggtgacactg atgctgccga tgctgtttat		660	
actggtactg aaccattaac accagaagat gttgctgaag tgggtgtttt tgcatttca		720	
agaaaacaaa ataccgttat tgctgatact ttgattttcc caaatcatca agcttctcca		780	
gatcatgttt atagaaaacc taattaa		807	
<210> 719			
<211> 268			
<212> PRT			
<213> Artificial Sequence			
<220><223> Candida dubliniensis XM_002419771			

<400> 719

Met Ser Phe Gly Arg Lys Ala Ala Glu Arg Leu Ala Asn Arg Ser Ile
1 5 10 15

Leu Ile Thr Gly Ala Ser Ser Gly Ile Gly Glu Ala Cys Ala Lys Val
20 25 30

Phe Ala Glu Ala Ser Asn Gly Gln Val Lys Leu Val Leu Gly Ala Arg
35 40 45

Arg Lys Glu Arg Leu Val Lys Leu Ser Asp Thr Leu Ile Lys Gln Tyr
50 55 60

Pro Asn Ile Lys Ile His His Asp Phe Leu Asp Val Thr Ile Lys Asp
65 70 75 80

Ser Ile Ser Lys Phe Ile Ala Gly Ile Pro His Glu Phe Glu Pro Asp
85 90 95

Val Leu Ile Asn Asn Ser Gly Lys Ala Leu Gly Lys Glu Glu Val Gly
100 105 110

Glu Leu Lys Asp Glu Asp Ile Thr Glu Met Phe Asp Thr Asn Val Ile
115 120 125

Gly Val Ile Arg Met Thr Gln Ala Val Leu Pro Leu Leu Lys Lys Lys
130 135 140

Pro Tyr Ala Asp Val Val Phe Ile Gly Ser Ile Ala Gly Arg Val Pro
145 150 155 160

Tyr Lys Asn Gly Gly Gly Tyr Cys Ala Ser Lys Ala Ala Val Arg Ser
165 170 175

Phe Thr Asp Thr Phe Arg Lys Glu Thr Ile Asn Thr Gly Ile Arg Val
180 185 190

Ile Glu Val Asp Pro Gly Ala Val Leu Thr Glu Phe Ser Val Val Arg
195 200 205

Tyr Lys Gly Asp Thr Asp Ala Ala Asp Ala Val Tyr Thr Gly Thr Glu
210 215 220

Pro Leu Thr Pro Glu Asp Val Ala Glu Val Val Val Phe Ala Ser Ser
225 230 235 240

Arg Lys Gln Asn Thr Val Ile Ala Asp Thr Leu Ile Phe Pro Asn His

245 250 255

Gln Ala Ser Pro Asp His Val Tyr Arg Lys Pro Asn

260 265

<210> 720

<

211> 822

<212> DNA

<213> Artificial Sequence

<220><223> Zygosaccharomyces rouxii XM_002494574

<400> 720

atgtcacaag gtgtcaaagc tgctgaaaga ctagctggta agactgtatt cattacaggt 60

gcttctgcag gtatcggta agcaactgca aaggaatatt tggatgcac caatggtaaa 120

attaaattga tcttggctgc aagaagatta gagaaattac acgagtttaa agaacaact 180

acaaagagtt acccaagcgc tcaagtcac attggtaaatt tggacgtcac tgcaattgac 240

accataaaac ctttttggga taaattacca aaggaatttc aagatgcga tttttgatc 300

aacaatgccg gtaaggcatt aggtactgat aaagttgggtg atattgcaga tgaagacgtg 360

gaaggtatgt tcgacaccaa tgtcttgggg ttaatcaaag ttactcaagc tgttttacct 420

atcttcaaaa gaaaaaattc tggatgatgc gttacatta gttcgggtgc tggtagagag 480

gcataccagc gtggttccat ttactgtgct actaaacacg ctgttaaggc attcactgaa 540

agtttgcgta aggaattagt cgatacaaaa atcagagtca tgagtattga tcctggtaat 600

gtagagaccg aattttctat ggtagattc cgtgggtgata cagaaaaggc aaagaagggt 660

taccaagaca ctgtccatt atatgcagat gacattgcag atttaatcgt ctatgcaacc 720

tctagaaagc aaaacactgt cattgtgac actttgatct tctcttctaa ccaggcatca 780

ccataccacc tctacagagg ctctcaagac aaaaccaatt ga 822

<210> 721

<211> 273

<212> PRT

<213> Artificial Sequence

<220><223> Zygosaccharomyces rouxii XM_002494574

<400> 721

Met Ser Gln Gly Val Lys Ala Ala Glu Arg Leu Ala Gly Lys Thr Val

1 5 10 15

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

260

265

270

Asn

<210> 722

<211> 804

<212> DNA

<213> Artificial Sequence

<220><223> Lachanea thermotolerans XM_002553230

<400> 722

atgtcacagg gaagaagagc agctgaaaga ctggcaggaa agactgtctt catcacaggc 60

gcatcagccg gtatcggta ggccactgca caagaatacc tggaagcatc cgaaggcaaa 120

atcaagttga tccttgcagc aagaagactc gacaagctgg aggaaatcaa agccaaggtt 180

tctaaagact tcctgaagc acaggtgcat atcgccagc tagatgtgac tcagacggac 240

aaaatccagc cttttgtcga caatttgcgc gaagagtca aagacatcga catcctgac 300

aacaacgcgg gcaaggcgct cggatccgac cccgtgggca caatcgaccc caatgatatt 360

caaggcatga tccagactaa cgttatcggg cttataaatg ttaccaagc cgttctgccc 420

atcttcaagg ccaaaaactc tggatgatac gtgaacctgg gttctgtcgc tggcagagaa 480

gcttaccta caggatctat ttactgcgct acgaagcacg cggatcgctgc tttcacccag 540

agcctgcgca aggaactgat caacacaaac atcagggtta ttgaggtcgc tccaggtaac 600

gtggagaccg agttttctct ggtagatac aaggcgact ctgagaaagc caagaaggtt 660

tacgaaggca cacaaccct ttacgtgac gatatcgag acctaactgt ttacgaacc 720

tcgagaaaac caaacaccgt catcgaggac gttttggtt tgccttcgaa ccaggcttcg 780

ccttaccaca tttaccgtgg ttag 804

<210> 723

<211> 267

<212> PRT

<213> Artificial Sequence

<220><223> Lachanea thermotolerans XM_002553230

<400> 723

Met Ser Gln Gly Arg Arg Ala Ala Glu Arg Leu Ala Gly Lys Thr Val

1 5 10 15

Phe Ile Thr Gly Ala Ser Ala Gly Ile Gly Gln Ala Thr Ala Gln Glu

20 25 30
 Tyr Leu Glu Ala Ser Glu Gly Lys Ile Lys Leu Ile Leu Ala Ala Arg
 35 40 45
 Arg Leu Asp Lys Leu Glu Glu Ile Lys Ala Lys Val Ser Lys Asp Phe
 50 55 60
 Pro Glu Ala Gln Val His Ile Gly Gln Leu Asp Val Thr Gln Thr Asp
 65 70 75 80
 Lys Ile Gln Pro Phe Val Asp Asn Leu Pro Glu Glu Phe Lys Asp Ile
 85 90 95
 Asp Ile Leu Ile Asn Asn Ala Gly Lys Ala Leu Gly Ser Asp Pro Val
 100 105 110
 Gly Thr Ile Asp Pro Asn Asp Ile Gln Gly Met Ile Gln Thr Asn Val
 115 120 125
 Ile Gly Leu Ile Asn Val Thr Gln Ala Val Leu Pro Ile Phe Lys Ala
 130 135 140
 Lys Asn Ser Gly Asp Ile Val Asn Leu Gly Ser Val Ala Gly Arg Glu
 145 150 155 160
 Ala Tyr Pro Thr Gly Ser Ile Tyr Cys Ala Thr Lys His Ala Val Arg
 165 170 175
 Ala Phe Thr Gln Ser Leu Arg Lys Glu Leu Ile Asn Thr Asn Ile Arg
 180 185 190
 Val Ile Glu Val Ala Pro Gly Asn Val Glu Thr Glu Phe Ser Leu Val
 195 200 205
 Arg Tyr Lys Gly Asp Ser Glu Lys Ala Lys Lys Val Tyr Glu Gly Thr
 210 215 220
 Gln Pro Leu Tyr Ala Asp Asp Ile Ala Asp Leu Ile Val Tyr Ala Thr
 225 230 235 240
 Ser Arg Lys Pro Asn Thr Val Ile Ala Asp Val Leu Val Phe Ala Ser
 245 250 255
 Asn Gln Ala Ser Pro Tyr His Ile Tyr Arg Gly
 260 265

<210> 724

<211> 807

<212> DNA

<213> Artificial Sequence

<220><223> Kluyveromyces lactis XM_451902

<400> 724

atgtctcaag gtagaaaggc tgctgaaaga ttgcaaaaca agacaatttt cattaccggt 60
gctttctgcag gtatttgtca agccacagca ttggaatata tagatgctgc taacggtaat 120
gtcaaattga tcttagcagc aagaagggtg gctaagttgg aagaattgaa ggaaaaaatc 180

aatgctgaat acccacaagc taaagtatat atcgggtcaat tggacgtcac tgaaactgag 240
aagattcaac ctttcattga taacttgccg gaagaattca aggatatacg tattttgatt 300
aacaatgccg gtaaagcttt gggatctgat gttgtcggta ccatcagtag cgaggacatc 360
aaaggatga tagatactaa cgttgttgcc cttatcaacg ttaccaagc tgttttgcct 420
attttcaaag caaagaattc cggtagacatc gttaacttag gttctgttgc cggtagagat 480
gcataatcaa ctggttctat ctattgtgct tcgaagcatg ctgtcagagc gttcactcag 540
tctttgagaa aagaattaat caatactggt attagggtca ttgagattgc tccaggtaat 600

gtcgaaactg agttctctct agttagatac aagggtgatg ccgatcgtgc taaacaggtt 660
tacaaggta ctactcctct atatgcagat gacattgctg acttgatcgt ttatgccact 720
tcaagaaaac ctaatactgt catcgtgat gttttggtat ttgcttccaa ccaagcatct 780
ccttaccaca ttaccgtgg cgaatag 807

<210> 725

<211> 268

<212> PRT

<213> Artificial Sequence

<220><223> Kluyveromyces lactis XM_451902

<400> 725

Met Ser Gln Gly Arg Lys Ala Ala Glu Arg Leu Gln Asn Lys Thr Ile

1 5 10 15

Phe Ile Thr Gly Ala Ser Ala Gly Ile Gly Gln Ala Thr Ala Leu Glu

20 25 30

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

35 40 45

Arg Leu Ala Lys Leu Glu Glu Leu Lys Glu Lys Ile Asn Ala Glu Tyr

50 55 60

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

65 70 75 80

Lys Ile Gln Pro Phe Ile Asp Asn Leu Pro Glu Glu Phe Lys Asp Ile

85 90 95

Asp Ile Leu Ile Asn Asn Ala Gly Lys Ala Leu Gly Ser Asp Val Val

100 105 110

Gly Thr Ile Ser Ser Glu Asp Ile Lys Gly Met Ile Asp Thr Asn Val

115 120 125

Val Ala Leu Ile Asn Val Thr Gln Ala Val Leu Pro Ile Phe Lys Ala

130 135 140

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

145 150 155 160

Ala Tyr Pro Thr Gly Ser Ile Tyr Cys Ala Ser Lys His Ala Val Arg

165 170 175

Ala Phe Thr Gln Ser Leu Arg Lys Glu Leu Ile Asn Thr Gly Ile Arg

180 185 190

Val Ile Glu Ile Ala Pro Gly Asn Val Glu Thr Glu Phe Ser Leu Val

195 200 205

Arg Tyr Lys Gly Asp Ala Asp Arg Ala Lys Gln Val Tyr Lys Gly Thr

210 215 220

Thr Pro Leu Tyr Ala Asp Asp Ile Ala Asp Leu Ile Val Tyr Ala Thr

225 230 235 240

Ser Arg Lys Pro Asn Thr Val Ile Ala Asp Val Leu Val Phe Ala Ser

245 250 255

Asn Gln Ala Ser Pro Tyr His Ile Tyr Arg Gly Glu

260 265

<210> 726

<

211> 807

<212> DNA

<213> Artificial Sequence

<220><223> Saccharomyces kluyveri SAKLOH04730

<400> 726

atgtctcaag gtagaagggc tgcagaaaga ctagcaaaca agaccgtttt tataactggc	60
gcctctgccg gcattggcca agctactgct ttggaatact gtgatgcttc taacggtaaa	120
ataaacttgg tgtaagtgc cagaaggctg gaaaaattgc aagagttaaa ggacaaaatc	180
accaaggagt atcctgaagc caaggtttat attggtgtgc ttgatgtgac cgaaacggaa	240
aaaatcaaac cattcttggg tggtttacca gaagaattta aagatattga catcttgatc	300
aataatgcag gcaaagcggt aggctctgat cctgttggta ccatcaaaac tgaagatatt	360
gaaggaatga tcaacaccaa tgtcttagct cttatcaata ttactcaagc tgtcttgcca	420
atcttcaaag ccaagaattt cggtgatata gtaaacttgg ggtctgtcgc tggtagagat	480
gcttatccaa ccggtgcaat ctactgtgct agcaaacaatg cagtcagagc ctactcaaa	540
agtttgagga aggaatttgt gaacaccaat atcagagtga ttgaaattgc tccgggtaat	600
gttgaaccg agttttcctt agttagatat aaagtgata cggaccgtgc taaaaaggtt	660
tatgaaggta ctaaccatt atatgcagat gacattgcag acctattgt gtatgctact	720
tctagaaagc ctaatactgt cattgcggat gttttggttt ttgcttcaa ccaagcatcc	780
ccttaccata tctatcgagg tgactaa	807

<210> 727

<211> 268

<212> PRT

<213> Artificial Sequence

<220><223> Saccharomyces kluyveri SAKLOH04730

<400> 727

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

65 70 75 80
 Lys Ile Lys Pro Phe Leu Asp Gly Leu Pro Glu Glu Phe Lys Asp Ile
 85 90 95
 Asp Ile Leu Ile Asn Asn Ala Gly Lys Ala Leu Gly Ser Asp Pro Val

 100 105 110
 Gly Thr Ile Lys Thr Glu Asp Ile Glu Gly Met Ile Asn Thr Asn Val
 115 120 125
 Leu Ala Leu Ile Asn Ile Thr Gln Ala Val Leu Pro Ile Phe Lys Ala
 130 135 140
 Lys Asn Phe Gly Asp Ile Val Asn Leu Gly Ser Val Ala Gly Arg Asp
 145 150 155 160
 Ala Tyr Pro Thr Gly Ala Ile Tyr Cys Ala Ser Lys His Ala Val Arg

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

225 230 235 240
 Ser Arg Lys Pro Asn Thr Val Ile Ala Asp Val Leu Val Phe Ala Ser
 245 250 255
 Asn Gln Ala Ser Pro Tyr His Ile Tyr Arg Gly Asp

 260 265

<210> 728

<211> 807

<212> DNA

<213> Artificial Sequence

<220><223> Yarrowia lipolytica XM_501554

<400> 728

atgtcttttcg gagataaagc tgctgctcga cttgcgggca agaccgtctt tgttaccggc

60

gcctcgtccg gcattggcca ggccactgtt ctgctctag ccgaagctgc caagggcgac 120

ctcaagtttg tgcttgctgc ccgacgaacc gaccgtctgg acgagctcaa gaagaagctg 180

gagaccgact acaagggtat ccaggtgctg cctttcaagc tggacgtgtc caaggtcgag 240

gagaccgaga acattgtgtc caagctgccc aaggagtgtt ccgaggtgga cgtgcttacc 300

aacaacgccg gcatggtcca cggcaccgaa aaggttggct ccatcaacca gaacgacatt 360

gagatcatgt tccacacaaa cgtgctcgga ctattttctg tctactcagca gtttgcggc 420

gagatcgcaa agcgaacaa gggcgacatt gtcaacattg gctccatcgc cggacgagag 480

ccctacgttg gaggaggaat ctactgtgcc accaaggccg ccgtgcgac tttactgag 540

actctccgaa aagagaacat cgacactcga atccgagtca ttgaggttga tcctggagcc 600

gttgagaccg agttctccgt cgtgcgattc cgaggagaca agtccaagc cgacgtgtt 660

tacgtggaa ccgagcctct ggctgctgac gatattgccg agttcatcac ctacactctc 720

actcgacgag agaattgtgt cattgccgat actctcattt tcccaacca ccaggcttct 780

cctactcacg tctaccgaaa gaactga 807

<210> 729

<211> 268

<212> PRT

<213> Artificial Sequence

<220><223> Yarrowia lipolytica XM_501554

<400> 729

Met Ser Phe Gly Asp Lys Ala Ala Ala Arg Leu Ala Gly Lys Thr Val

1 5 10 15

Phe Val Thr Gly Ala Ser Ser Gly Ile Gly Gln Ala Thr Val Leu Ala

20 25 30

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

35 40 45

Arg Thr Asp Arg Leu Asp Glu Leu Lys Lys Lys Leu Glu Thr Asp Tyr

50 55 60

Lys Gly Ile Gln Val Leu Pro Phe Lys Leu Asp Val Ser Lys Val Glu

65 70 75 80

Glu Thr Glu Asn Ile Val Ser Lys Leu Pro Lys Glu Phe Ser Glu Val

85 90 95

Asp Val Leu Ile Asn Asn Ala Gly Met Val His Gly Thr Glu Lys Val

100 105 110

Gly Ser Ile Asn Gln Asn Asp Ile Glu Ile Met Phe His Thr Asn Val

115 120 125

Leu Gly Leu Ile Ser Val Thr Gln Gln Phe Val Gly Glu Met Arg Lys

130 135 140

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

145 150 155 160

Pro Tyr Val Gly Gly Gly Ile Tyr Cys Ala Thr Lys Ala Ala Val Arg

165 170 175

Ser Phe Thr Glu Thr Leu Arg Lys Glu Asn Ile Asp Thr Arg Ile Arg

180 185 190

Val Ile Glu Val Asp Pro Gly Ala Val Glu Thr Glu Phe Ser Val Val

195 200 205

Arg Phe Arg Gly Asp Lys Ser Lys Ala Asp Ala Val Tyr Ala Gly Thr

210 215 220

Glu Pro Leu Val Ala Asp Asp Ile Ala Glu Phe Ile Thr Tyr Thr Leu

225 230 235 240

Thr Arg Arg Glu Asn Val Val Ile Ala Asp Thr Leu Ile Phe Pro Asn

245 250 255

His Gln Ala Ser Pro Thr His Val Tyr Arg Lys Asn

260 265

<210> 730

<211> 780

<212> DNA

<213> Artificial Sequence

<220><223> Schizosaccharomyces pombe NM_001018495

<400> 730

atgagccgtt tggatggaaa aacgatttta atcactggtg cctcttctgg aattggaaaa 60

agcactgctt ttgaaattgc caaagttgcc aaagtaaaac ttattttggc tgctcgcaga 120

ttttctaccg ttgaagaaat tgcaaaggag ttagaatcga aatatgaagt atcggttctt 180

cctcttaaat tggatgtttc tgatttgaag tctattcctg gggttaattga gtcattgcc 240

aaggaatttg ctgatacga tgtcttgatt aataatgctg gacttgctct aggtaccgat 300

aaagtcattg atcttaatat tgatgacgcc gttaccatga ttactaccaa tgttcttggt 360

atgatggcta tgactcgtgc gggttcttct atattctaca gcaaaaacaa gggtgatatt 420

ttgaacgttg gcagtattgc cggcagagaa tcatacgtag gcggctccgt ttactgctct 480

accaagtctg cccttgctca attcacttcc gctttgcgta aggagactat tgacactcgc 540

attcgtatta tggaggttga tcctggcttg gtcgaaaccg aattcagcgt tgtgagattc 600

cacggagaca aacaaaaggc tgataatggt tacaaaaata gtgagccttt gacacccgaa 660

gacattgctg aggtgattct ttttgcctc actgcagag aaaacgtcgt tattgccgat 720

acacttgttt tcccatccca tcaaggtagt gccaatcatg tgtacagaaa gcaagcgtag 780

<210> 731

<211> 259

<212> PRT

<213> Artificial Sequence

<220><223> Schizosaccharomyces pombe NM_001018495

<400> 731

Met Ser Arg Leu Asp Gly Lys Thr Ile Leu Ile Thr Gly Ala Ser Ser

1 5 10 15

Gly Ile Gly Lys Ser Thr Ala Phe Glu Ile Ala Lys Val Ala Lys Val

20 25 30

Lys Leu Ile Leu Ala Ala Arg Arg Phe Ser Thr Val Glu Glu Ile Ala

35 40 45

Lys Glu Leu Glu Ser Lys Tyr Glu Val Ser Val Leu Pro Leu Lys Leu

50 55 60

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

65 70 75 80

Lys Glu Phe Ala Asp Ile Asp Val Leu Ile Asn Asn Ala Gly Leu Ala

85 90 95

Leu Gly Thr Asp Lys Val Ile Asp Leu Asn Ile Asp Asp Ala Val Thr

100 105 110

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

115 120 125

Leu Pro Ile Phe Tyr Ser Lys Asn Lys Gly Asp Ile Leu Asn Val Gly
 130 135 140
 Ser Ile Ala Gly Arg Glu Ser Tyr Val Gly Gly Ser Val Tyr Cys Ser
 145 150 155 160
 Thr Lys Ser Ala Leu Ala Gln Phe Thr Ser Ala Leu Arg Lys Glu Thr
 165 170 175
 Ile Asp Thr Arg Ile Arg Ile Met Glu Val Asp Pro Gly Leu Val Glu
 180 185 190
 Thr Glu Phe Ser Val Val Arg Phe His Gly Asp Lys Gln Lys Ala Asp
 195 200 205
 Asn Val Tyr Lys Asn Ser Glu Pro Leu Thr Pro Glu Asp Ile Ala Glu
 210 215 220
 Val Ile Leu Phe Ala Leu Thr Arg Arg Glu Asn Val Val Ile Ala Asp

225 230 235 240
 Thr Leu Val Phe Pro Ser His Gln Gly Gly Ala Asn His Val Tyr Arg
 245 250 255
 Lys Gln Ala

<210> 732

<211> 1521

<212> DNA

<213> Artificial Sequence

<220><223> ALD2 from *S. cerevisiae*

<400> 732

atgcctacct tgtatactga tatcgaaatc ccacaattga aaatctcttt aaagcaaccg	60
ctagggtttgt ttatcaacaa tgagttttgt ccatcatcag atggaaagac catcgaaact	120
gtgaaccacg ctactggcga accgataaca tccttccaag cagctaacga aaaggatgta	180
gacaaagctg tgaaagctgc cagggtctgt tttgataacg tttggtcgaa gacatcttct	240
gagcaacgtg gtatttatct ttcaactta ttaaaactta ttgaggagga gcaagacaca	300
cttgccgcgt tagagacttt agacgttgga aagccttacc attcaaatgc caaagtgat	360
ttggcacaaa ttttacagct taccagatat tttgctgggt ccgctgataa gtttgacaaa	420
ggtgcaacca taccattgac ttttaacaag ttgcatata ctctaaaagt tccttttggc	480

gttgttgctc aaatcggtcc atggaattat cctctagcta tggcttggtg gaaattgcaa 540
 ggtagccttag cagccggtaa cagggttatc atcaaacctg ctgagaatac ctctctatct 600

ctactttatt ttgctacttt aattaataaaa gcagggtttc cacctgggtg tgtcaatata 660
 gtacctgggtt atggatcact ttagggccaa gccctagcat ctacatgga tatcgacaaa 720
 atatctttta cgggaagcac caaggtcggg ggatttggtg tggaaagctc cggccaatcg 780
 aacctaaag acgttacact agaatgcggg ggtaagtctc ctgctctcgt atttgaagat 840
 gcagaccttg ataaggctat cgattggata gcagctggca ttttctacaa ttcaggacag 900
 aattgtaccg caaactcaag agtttatgtt caaagtctga tctacgacaa gtttgttgaa 960
 aagtttaaag aaactgcaaa gaaggagtgg gatgttcag gaaaatttga tccgtttgat 1020

gagaaatgca tcgttggtcc agttatatca agtacacagt atgaccgcat caaaagttac 1080
 atagaacgtg gtaaaaggga ggaaaagttg gacatgttcc agacctctga atttctatt 1140
 ggtggagcta aaggctactt cattccccca accatcttca ctgatgtccc gcaaaccatcg 1200
 aaactgttac aggatgagat atttggcccg gttgtggttg ttagcaagtt cacaaattat 1260
 gatgacgctc tgaagctggc taatgatact tgctacgggc tcgcctctgc ggtcttcaca 1320
 aaagatgtca agaaagcga catgtttgct cgcgatatta aagcaggaac tgtttggatc 1380
 aactcatcta acgatgaaga tgttaccgtt ctttttggcg ggtttaaaat gagtgttatt 1440

ggtagagaac tggggcaaag tgggtttgat acctatcttc aaacaaaagc agttcacata 1500
 aatctctctt tggacaacta a 1521

<210> 733
 <211> 506
 <212> PRT
 <213> Artificial Sequence
 <220><223> ALD2 from *S. cerevisiae*
 <400> 733

Met Pro Thr Leu Tyr Thr Asp Ile Glu Ile Pro Gln Leu Lys Ile Ser
 1 5 10 15
 Leu Lys Gln Pro Leu Gly Leu Phe Ile Asn Asn Glu Phe Cys Pro Ser
 20 25 30

Ser Asp Gly Lys Thr Ile Glu Thr Val Asn Pro Ala Thr Gly Glu Pro
 35 40 45
 Ile Thr Ser Phe Gln Ala Ala Asn Glu Lys Asp Val Asp Lys Ala Val

50				55				60							
Lys	Ala	Ala	Arg	Ala	Ala	Phe	Asp	Asn	Val	Trp	Ser	Lys	Thr	Ser	Ser
65				70				75				80			
Glu	Gln	Arg	Gly	Ile	Tyr	Leu	Ser	Asn	Leu	Leu	Lys	Leu	Ile	Glu	Glu
85				90				95							
Glu	Gln	Asp	Thr	Leu	Ala	Ala	Leu	Glu	Thr	Leu	Asp	Ala	Gly	Lys	Pro
100				105				110							
Tyr	His	Ser	Asn	Ala	Lys	Gly	Asp	Leu	Ala	Gln	Ile	Leu	Gln	Leu	Thr
115				120				125							
Arg	Tyr	Phe	Ala	Gly	Ser	Ala	Asp	Lys	Phe	Asp	Lys	Gly	Ala	Thr	Ile
130				135				140							
Pro	Leu	Thr	Phe	Asn	Lys	Phe	Ala	Tyr	Thr	Leu	Lys	Val	Pro	Phe	Gly
145				150				155				160			
Val	Val	Ala	Gln	Ile	Val	Pro	Trp	Asn	Tyr	Pro	Leu	Ala	Met	Ala	Cys
165				170				175							
Trp	Lys	Leu	Gln	Gly	Ala	Leu	Ala	Ala	Gly	Asn	Thr	Val	Ile	Ile	Lys
180				185				190							
Pro	Ala	Glu	Asn	Thr	Ser	Leu	Ser	Leu	Leu	Tyr	Phe	Ala	Thr	Leu	Ile
195				200				205							
Lys	Lys	Ala	Gly	Phe	Pro	Pro	Gly	Val	Val	Asn	Ile	Val	Pro	Gly	Tyr
210				215				220							
Gly	Ser	Leu	Val	Gly	Gln	Ala	Leu	Ala	Ser	His	Met	Asp	Ile	Asp	Lys
225				230				235				240			
Ile	Ser	Phe	Thr	Gly	Ser	Thr	Lys	Val	Gly	Gly	Phe	Val	Leu	Glu	Ala
245				250				255							
Ser	Gly	Gln	Ser	Asn	Leu	Lys	Asp	Val	Thr	Leu	Glu	Cys	Gly	Gly	Lys
260				265				270							
Ser	Pro	Ala	Leu	Val	Phe	Glu	Asp	Ala	Asp	Leu	Asp	Lys	Ala	Ile	Asp
275				280				285							
Trp	Ile	Ala	Ala	Gly	Ile	Phe	Tyr	Asn	Ser	Gly	Gln	Asn	Cys	Thr	Ala
290				295				300							

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

 Gln Tyr Asp Arg Ile Lys Ser Tyr Ile Glu Arg Gly Lys Arg Glu Glu
 355 360 365
 Lys Leu Asp Met Phe Gln Thr Ser Glu Phe Pro Ile Gly Gly Ala Lys
 370 375 380
 Gly Tyr Phe Ile Pro Pro Thr Ile Phe Thr Asp Val Pro Gln Thr Ser
 385 390 395 400
 Lys Leu Leu Gln Asp Glu Ile Phe Gly Pro Val Val Val Val Ser Lys
 405 410 415

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

 Gly Arg Glu Leu Gly Gln Ser Gly Val Asp Thr Tyr Leu Gln Thr Lys
 485 490 495
 Ala Val His Ile Asn Leu Ser Leu Asp Asn
 500 505

 <210> 734
 <211> 1521
 <212> DNA
 <213> Artificial Sequence
 <220><223> ALD3 from *S. cerevisiae*
 <400> 734

atgcctacct tgiatactga tatcgaaatc ccacaattga aaatctcttt aaagcaaccg 60

ctagggttgt ttatcaacaa tgagttttgt ccatcatcag atggaaagac catcgaaact 120

gtgaaccag ctactggcga accgataaca tccttccaag cagctaacga aaaggatgta 180

gacaaagctg tgaagctgc cagggtgct ttgataacg tttggtcgaa gacatcttct 240

gagcaacgtg gtatttatct ttcaaaactta ttaaaactta ttgaggagga gcaagacaca 300

cttgccgcat tagagacttt agacgttgtt aagcctttcc attccaatgc taaacaagac 360

ttagcccaga ttatagaact tacaagatac tatgcggggg cggtcgacaa gttcaatatg 420

ggtgaaacca ttccattgac ttttaacaag ttgcatata ctctaaaagt tccttttggc 480

gttgttgctc aaatcgttcc atggaattat cctctagcta tggctttagt aaaaatgcaa 540

ggtgccttag cggccggtaa cacggttatc atcaaacctg ctgaaaatac ctctctatct 600

ctactttatt ttgctacttt aattaaataa gcaggttttc cacctggtgt tgtcaatgtc 660

attcctggtt atggttccgt tgtgggaaa gctttaggaa cccacatgga tatcgacaaa 720

atatctttta cgggaagtac taagtttggc ggctcagtat tggaaagctt cggccaatcg 780

aaccttaagg atatacact agaatgcggt ggtaagtctc ctgctcttgt atttgaagat 840

gcagaccttg ataaggctat agaattgggt gcaaatggta ttttttttaa ttcgggacag 900

atctgcactg caaactcaag agtttatgtt caaagtctga tctacgacaa gtttgttgaa 960

aagtttaaag aaactgcaaa gaaggagtgg gatgttcag gaaaatttga tccgtttgat 1020

gagaaatgca tcgttgggcc agttatatca agtacacagt atgaccgcat caaaagttac 1080

atagaacgtg gtaaaaagga ggaaaagttg gacatgttcc agacctctga atttctatt 1140

ggtggagcta aaggctactt cattccccca accatcttca ctgatgtacc agaaacatct 1200

aagttgtctg gtgatgaaat atttggcccg gttgtggttg ttagcaagtt cacaaattat 1260

gatgacgtc tgaagctggc taatgatact tgetacgggc tgcctctgc ggtcttcacc 1320

aaagatgtca agaaagcga catgtttgct cgcgatatta aagcaggaac tgtttggatc 1380

aatcaaacca atcaagaaga agctaaagtt ccttttggcg gatttaagat gagtgttatt 1440

ggtagagaat caggcgacac cggcgttgat aactatttac aaataaaatc agtccatgtg 1500

gatctttcat tggataaata a 1521

<210> 735

<211> 506

<212> PRT

<213> Artificial Sequence

<220><223> ALD3 from *S. cerevisiae*

<400> 735

Met Pro Thr Leu Tyr Thr Asp Ile Glu Ile Pro Gln Leu Lys Ile Ser
1 5 10 15

Leu Lys Gln Pro Leu Gly Leu Phe Ile Asn Asn Glu Phe Cys Pro Ser
20 25 30

Ser Asp Gly Lys Thr Ile Glu Thr Val Asn Pro Ala Thr Gly Glu Pro
35 40 45

Ile Thr Ser Phe Gln Ala Ala Asn Glu Lys Asp Val Asp Lys Ala Val
50 55 60

Lys Ala Ala Arg Ala Ala Phe Asp Asn Val Trp Ser Lys Thr Ser Ser
65 70 75 80

Glu Gln Arg Gly Ile Tyr Leu Ser Asn Leu Leu Lys Leu Ile Glu Glu
85 90 95

Glu Gln Asp Thr Leu Ala Ala Leu Glu Thr Leu Asp Ala Gly Lys Pro
100 105 110

Phe His Ser Asn Ala Lys Gln Asp Leu Ala Gln Ile Ile Glu Leu Thr
115 120 125

Arg Tyr Tyr Ala Gly Ala Val Asp Lys Phe Asn Met Gly Glu Thr Ile
130 135 140

Pro Leu Thr Phe Asn Lys Phe Ala Tyr Thr Leu Lys Val Pro Phe Gly
145 150 155 160

Val Val Ala Gln Ile Val Pro Trp Asn Tyr Pro Leu Ala Met Ala Cys
165 170 175

Arg Lys Met Gln Gly Ala Leu Ala Ala Gly Asn Thr Val Ile Ile Lys
180 185 190

Pro Ala Glu Asn Thr Ser Leu Ser Leu Leu Tyr Phe Ala Thr Leu Ile
195 200 205

Lys Lys Ala Gly Phe Pro Pro Gly Val Val Asn Val Ile Pro Gly Tyr
210 215 220

Gly Ser Val Val Gly Lys Ala Leu Gly Thr His Met Asp Ile Asp Lys
225 230 235 240

Ile Ser Phe Thr Gly Ser Thr Lys Val Gly Gly Ser Val Leu Glu Ala
245 250 255

Ser Gly Gln Ser Asn Leu Lys Asp Ile Thr Leu Glu Cys Gly Gly Lys
260 265 270

Ser Pro Ala Leu Val Phe Glu Asp Ala Asp Leu Asp Lys Ala Ile Glu
275 280 285

Trp Val Ala Asn Gly Ile Phe Phe Asn Ser Gly Gln Ile Cys Thr Ala
290 295 300

Asn Ser Arg Val Tyr Val Gln Ser Ser Ile Tyr Asp Lys Phe Val Glu
305 310 315 320

Lys Phe Lys Glu Thr Ala Lys Lys Glu Trp Asp Val Ala Gly Lys Phe
325 330 335

Asp Pro Phe Asp Glu Lys Cys Ile Val Gly Pro Val Ile Ser Ser Thr
340 345 350

Gln Tyr Asp Arg Ile Lys Ser Tyr Ile Glu Arg Gly Lys Lys Glu Glu
355 360 365

Lys Leu Asp Met Phe Gln Thr Ser Glu Phe Pro Ile Gly Gly Ala Lys
370 375 380

Gly Tyr Phe Ile Pro Pro Thr Ile Phe Thr Asp Val Pro Glu Thr Ser
385 390 395 400

Lys Leu Leu Arg Asp Glu Ile Phe Gly Pro Val Val Val Val Ser Lys
405 410 415

Phe Thr Asn Tyr Asp Asp Ala Leu Lys Leu Ala Asn Asp Thr Cys Tyr
420 425 430

Gly Leu Ala Ser Ala Val Phe Thr Lys Asp Val Lys Lys Ala His Met
435 440 445

Phe Ala Arg Asp Ile Lys Ala Gly Thr Val Trp Ile Asn Gln Thr Asn
450 455 460

Gln Glu Glu Ala Lys Val Pro Phe Gly Gly Phe Lys Met Ser Gly Ile
465 470 475 480

Gly Arg Glu Ser Gly Asp Thr Gly Val Asp Asn Tyr Leu Gln Ile Lys

	485	490	495
Ser Val His Val Asp Leu Ser Leu Asp Lys			
	500	505	
<210> 736			
<211> 1560			
<212> DNA			
<213> Artificial Sequence			
<220><223> ALD4 from <i>S. cerevisiae</i>			
<400> 736			
atgttcagta gatctacgt ctgcttaaag acgtctgcat cctccattgg gagacttcaa			60
ttgagatatt tctcacacct tctatgaca gtgcctatca agctgcccaa tgggttggaa			120
tatgagcaac caacggggtt gttcatcaac aacaagtttg ttccttctaa acagaacaag			180
accttcgaag tcattaaccc ttccacggaa gaagaaatat gtcataatga tgaagtaga			240
gaggacgatg tggaagagge cgtgcaggcc gccgaccgtg ccttctctaa tgggtcttgg			300
aacggatcgc accctattga caggggtaag gctttgtaca ggtagccga attaattgaa			360
caggacaagg atgtcattgc ttccatcgag actttggata acggtaaagc tatctcttcc			420
tcgagaggag atgttgattt agtcatcaac tatttgaaat cttctgctgg ctttctgat			480
aaaattgatg gtagaatgat tgatactggt agaaccatt tttcttacac taagagacag			540
cctttgggtg tttgtgggca gattattcct tggaatttcc cactgttgat gtgggcctgg			600
aagattgccc ctgctttggt caccggtaac accgtcgtgt tgaagactgc cgaatccacc			660
ccattgtccg ctttgtatgt gtctaaatac atcccacagg cgggtattcc acctggtgtg			720
atcaacattg tatccgggtt tggttaagatt gtgggtgagg ccattacaaa ccatccaaaa			780
atcaaaaagg ttgccttcac aggggtccacg gctacgggta gacacattta ccagtccgca			840
gccgcaggct tgaaaaaagt gactttggag ctgggtggta aatcaccaaa cattgtcttc			900
gcgagacgcc agttgaaaaa agccgtgcaa aacattatcc ttggtatcta ctacaattct			960
ggtgaggctc gtgtgctggg ttcaagggtg tatgttgaag aatctattta cgacaaattc			1020
attgaagagt tcaaagccgc ttctgaatcc atcaaggtgg gcgaccatt cgatgaatct			1080
actttccaag gtgcacaaac ctctcaaatg caactaaaca aaatcttgaa atacgttgac			1140
attggtgaaga atgaagggtc tactttgatt accggtgggtg aaagattagg tagcaagggt			1200
tacttcatta agccaactgt ctttgggtgac gtttaaggaag acatgagaat tgtcaaagag			1260
gaaatctttg gccctgttgt cactgtaacc aaattcaaat ctgccgacga agtcattaac			1320
atggcgaacg attctgaata cgggttggct gctggtattc acacctctaa tattaatacc			1380

gccttaaaag tggctgatag agttaatgcg ggtacggtct ggataaacac ttataacgat 1440

ttccaccacg cagttccttt cggtaggggttc aatgcatctg gtttgggcag ggaaatgtct 1500

gttgatgctt tacaaaacta cttgcaagtt aaagcggtcc gtgccaaatt ggacgagtaa 1560

<210> 737

<211> 519

<212> PRT

<213> Artificial Sequence

<220><223> ALD4 from *S. cerevisiae*

<400> 737

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> 738

<211> 1563

<212> DNA

<213> Artificial Sequence

<220><223> ALD5 from *S. cerevisiae*

<400> 738

atgctttctc gcacaagagc tgcagctccg aattccagaa tattcactag aagcttgtaa 60

cgcttttatt ctcaagcacc attacgcgtt ccaattactc ttccaaatgg tttcacctac 120

gaacagccaa cagggttatt catcaatggt gaatttggtg cctcgaagca aaagaaaacg 180

tttgactgta tcaatccatc taacgaagaa aagataacaa ctgtatacaa ggctatggaa 240

gatgatgttg atgaagccgt tgcagcggct aaaaaagctt ttgaaacgaa gtggtctatt 300

gtagagccgg aggttcgcgc taaagcttta ttcaatctcg ctgacttggt tgagaaacac 360

caagaaacac tggctgccat tgagicaatg gataatggta agtcattgtt ttgtgcgcgc 420

ggtgacgtcg ctttagtata taaatacttg cgttcttgcg gtggttgggc agataaaatc 480

tacggtaacg ttattgacac aggtaaaaac cattttacct actcaattaa ggaaccatta 540

ggcgtttgcg gccaaataat cccttggaaac ttccctttat tgatgtggtc atggaaaatt 600

gggcctgctc tggtacagg taacaccgtc gtattgaaac ccgctgaaac aacaccttta 660

tctgcccttt tcgcttccca gttgtgtcag gaagcaggca taccgctgg ttagtcaat 720

atccttccgg gttccggtag agttgttgga gaaagattga gtgcacaccc agacgtgaag 780

aagattgctt ttacaggtct tactgccacc ggccgccata ttatgaaggt cgctgccgat 840
actgtcaaga aagtcacttt ggagctggga ggtaaatcac caaatattgt gtttctgac 900

gctgatctag ataaagccgt caagaacatt gccttcggta tttttacaa ctctggtgaa 960
gtttgctgcg ctggttcag aatatacatt caagatacag tatacgagga ggtgttgcaa 1020
aaactaaagg attacaccga gtcactaaag gtcggtgacc catttgatga ggaagttttc 1080
caagggtgctc aaacatctga caaacagctg cataaaattt tagactatgt cgatgtagca 1140
aatcagagg gggctcgtct tgtgactgga ggggccagac atggcagtaa aggttatattt 1200
gtcaagccaa cagtgtttgc tgatgtcaaa ggagatatga gaattgttaa ggaggaagtg 1260
tttgggccca ttgtaactgt atccaagttt tctactgttg atgaagtgat tgctatggca 1320

aatgattctc aatatgggtt agccgcaggt attcacacta acgatattaa caaggctgtt 1380
gatgtgtcca aaagagtga agctgggtact gtttgataa atacctataa caacttcac 1440
caaatgttc ctttcgggtg cttcgccag tcaggtattg gccgtgaaat gggtaggct 1500
gctttaagta actacactca aacaaaatct gtcagaattg ccattgacaa gccaattcgt 1560
tga 1563

<210> 739

<211> 520

<212> PRT

<213> Artificial Sequence

<220><223> ALD5 from *S. cerevisiae*

<400> 739

Met Leu Ser Arg Thr Arg Ala Ala Ala Pro Asn Ser Arg Ile Phe Thr

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

65	70	75	80
Asp	Asp	Val	Asp
Glu	Ala	Val	Ala
Ala	Ala	Ala	Lys
Lys	Ala	Phe	Glu
Thr			

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

Glu Val Leu Gln Lys Leu Lys Asp Tyr Thr Glu Ser Leu Lys Val Gly
 340 345 350
 Asp Pro Phe Asp Glu Glu Val Phe Gln Gly Ala Gln Thr Ser Asp Lys
 355 360 365
 Gln Leu His Lys Ile Leu Asp Tyr Val Asp Val Ala Lys Ser Glu Gly
 370 375 380
 Ala Arg Leu Val Thr Gly Gly Ala Arg His Gly Ser Lys Gly Tyr Phe

385 390 395 400
 Val Lys Pro Thr Val Phe Ala Asp Val Lys Gly Asp Met Arg Ile Val
 405 410 415
 Lys Glu Glu Val Phe Gly Pro Ile Val Thr Val Ser Lys Phe Ser Thr
 420 425 430
 Val Asp Glu Val Ile Ala Met Ala Asn Asp Ser Gln Tyr Gly Leu Ala
 435 440 445
 Ala Gly Ile His Thr Asn Asp Ile Asn Lys Ala Val Asp Val Ser Lys

450 455 460
 Arg Val Lys Ala Gly Thr Val Trp Ile Asn Thr Tyr Asn Asn Phe His
 465 470 475 480
 Gln Asn Val Pro Phe Gly Gly Phe Gly Gln Ser Gly Ile Gly Arg Glu
 485 490 495
 Met Gly Glu Ala Ala Leu Ser Asn Tyr Thr Gln Thr Lys Ser Val Arg
 500 505 510
 Ile Ala Ile Asp Lys Pro Ile Arg
 515 520

<210> 740

<211> 1503

<212> DNA

<213> Artificial Sequence

<220><223> ALD6 from *S. cerevisiae*

<400> 740

atgactaagc tacactttga cactgctgaa ccagtcaaga tcacacttcc aaatggtttg 60

acatacgagc aaccaaccgg tctattcatt aacaacaagt ttatgaaagc tcaagacggt 120

aagacctatc ccgtcgaaga tccttccact gaaaacaccg tttgtgaggt ctcttctgcc 180
accactgaag atgttgaata tgctatcgaa tgtgccgacc gtgctttcca cgacactgaa 240
tgggctaccc aagaccaag agaaagaggc cgtctactaa gtaagtggc tgacgaattg 300
gaaagccaaa ttgacttggg ttcttccatt gaagctttgg acaatggtaa aactttggcc 360

ttagcccggtg gggatgttac cattgcaatc aactgtctaa gagatgctgc tgcctatgcc 420
gacaaagtca acggtagaac aatcaacacc ggtgacggct acatgaactt caccacctta 480
gagccaatcg gtgtctgtgg tcaaattatt ccatggaact ttccaataat gatgttggct 540
tggaagatcg cccagcatt ggccatgggt aacgtctgta tcttgaaacc cgctgctgac 600
acacctttaa atgccctata ctttgccttct ttatgtaaga aggttgggtat tccagctggg 660
gtcgtcaaca tcgttccagg tcctggtaga actgttgggt ctgctttgac caacgacca 720
agaatcagaa agctggcttt taccggttct acagaagtcg gtaagagtgt tgctgtcgac 780

tcttctgaat ctaacttgaa gaaaatcact ttggaactag gtggtaagtc cgccatttg 840
gtctttgacg atgctaacat taagaagact ttaccaaacc tagtaaaccg tattttcaag 900
aacgtgggc aaatttgttc ctctgggttct agaatttacg ttcaagaagg tatttacgac 960
gaactattgg ctgctttcaa ggcttacttg gaaaccgaaa tcaaagttgg taatccattt 1020
gacaaggcta acttccaagg tgctatcact aaccgtcaac aattcgacac aattatgaac 1080
tacatcgata tcggtaagaa agaaggcgcc aagatcttaa ctggaggcga aaaagttggg 1140
gacaagggtt attcatcag accaaccgtt ttctacgatg ttaatgaaga catgagaatt 1200

gttaaggaag aaatttttgg accagtgtc actgtcgcaa agttcaagac tttagaagaa 1260
ggtgtcgaaa tggctaacag ctctgaattc ggtctagggt ctggatcga aacagaatct 1320
ttgagcacag gtttgaaggt ggccaagatg ttgaaggccg gtaccgtctg gatcaacaca 1380
tacaacgatt ttgactccag agttccattc ggtgggttta agcaatctgg ttacggtaga 1440
gaaatgggtg aagaagtcta ccatgcatac actgaagtaa aagctgtcag aattaagttg 1500
taa 1503

<210> 741

<211> 500

<212> PRT

<213> Artificial Sequence

<220><223> ALD6 from *S. cerevisiae*

<400> 741

Met Thr Lys Leu His Phe Asp Thr Ala Glu Pro Val Lys Ile Thr Leu

1 5 10 15
 Pro Asn Gly Leu Thr Tyr Glu Gln Pro Thr Gly Leu Phe Ile Asn Asn
 20 25 30
 Lys Phe Met Lys Ala Gln Asp Gly Lys Thr Tyr Pro Val Glu Asp Pro
 35 40 45
 Ser Thr Glu Asn Thr Val Cys Glu Val Ser Ser Ala Thr Thr Glu Asp

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

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

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

 245 250 255

Val Ala Val Asp Ser Ser Glu Ser Asn Leu Lys Lys Ile Thr Leu Glu
260 265 270

Leu Gly Gly Lys Ser Ala His Leu Val Phe Asp Asp Ala Asn Ile Lys
275 280 285

Lys Thr Leu Pro Asn Leu Val Asn Gly Ile Phe Lys Asn Ala Gly Gln
290 295 300

Ile Cys Ser Ser Gly Ser Arg Ile Tyr Val Gln Glu Gly Ile Tyr Asp

305 310 315 320

Glu Leu Leu Ala Ala Phe Lys Ala Tyr Leu Glu Thr Glu Ile Lys Val
325 330 335

Gly Asn Pro Phe Asp Lys Ala Asn Phe Gln Gly Ala Ile Thr Asn Arg
340 345 350

Gln Gln Phe Asp Thr Ile Met Asn Tyr Ile Asp Ile Gly Lys Lys Glu
355 360 365

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

370 375 380

Phe Ile Arg Pro Thr Val Phe Tyr Asp Val Asn Glu Asp Met Arg Ile
385 390 395 400

Val Lys Glu Glu Ile Phe Gly Pro Val Val Thr Val Ala Lys Phe Lys
405 410 415

Thr Leu Glu Glu Gly Val Glu Met Ala Asn Ser Ser Glu Phe Gly Leu
420 425 430

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

435 440 445

Lys Met Leu Lys Ala Gly Thr Val Trp Ile Asn Thr Tyr Asn Asp Phe
450 455 460

Asp Ser Arg Val Pro Phe Gly Gly Val Lys Gln Ser Gly Tyr Gly Arg
465 470 475 480

Glu Met Gly Glu Glu Val Tyr His Ala Tyr Thr Glu Val Lys Ala Val
485 490 495

Arg Ile Lys Leu
500

<210> 742

<211> 1569

<212> DNA

<213> Artificial Sequence

<220><223> KLLAOF00440 from *Kluyveromyces lactis*

<400> 742

atgttattac gtactaccac tgctggtact agactagctt tcaagtcaca aactgcggca	60
ttcgcacatg ctttacgtta ctattcgcat ttaccattga aggtcaaggt tgatttgcca	120
aatggatatt catacgaaca accaactggt ttgttcatca ataaccaatt tgttcttcc	180
aagcaaggga agacttttga agtgttagt ccttcaaccg aagaagaaat tactcatgtt	240
tacgaaggtc gggaggaaga cgtggatgcg gctgttgagg ctgcagacaa agctttcaag	300
tccggtgtct ggictaaggt tgaccactia gtcagatctc agtgtttgag caagttagct	360
gattacattg aaaaggattt ggataccatt gcttctatcg aatctctgga caatggtaaa	420
gctatctcat ctgctaaagg tgatgttcaa ctagtgttga actacctaag atcagctgct	480
ggctacgctg ataagggtga tggtagagtc attgatactg gttcgactca tttctcttac	540
acagtttagag aaccattagg tgtctgtggt caaattatcc catggaactt cccattgttg	600
atgtgggctt ggaagattgc tccagctctt gctaccggtg atactgttgt ctgaagacc	660
gctgaatcca cacctttgtc tgcactatac gtttctcaat acattccaaa agcgggtatc	720
ccagacggtg ttgtgaacat cgtttctggt ttcggtaaaa ttgttggtga ggctattaca	780
actcatcaa agattaagaa ggttgcattc actgggtcca ctgctactgg tcgtcatatc	840
tacaaaaatg ccggtgtcaa ttgaagaag gtcacattgg agctaggtgg taagtcacca	900
aacgttgttt tcgctgatgc aaacttgaag tccgctgttc aaaactcttt gtgggtatc	960
tactataact ccggtgaagt ctgctgtgct ggttccaggt tgtatgttga agaactctatt	1020
tatgacgatt tcttgaagga aattaagtct gcttcagaag ctctaaaggt tggatgacct	1080
ttcgatgcag aaactttcca ggggtctcaa acttcgcaa atcaattgga caagatcttg	1140
aaatatgtgg agattggtaa gaatgaaggt gctactttaa ttactggtgg tgaaagaatt	1200
ggtgacaagg gttactttgt tagaccaact atttctctg atgttgaaga aagcatgaga	1260
atcgtaagg aagaaatctt cggccctgtt gtcactgtca ccaagtttaa aaccattgat	1320
gaattgattg aaaaggccaa cgactctgaa tatggattgg ctgcaggaat tcacacatct	1380
aacatcaata ctgccatcca ggtttcaaac agattgaatg ctggtagcgt ctggattaac	1440
acatacaacg acttccacca ctgtgttccg ttcgggtggt tcaatgcttc tggattggt	1500

agagaaatgg gtctgaagc tctagacaat tacactcaag tcaaggctgt ccgtgttaaa 1560
ttggagtaa 1569

<210> 743

<211> 522

<212> PRT

<213> Artificial Sequence

<220><223> KLLA0F00440 from Kluyveromyces lactis

<400> 743

Met Leu Leu Arg Thr Thr Thr Ala Gly Thr Arg Leu Ala Phe Lys Ser

1 5 10 15

Gln Thr Ala Ala Phe Ala His Ala Leu Arg Tyr Tyr Ser His Leu Pro

20 25 30

Leu Lys Val Lys Val Asp Leu Pro Asn Gly Ile Ser Tyr Glu Gln Pro

35 40 45

Thr Gly Leu Phe Ile Asn Asn Gln Phe Val Pro Ser Lys Gln Gly Lys

50 55 60

Thr Phe Glu Val Leu Ser Pro Ser Thr Glu Glu Glu Ile Thr His Val

65 70 75 80

Tyr Glu Gly Arg Glu Glu Asp Val Asp Ala Ala Val Glu Ala Ala Asp

85 90 95

Lys Ala Phe Lys Ser Gly Val Trp Ser Lys Val Asp Pro Leu Val Arg

100 105 110

Ser Gln Cys Leu Ser Lys Leu Ala Asp Tyr Ile Glu Lys Asp Leu Asp

115 120 125

Thr Ile Ala Ser Ile Glu Ser Leu Asp Asn Gly Lys Ala Ile Ser Ser

130 135 140

Ala Lys Gly Asp Val Gln Leu Val Val Asn Tyr Leu Arg Ser Ala Ala

145 150 155 160

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

165 170 175

His Phe Ser Tyr Thr Val Arg Glu Pro Leu Gly Val Cys Gly Gln Ile

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

Val Thr Lys Phe Lys Thr Ile Asp Glu Leu Ile Glu Lys Ala Asn Asp
 435 440 445
 Ser Glu Tyr Gly Leu Ala Ala Gly Ile His Thr Ser Asn Ile Asn Thr
 450 455 460
 Ala Ile Gln Val Ser Asn Arg Leu Asn Ala Gly Thr Val Trp Ile Asn
 465 470 475 480
 Thr Tyr Asn Asp Phe His His Cys Val Pro Phe Gly Gly Phe Asn Ala
 485 490 495

Ser Gly Ile Gly Arg Glu Met Gly Ser Glu Ala Leu Asp Asn Tyr Thr
 500 505 510
 Gln Val Lys Ala Val Arg Val Lys Leu Glu
 515 520

<210> 744

<211> 1524

<212> DNA

<213> Artificial Sequence

<220><223> KLLA0E23057 from Kluyveromyces lactis

<400> 744

atgtcctcta caattgctga gaaattgaac ctcaagatcg tcgaacaaga cgctgttagc	60
atcactttgc caaacggttt gacttaccaa caaccaactg gtttgttcat caacaatcag	120
ttcatcaagt ctcaagacgg taagactttg aagggtgaaa acccatctac tgaggaaatc	180
attgtcgaag tccaatctgc tactttctcaa gacgtcgagt acgccgttga agctgccgat	240
gtctgtttca actccgaatg gtctactatg gacccaaaa agcgtgggtc tttgttgttt	300
aagttggctg acttgattga agctcaaaag gaattgattg cttctatcga atctgtctgac	360
aacggtaaga ctttggccct agccagaggt gatgttggtt tggtcattga ctacatcaga	420
tctgtctgtg gttatgctga caagttgggt ggtagaacta tcaacactgg tgatggttac	480
gctaacttca cttacaagga acctctaggt gtctgtggtc aaatcatccc atggaacttc	540
ccattgatga tgccttcttg gaagatgcc cctgcttggg ttgctggtaa caccgttata	600
ttgaagccag cttccccaac cccattgaac gctttgttct ttgcttcttt gtgtaaggaa	660
gcaggtatcc cagctgggtg cgtaaatac gtccaggtc caggtagatc cgttggtgac	720
accatcacca accatccaaa aattagaaag attgccttca ctggttcac tgacattggt	780
agagacgttg ctatcaaggc tgcccaatct aacttgaaga aggtcacctt ggaattgggt	840

ggtaaatccg ctcatattggt ctttgaagat gccaacatta agaagactat tccaaacttg 900
gtcaacggta ttttcaagaa tgctgggtcaa atttggttct ctgggtccag aatctatgtc 960
caagacacca tctacgatca actattgtct gaattcaaga cttacctgga aactgaaatt 1020

aaggtcgggt ccccatcga tgaatctaac ttccaagctg ctatcaacaa caaggtcaa 1080
ttcgaaacta tcttgaacta catcgacatc ggtaagaagg aaggtgcttc tatcttgact 1140
ggtggtgaaa gagtaggcaa caagggttac ttcattaaac caactgtatt ctacaacgtt 1200
aaggaagata tgagaatcgt caaggaagaa atctttggtc ctgtcgtcac catctccaag 1260
ttctctactg tcgacgaagc tgctgctttg gctaacgact ccgaattcgg tttgggtgct 1320
ggtatcgaaa ctgaaaacat ctccgttgcc ttgaaggtcg caaagagact aaaagctggt 1380
accgtctgga tcaacactta caacgatttc gacgtgccg ttccattcgg tggttacaag 1440

caatctgggt acggtagaga aatgggtgaa gaagctttcg aatcttacac tcaaatcaag 1500
gccgtcagga tcaagttgga ttaa 1524

<210> 745

<211> 507

<212> PRT

<213> Artificial Sequence

<220><223> KLLA0E23057 from Kluyveromyces lactis

<400> 745

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

1 5 10 15

Asp Ala Val Ser Ile Thr Leu Pro Asn Gly Leu Thr Tyr Gln Gln Pro

20 25 30

Thr Gly Leu Phe Ile Asn Asn Gln Phe Ile Lys Ser Gln Asp Gly Lys

35 40 45

Thr Leu Lys Val Glu Asn Pro Ser Thr Glu Glu Ile Ile Val Glu Val

50 55 60

Gln Ser Ala Thr Ser Gln Asp Val Glu Tyr Ala Val Glu Ala Ala Asp

65 70 75 80

Ala Ala Phe Asn Ser Glu Trp Ser Thr Met Asp Pro Lys Lys Arg Gly

85 90 95

Ser Leu Leu Phe Lys Leu Ala Asp Leu Ile Glu Ala Gln Lys Glu Leu

100 105 110
 Ile Ala Ser Ile Glu Ser Ala Asp Asn Gly Lys Thr Leu Ala Leu Ala
 115 120 125
 Arg Gly Asp Val Gly Leu Val Ile Asp Tyr Ile Arg Ser Ala Ala Gly
 130 135 140
 Tyr Ala Asp Lys Leu Gly Gly Arg Thr Ile Asn Thr Gly Asp Gly Tyr
 145 150 155 160

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

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

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

Ala Ala Ile Asn Asn Lys Ala Gln Phe Glu Thr Ile Leu Asn Tyr Ile
 355 360 365
 Asp Ile Gly Lys Lys Glu Gly Ala Ser Ile Leu Thr Gly Gly Glu Arg
 370 375 380
 Val Gly Asn Lys Gly Tyr Phe Ile Lys Pro Thr Val Phe Tyr Asn Val
 385 390 395 400
 Lys Glu Asp Met Arg Ile Val Lys Glu Glu Ile Phe Gly Pro Val Val
 405 410 415

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

Gln Ser Gly Tyr Gly Arg Glu Met Gly Glu Glu Ala Phe Glu Ser Tyr
 485 490 495
 Thr Gln Ile Lys Ala Val Arg Ile Lys Leu Asp
 500 505

<210> 746

<211> 1512

<212> DNA

<213> Artificial Sequence

<220><223> KLLA0D10021 from Kluyveromyces lactis

<400> 746

atgtccttta ccagcatcaa actagcagat ggtactgagt accaacaacc ttaggttta 60

ttcattaaca atgaatatgt tgaagcacat ggttcggaga ctttaacatc atataaccca 120

agttctggtg aagaaatctg caaagttcaa tgtgcagacc caaagattga tgtcgatttg 180

gctgttaagg ctgcaagaga cgcttttcct ggttggaag cgactccagc tatagaaaag 240

agagacttgt tctacaaatt ggccggcttg ttggaaagag atattgacat attaagtaaa 300

atcgagtcctc tggacactgg aaagccaatt gccaacaatg ccaagtacga cataactagaa 360

gtggttgagg tcttaaagta ctatgctgga tggatcgata aggcaggtgg tcaatctttc 420
gctccaagca gtgaaaagtt agcttatact ttacatcagc cattaggcgt ggttggaagt 480
attatcccat tcaactatcc attggctatg atgtcttga aatggactgc tatcgccact 540

gggaacacaa ctgtttttaa atctgcggat cagactccat tatccatcct atatttcgcc 600
aaattggtga aggaagcagg gtccctgca ggtgttttca atgtcttate cggtagcagg 660
gcctctgttg gtagtgccat tgtaaacac actggtgtag acaagattgc attcacagg 720
tccactgcag ttggtcaaat tatccaacaa caatctgcag taaacttgaa acctatcaca 780
ttggaatgtg gtggtaaatc tgcgaacatg atctttgaag attgtgattt agaacaggcc 840
gtgaaatgga gtgcatgggg tattttcaac aatatgggtc aaatctgttc cggtagctca 900
agagtcttcg tgcaagaatc tatttacgag gaattttag cacaattggc tgaccatgtc 960

aagaaaactt atatcgtggg tgctccttca gatgatgccg tcatcgtcgg acctcaagtc 1020
tccaaaaaac aacaagaacg tgctccttga tacattaaga tcggtatcga agaaggatgc 1080
aaagtcccat tgggtggtct tggagcacct gagccaattg cgaatgatcc agcattagct 1140
aacggttact acgtcaagcc aaccatttta gcagatgtca aagcgtcgta cagaatcgcc 1200
aaggaagaaa tctttggtcc cgtcatagtc gtaggctctt tcaaggacta caaagaagct 1260
ttagacattg ccaacgattc agactacggt cttggttcgg caattttctc aaaagacatc 1320
gtaacatgtc acaaatttgc ccaggacgtg gaagcaggta tggctcggat caactcttct 1380

aacgatgttg acatcaacgt cccattcggg ggtgtaaaaa tgtccggtca tggccgtgaa 1440
ttaggtgaat acggttttagc taacttcaact aacgttaaat cagttcatgt taacctagga 1500
agaaagtgtg aa 1512

<210> 747

<211> 503

<212> PRT

<213> Artificial Sequence

<220><223> KLLA0D10021 from Kluyveromyces lactis

<400> 747

Met Ser Phe Thr Ser Ile Lys Leu Ala Asp Gly Thr Glu Tyr Gln Gln

1 5 10 15

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

20 25 30

Glu Thr Leu Thr Ser Tyr Asn Pro Ser Ser Gly Glu Glu Ile Cys Lys

35 40 45
 Val Gln Cys Ala Asp Pro Lys Ile Asp Val Asp Leu Ala Val Lys Ala
 50 55 60
 Ala Arg Asp Ala Phe Pro Gly Trp Lys Ala Thr Pro Ala Ile Glu Lys
 65 70 75 80
 Arg Asp Leu Phe Tyr Lys Leu Ala Gly Leu Leu Glu Arg Asp Ile Asp

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

 145 150 155 160
 Ile Ile Pro Phe Asn Tyr Pro Leu Ala Met Met Ser Trp Lys Trp Thr
 165 170 175
 Ala Ile Ala Thr Gly Asn Thr Thr Val Phe Lys Ser Ala Asp Gln Thr
 180 185 190
 Pro Leu Ser Ile Leu Tyr Phe Ala Lys Leu Val Lys Glu Ala Gly Phe
 195 200 205
 Pro Ala Gly Val Phe Asn Val Leu Ser Gly Thr Gly Ala Ser Val Gly

 210 215 220
 Ser Ala Ile Val Lys His Thr Gly Val Asp Lys Ile Ala Phe Thr Gly
 225 230 235 240
 Ser Thr Ala Val Gly Gln Ile Ile Gln Gln Gln Ser Ala Val Asn Leu
 245 250 255
 Lys Pro Ile Thr Leu Glu Cys Gly Gly Lys Ser Ala Asn Met Ile Phe
 260 265 270
 Glu Asp Cys Asp Leu Glu Gln Ala Val Lys Trp Ser Ala Trp Gly Ile

 275 280 285

Phe Asn Asn Met Gly Gln Ile Cys Ser Gly Thr Ser Arg Val Phe Val
 290 295 300
 Gln Glu Ser Ile Tyr Glu Glu Phe Val Ala Gln Leu Ala Asp His Val
 305 310 315 320
 Lys Lys Thr Tyr Ile Val Gly Ala Pro Ser Asp Asp Ala Val Ile Val
 325 330 335
 Gly Pro Gln Val Ser Lys Lys Gln Gln Glu Arg Val Leu Glu Tyr Ile
 340 345 350
 Lys Ile Gly Ile Glu Glu Gly Cys Lys Val Pro Leu Gly Gly Leu Gly
 355 360 365
 Ala Pro Glu Pro Ile Ala Asn Asp Pro Ala Leu Ala Asn Gly Tyr Tyr
 370 375 380
 Val Lys Pro Thr Ile Leu Ala Asp Val Lys Ala Ser Tyr Arg Ile Ala
 385 390 395 400
 Lys Glu Glu Ile Phe Gly Pro Val Ile Val Val Gly Ser Phe Lys Asp
 405 410 415
 Tyr Lys Glu Ala Leu Asp Ile Ala Asn Asp Ser Asp Tyr Gly Leu Gly
 420 425 430
 Ser Ala Ile Phe Ser Lys Asp Ile Val Thr Cys His Lys Phe Ala Gln
 435 440 445
 Asp Val Glu Ala Gly Met Val Trp Ile Asn Ser Ser Asn Asp Val Asp
 450 455 460
 Ile Asn Val Pro Phe Gly Gly Val Lys Met Ser Gly His Gly Arg Glu
 465 470 475 480
 Leu Gly Glu Tyr Gly Leu Ala Asn Phe Thr Asn Val Lys Ser Val His
 485 490 495
 Val Asn Leu Gly Arg Lys Leu
 500
 <210> 748
 <211> 1506
 <212> DNA
 <213> Artificial Sequence

<220><223> KLLA0D09999 from Kluyveromyces lactis

<400> 748

atgtcattgt acgaaacggt taacattcca cagttgaaca ttacatatga gcagccattg	60
gggttgttca tcaacaacga gtttgtcaag tccagtgcag gtaagaaaat tgaaagcatc	120
aaccaagta ctgaggaatc gattgtctct ttttttcag cttctgaaaa agatgtcaat	180
agtgcgttta cagcagcaag agatgctttt gaaaacaatt ggtccaagac atcaggcgaa	240
cagagaggaa ttcttttgc gaagttattt gatttgattg agagagataa gttattgcta	300
gctgcgattg agacgttaga ttccgggaag cttaccact ctaatagtct tgctgatctt	360
gaacaaatct tgcaattaac tcgttactat gctggcgggt ctgataagtt caccggaggg	420
gaatatattc cattaaacca agacaagtat gcgtatacca tgaaggtagc ttatggtgtt	480
gttgcacaaa ttgttccatg gaattatccc ttggctatgg ctagtggaa aatccaatcg	540
tgtttggctg ctggtaacac tattgttatc aagccagcag agaacacttc ttgtctttg	600
ctttacttcg cacaacttgt caaagaggct ggctttcctc cagggtgtgtt gaatatctta	660
ccaggttatg gttccgttac gggaactgca atggctagcc aactgacat agacaagatt	720
gcgtttactg gttcgaccgc tgttggacaa aaagtcattg cacttgctgc ccaatcaaac	780
ttgaaagctg ttacttttga atgcggtggg aagtcctcat gtctaatttt tgctgatgcc	840
aaccttgatg aagccattcc ttgggttgcg ggaggtatct tctataactc agggcagaat	900
tgtacagcaa actcgaggat ctacgtccac gaggatattt acgaagagtt tttaactaaa	960
tttgaagaat acgctaaatc aaactggaag ttccgagcta aattcgatcc ttctgataaa	1020
gactgtactc tagggcctgt tatttcaaaa gtccaatatg atcgtatcca atcctatgtt	1080
gaacatggta aagcctctga gaaactaggc ttcaaggagt ttgccgaata tccttctaag	1140
ggttatttta tccacctac tatattttca gatgttccaa actcttcgaa actctctcaa	1200
gaagaaatth ttggaccggt ggggtttgtg gcatcattht caacttatga ggaagccatc	1260
aagttggcaa atgactctc ttacggtttg gcttcttgtg ttttctctga aaatgccga	1320
acaacagcac aatttgctcg cgatattagg gcaggtaccg tttggatcaa ttcttccaat	1380
gacgaagaaa tttctgttcc atttgggtggg ttcaaatgt ccggtattgg tcgtgaattg	1440
gggcgtaacg gtgttgactc ttatttgcaa acaaaagctg ttacagtaaa tattgcaaag	1500
ctgtga	1506

<210> 749

<211> 501

<212> PRT

<213> Artificial Sequence

<220><223> KLLA0D09999 from Kluyveromyces lactis

<400> 749

Met Ser Leu Tyr Glu Thr Val Asn Ile Pro Gln Leu Asn Ile Thr Tyr

1 5 10 15

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

20 25 30

Asp Gly Lys Lys Ile Glu Ser Ile Asn Pro Ser Thr Glu Glu Ser Ile

35 40 45

Val Ser Phe Phe Ala Ala Ser Glu Lys Asp Val Asn Ser Ala Val Thr

50 55 60

Ala Ala Arg Asp Ala Phe Glu Asn Asn Trp Ser Lys Thr Ser Gly Glu

65 70 75 80

Gln Arg Gly Ile Leu Leu Ser Lys Leu Phe Asp Leu Ile Glu Arg Asp

85 90 95

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

100 105 110

His Ser Asn Ser Leu Ala Asp Leu Glu Gln Ile Leu Gln Leu Thr Arg

115 120 125

Tyr Tyr Ala Gly Gly Ala Asp Lys Phe Thr Gly Gly Glu Tyr Ile Pro

130 135 140

Leu Asn Gln Asp Lys Tyr Ala Tyr Thr Met Lys Val Pro Tyr Gly Val

145 150 155 160

Val Ala Gln Ile Val Pro Trp Asn Tyr Pro Leu Ala Met Ala Ser Trp

165 170 175

Lys Ile Gln Ser Cys Leu Ala Ala Gly Asn Thr Ile Val Ile Lys Pro

180 185 190

Ala Glu Asn Thr Ser Leu Ser Leu Leu Tyr Phe Ala Gln Leu Val Lys

195 200 205

Glu Ala Gly Phe Pro Pro Gly Val Leu Asn Ile Leu Pro Gly Tyr Gly

210 215 220

Ser Val Thr Gly Thr Ala Met Ala Ser His Thr Asp Ile Asp Lys Ile
 225 230 235 240
 Ala Phe Thr Gly Ser Thr Ala Val Gly Gln Lys Val Met Ser Leu Ala
 245 250 255
 Ala Gln Ser Asn Leu Lys Ala Val Thr Leu Glu Cys Gly Gly Lys Ser
 260 265 270
 Pro Cys Leu Ile Phe Ala Asp Ala Asn Leu Asp Glu Ala Ile Pro Trp
 275 280 285
 Val Ala Gly Gly Ile Phe Tyr Asn Ser Gly Gln Asn Cys Thr Ala Asn
 290 295 300
 Ser Arg Ile Tyr Val His Glu Asp Ile Tyr Glu Glu Phe Leu Thr Lys
 305 310 315 320
 Phe Glu Glu Tyr Ala Lys Ser Asn Trp Lys Phe Gly Ala Lys Phe Asp
 325 330 335
 Pro Phe Asp Lys Asp Cys Thr Leu Gly Pro Val Ile Ser Lys Val Gln
 340 345 350
 Tyr Asp Arg Ile Gln Ser Tyr Val Glu His Gly Lys Ala Ser Glu Lys
 355 360 365
 Leu Gly Phe Lys Glu Phe Ala Glu Tyr Pro Ser Lys Gly Tyr Phe Ile
 370 375 380
 Pro Pro Thr Ile Phe Ser Asp Val Pro Asn Ser Ser Lys Leu Ser Gln
 385 390 395 400
 Glu Glu Ile Phe Gly Pro Val Gly Val Val Ala Ser Phe Ser Thr Tyr
 405 410 415
 Glu Glu Ala Ile Lys Leu Ala Asn Asp Ser Ser Tyr Gly Leu Ala Ser
 420 425 430
 Cys Val Phe Ser Glu Asn Ala Arg Thr Thr Ala Gln Phe Ala Arg Asp
 435 440 445
 Ile Arg Ala Gly Thr Val Trp Ile Asn Ser Ser Asn Asp Glu Glu Ile
 450 455 460
 Ser Val Pro Phe Gly Gly Phe Lys Met Ser Gly Ile Gly Arg Glu Leu

465 470 475 480
 Gly Arg Asn Gly Val Asp Ser Tyr Leu Gln Thr Lys Ala Val His Val
 485 490 495
 Asn Ile Ala Lys Leu
 500
 <210> 750
 <211> 1506
 <212> DNA
 <213> Artificial Sequence
 <220><223> ALD2 from Pichia stipits
 <400> 750
 ctacaacttg ttgccaagt taactgaat ggcctttgcc tgggtgtaga tggacaatcc 60
 glattcacc aattctctac ctacaccaga catcttgaca ccaccgaatg gaatgtggac 120
 atcagagtgc ttggaagagt tgatccacac cataccagct tgtatttcag cagccatagt 180
 gtgggccacg gtgatgtcct tgggaagat ggcagcacc aaaccgtatt gggctctggtt 240
 ggcgtaagt atgacttctt cgtctgacga gaacttacc accactaca cagggccaaa 300
 tatttctcgc ttgacgattc tcatctcagg ctgtgatgca gcaaaaatag ttggcttgat 360
 glagaatcct ttggataatt caccctccga gttcttttca cctcccaaaa gcaagcgggc 420
 accttctttc ttaccgattt cgatgtagct caacacttta tcacgtgaa ccttcgaaac 480
 ttgtggacca acgactgctt cagaatcgaa catactacce tgtttgtagg cttctttcac 540
 gtgtgcagtg tattctcca agaactgtc gtagacagat tcgtgaacgt acactctgga 600
 agtggacgta cagatctgac cttggttact catgataccg atggcagccc acttgacagc 660
 ttgttccaag tcagcgtctt cacggatgat caaaggcgac ttaccaccac actctaaggt 720
 cacagccttc aagttagaag cagcagcctg ttgtataagc ttaccagtag cagtagaacc 780
 agtgaaagcg atcttgtcga catctaaatg ggatgaaaga gccttaccgg caacagcacc 840
 gaagccggag atgatgttaa ccacaccggc tgggaaacca gcgtccttga agaggcgagc 900
 tacgtacaag agagacaatg gtgtaatttc tgaagttttc aacacaacta cgttaccagc 960
 agccaaagct ggagccaact tccaggctgc catagccaat gggtagttcc atgggatgat 1020
 ctgaccacaa acaccaagtg gttcatggat ggtgtaggcc aacttctttg gatcgttctg 1080
 gatcacctta ccttgatttt tgtagccca gccagcaaag tatctgaaac aactgataca 1140
 ttcgtcaata tcgtatacgg cgttttgttc cttggcttta ccagagtccc aggcttcaat 1200
 ggcaccaatc aagtctcttt gggcttcaaa ggcacacg accttgttca acaagattcc 1260

tctgtccaca ccggtgacct tcttccagtc cttaaagcg gctctggcag ccttgacggc 1320
gatatcgaca tcttcttctt cggcacagta aacggtacca ttgacttcac tggtagatgg 1380

gttgatggag tccaacgtct ttccggaaac agacttgaca aactcgttgt tgatgaacaa 1440
accagtaggc tgtgtgaatg tgttgccgtc tggaagggtacc accaggtatt ctaatgggag 1500
agacat 1506

<210> 751
<211> 501
<212> PRT
<213> Artificial Sequence
<220><223> ALD2 from Pichia stipits
<400> 751

Met Ser Leu Pro Leu Glu Tyr Ser Val Thr Leu Pro Asp Gly Asn Thr
1 5 10 15
Phe Thr Gln Pro Thr Gly Leu Phe Ile Asn Asn Glu Phe Val Lys Ser

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

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

145 150 155 160

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

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

465 470 475 480
 Glu Tyr Gly Leu Ser Ile Tyr Thr Gln Ala Lys Ala Ile His Val Asn
 485 490 495
 Leu Gly Asn Lys Leu

500

<210> 752

<211> 1434

<212> DNA

<213> Artificial Sequence

<220><223> ALD3 from *Pichia stipitis*

<400> 752

atgtccgcta ttagtagaac attccccagt attatcgagg gaaaggattt ccacagcaat 60

gaaaaacatc cagtatactc ccatgtcacc cagaaagaag caatccacta tttctcctac 120

ttaactgata tcaagaaggc agtttcgaaa atcgtgctg atgctgacga agggtttgaa 180

gaatggctgt ccatggccta tcaagaaaga gtgaagattt tcgagaaggc tgctgcttta 240

cttgctgaaa gaagagaaga gttgatagct tctcacaaaa atatcggagg tcctacctgg 300

ttttcccatg tgaatgctga tgaaatcatc tcgcaattga aggaatatac ttcacttctc 360

tctagaccta ctggtttagt agcccaatct gctcattctg atctcgact cgtcgtcaag 420

cagccactag gtctgttctt cgccattgct ccttggaacg ctctgttctt tttggcaggt 480

agagccatag tggtccgtt ggctcgggc tgttcggtca tctaaaaagc ttctgaaaag 540

gctccagaat ccgcatacct tggtgtgaag accttcattg atgctggtat cccgtcaaaa 600

gcattgcaat tggcttctcat caaaccagat gacaatccag aattcatcaa ctccatcttg 660

gacactgggt tgatcaaaaa ggtcaacttc actggctcta caatcgtggg caagaagatc 720

gctgaagctg ccagtaaaca cttagtagca tatcttatgg aattaggcgg aaagaatgtg 780

tctattgttg aaaaggatgc tgacttgggtg agggccgttg aaactataat ctggagttcg 840
 tggtcgcaca aaggtcaaat atgcatgagt actgacaagg tcttcgttga tgaaagcatc 900
 tacgacaaat tcgttgctca attaaaagta tccgccaatg agatcgtcaa ggaccccgac 960

tacgcaattt ctcaaagaga tattacattc aagagaaacc ttgttaagtt ggtaagaat 1020
 gcattagatt tgggtgcaaa ttgatattt ggtaaattaa atgaccattt ggacagcggc 1080
 tccttcagtc cattgatctt ggaaaatgtc acttcaaaca tgttgcttga ttctaccgaa 1140
 tcattcggac ctttgttcgc tgtatataag tattcagata caatcaaact tgtcaaggaa 1200
 ttaaacagag ctgattatgg attgaaggcc tcaatttggc cccaaaatgt ttgcaagca 1260
 ttggaaacag ctaaaaaaat ccacgtaggt ggtgtacata tcaatagttc tacgattcac 1320
 gacgaagcga ctctaccaca tggcgggtgtt aagtcaagtg gtgctggaag attcaactcc 1380

atatggggta ttgacgattt ttccattacc aagacaatta ctcttagtca gtaa 1434

<210> 753

<211> 477

<212> PRT

<213> Artificial Sequence

<220><223> ALD3 from Pichia stipitis

<400> 753

Met Ser Ala Ile Ser Arg Thr Phe Pro Ser Ile Ile Glu Gly Lys Asp

1 5 10 15

Phe His Ser Asn Glu Lys His Pro Val Tyr Ser His Val Thr Gln Lys

20 25 30

Glu Ala Ile His Tyr Phe Ser Tyr Leu Thr Asp Ile Lys Lys Ala Val

35 40 45

Ser Lys Ile Ala Ala Asp Ala Asp Glu Gly Phe Glu Glu Trp Ser Ser

50 55 60

Met Ala Tyr Gln Glu Arg Val Lys Ile Phe Glu Lys Ala Ala Ala Leu

65 70 75 80

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

85 90 95

Gly Pro Thr Trp Phe Ser His Val Asn Ala Asp Glu Ile Ile Ser Gln

100 105 110

Leu Lys Glu Tyr Thr Ser Leu Leu Ser Arg Pro Thr Gly Leu Val Ala
 115 120 125
 Gln Ser Ala His Ser Asp Leu Ala Leu Val Val Lys Gln Pro Leu Gly
 130 135 140
 Pro Val Leu Ala Ile Ala Pro Trp Asn Ala Pro Val Leu Leu Ala Gly
 145 150 155 160
 Arg Ala Ile Val Ala Pro Leu Ala Ala Gly Cys Ser Val Ile Leu Lys
 165 170 175
 Ala Ser Glu Lys Ala Pro Glu Ser Ala Tyr Leu Val Val Lys Thr Phe
 180 185 190
 Ile Asp Ala Gly Ile Pro Ser Lys Ala Leu Gln Leu Val Phe Ile Lys
 195 200 205
 Pro Asp Asp Asn Pro Glu Phe Ile Asn Ser Ile Leu Asp Thr Gly Leu
 210 215 220
 Ile Lys Lys Val Asn Phe Thr Gly Ser Thr Ile Val Gly Lys Lys Ile
 225 230 235 240
 Ala Glu Ala Ala Ser Lys His Leu Val Pro Tyr Leu Met Glu Leu Gly
 245 250 255
 Gly Lys Asn Val Ser Ile Val Glu Lys Asp Ala Asp Leu Val Arg Ala
 260 265 270
 Val Glu Thr Ile Ile Trp Ser Ser Trp Ser His Lys Gly Gln Ile Cys
 275 280 285
 Met Ser Thr Asp Lys Val Phe Val Asp Glu Ser Ile Tyr Asp Lys Phe
 290 295 300
 Val Ala Gln Leu Lys Val Ser Ala Asn Glu Ile Val Lys Asp Pro Asp
 305 310 315 320
 Tyr Ala Ile Ser Gln Arg Asp Ile Thr Phe Lys Arg Asn Leu Val Lys
 325 330 335
 Leu Val Lys Asn Ala Leu Asp Leu Gly Ala Asn Leu Ile Phe Gly Lys
 340 345 350
 Leu Asn Asp His Leu Asp Ser Gly Ser Phe Ser Pro Leu Ile Leu Glu

355 360 365
 Asn Val Thr Ser Asn Met Leu Leu Asp Ser Thr Glu Ser Phe Gly Pro
 370 375 380
 Leu Phe Ala Val Tyr Lys Tyr Ser Asp Thr Ile Lys Leu Val Lys Glu
 385 390 395 400
 Leu Asn Arg Ala Asp Tyr Gly Leu Lys Ala Ser Ile Trp Ser Gln Asn
 405 410 415
 Val Leu Gln Ala Leu Glu Thr Ala Lys Lys Ile His Val Gly Gly Val

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

<210> 754

<211> 1434

<212> DNA

<213> Artificial Sequence

<220

><223> ALD4 from Pichia stipitis

<400> 754

atggcaactg cttcaacaac tttctcgagc attatcgaag gtaaggtttt ctacactgaa	60
ggaaaacatc ctgtctactc acacgataac catgaggaaa taattcacca cttttcttac	120
ttgactgaca tcaagaagtc agtcccccaa attgcagctg cagctgagca aggttttgag	180
gaatgggtcat ctttcgcata tgcagaaaaa gtgaagatit ttgacaaagc tgctatcttg	240
cttactgaaa ggagagacga attgatcgcc tcgcacaaga acatcggagg tcctacatgg	300
ttctctcagc tcaacgcaga tgagtcgctc tcacaattga aggaatacac tggacttctc	360
tcaagaccta cgggtgttgt agctcagctc attcacactg acctgcact tactgtaaga	420
caaccaattg gcccagttct tgctatttct ccttggaatg cacctgttgt attagcaagt	480
agagccatct tggtccatt ggctgctggt tgttctgtta tcttgaaggc ttccgagaag	540
gtccagaat ctgcctacct catagtcaga accttcattg atgcaggagt accacctaag	600
gcattgcaat tagtatttat caagcccag gacaacgcta cttcattag ctcgattttg	660

cacactggtt caatcagaaa gatcaacttc actggatcaa ctctcgttgg gaagaaaatt 720
gcagaagttg ccagtaagta cttagtcca tatcttatgg aattaggagg gaagaacgtc 780
tctattgttg aaaaggatgc tgacttaacc aaggcagctg gtagtattat ctggagctcg 840

tgggtccaca aagggtcaaat atgtatgagc acagacaagg tttttgttga agattctata 900
tacgaccact ttattgccca attaaaagag ctggctgcgg atatcgtgaa agaccctgac 960
taccaaatat ctcaaagaga tattacattc aaaagaaagc ttgttgagtt ggtaaaggat 1020
gccctagaca aggggtgctaa cttgggtgttt ggtgaactca attctgatgt caacagcagc 1080
tcattcagtc cattgatatt agagaatgtc acttcagaca tgttgcttga ttccactgaa 1140
tccttcggac ctttgttttc tgtgcataag tatgccaata cgaacaaact agtcaaggaa 1200
ttaaacagag ttgactatgg gttgaaggct tctatttggg ccataaatgt tttggaagcc 1260

ttagaaaccg caaagaagat ccatgtcggg ggtgtacata tcaatagtcc aactatccat 1320
gatgaaccaa ctcttcaca cgggtggtgtt aaatcaagtg gaactggaag attcaactcg 1380
acatggggaa ttgatgactt ttctatcaca aaaactatta ccctaagcca gtga 1434

<210> 755

<211> 477

<212> PRT

<213> Artificial Sequence

<220><223> ALD4 from Pichia stipitis

<400> 755

Met Ala Thr Ala Ser Thr Thr Phe Ser Ser Ile Ile Glu Gly Lys Val

1 5 10 15

Phe Tyr Thr Glu Gly Lys His Pro Val Tyr Ser His Asp Asn His Glu

20 25 30

Glu Ile Ile His His Phe Ser Tyr Leu Thr Asp Ile Lys Lys Ser Val

35 40 45

Pro Gln Ile Ala Ala Ala Ala Glu Gln Gly Phe Glu Glu Trp Ser Ser

50 55 60

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

65 70 75 80

Leu Thr Glu Arg Arg Asp Glu Leu Ile Ala Ser His Lys Asn Ile Gly

85 90 95

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

 145 150 155 160
 Arg Ala Ile Leu Ala Pro Leu Ala Ala Gly Cys Ser Val Ile Leu Lys
 165 170 175
 Ala Ser Glu Lys Ala Pro Glu Ser Ala Tyr Leu Ile Val Arg Thr Phe
 180 185 190
 Ile Asp Ala Gly Val Pro Pro Lys Ala Leu Gln Leu Val Phe Ile Lys
 195 200 205
 Pro Glu Asp Asn Ala Thr Phe Ile Ser Ser Ile Leu His Thr Gly Ser

 210 215 220
 Ile Arg Lys Ile Asn Phe Thr Gly Ser Thr Leu Val Gly Lys Lys Ile
 225 230 235 240
 Ala Glu Val Ala Ser Lys Tyr Leu Val Pro Tyr Leu Met Glu Leu Gly
 245 250 255
 Gly Lys Asn Val Ser Ile Val Glu Lys Asp Ala Asp Leu Thr Lys Ala
 260 265 270
 Ala Gly Ser Ile Ile Trp Ser Ser Trp Ser His Lys Gly Gln Ile Cys

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

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

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

465 470 475

<210> 756

<211> 1488

<212> DNA

<213> Artificial Sequence

<220><223> ALD5 from *Pichia stipitis*

<400> 756

atgtccttgc cattgttctg tcccatcaaa ttgccaaacg gcactaccta tgagcagcct	60
accggtttgt tcatcaacaa cgagtttctc cagtccaagt ccaagaagac ctttggaaact	120
gtttctcctt ctactgaaga agagatcacc caagtctacg aagccttctc cgaagatatac	180
gatgatgccg ttgaagctgc tactgctgct ttccactcat cgtggtctac ttccgaccct	240
caagtcagaa tgaaggttct ctacaaattg gctgacttga tcgacgaaca cgccgacact	300
ttggcccaca ttgaagctct tgataacggt aagtcgttga tgtgttccaa gggtagcgtt	360
gctttgactg cgcctactt cagatcctgt gctggctgga cgcacaagat taagggttct	420
gttatcgaaa ctggtgacac ccactttaac tacaccagaa gagaaccaat tgggtgttgt	480
ggccagatca tcccatggaa ctcccattg ttgatggctt cctggaagtt aggaccagtc	540

ctttgtaccg gttgcaccac tgttttgaag actgctgaat cgactccttt gtctgtctctt 600
tacttggtctt ctttgatcaa ggaagccggt gctccacctg gtgttgtcaa tgttgtttcc 660
ggtttcggtc caactgccgg tgcctcaatc tcttctcacc caaagatcaa gaaggttgcc 720

ttcactgggtt ctaccgtac cggtagacac atcatgaagg ctgctgctga atccaacttg 780
aagaaggtca ctttggaatt ggggtggaag tctccaaaca ttgtctttga cgatgctgac 840
gttaagtcca ccaccaaca cttggttact ggtatcttct acaacactgg tgaagtgtgc 900
tgtgccggtt ccagaatcta cgtccaggaa ggtatctacg acaagattgt ttccgagttc 960
aagaatgcag ccgaatcact taagatcggt gaccattca aggaagacac cttcatgggt 1020
gccccaaact ctcaattgca attggacaag atcttgaagt acatcgatat cggtagaag 1080
gaaggtgccca ccgttatcac tgggtggtgaa agattcggt acaagggtta ctttatcaag 1140

ccaaccatct tcggagatgt caaggaagac caccagatcg tcagggatga aatcttcggt 1200
cctgttgtca ccaccacaa gttcaagacc gtcgaagaag tcctgcctt ggccaacgac 1260
tccgagtacg gtttggtgc tgggtgtcac accactaacc tcagcactgc catcagtgtt 1320
tccaacaaga tcaattccgg taccatttgg gtcaacacct acaatgactt ccaccaatg 1380
gttcattcgc gaggttactc ccaatctggt attggtagag aaatgggaga agaagccttg 1440
gacaactaca ccaggtcaa ggcagtcaga attggtctta gccagtaa 1488

<210> 757

<211> 495

<212> PRT

<213> Artificial Sequence

<220><223> ALD5 from Pichia stipitis

<400> 757

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

1 5 10 15

Tyr Glu Gln Pro Thr Gly Leu Phe Ile Asn Asn Glu Phe Val Gln Ser

20 25 30

Lys Ser Lys Lys Thr Phe Gly Thr Val Ser Pro Ser Thr Glu Glu Glu

35 40 45

Ile Thr Gln Val Tyr Glu Ala Phe Ser Glu Asp Ile Asp Asp Ala Val

50 55 60

Glu Ala Ala Thr Ala Ala Phe His Ser Ser Trp Ser Thr Ser Asp Pro

65					70				75				80			
Gln	Val	Arg	Met	Lys	Val	Leu	Tyr	Lys	Leu	Ala	Asp	Leu	Ile	Asp	Glu	
				85				90				95				
His	Ala	Asp	Thr	Leu	Ala	His	Ile	Glu	Ala	Leu	Asp	Asn	Gly	Lys	Ser	
				100				105				110				
Leu	Met	Cys	Ser	Lys	Gly	Asp	Val	Ala	Leu	Thr	Ala	Ala	Tyr	Phe	Arg	
				115				120				125				
Ser	Cys	Ala	Gly	Trp	Thr	Asp	Lys	Ile	Lys	Gly	Ser	Val	Ile	Glu	Thr	
				130				135				140				
Gly	Asp	Thr	His	Phe	Asn	Tyr	Thr	Arg	Arg	Glu	Pro	Ile	Gly	Val	Cys	
				145				150				155				
				160				165				170				
Gly	Gln	Ile	Ile	Pro	Trp	Asn	Phe	Pro	Leu	Leu	Met	Ala	Ser	Trp	Lys	
				165				170				175				
Leu	Gly	Pro	Val	Leu	Cys	Thr	Gly	Cys	Thr	Thr	Val	Leu	Lys	Thr	Ala	
				180				185				190				
Glu	Ser	Thr	Pro	Leu	Ser	Ala	Leu	Tyr	Leu	Ala	Ser	Leu	Ile	Lys	Glu	
				195				200				205				
Ala	Gly	Ala	Pro	Pro	Gly	Val	Val	Asn	Val	Val	Ser	Gly	Phe	Gly	Pro	
				210				215				220				
Thr	Ala	Gly	Ala	Pro	Ile	Ser	Ser	His	Pro	Lys	Ile	Lys	Lys	Val	Ala	
				225				230				235				
				240				245				250				
Phe	Thr	Gly	Ser	Thr	Ala	Thr	Gly	Arg	His	Ile	Met	Lys	Ala	Ala	Ala	
				245				250				255				
Glu	Ser	Asn	Leu	Lys	Lys	Val	Thr	Leu	Glu	Leu	Gly	Gly	Lys	Ser	Pro	
				260				265				270				
Asn	Ile	Val	Phe	Asp	Asp	Ala	Asp	Val	Lys	Ser	Thr	Ile	Gln	His	Leu	
				275				280				285				
Val	Thr	Gly	Ile	Phe	Tyr	Asn	Thr	Gly	Glu	Val	Cys	Cys	Ala	Gly	Ser	
				290				295				300				
Arg	Ile	Tyr	Val	Gln	Glu	Gly	Ile	Tyr	Asp	Lys	Ile	Val	Ser	Glu	Phe	
				305				310				315				
				320				325				330				

Lys Asn Ala Ala Glu Ser Leu Lys Ile Gly Asp Pro Phe Lys Glu Asp
 325 330 335
 Thr Phe Met Gly Ala Gln Thr Ser Gln Leu Gln Leu Asp Lys Ile Leu
 340 345 350
 Lys Tyr Ile Asp Ile Gly Lys Lys Glu Gly Ala Thr Val Ile Thr Gly
 355 360 365
 Gly Glu Arg Phe Gly Asn Lys Gly Tyr Phe Ile Lys Pro Thr Ile Phe

370 375 380
 Gly Asp Val Lys Glu Asp His Gln Ile Val Arg Asp Glu Ile Phe Gly
 385 390 395 400
 Pro Val Val Thr Ile Thr Lys Phe Lys Thr Val Glu Glu Val Ile Ala
 405 410 415
 Leu Ala Asn Asp Ser Glu Tyr Gly Leu Ala Ala Gly Val His Thr Thr
 420 425 430
 Asn Leu Ser Thr Ala Ile Ser Val Ser Asn Lys Ile Asn Ser Gly Thr

435 440 445
 Ile Trp Val Asn Thr Tyr Asn Asp Phe His Pro Met Val Pro Phe Gly
 450 455 460
 Gly Tyr Ser Gln Ser Gly Ile Gly Arg Glu Met Gly Glu Glu Ala Leu
 465 470 475 480
 Asp Asn Tyr Thr Gln Val Lys Ala Val Arg Ile Gly Leu Ser Gln
 485 490 495

<210> 758

<211> 1680

<212> DNA

<213> Artificial Sequence

<220><223> ALD7 from Pichia stipitis

<400> 758

atgactcctc cttcagcaaa caccctcaag gactccgttt cgtccgtcga atcggtcaca	60
aagatcgaca aggccacccc ctcggttcc accaaggett ctgagtcctt cttgaagtac	120
actgctttgg acgacattcc tgttggagtc aagaagctta ctgactcgtt ccacaagaat	180
ggaaagactc actccatcca gtacagattg aaccagtga gaaacctcta tttgtcatc	240

aaggataacc aggatgccat ctgcgatgcc ctcttcaagg atttcggcag agtgcctact 300
gagtcgcaga tcttggaat cgatggaagc ttgaacgagt tgggtcacac catggccaac 360

ttgcacaact ggttgaaacc cgaacctgtc aaggacttgc ccattacatt gaagctgaat 420
cctatctaca ttgaaagaat tccctacgga gttgtgttga tcatctctcc cttcaactac 480
ccattcttcc tctcattttc agccatcatt ggggccattg ctgctggtaa cgtggttgtg 540
ttcaagccct ccgaattgac cccacacttt tcgcagtgtg tactaactt gttgtctgag 600
gctcttgatg atgacctttt gttcaccata aacggttcga tccccgagac cacgaaggcc 660
ttggaacaga agtacgacaa gatcatgtac accggttaata acgctgtagg aaccattatc 720
gccaagaagg ctgccgagac tttaactcct gtcactcttag agttgggagg aaagtcacca 780

gcgtttgtct tggacgatct ttctgaaaag gagttgaccg tcgtagccag aagaatcgca 840
tggggtgatg ttaccaatgc tgggtcaaacc tgtgtggccg tggactatgt tttggctcat 900
gactctatca aagagagatt ggttcgcgag atcgtcaagg tgggtcaagga ggagtctat 960
ccagaactta acaaagacaa caaggacttc acccacatta tccacgaccg tgctgtatcc 1020
aacttatcca agatcatcaa gaccaccaag ggtaagattg ttgttgagg tgatgtagac 1080
gaagcctcga ggtacgtggc tccaactgtc gtggacaatg tagactggga cgactctacc 1140
atgaagggcg agatctttgg tctatcttg cctatcttgt cgtacagctc tttggatgaa 1200

gctttgtcaa gattgcagag tcgtcatgac actcctttgg cacagtatat cttcactggt 1260
ggatcgactt cccgtgctaa gaacctaaa cttaacaaga tctcacagca gatcagatcc 1320
ggtggtgctg tcattaacga tgtcttgatg cacgtagcct taactaatgc cccattcgga 1380
ggtattggta gticcgggtc gggatcgtag catggttggg tttctttccg tgcatttact 1440
catgagagaa ccaccatgga acaaaagtgt tggaacgact ttgttttgaa ggctagatac 1500
cctcctttca cagaaaagaa ccagaagtgt gtcagtgtt cacagacaga ctacaacggt 1560
aaggtctggt ttgatagaca aggggacgtc agagtcaagg gtccatcggg cttgttcagc 1620

acttggaact ctgttgcgg tgctgcagcc ttgatctact actttgttgg taattttag 1680

<210> 759
<211> 559
<212> PRT
<213> Artificial Sequence
<220><223> ALD7 from Pichia stipitis
<400> 759

Met Thr Pro Pro Ser Ala Asn Thr Leu Lys Asp Ser Val Ser Ser Val

1 5 10 15
 Glu Ser Phe Thr Lys Ile Asp Lys Ala Thr Pro Ser Ala Ser Thr Lys
 20 25 30
 Ala Ser Glu Ser Ile Leu Lys Tyr Thr Ala Leu Asp Asp Ile Pro Val

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

 100 105 110
 Glu Leu Val His Thr Met Ala Asn Leu His Asn Trp Leu Lys Pro Glu
 115 120 125
 Pro Val Lys Asp Leu Pro Ile Thr Leu Lys Ser Asn Pro Ile Tyr Ile
 130 135 140
 Glu Arg Ile Pro Tyr Gly Val Val Leu Ile Ile Ser Pro Phe Asn Tyr
 145 150 155 160
 Pro Phe Phe Leu Ser Phe Ser Ala Ile Ile Gly Ala Ile Ala Ala Gly

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

 225 230 235 240
 Ala Lys Lys Ala Ala Glu Thr Leu Thr Pro Val Ile Leu Glu Leu Gly
 245 250 255

Gly Lys Ser Pro Ala Phe Val Leu Asp Asp Leu Ser Glu Lys Glu Leu
 260 265 270
 Thr Val Val Ala Arg Arg Ile Ala Trp Gly Arg Phe Thr Asn Ala Gly
 275 280 285
 Gln Thr Cys Val Ala Val Asp Tyr Val Leu Ala His Asp Ser Ile Lys
 290 295 300
 Glu Arg Leu Val Arg Glu Ile Val Lys Val Val Lys Glu Glu Phe Tyr
 305 310 315 320
 Pro Glu Leu Asn Lys Asp Asn Lys Asp Phe Thr His Ile Ile His Asp
 325 330 335
 Arg Ala Val Ser Asn Leu Ser Lys Ile Ile Lys Thr Thr Lys Gly Lys
 340 345 350
 Ile Val Val Gly Gly Asp Val Asp Glu Ala Ser Arg Tyr Val Ala Pro
 355 360 365
 Thr Val Val Asp Asn Val Asp Trp Asp Asp Ser Thr Met Lys Gly Glu
 370 375 380
 Ile Phe Gly Pro Ile Leu Pro Ile Leu Ser Tyr Ser Ser Leu Asp Glu
 385 390 395 400
 Ala Leu Ser Arg Leu Gln Ser Arg His Asp Thr Pro Leu Ala Gln Tyr
 405 410 415
 Ile Phe Thr Gly Gly Ser Thr Ser Arg Ala Lys Asn Pro Lys Leu Asn
 420 425 430
 Lys Ile Ser Gln Gln Ile Arg Ser Gly Gly Ala Val Ile Asn Asp Val
 435 440 445
 Leu Met His Val Ala Leu Thr Asn Ala Pro Phe Gly Gly Ile Gly Ser
 450 455 460
 Ser Gly Ser Gly Ser Tyr His Gly Trp Phe Ser Phe Arg Ala Phe Thr
 465 470 475 480
 His Glu Arg Thr Thr Met Glu Gln Lys Leu Trp Asn Asp Phe Val Leu
 485 490 495
 Lys Ala Arg Tyr Pro Pro Phe Thr Glu Lys Asn Gln Lys Leu Val Ser

500 505 510
Ala Ser Gln Thr Asp Tyr Asn Gly Lys Val Trp Phe Asp Arg Gln Gly
515 520 525
Asp Val Arg Val Lys Gly Pro Ser Gly Leu Phe Ser Thr Trp Thr Ser
530 535 540
Val Ala Gly Val Ala Ala Leu Ile Tyr Tyr Phe Val Gly Asn Leu

545 550 555

<210> 760

<211> 1440

<212> DNA

<213> Artificial Sequence

<220><223> aldA from E. coli

<400> 760

atgtcagtac ccgttcaaca tcctatgtat atcgatggac agtttgttac ctggcgtgga	60
gacgcatgga ttgatgtggt aaacctgct acagaggctg tcatttcccg cataccgat	120
ggtcaggccg aggatgcccg taaggcaatc gatgcagcag aacgtgcaca accagaatgg	180
gaagcgttgc ctgctattga acgcgccagt tggttgcgca aaatctccgc cgggatccgc	240
gaacgcgcca gtgaaatcag tgcgctgatt gttgaagaag ggggcaagat ccagcagctg	300
gctgaagtcg aagtggcttt tactgccgac tatatcgatt acatggcgga gtgggcacgg	360
cgttacgagg gcgagattat tcaaagcgat cgtccaggag aaaatattct ttgtttaaa	420
cgtgcgcttg gtgtgactac cggcattctg ccgtggaact tcccgttctt cctcattgcc	480
cgcaaaatgg ctcccgtctt ttgaccggt aataccatcg tcattaaacc tagtgaattt	540
acgccaaca atgcgattgc attcgcaaaa atcgtcgatg aaataggcct tccgcgcggc	600
gtgtttaacc ttgtactggg gcgtggtgaa accgttgggc aagaactggc gggtaaccca	660
aaggtcgcaa tggtcagtat gacaggcagc gtctctgcag gtgagaagat catggcgact	720
gcggcgaaaa acatcaccaa agtgtgtctg gaattggggg gtaaagcacc agctatcgta	780
atggacgatg ccgatcttga actggcagtc aaagccatcg ttgattcacg cgtcattaat	840
agtgggcaag tgtgtaactg tgcagaacgt gtttatgtac agaaaggcat ttatgatcag	900
ttcgtcaatc ggctgggtga agcgatgcag gcggttcaat ttggtaaccc cgctgaacgc	960
aacgacattg cgatggggcc gttgattaac gccgcggcgc tggaaagggt cgagcaaaaa	1020
gtggcgcgcg cagtagaaga aggggcgaga gtggcggttcg gtggcaaagc ggtagagggg	1080

aaaggatatt attatccgcc gacattgctg ctggatgttc gccaggaaat gtcgattatg 1140

catgaggaaa cctttggccc ggtgctgccca gttgtcgcac ttgacacgct ggaagatgct 1200

atctcaatgg ctaatgacag tgattacggc ctgacctcat caatctatac ccaaaatctg 1260

aacgtcgcga tgaagccat taaagggtg aagtttggtg aaacttacat caaccgtgaa 1320

aacttcgaag ctatgcaagg cttccacgcc ggatggcgta aatccggtat tggcggcgca 1380

gatggtaaac atggcttgca tgaatatctg cagaccagg tggtttattt acagtcttaa 1440

<210> 761

<211> 479

<212> PRT

<213> Artificial Sequence

<220><223> aldA from E. coli

<400> 761

Met Ser Val Pro Val Gln His Pro Met Tyr Ile Asp Gly Gln Phe Val

1 5 10 15

Thr Trp Arg Gly Asp Ala Trp Ile Asp Val Val Asn Pro Ala Thr Glu

20 25 30

Ala Val Ile Ser Arg Ile Pro Asp Gly Gln Ala Glu Asp Ala Arg Lys

35 40 45

Ala Ile Asp Ala Ala Glu Arg Ala Gln Pro Glu Trp Glu Ala Leu Pro

50 55 60

Ala Ile Glu Arg Ala Ser Trp Leu Arg Lys Ile Ser Ala Gly Ile Arg

65 70 75 80

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

85 90 95

Ile Gln Gln Leu Ala Glu Val Glu Val Ala Phe Thr Ala Asp Tyr Ile

100 105 110

Asp Tyr Met Ala Glu Trp Ala Arg Arg Tyr Glu Gly Glu Ile Ile Gln

115 120 125

Ser Asp Arg Pro Gly Glu Asn Ile Leu Leu Phe Lys Arg Ala Leu Gly

130 135 140

Val Thr Thr Gly Ile Leu Pro Trp Asn Phe Pro Phe Phe Leu Ile Ala

145 150 155 160
 Arg Lys Met Ala Pro Ala Leu Leu Thr Gly Asn Thr Ile Val Ile Lys
 165 170 175
 Pro Ser Glu Phe Thr Pro Asn Asn Ala Ile Ala Phe Ala Lys Ile Val
 180 185 190
 Asp Glu Ile Gly Leu Pro Arg Gly Val Phe Asn Leu Val Leu Gly Arg

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

 260 265 270
 Ile Val Asp Ser Arg Val Ile Asn Ser Gly Gln Val Cys Asn Cys Ala
 275 280 285
 Glu Arg Val Tyr Val Gln Lys Gly Ile Tyr Asp Gln Phe Val Asn Arg
 290 295 300
 Leu Gly Glu Ala Met Gln Ala Val Gln Phe Gly Asn Pro Ala Glu Arg
 305 310 315 320
 Asn Asp Ile Ala Met Gly Pro Leu Ile Asn Ala Ala Ala Leu Glu Arg

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

 385 390 395 400

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

450 455 460
 Gly Leu His Glu Tyr Leu Gln Thr Gln Val Val Tyr Leu Gln Ser
 465 470 475

<210> 762

<211> 1539

<212> DNA

<213> Artificial Sequence

<220><223> aldB from E. coli

<400> 762

atgaccaata atcccccttc agcacagatt aagccccggcg agtatggttt ccccccaag	60
ttaaaagccc gciatgacaa ctttattggc ggcgaatggg tagcccctgc cgacggcgag	120
tattaccaga atctgacgcc ggtgaccggg cagctgctgt gcgaagtggc gtcttcgggc	180
aaacgagaca tcgatctggc gctggatgct gcgcacaaag tgaaagataa atgggcgcac	240
acctcggtgc aggatcgtgc ggcgattctg ttttaagattg cccgatcgaat ggaacaaaac	300
ctcgagctgt tagcgacagc tgaaacctgg gataacggca aaccattcgc cgaaaccagt	360
gctgcggatg taccgctggc gattgacat ttcgctatt tcgcctcgtg tattcgggcg	420
caggaaggtg ggatcagtga agttgatagc gaaacctggg cctatcattt ccatgaaccg	480
ttaggcgtgg tggggcagat tatcccgctg aacttcccgc tgctgatggc gagctggaaa	540
atggctcccg cgctggcggc gggcaactgt gtggtgctga aaccgcacg tcttaccg	600
ctttctgtac tgctgctaata ggaaattgtc ggtgatttac tgccgccggg cgtggtgaac	660
gtggtcaatg gcgcaggctg ggtaattggc gaatatctgg cgacctgaa acgcatcgcc	720
aaagtggcgt ttaccggctc aacggaagtg ggccaacaaa ttatgcaata cgcaacgcaa	780
aacattattc cggtagcgt ggagttgggc ggtaagtgc caaatatctt ctttctgat	840
gtgatggatg aagaagatgc ctttttcgat aaagcgctgg aaggctttgc actgtttgcc	900
tttaaccagg gcgaagtgtg cacctgtccg agtcgtgctt tagtgcagga atctatctac	960

gaacgcttta tggaacgcgc catccgccgt gtcgaaagca ttcgtagcgg taaccgcctc 1020

gacagcgtga cgcaaatggg cgcgcaggtt tctcacgggc aactggaaac catcctcaac 1080

tacattgata tcggtaaaaa agagggcgct gacgtgctca caggcgggcg gcgcaagctg 1140

ctggaaggtg aactgaaaga cggctactac ctggaaccga cgattctgtt tggtcagaac 1200

aatatgcggg tgttcagga ggagattttt ggcccgtgc tggcggtgac caccttcaaa 1260

acgatggaag aagcgtgga gctggcgaac gatacgcaat atggcctggg cgcgggcgtc 1320

tggagccgca acggtaatct ggcctataag atggggcgcg gcatacaggc tgggcgcgtg 1380

tggaccaact gttatcacgc ttaccggca catgcggcgt ttggtggcta caaacaatca 1440

ggtatcggtc gcgaaacca caagatgatg ctggagcatt accagcaaac caagtcctg 1500

ctggtgagct actcggataa accgttgggg ctgttctga 1539

<210> 763

<211> 512

<212> PRT

<213> Artificial Sequence

<220><223> aldB from E. coli

<400> 763

Met Thr Asn Asn Pro Pro Ser Ala Gln Ile Lys Pro Gly Glu Tyr Gly

1 5 10 15

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

20 25 30

Trp Val Ala Pro Ala Asp Gly Glu Tyr Tyr Gln Asn Leu Thr Pro Val

35 40 45

Thr Gly Gln Leu Leu Cys Glu Val Ala Ser Ser Gly Lys Arg Asp Ile

50 55 60

Asp Leu Ala Leu Asp Ala Ala His Lys Val Lys Asp Lys Trp Ala His

65 70 75 80

Thr Ser Val Gln Asp Arg Ala Ala Ile Leu Phe Lys Ile Ala Asp Arg

85 90 95

Met Glu Gln Asn Leu Glu Leu Leu Ala Thr Ala Glu Thr Trp Asp Asn

100 105 110

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

115	120	125	
Asp His Phe Arg Tyr Phe Ala Ser Cys Ile Arg Ala Gln Glu Gly Gly			
130	135	140	
Ile Ser Glu Val Asp Ser Glu Thr Val Ala Tyr His Phe His Glu Pro			
145	150	155	160
Leu Gly Val Val Gly Gln Ile Ile Pro Trp Asn Phe Pro Leu Leu Met			
165	170	175	
Ala Ser Trp Lys Met Ala Pro Ala Leu Ala Ala Gly Asn Cys Val Val			
180	185	190	
Leu Lys Pro Ala Arg Leu Thr Pro Leu Ser Val Leu Leu Leu Met Glu			
195	200	205	
Ile Val Gly Asp Leu Leu Pro Pro Gly Val Val Asn Val Val Asn Gly			
210	215	220	
Ala Gly Gly Val Ile Gly Glu Tyr Leu Ala Thr Ser Lys Arg Ile Ala			
225	230	235	240
Lys Val Ala Phe Thr Gly Ser Thr Glu Val Gly Gln Gln Ile Met Gln			
245	250	255	
Tyr Ala Thr Gln Asn Ile Ile Pro Val Thr Leu Glu Leu Gly Gly Lys			
260	265	270	
Ser Pro Asn Ile Phe Phe Ala Asp Val Met Asp Glu Glu Asp Ala Phe			
275	280	285	
Phe Asp Lys Ala Leu Glu Gly Phe Ala Leu Phe Ala Phe Asn Gln Gly			
290	295	300	
Glu Val Cys Thr Cys Pro Ser Arg Ala Leu Val Gln Glu Ser Ile Tyr			
305	310	315	320
Glu Arg Phe Met Glu Arg Ala Ile Arg Arg Val Glu Ser Ile Arg Ser			
325	330	335	
Gly Asn Pro Leu Asp Ser Val Thr Gln Met Gly Ala Gln Val Ser His			
340	345	350	
Gly Gln Leu Glu Thr Ile Leu Asn Tyr Ile Asp Ile Gly Lys Lys Glu			
355	360	365	

Gly Ala Asp Val Leu Thr Gly Gly Arg Arg Lys Leu Leu Glu Gly Glu
 370 375 380
 Leu Lys Asp Gly Tyr Tyr Leu Glu Pro Thr Ile Leu Phe Gly Gln Asn
 385 390 395 400
 Asn Met Arg Val Phe Gln Glu Glu Ile Phe Gly Pro Val Leu Ala Val
 405 410 415

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

Gly Ile Gly Arg Glu Thr His Lys Met Met Leu Glu His Tyr Gln Gln
 485 490 495
 Thr Lys Cys Leu Leu Val Ser Tyr Ser Asp Lys Pro Leu Gly Leu Phe
 500 505 510

<210> 764

<211> 1488

<212> DNA

<213> Artificial Sequence

<220><223> aldH(puC) from E. coli

<400> 764

atgaattttc atcatctggc ttactggcag gataaagcgt taagtctcgc cattgaaaac 60
 cgcttattta ttaacgggtga atatactgct gcggcggaaa atgaaacctt tgaaaccgtt 120

gatccggtca cccaggcacc gctggcgaaa attgcccgcg gcaagagcgt cgatatcgac 180
 cgtgcgatga gcgcagcacg cggcgtatit gaacgcggcg actggtcact ctcttctccg 240
 gctaaacgta aagcgggtact gaataaactc gccgatttaa tggaagccca cgccgaagag 300
 ctggcactgc tggaacctct gcacaccggc aaaccgattc gtcacagtct gcgtgatgat 360
 attcccggcg cggcgcgcgc cattcgctgg tacgccgaag cgatcgacaa agtgtatggc 420
 gaagtggcga ccaccagtag ccatgagctg gcgatgatcg tgcgtgaacc ggtcggcgtg 480

attgccgcca tcgtgccgtg gaacttcccg ctgttgctga cttgctggaa actcggcccc 540

gcgctggcgg cgggaaacag cgtgattcta aaaccgtctg aaaaatcacc gtcagtgcg 600

attcgtctcg cggggtggc gaaagaagca ggcttgccgg atggtgtgtt gaacgtggtg 660

acgggttttg gtcatgaagc cgggcaggcg ctgtcgcgtc ataacgatat cgacgccatt 720

gcctttaccg gttcaaccgg taccgggaaa cagctgctga aagatgcggg cgacagcaac 780

atgaaacgcg tctggctgga agcgggcggc aaaagcgcca acatcgtttt cgctgactgc 840

ccggatttgc aacaggcggc aagcgccacc gcagcaggca ttttctaca ccagggacag 900

gtgtgcatcg ccggaacgcg cctgttgctg gaagagagca tcgccgatga attcttagcc 960

ctgttaaaac agcaggcgca aaactggcag ccgggccatc cacttgatcc cgcaaccacc 1020

atgggcacct taatcgactg cgcccacgcc gactcgttcc atagctttat tcgggaaggc 1080

gaaagcaaag ggcaactgtt gttggatggc cgtaacgccg ggctggctgc cgccatcggc 1140

ccgaccatct ttgtggatgt ggaccgcaat gcgtccttaa gtcgcgaaga gattttcggg 1200

ccggtgctgg tggtcacgcg ttccacatca gaagaacagg cgctacagct tgccaacgac 1260

agccagtacg gccttggcgc ggcgttatgg acgcgcgacc tctccgcgc gcaccgcatg 1320

agccgacgcc tgaaagccgg ttccgtcttc gtcaataact acaacgacgg cgatatgacc 1380

gtgccgtttg gcggctataa gcagagcggc aacggtcgcg aaaaatccct gcatgccctt 1440

gaaaaattca ctgaactgaa aaccatctgg ataagcctgg aggcctga 1488

<210> 765

<211> 495

<212> PRT

<213> Artificial Sequence

<220><223> aldH(puuC) from E. coli

<400> 765

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
 385 390 395 400
 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
 465 470 475 480

Glu Lys Phe Thr Glu Leu Lys Thr Ile Trp Ile Ser Leu Glu Ala
 485 490 495

<210> 766

<211> 725

<212> DNA

<213> Artificial Sequence

<220><223> UAS (PKG1)-FAB1

<400> 766

aattaccgtc gctcgtgatt tgtttgcaaa aagaacaaaa ctgaaaaaac ccagacacgc 60
 tcgacttcct gtcttcctat tgattgcagc ttccaatttc gtcacacaac aaggtcctgt 120
 cgacgcctac ttggcttcac atacgttgca tacgtcgata tagataataa tgataatgac 180

agcaggatta tcgtaatacgt taatagttga aaatctcaaa aatgtgtggg tcattacgta 240

aataatgata ggaatgggat tcttctatct ttcctttttc cattctagca gccgtcggga 300

aaacgtggca tctctctctt cgggctcaat tggagtcacg ctgccgtgag catcctctct 360

ttccatatct aacaactgag cacgtaacca atggaaaagc atgagcttag cgttgctcca 420

aaaaagtatt ggatgggtta taccatttgt ctgttctctt ctgactttga ctctcctcaaaa 480

aaaaaaaaatc tacaatcaac agatcgcttc aattacgccc tcacaaaaac ttttttcctt 540

cttcttcgcc cacgttaaat tttatccctc atgttgtcta acggatttct gcacttgatt 600

tattataaaa agacaaagac ataatacttc tctatcaatt tcagttattg ttcttccttg 660

cgttattctt ctgttctctt ttttcttttg tcatatataa ccataaccaa gtaatacata 720

ttcaa 725

<210> 767

<211> 47

<212> DNA

<213> Artificial Sequence

<220><223> T-FBA1(SalI) primer

<400> 767

attctacgta cgtcgacgcc tacttggctt cacatacgtt gcatacg 47

<210> 768

<211> 49

<212> DNA

<213> Artificial Sequence

<220><223> B-FBA1(SpeI) primer

<400> 768

glatcaata ctagtttgaa tatgtattac ttggttatgg ttatatatg 49

<210> 769

<211> 10494

<212> DNA

<213> Artificial Sequence

<220><223> pWS358-PGK1p-GUS plasmid

<400> 769

tcgcgcgttt cggatgatgac ggtgaaaacc tctgacacat gcagctccc gagacgggtca 60

cagcttgtct gtaagcggat gccgggagca gacaagccc tcagggcgcg tcagcgggtg 120

ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc	180
accataaatt cccgttttaa gagcttgggtg agcgctagga gtcactgccca ggtatcgttt	240
gaacacggca ttagtcaggg aagtcataac acagtccttt cccgcaattt tctttttcta	300
ttactcttgg cctcctctag tacactctat atttttttat gcctcggtaa tgattttcat	360
tttttttttt cccctagcgg atgactcttt ttttttctta gcgattggca ttatcacata	420
atgaattata cattatataa agtaatgtga tttcttcgaa gaatatacta aaaaatgagc	480
aggcaagata aacgaaggca aagatgacag agcagaaagc cctagtaaag cgtattacaa	540
atgaaaccaa gattcagatt gcgactcttt taaagggtgg tcccctagcg atagagcact	600
cgatcttccc agaaaaagag gcagaagcag tagcagaaca ggccacacaa tcgcaagtga	660
ttaacgtcca cacaggata gggtttctgg accatatgct agggattcat aaccttttc	720
tcaatcgaat tacacagaac acaccgtaca aacctctcta tcataactac ttaatagtca	780
cacacgtact cgictaaata cacatcatcg tcctacaagt tcatcaaagt gttggacaga	840
caactatacc agcatggatc tcttgatcgc gttcttttct cccgctctct cgcaataaca	900
atgaacactg ggtcaatcat agcctacaca ggtgaacaga gtagcgttta tacagggttt	960
atacggatgat tcctacggca aaaatttttc atttctaaaa aaaaaagaa aaatttttct	1020
ttccaacgct agaaggaaaa gaaaaatcta attaaattga tttggtgatt ttctgagagt	1080
tccctttttc atatatcgaa ttttgaatat aaaaggagat cgaaaaaatt ttctattca	1140
atctgttttc tggttttatt tgatagtttt tttgtgtatt attattatgg attagtactg	1200
gtttatatgg gtttttctgt ataacttctt tttatttttag tttgtttaat ctatttttga	1260
gttacattat agttccctaa ctgcaagaga agtaacatta aaactcgaga tgggtaagga	1320
aaagactcac gtttcgagcg cgcgattaaa ttccaacatg gatgctgatt tatatgggta	1380
taaatgggct cgcgataatg tcgggcaatc aggtgcgaca atctatcgat tgtatgggaa	1440
gcccgatcgc ccagagtgtg ttctgaaaca tggcaaagggt agcgttgcca atgatgttac	1500
agatgagatg gtcagactaa actggctgac ggaatttatg cctcttccga ccatcaagca	1560
ttttatccgt actcctgatg atgcatggtt actcaccact gcgatccccg gcaaaacagc	1620
attccaggta ttagaagaat atcctgattc aggtgaaaat attgttgatg cgctggcagt	1680
gttctcgcgc cggttgcatc cgattcctgt ttgtaattgt ccttttaaca gcgatcgcgt	1740
atttcgtctc gtcaggcgc aatcacgaat gaataacgggt ttgggtgatg cgagtgattt	1800
tgatgacgag cgtaatggct ggcctgttga acaagtctgg aaagaaatgc ataagctttt	1860
gccattctca ccggattcag tcgtcactca tggtgatttc tcacttgata accttatttt	1920
tgacgagggg aaattaatag gttgtattga tgttgacga gtcggaatcg cagaccgata	1980

ccaggatctt gccatcctat ggaactgcct cggtagagttt tctccttcat tacagaaacg 2040

gctttttcaa aaatatggta ttgataatcc tgatatgaat aaattgcagt ttcatttgat 2100

gctcgatgag tttttctaag ttttaactga tactactaga ttttttctct tcatttataa 2160

aatttttggg tataattgaa gctttagaag tatgaaaaaa tccitttttt tcattctttg 2220

caacaaaaat aagaagcttc ttttattcat tgaatgatg aatataaacc taacaaaaga 2280

aaaagactcg aatatcaaac attaaaaaaa aataaaagag gttatctgtt ttcccattta 2340

gttggagttt gcattttcta atagatagaa ctctcaatta atgtggattt agtttctctg 2400

ttcgctgcag catacgatat atatacatgt gtatatatgt atacctatga atgtcagtaa 2460

glatgtatac gaacagiatg atactgaaga tgacaaggta atgcatcatt ctatacgtgt 2520

cattctgaac gaggcgcgtt ttcctttttt ctttttgctt tttctttttt tttctcttga 2580

actcgacgga tctatgcggt gtgaaatacc gcacagatgc gtaaggagaa aataccgcat 2640

caggaaattg taaacgttaa tattttgtaa aaattcgcgt taaatttttg ttaaatcagc 2700

tcatttttta accaataggc cgaaatcggc aaaatccctt ataaatcaaa agaatagacc 2760

gagatagggt tgagtgttgt tccagtttgg aacaagagtc cactattaaa gaacgtggac 2820

tccaactgca aagggcgaaa aaccgtctat cagggcgatg gccactacg tgaaccatca 2880

ccctaataca gttttttggg gtcgaggtgc cgtaaagcac taaatcggaa ccctaaaggg 2940

agcccccgat ttagagcttg acggggaaaag ccggcgaacg tggcgagaaa ggaagggaag 3000

aaagcgaaag gagcgggcgc tagggcgcgt gcaagtgtag cggtcacgct gcgcgtaacc 3060

accacacccg ccgcgcttaa tgcgccgcta cagggcgcgt cgcgccattc gccattcagg 3120

ctgcgcaact gtigggaagg gcgatcgggt cgggcctctt cgctattacg ccagctggcg 3180

aaagggggat gtgctgcaag gcgattaagt tgggtaacgc cagggttttc ccagtcacga 3240

cgttgtaaaa cgacggccag tgagcgcgcg taatacgact cactataggg cgaattgggt 3300

accgggcccc cctcagagt cgacgtgagt aaggaaagag tgaggaacta tcgcatacct 3360

gcattttaaag atgccgattt gggcgcgaat cttttatttt ggcttcaccc tcatactatt 3420

atcagggccca gaaaaaggaa gtgtttccct ctttcttgaa ttgatgttac ctcataaag 3480

cacgtggcct cttatcgaga aagaaattac cgtcgcctgt gatttgtttg caaaaagaac 3540

aaaactgaaa aaaccagac acgtcgcact tctgtcttc ctattgattg cagcttccaa 3600

tttcgtcaca caacaaggtc ctacgcacgg ctacacaggtt ttgtaacaag caatcgaagg 3660

ttctggaatg gcgggaaaagg gtttagtacc acatgctatg atgccactg tgatctccag 3720

agcaaagttc gticgatcgt actgttactic tctctctttc aaacagaatt gtccgaatcg 3780

tgtgacaaca acagcctgtt ctcacacact cttttcttct aaccaagggg gtggtttagt 3840

ttagtagaac ctctgaaac ttacatttac atatatataa acttgcataa attggtcaat 3900

gcaagaaata catatttggg cttttctaata tcgtagtttt tcaagttctt agatgctttc 3960

tttttctctt ttttacagat catcaaggaa gtaattatct actttttaca acaaatataa 4020

aacaactagt atggtagtc ctgtagaaac cccaaccgt gaaatcaaaa aactcgacgg 4080

cctgtgggca ttcagctctgg atcgcgaaaa ctgtggaatt gatcagcgtt ggtgggaaag 4140

cgcgttacaa gaaagccggg caattgctgt gccaggcagt tttaacgatac agttcgccga 4200

tgcagatatt cgtaattatg cgggcaacgt ctggtatcag cggaagtct ttataccgaa 4260

aggttgggca ggccagcgta tcgtgctgcg tttcgatgcg gtcactcatt acggcaaagt 4320

gtgggtcaat aatcaggaag tgatggagca tcagggcggc tatactccat ttgaagccga 4380

tgtcacgccg taigtatttg ccgggaaaag tgiacgtatc accgtttgtg tgaacaacga 4440

actgaactgg cagactatcc cgccgggaat ggtgattacc gacgaaaacg gcaagaaaaa 4500

gcagtcttac ttccatgatt tctttaacta tgcgggaate catcgagcg taatgctcta 4560

caccacgccg aacacctggg tggacgatat caccgtggtg acgcatgtcg cgcaagactg 4620

taaccacgcg tctgttgact ggacggtggt ggccaatggt gatgtcagcg ttgaactgcg 4680

tgatgcggat caacaggttg ttgcaactgg acaaggcact agcgggactt tgcaagtggg 4740

gaatccgcac ctctggcaac cgggtgaagg ttatctctat gaactgtgcg tcacagccaa 4800

aagccagaca gagtgtgata tctaccgcgt tcgcgtcggc atccggtcag tggcagtga 4860

ggcggaacag ttcttgatta accacaaacc gttctacttt actggctttg gtcgtcatga 4920

agatgcggac ttgcgtggca aaggattcga taactgtctg atggtgcacg accacgcatt 4980

aatggactgg attggggcca actcctaccg tacctcgcgt tacccttacg ctgaagagat 5040

gctcgactgg gcagatgaac atggcatcgt ggtgattgat gaaactgctg ctgtcggctt 5100

taacctctct ttaggcattg gtttcgaagc gggcaacaag ccgaaagaac tgtacagcga 5160

agaggcagtc aacggggaaa ctacgaagc gcacttacag gcgattaaag agctgatagc 5220

gcgtgacaaa aaccacccaa gcgtggtgat gtggagtatt gccaacgaac cggatacccg 5280

tccgcaaggt gcacgggaat atttcgcgcc actggcgga gcaacgcgta aactcgaccc 5340

gacgcgtccg atcacctgcg tcaatgtaat gttctgcgac gtcacacccg ataccatcag 5400

cgatctcttt gatgtgctgt gcctgaaccg ttattacgga tggatgtcc aaagcggcga 5460

tttggaacg gcagagaagg tactggaaaa agaacttctg gcctggcagg agaaactgca 5520

tcagccgatt atcatcaccg aatacggcgt ggatacgtta gccgggctgc actcaatgta 5580

caccgacatg tggagtgaag agtatcagtg tgcattggctg gatattgtatc accgcgtctt 5640
 tgatcgctgc agcgccgctg tccgtgaaca ggtatggaat ttcgccgatt ttgcgacctc 5700
 gcaaggcata ttgcgcgttg gcggttaacaa gaaagggatc ttactcgcg accgcaaacc 5760
 gaagtcggcg gcttttctgc tgcaaaaacg ctggactggc atgaacttcg gtgaaaaacc 5820

gcagcaggga ggcaacaat gattaattaa ctagagcggc cgcgttaatt caaattaatt 5880
 gatatagttt ttaaatgagt attgaatctg tttagaaata atggaatatt atttttattt 5940
 atttatttat attattggtc ggctcttttc ttctgaaggt caatgacaaa atgatatgaa 6000
 ggaaataatg atttctaaaa ttctacaacg taagatattt ttacaaaagc ctagctcatc 6060
 ttttgcctg cactatttta ctacgccttg aaattaacgg ccagtcact gcggagtcac 6120
 ttcaaagtca tctaatcga tctatcgtt ttgatagctc attttggagt tcgcgattgt 6180
 cttctgttat tcacaactgt ttttaattttt atttcattct ggaactcttc gatttctttg 6240

taaagtcctt catagtagct tactttatcc tccaacatat ttaacttcat gtcaatttcg 6300
 gctcttaaat tttccacatc atcaagtcca acatcatctt ttaacttgaa tttattctct 6360
 agctcttcca accaagctc attgtctctt gatttactgg tgaaaagtga tacactttgc 6420
 gcgcaatcca ggtcaaaact ttctgcaaa gaattcacca atttctgac atcatagtac 6480
 aatttgcttt gttctccat cacaatttaa tatacctgat ggattcttat gaagcgttg 6540
 gtaatggacg tgcactcta ctctgccttt ttccctactc cttttagtagc ggaagacaat 6600
 gctaataaat aagagggtaa taataatatt attaatcggc aaaaaagatt aaacccaag 6660

cgtttaatta tcagaaagca aacgtcgtac caatcctga atgtctcca attgtatatt 6720
 aagagtcac acagcaacat attcttgtta ttaaatatatt tattattgat ttttgatatt 6780
 gtataaaaa accaatatg tataaaaaaa gtgaataaaa aataccaagt atggagaaat 6840
 atattagaag tctatacgtt aaaccaccgc ggtggagctc cagcttttgt tccctttagt 6900
 gagggtaat tgcgcgcttg gcgtaatcat ggtcatagct gtttcctgtg tgaaattgtt 6960
 atccgctcac aattccacac aacataggag ccggaagcat aaagtgtaaa gcctggggtg 7020
 cctaattgagt gaggttaactc acattaattg cgttgcgctc actgcccgtt ttccagtcgg 7080

gaaacctgtc gtgccagctg cattaatgaa tcggccaacg cgcggggaga ggcggtttgc 7140
 gtattggcg ctcttcgct tctcgtcga ctgactcgt gcgctcggtc gttcggctgc 7200
 ggcgagcgt atcagctcac tcaaaggcgg taatacgtt atccacagaa tcaggggata 7260
 acgcaggaaa gaacatgtga gcaaaaggcc agcaaaaggc caggaaccgt aaaaaggccg 7320
 cgttgcctggc gtttttccat aggtctcgcc cccctgacga gcatacaciaa aatcgacgt 7380
 caagtcagag gtggcgaaac ccgacaggac tataaagata ccaggcgttt cccctggaa 7440

gctccctcgt gcgctctcct gttccgaccc tgccgcttac cggataacctg tccgcctttc	7500
tcccttcggg aagcgtggcg ctttctcata gctcacgctg taggtatctc agttcgggtg	7560
aggtcgttcg ctccaagctg ggctgtgtgc acgaaccccc cgttcagccc gaccgctgcg	7620
ccttatccgg taactatcgt cttgagtcca acccggttaag acacgactta tcgccactgg	7680
cagcagccac tggtaacagg attagcagag cgaggatatgt aggcggtgct acagagtctt	7740
tgaagtgggt gcctaactac ggctacacta gaaggacagt atttggatc tgcgctctgc	7800
tgaagccagt taccttcgga aaaagagttg gtatgtcttg atccggcaaa caaaccaccg	7860
ctggtagcgg tggttttttt gtttgcaagc agcagattac gcgcagaaaa aaaggatctc	7920
aagaagatcc ttigatcttt tctacggggt ctgacgtca gtggaacgaa aactcacgtt	7980
aagggatctt ggtcatgaga ttatcaaaaa ggatcttcac ctagatcctt ttaaattaaa	8040
aatgaagttt taaatcaatc taaagtatat atgagtaaac ttggtctgac agttaccaat	8100
gcttaatcag tgaggcacct atctcagcga tctgtctatt tcgttcaccc atagtgcct	8160
gactccccgt cgtgtagata actacgatac gggagggtct accatctggc cccagtctg	8220
caatgatacc gcgagaccca cgctcacccg ctccagattt atcagcaata aaccagccag	8280
ccggaagggc cgagcgcaga agtggtcctg caactttatc cgcctccatc cagtctatta	8340
attgttgccg ggaagctaga gtaagtagtt cgccagttaa tagtttgccg aacgttggtg	8400
ccattgctac aggcacgtg gtgtcacgct cgtcgtttgg tatggcttca ttcagctccg	8460
gttccaacg atcaaggcga gttacatgat ccccatgtt gtgcaaaaa gcggttagct	8520
ccttcggtec tccgatcgtt gtcagaagta agttggccgc agtgttatca ctcatggtta	8580
tggcagcact gcataattct cttactgtca tgccatccgt aagatgcttt tctgtgactg	8640
gtgagtactc aaccaagtca ttctgagaat agtgtatgcg gcgaccgagt tgctcttgcc	8700
cggcgtcaat acgggataat accgcgccac atagcagaac tttaaaagtg ctcatcattg	8760
gaaaacgttc ttcggggcga aaactctcaa ggatcttacc gctgttgaga tccagttcga	8820
tgtaaccac tcgtgcacc aactgatctt cagcatcttt tactttcacc agcgtttctg	8880
ggtgagcaaa aacaggaagg caaaatgccg caaaaaaggg aataaggcg acacggaaat	8940
gttgaatact catactcttc ctttttcaat attattgaag catttatcag ggttattgtc	9000
tcatgagcgg atacatatgt gaatgtatgt agaaaaataa acaaataggg gttccgcgca	9060
catttccccg aaaagtgcga cctgaacgaa gcactctgtc ttcattttgt agaacaaaaa	9120
tgcaacgcga gagcgctaata ttttcaaaca aagaatctga gctgcatttt tacagaacag	9180

aaatgcaacg cgaaagcgct attttaccac cgaagaatct gtgcttcatt ttgtgaaaac 9240
 aaaaatgcaa cgcgagagcg ctaatttttc aaacaaagaa tctgagctgc atttttacag 9300
 aacagaaatg caacgcgaga gcgctatfff accaacaag aatctatact tcttttttgt 9360
 tctacaaaaa tgcatcccga gagcgctatt tttctaaca agcatcttag attacttttt 9420
 ttctcctttg tgcgctctat aatgcagtct ctfgataact ttttgcactg taggtccgtt 9480
 aaggtagaa gaaggctact ttgggtgcta ttttctcttc cataaaaaaa gcctgactcc 9540
 acttcccgcg ttiactgatt actagcgaag ctgcgggtgc attttttcaa gataaaggca 9600

tccccgatta tattctatac cgatgtggat tgcgcatact ttgtgaacag aaagtgatag 9660
 cgttgatgat tcttcattgg tcagaaaatt atgaacgggt tcttctatft tgtctctata 9720
 tactacgtat aggaaatgtt tacattttcg tattgttttc gattcactct atgaatagtt 9780
 ctactacaa tttttttgtc taaagagtaa tactagagat aaacataaaa aatgtagagg 9840
 tcgagtttag atgcaagtgc aaggagcgaa aggtggatgg gtaggttata tagggatata 9900
 gcacagagat atatagcaaa gagatacttt tgagcaatgt ttgtggaagc ggtattcgca 9960
 atattttagt agctcgttac agtccgggtgc gtttttggtt ttttgaaagt gcgtcttcag 10020

agcgcttttg gttttcaaaa gcgctctgaa gttcctatac tttctagaga ataggaactt 10080
 cggaatagga acttcaaagc gtttccgaaa acgagcgctt ccgaaaatgc aacgcgagct 10140
 gcgcacatac agctcactgt tcacgtcgca cctatatctg cgtgttcct gtatatatat 10200
 atacatgaga agaacggcat agtgcgtgtt tatgcttaaa tgcgtactta tatgcgtcta 10260
 tttatgtagg atgaaaggta gtctagtacc tcctgtgata ttatcccatt ccatgcgggg 10320
 tatcgtatgc ttccittcage actacccttt agctgttcta tatgttgcca ctccctaatt 10380
 ggattagtct catccttcaa tgctatcatt tcctttgata ttggatcatc taagaaacca 10440

ttattatcat gacattaacc tataaaaaata ggcgatatcac gaggcccttt cgtc 10494

<210> 770

<211> 43

<212> DNA

<213> Artificial Sequence

<220><223> T-U/PGK1(KpnI) primer

<400> 770

actacagatg gtaccaatta ccgtcgtctg tgatttgttt gca 43

<210> 771

<211> 45

<212> DNA

<213> Artificial Sequence

<220><223> B-U/PGK1(SalI) primer

<400> 771

agcatccttg tcgacaggac cttgttgtgt gacgaaattg gaagc 45

<210> 772

<211> 7533

<212> DNA

<213> Artificial Sequence

<220><223> pYZ067DkivDDilvD plasmid

<400> 772

tcgcgcgttt cggatgatgac ggtgaaaacc tctgacacat gcagctcccg gagacgggtca 60

cagcttgtct gtaagcggat gccgggagca gacaagcccg tcagggcgcg tcagcgggtg 120

ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180

accataaatt cccgttttaa gagcttggtg agcgcctagga gtcactgcca ggtatcgttt 240

gaacacggca ttagtcaggg aagtcataac acagtccttt cccgcaattt tctttttcta 300

ttactcttgg cctcctctag tacactctat atttttttat gcctcggtaa tgattttcat 360

tttttttttt ccacctagcg gatgactctt tttttttctt agcgattggc attatcacat 420

aatgaattat acattatata aagtaatgtg atttcttcga agaataact aaaaaatgag 480

caggcaagat aaacgaaggc aaagatgaca gagcagaaag ccctagtaaa gcgtattaca 540

aatgaaacca agattcagat tgcgatctct ttaaagggtg gtcccctagc gatagagcac 600

tcgatcttcc cagaaaaaga ggcagaagca gtagcagaac aggccacaca atcgcaagtg 660

attaacgtcc acacaggat agggtttctg gaccatatga tacatgctct ggccaagcat 720

tccggctggt cgctaactgt tgagtgcatt ggtgacttac acatagacga ccatcacacc 780

actgaagact gcgggattgc tctcggtaaa gctttttaaag aggccttagg ggccgtgcgt 840

ggagtaaaaa ggtttggatc aggatttgcg cctttggatg aggcactttc cagagcgggtg 900

gtagatcttt cgaacaggcc gtacgcagtt gtcgaacttg gtttgcaaag ggagaaagta 960

ggagatctct cttgcgagat gatcccgcat tttcttgaaa gctttgcaga ggctagcaga 1020

attaccctcc acgttgattg tctgcgaggc aagaatgac atcaccgtag tgagagtgcg 1080

ttcaaggctc ttgcggttgc cataagagaa gccacctcgc ccaatggtac caacgatgtt 1140

ccctccacca aaggtgttct tatgtagtga caccgattat ttaaagctgc agcatacgt 1200

atatacatat gtgtatatat gtatacctat gaatgtcagt aagtatgtat acgaacagta 1260

tgatactgaa gatgacaagg taatgcatca ttctatacgt gtcattctga acgaggcgcg	1320
ctttcctttt ttctttttgc tttttctttt tttttctctt gaactcgacg gatctatgcg	1380
gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcaggaaat tgtaagcggt	1440
aatatattgt taaaattcgc gttaaatttt tgttaaata gtcatttttt taaccaatag	1500
gccgaaatcg gcaaaatccc ttataaatca aaagaataga ccgagatagg gttgagtgtt	1560
gttccagttt ggaacaagag tccactatta aagaacgtgg actccaacgt caaagggcga	1620
aaaaccgtct atcagggcga tggcccaacta cgtggccggg atccactagt tctagagcgg	1680
ccgctctaga actagtacca cagggtgtgt cctctgagga cataaaatac acaccgagat	1740
tcatcaactc attgctggag ttagcatatc tacaattggg tgaaatgggg agcgatttgc	1800
aggcatttgc tcggcatgcc ggtagagggt tggtaataa gagcgacctc atgctatacc	1860
tgagaaagca acctgacctc caggaaagag ttactcaaga ataagaattt tcgttttaaa	1920
acctaagagt cactttaaaa ttgtatata cttatttttt ttataactta ttaataata	1980
aaaatcataa atcataagaa attcgcttac tcttaattaa tcaaaaagtt aaaattgtac	2040
gaatagattc accacttctt aacaaatcaa acccttcatt gattttctcg aatggcaata	2100
catgtgtaat taaaggatca agagcaaact tcttcgcat aaagtcggca acaagttttg	2160
gaacactatc cttgctctta aaaccgcaa atatagctcc cttccatgta cgaccgctta	2220
gcaacagcat aggattcatc gacaaatttt gtgaatcagg aggaacacct acgatcacac	2280
tgactccata tgcctcttga cagcaggaca acgcagttac catagtatca agacggccta	2340
taacttcaaa agagaaatca actccaccgt ttgacatttc agtaaggact tcttgtattg	2400
gtttcttata atcttgaggg ttaacacatt cagtagcccc gacctcctta gcttttgcaa	2460
atttgctctt attgatgtct acacctataa tcctcgctgc gcctgcagct ttacacccca	2520
taataacgct tagtctact cctcctaaac cgaatactgc acaagtcgaa ccctgtgtaa	2580
cctttgcaac tttaactcgc gaaccgtaac cggtggaata tccgcaccct atcaagcaaa	2640
ctttttccag tggatgaagct gcatcgattt tagcgacaga tatctcgtcc accactgtgt	2700
attgggaaaa tglagaagta ccaaggaaat ggtgtatagg tttccctctg catgtaaatc	2760
tgcttgatcc atcctgcata gtacctctag gcatagacaa atcattttta aggcagaaat	2820
tacctcagg atgtttgcag actctacact taccacattg aggagtgaac agtgggatca	2880
ctttatcacc aggacgaaca gtggtaaac cttcacctat ggattcaacg attccggcag	2940
cctcgtgtcc cgcgattact ggcaaaggag taactagagt gccactcacc acatggtcgt	3000
cggatctaca gattccggtg gcaaccatct tgattctaac ctctgtgtgt tttgggtggcg	3060
ctacttctac ttcttctatg ctaaaccggt ttttctcttc ccacaaaact gccgctttac	3120

acttaataac ttiaccggct gttgacatcc tcagctagct attgtaatat gtgtgtttgt	3180
ttggattatt aagaagaata attacaaaa aaattacaaa ggaaggtaat tacaacagaa	3240
ttaagaaagg acaagaagga ggaagagaat cagttcatta tttcttcttt gttatataac	3300
aaaccaagt agcgatttgg ccatacatta aaagttaga accaccctcc ctggcaacag	3360
ccacaactcg ttaccattgt tcatcacgat catgaaactc gctgtcagct gaaatttcac	3420
ctcagtggat ctctcttttt attcttcac gttccactaa cctttttcca tcagctggca	3480
gggaacggaa agtggaatcc catttagcga gcttcctctt tcttcaaga aaagacgaag	3540
cttgtgtgtg ggtgcgcgcg ctagtatctt tccacattaa gaaatatacc ataaaggtta	3600
cttagacatc actatggcta tataatatata tatatatata tgtaacttag caccatcgcg	3660
cgtgcatcac tgcatgtgtt aaccgaaaag ttggcgaac acttcaccga cacggtcatt	3720
tagatctgtc gtctgcattg cacgtccctt agccttaaat cctaggcggg agcattctcg	3780
tgtaattgtg cagcctgcgt agcaactcaa catagcgtag tctaccaggt ttttcaaggg	3840
tttatcgta gaagattctc ccttttcttc ctgctcacia atcttaaagt catacattgc	3900
acgactaaat gcaagcgcca gcttttgttc cctttagtga gggtaattg cgcgcttggc	3960
glaatcatgg tcatagctgt ttctgtgtg aaattgttat ccgctcacia ttccacacia	4020
catacgagcc ggaagcataa agtgtaaagc ctggggtgcc taatgagtga gctaactcac	4080
attaattgcg ttgcgctcac tgcccgtttt ccagtcggga aacctgtcgt gccagctgca	4140
ttaatgaatc ggccaacgcg cggggagagg cggtttgcgt attgggcgct ctccgcttc	4200
ctcgtcact gactcgtgc gctcggctgt tcggctcggg cgagcggat cagctcactc	4260
aaaggcggta atacggttat ccacagaatc aggggataac gcaggaaaga acatgtgagc	4320
aaaaggccag caaaaggcca ggaaccgtaa aaaggccgcg ttgctggcgt ttttccatag	4380
gtccgcccc cctgacgagc atcacaaaaa tcgacgtca agtcagaggt ggcgaaaccc	4440
gacaggacta taaagatacc aggcgttttc ccttgaagc tccctcgtgc gctctcctgt	4500
tccgaccctg ccgcttaccg gatactgtc cgcctttctc ccttcgggaa gcgtggcgt	4560
ttctcatagc tcacgtgta ggtatctcag ttcggtgtag gtcgttcgt ccaagctggg	4620
ctgtgtgcac gaacccccg ttcagccga ccgctgcgc ttatccgta actatcgtct	4680
tgagtccaac ccgtaagac acgacttct gccactggca gcagccactg gtaacaggat	4740
tagcagagcg aggtatgtag gcggtgctac agagtctctg aagtgggtgc ctaactacgg	4800
ctacactaga agaacagtat ttggtatctg cgctctgctg aagccagtta ccttcggaaa	4860
aagagtgggt agctcttgat ccggcaaaaca aaccaccgct ggtagcgggtg gttttttgt	4920

ttgcaagcag cagattacgc gcagaaaaaa aggatctcaa gaagatcctt tgatcttttc	4980
tacggggctct gacgctcagt ggaacgaaaa ctacagttaa gggattttgg tcatgagatt	5040
atcaaaaagg atcttcacct agatcctttt aaattaaaaa tgaagtttta aatcaatcta	5100
aagtatataat gagtaaacctt ggtctgacag ttaccaatgc ttaatcagtg aggcacctat	5160
ctcagcgatc tgtctatttc gtctcatccat agttgcctga ctccccgtcg ttagataaac	5220
tacgatacgg gagggtttac catctggccc cagtgtctga atgataccgc gagaccacg	5280
ctcaccggct ccagatttat cagcaataaa ccagccagcc ggaagggccg agcgcagaag	5340
tggctctgca actttatccg cctccatcca gtctattaat tgttgccggg aagctagagt	5400
aagtagttcg ccagttaata gtttgcgcaa cgttgttgcc attgtctacag gcatcgtggt	5460
gtcacgctcg tcgtttggta tggcttcatt cagctccggt tccaacgat caaggcgagt	5520
tacatgatcc cccatgttgt gcaaaaaagc ggtagctcc ttcggctctc cgatcgttgt	5580
cagaagtaag ttggccgcag tgttatcact catggttatg gcagcactgc ataattctct	5640
tactgtcatg ccatccgtaa gatgttttc tgtgactggt gactactcaa ccaagtcatt	5700
ctgagaatag tgtatgcggc gaccgagttg ctcttgcccc gcgtcaatac gggataatac	5760
cgcgccacat agcagaactt taaaagtgtc catcattgga aaacgttctt cggggcgaaa	5820
actctcaagg atcttacgcg tgttgagatc cagttcgtatg taaccactc gtgcacccaa	5880
ctgatcttca gcatctttta ctttcaccag cgtttctggg tgagcaaaaa caggaaggca	5940
aaatgccgca aaaaagggaa taaggcgac acggaaatgt tgaatactca tactcttct	6000
ttttcaatat tattgaagca tttatcaggg ttattgtctc atgagcggat acatatttga	6060
atgtatttag aaaaataaac aaataggggt tccgcgcaca tttccccgaa aagtgccacc	6120
tgaacgaagc atctgtgctt cttttgttag aacaaaaatg caacgcgaga gcgctaattt	6180
ttcaaacaaa gaatctgagc tgcattttta cagaacagaa atgcaacgcg aaagcgtat	6240
tttaccaacg aagaatctgt gcttcatttt tgtaaaacaa aaatgcaacg cgagagcgct	6300
aatttttcaa acaagaatc tgagctgcat ttttacagaa cagaaatgca acgcgagagc	6360
gctattttac caacaaagaa tctatacttc tttttgttc tacaaaaatg catcccgaga	6420
gcgctatttt tctaacaag catcttagat tacttttttt ctctttgtg cgctctataa	6480
tgcagtctct tgataacttt ttgcactgta ggtccgttaa ggtagaaga aggctacttt	6540
ggtgtctatt ttctcttcca taaaaaagc ctgactccac ttcccgcgtt tactgattac	6600
tagcgaagct gcgggtgcat tttttcaaga taaaggcatc cccgattata ttctataccg	6660

atgtggattg cgcatacttt gtgaacagaa agtgatagcg ttgatgattc ttcattggtc 6720
 aaaaaattat gaacggtttc ttctattttg tctctatata ctacgtatag gaaatgttta 6780
 ctttttcgta ttgttttcga ttcactctat gaatagtctt tactacaatt tttttgtcta 6840
 aagagtaata cttagagataa acataaaaaa tgtagaggtc gagtttagat gcaagttcaa 6900
 ggagcgaaag gtggatgggt aggttatata gggatatagc acagagatat atagcaaaga 6960
 gatacttttg agcaatgttt gtggaagcgg taticgcaat attttagtag ctcgttacag 7020
 tccggtgcgt ttttggtttt ttgaaagtgc gtcttcagag cgcttttggt tttcaaaagc 7080

gctctgaagt tcctatactt tctagagaat aggaacttcg gaataggaac ttcaaagcgt 7140
 ttccgaaaac gagcgcttcc gaaaatgcaa cgcgagctgc gcacatacag ctcaactgttc 7200
 acgtcgcacc tatatctgcg tgttgccctgt atatatatat acatgagaag aacggcatag 7260
 tgcgtgttta tgcttaaatg cgtacttata tgcgtctatt tatgtaggat gaaaggtagt 7320
 ctagtacctc ctgtgatatt atcccattcc atcggggta tcgtatgctt ccttcagcac 7380
 taccctttag ctgttctata tgctgccact cctcaattgg attagtctca tccttcaatg 7440
 ctatcatttc ctttgatatt ggatcatact aagaaacat tattatcatg acattaacct 7500

ataaaaatag gcgtatcacg aggccctttc gtc 7533

<210> 773

<211> 12518

<212> DNA

<213> Artificial Sequence

<220><223> pK9G9.0LE1p.ilvD plasmid

<400> 773

tccattacc gacatttggg cgctatacgt gcatatgttc atgtatgtat ctgtatttaa 60
 aacacttttg tattattttt cctcatatat gtgtataggt ttatacggat gatttaatta 120
 ttacttcacc accctttatt tcaggctgat atcttagcct tgttactaga ttaatcatgt 180
 aattagttaa gtcacgctta cattcacgcc ctccccccac atccgctcta accgaaaagg 240
 aaggagttag acaacctgaa gtctaggtcc ctatttattt tttttagtt atgttagtat 300

taagaacgtt atttatattt caaatttttc ttttttttct gtacagacgc gtgtacgcat 360
 gtaacattat actgaaaacc ttgcttgaga aggttttggg acgctcgaag gctttaattt 420
 gcgggcgccc gcacctggtg aaacctctag tggagtagta gatglaatca atgaagcggg 480
 agccaaaaga ccagagtaga ggcctataga agaaactgcg ataccttttg tgatggctaa 540
 acaaacagac atctttttat atgtttttac ttctgtatat cgtgaagtag taagtataa 600
 gcgaatttgg ctaagaacgt tgtaagttaa caaggacct cttttgcctt tcaaaaaagg 660

attaaatgga gttaatcatt gagatttagt tttcgttaga tttgtatcc ctaaataact	720
cccttaccg acgggaaggc acaaaagact tgaataatag caaacggcca gtagccaaga	780
ccaaataata ctagagttaa ctgatggctt taaacaggca ttacgtgggtg aactccaaga	840
ccaatataca aaatatcgat aagttattct tgcccaccaa ttttaaggagc ctacatcagg	900
acagtagtac cattcctcag agaagaggta tacataacaa gaaaatcgcg tgaacacctt	960
atataactta gcccgttatt gagctaaaaa accttgcaaa atttcctatg aataagaata	1020
cttcagacgt gataaaaatt tactttctaa ctcttctcac gctgcccta tctgttcttc	1080
cgctctaccg tgagaaataa agcatcgagt acggcagttc gctgtcactg aactaaaaca	1140
ataaggctag ttcgaatgat gaacttgctt gctgtcaaac tttcagttg ccgctgatgt	1200
gacactgtga caataaatc aaaccggtta tagcggcttc ctccgtacc ggttctgcca	1260
cctccaatag agctcagtag gagtcagaac ctctcggtg gctgtcagtg actcatccgc	1320
gtttcgtaag ttgtgcgct gcacatttcg cccgttcccg ctcatcttg agcaggcgga	1380
aattttcatc acgctgtagg acgcaaaaaa aaaataatta atcgtacaag aatcttgaa	1440
aaaaaatga aaaaattgt ataaaaggga tgacctaaact tgactcaatg gcttttacac	1500
ccagtatttt ccttttctt gtttgttaca attatagaag caagacaaaa acatatagac	1560
aacctattcc taggagttat atttttttac cctaccagca atataagtaa aaaactgttt	1620
aaacagtatg gaagaatgta agatggctaa gatttactac caagaagact gtaactgtc	1680
cttgttggat ggtaagacta tcgccgttat cggttacggt tctcaaggtc acgctcatgc	1740
cctgaatgct aaggaatccg gttgtaacgt tatcattggt ttatacgaag gtgcgaagga	1800
gtggaaaaga gctgaagaac aaggtttcga agtctacacc gctgctgaag ctgctaagaa	1860
ggctgacatc attatgatct tgatcaacga tgaagaagcag gctaccatgt aaaaaacga	1920
catcgaacca aacttggaag cggtaacat gttgatgttc gctcacggtt tcaacatcca	1980
tttcggttgt attgttcac caaaggacgt tgatgtcact atgatcgctc caaagggtcc	2040
aggtcacacc gttagatccg aatacgaaga aggtaaagggt gtcccatgct tggttgctgt	2100
cgaacaagac gctactggca aggctttgga tatggctttg gcctacgctt tagccatcgg	2160
tgggtctaga gccggtgtct tggaaactac cttcagaacc gaaactgaaa ccgacttgtt	2220
cggtaacaa gctgttttat ttgggtgggt ctgcgtttg atgcaggccg gttttgaaac	2280
cttgggtgaa gccggttacg acccaagaaa cgcttacttc gaatgtatcc acgaaatgaa	2340
gttgatcggt gacttgatct accaatctgg tttctccggt atgcgttact ctatctccaa	2400

cactgctgaa tacggtgact acattaccgg tccaaagatc attactgaag ataccaagaa	2460
ggctatgaag aagattttgt ctgacattca agatggtacc tttgccaagg acttcttgg	2520
tgacatgtct gatgctggtt cccaggtcca cttcaaggct atgagaaaagt tggcctccga	2580
acaccagct gaagtgtcg gtgaagaaat tagatccttg tactcctggt cgcacgaaga	2640
caagttgatt aacaactgag gccctgcagg ccagaggaaa ataatatcaa gtgctggaaa	2700
ctttttctct tggaattttt gcaacatcaa gtcatagtca attgaattga cccaatttca	2760
catttaagat tttttttttt tcatccgaca tacatctgta cactaggaag ccctgttttt	2820
ctgaagcagc ttcaaatata tatatttttt acatatttat tatgattcaa tgaacaatct	2880
aattaaatcg aaaacaagaa ccgaaacgag aataaataat ttatttagat ggtgacaagt	2940
glataagtcc tcatcgggac agctacgatt tctctttcgg ttttggctga gctactggtt	3000
gctgtgacgc agcggcatta gcgcggcggt atgagctacc ctcgtggcct gaaagatggc	3060
gggaataaag cggaactaaa aattactgac tgagccatat tgaggtaaat ttgtcaactc	3120
gtcaagtcac gtttgggtga cggccctttt ccaacgaatc gtatatacta acatgcgcgc	3180
gcttcctata tacacatata catatatata tatatatata tgtgtgcgtg tatgtgtaca	3240
ctgtatttta atttccttac tcgcgggttt ttcttttttc tcaattcttg gcttcctctt	3300
tctcgagcgg accggatcct cgcgaaactcc aaaatgagct atcaaaaacg atagatcgat	3360
taggatgact ttgaaatgac tccgcagtgg actggccgtt aatttcaagc gtgagtaaaa	3420
tagtgcatga caaaagatga gctaggcttt tgtaaaaata tcttacgttg taaaatttta	3480
gaaatcatta ttctcttcat atcattttgt cattgacctt cagaagaaaa gagccgacca	3540
ataatataaa taaataaata aaaataatat tccattatit ctaaacagat tcaatactca	3600
ttaaaaaact atatcaatta atttgaatta acgcggccgc ttaaccacag caaccaggac	3660
aacatttttt gccagtttct tcaggcttcc aaaagtctgt tacggctccc ctagaagcag	3720
acgaaacgat gtgagcatat ttaccaagga taccgctga atagagcggg ggcaattcaa	3780
tggctctctg acgatgtttt aactcttcat cggagatata aaagtgtaat tccttagtgt	3840
cttggtaaat agtgactatg tctctgtttt gcaggtaggc gattggaccg ccatcttgtg	3900
cttcaggagc gatatgaccc acgacaagac cataagtacc acctgagaag cggccatctg	3960
tcagaagggc aactttttca ccttgccctt taccaacaat cattgatgaa agggaaagca	4020
tttcaggcat accaggaccg ccctttgggtc ctacaaaacg tacgacaaca acatcaccat	4080
caacaatata atcattcaag acagcttcaa tggtctcttc ttcagaatta aagaccttag	4140
caggaccgac atgacgacgc actttttacac cagaaacttt ggcaacggca cgtctggag	4200
ccaagttacc atggagaata atgagcggac catcttcacg tttaggattt tcaagcggca	4260

taataacctt ttgaccaggt gttaaatacat caaaagcctt caaatTTTca gcgactgttt	4320
tgccagtaca agtgatacgg tcaccatgaa ggaagccatt tTTaaggaga tatttcataa	4380
ctgctggtag ccttccgacc ttgtaaaggt cttggaatac atattgacca gaaggtttca	4440
aatcagccaa atgaggaact ttttcttTga aagtattgaa atcatcaagt gtcaattcca	4500
cattagcagc atgggcaata gctaagaggt gaagggttga gttggttgaa cctcccagag	4560
ccatagttag agtaatagca tcttcaaaag cttcacgcgt taaaatgtca gaaggtttta	4620
agcccatTTc gagcatTTtg acaacagcgc gaccagcttc ttcaatatct gctttctttt	4680
ctgcggattc agccgggtga gaagatgaac ccggaaggct aagtcccaaa acttcaatag	4740
ctgtcgccat tTgttttagca gtatacatac caccgcagcc tccaggaccg ggacaagcat	4800
tacattccaa agctTTaact tcttctttgg tcatatcgcc gtggttccaa tggccgacac	4860
cttcaaagac agagactaaa tcgatatctt tgccgtctaa attaccaggt gcaattgttc	4920
cgccgtaagc aaaaatggct gggatatcca tTtagccat agcgataaca gaaccgggca	4980
tgtttttatc acaaccgcca atggctacaa aagcatccgc attatgacct cccatggctg	5040
cttcaataga atctgcaata atatcacgag atgtcaagga gaaacgcatt ccttgggttc	5100
ccatggcgat tccatcagaa accgtgattg ttccgaactg aactggccaa gcaccagctt	5160
ccttaacacc gactttggct agtttaccaa agtcatgtaa gtggatatta caaggtgtgt	5220
tttcagccca agttgaaatg acaccgacga taggtttttc aaagcttTca tcttgcatac	5280
cagttgcacg caacatagca cgattaggTg atttaacat tgaatcgtaa acagaactac	5340
gatttcttaa gtctttaaga gttttttTgt cagtcatact cacgtgcttt gtTgtaatgt	5400
tttagtgctg ttTataatat gatcaccaca actatctatt actatgatgt tctattctac	5460
gtaatacaaa atataaacgg aaacagaagt aggaaagatg gaaatagaac aataaatgaa	5520
tcaagatctg ccccatata tatatgtata tTctgatttg caagactcga tgagccagga	5580
gccgatgatt tTctgcatat attgtTaaact actattatTT ccacctttgt gtgcatccc	5640
catagccgta acaataggga taggtgtgtc tgagttagca agactcgtag aagcacacct	5700
ggttgggcac tagataaggt ttgttgagtTg ttcaacgtcc gaaagaaagc tgccgactat	5760
gcgaagagaa ccttaagccg ttattacctc tgctgtcac aggcgatgtg atgctaacga	5820
acagcaccag agccaagcca actggggcgTg tctgcagaga aggtgggat acccgaaata	5880
gctcgctcaa cagctTTTTt tcttctacgg aagcccacca gataagcgcc tttgttgggc	5940
ccgctaacc cgggacatgc ccgggctcgg agttagtttt tgcacggccg gcagatctat	6000
ttaaatggcg cgccgacgtc aggtggcaact tttcggggaa atgtgcgcgg aaccttatt	6060
tgtttatttt tctaaataca ttcaaatatg tatccgtca tgagacaata accctgataa	6120

atgcttcaat aatattgaaa aaggaagagt atgagtattc aacatttccg tgcgcctt 6180

attccctttt ttgcggcatt ttgccttctt gtttttctc acccagaaac gctggtgaaa 6240

gtaaaagatg ctgaagatca gttgggtgca cgagtgggtt acatcgaact ggatctcaac 6300

agcggtaaga tccttgagag ttttcgcccc gaagaacgtt ttccaatgat gagcactttt 6360

aaagtcttgc tatgtggcgc ggtattatcc cgtattgacg cgggcaaga gcaactcgg 6420

cgccgcatc actatttca gaatgacttg gttgagtact caccagtcac agaaaagcat 6480

cttacggatg gcatgacagt aagagaatta tgcagtgtg ccataacatc gtagtataac 6540

actgcggcca acttacttct gacaacgatc ggaggaccga aggagctaac cgctttttt 6600

cacaacatgg gggatcatgt aactcgctt gatcgttggg aaccggagct gaatgaagcc 6660

ataccaaacg acgagcgtga caccacgatg cctgtagcaa tggcaacaac gttgcgcaaa 6720

ctattaactg gcgaactact tactctagct tcccggcaac aattaataga ctggatggag 6780

gcggataaag ttgcaggacc acttctgcgc tcggcccttc cggctggctg gtttattgct 6840

gataaatctg gagccgtga gcgtgggtct cgcggtatca ttgcagcact ggggccagat 6900

ggtaagccct cccgtatcgt agttatctac acgacgggga gtcaggcaac tatggatgaa 6960

cgaaatagac agatcgctga gataggtgcc tcaactgatta agcattggta actgtcagac 7020

caagtttact catatatact ttagattgat ttaaaacttc atttttaatt taaaaggatc 7080

taggtgaaga tcctttttga taatctcatg accaaaatcc cttaacgtga gttttcgctt 7140

cactgagcgt cagacccctg agaaaagatc aaaggatctt cttgagatcc ttttttctg 7200

cgcgtaatct gctgcttga aacaaaaaaa ccaccgtac cagcgggtgt ttgtttgccg 7260

gatcaagagc taccaactct tttccgaag gtaactggct tcagcagagc gcagatacca 7320

aatactgttc ttctagtga gccgtagtta ggccaccact tcaagaactc ttagcaccg 7380

cctacatacc tcgtctgct aatcctgtta ccagtggctg ctgccagtgg cgataagtcg 7440

tgtcttaccg ggttggactc aagacgatag ttaccggata aggcgcagcg gtcgggctga 7500

acgggggggtt cgtgcacaca gccagcttg gagcgaacga cctacaccga actgagatac 7560

ctacagcgtg agctatgaga aagcggcacg cttcccgaag ggagaaaggc ggacaggtat 7620

ccggtaaagc gcagggtcgg aacaggagag cgcacagagg agcttccagg gggaaacgcc 7680

tggtatcttt atagtctgt cgggtttcgc cacctctgac ttgagcgtcg atttttgtga 7740

tgctcgtcag gggggcggag cctatgaaa aacgccagca acgcggcctt ttacggttc 7800

ctggcctttt gctggccttt tgctcacatg ttctttctg cgttatcccc tgattctgtg 7860

gataaccgta ttaccgcctt tgagttagct gataccgctc gccgcagccg aacgaccgag	7920
cgcagcgagt cagttagcga ggaagcggaa gagcgcccaa tacgcaaacc gcctctcccc	7980
gcgcgttggc cgattcatta atgcagctgg cagcagaggt ttcccactg gaaagcgggc	8040
agttagcgca acgcaattaa tgtaggttag ctactcatt aggcaccca ggctttacac	8100
tttatgcttc cggtcgtat gttgtgtgga attgtgagcg gataacaatt tcacacagga	8160
aacagctatg accatgatta cgccaagctt tttctttcca atttttttt tttcgtcatt	8220
ataaaaatca ttacgaccga gattccccggg taataactga tataattaaa ttgaagctct	8280
aatttgtgag tttagtatac atgcatttac ttataatata gttttttagt tttgctggcc	8340
gcatcttctc aaatatgctt cccagcctgc ttttctgtaa cgttcacctt ctaccttagc	8400
atcccttccc ttgcaaata gtcctcttcc aacaataata atgtcagatc ctgtagagac	8460
cacatcatcc acggttctat actgttgacc caatgcgtct ccttgtcat ctaaaccac	8520
accgggtgtc ataatacaacc aatcgtaacc ttcattcttt ccacccatgt ctcttgagc	8580
aataaagccg ataacaaat ctttctgct cttcgcaatg tcaacagtac ctttagtata	8640
ttctccagta gataggagc ctttgcata caattctgct aacatcaaaa ggctcttagg	8700
ttcctttgtt acttctctg cgcctgctt caaacgcta acaatactg ggcccaccac	8760
accgtgtgca ttcgtaatgt ctgccattc tgctattctg tatacacccg cagagtactg	8820
caatttgact gtattacaa tgtcagcaaa ttttctgtct tcgaagagta aaaaattgta	8880
cttggcggat aatgccttta gcggttaac tgtgccctcc atggaaaaat cagtcaagat	8940
atccacatgt gtttttagta aacaaatttt gggaacctaat gcttcaacta actccagtaa	9000
ttccttgggt gtacgaacat ccaatgaagc acacaagttt gtttgctttt cgtgcatgat	9060
attaaatagc ttggcagcaa caggactagg atgagtagca gcacgttcct tatatgtagc	9120
tttcgacatg atttatcttc gtttcctgca ggtttttgtt ctgtgcagtt ggggttaagaa	9180
tactgggcaa tticatgttt cttcaacact acatatgcgt atatatacca atctaagtct	9240
gtgtctcttc cttcgttctt ctttctgttc ggagattacc gaatcaaaaa aatttcaagg	9300
aaaccgaaat caaaaaaag aataaaaaaa aatgatgaa ttgaaaagct tgcattgcctg	9360
caggtcgact ctagtatact ccgtctactg tacgatacac ttccgctcag gtccttgctc	9420
tttaacgagg ctttaccact cttttgttac tctattgac cagctcagca aaggcagtg	9480
gatctaagat tctatcttcg cgatgtagta aaactagcta gaccgagaaa gagactagaa	9540
atgcaaaagg cacttctaca atggctgcca tcattattat ccgatgtgac gctgcatttt	9600
ttttttttt ttttttttt ttttttttt ttttttttt tttttttgt acaaatatca	9660
taaaaaaga gaatctttt aagcaaggat tttcttaact tcttcggcga cagcatcacc	9720

gacttcggtg gtactgttgg aaccacctaa atcaccagtt ctgatactg catccaaaac	9780
ctttttaact gcatcttcaa tggctttacc ttcttcaggc aagttcaatg acaatttcaa	9840
catcatgtca gcagacaaga tagtggcgat aggggttgacc ttattctttg gcaaactctg	9900
agcggaacca tggcatgggt cgtacaaacc aaatgcggtg ttcttgtctg gcaaagaggc	9960
caaggacgca gatggcaaca aaccaagga gcctgggata acggaggctt catcggagat	10020
gatatcacca aacatgttgc tgggtattat aataccatit aggtgggttg ggttcttaac	10080
taggatcatg gcggcagaat caatcaattg atgttgaact ttcaatgtag ggaattcgtt	10140
cttgatgggt tcttccacag tttttctcca taatcttgaa gaggccaaaa cattagcttt	10200
atccaaggac caaataggca atgggtggctc atgttgtagg gccatgaaag cggccattct	10260
tgtgattctt tgcacttctg gaacgggtga ttgttcacta tccaagcga caccatcacc	10320
atcgtcttcc tttctcttac caaagtaa atacctccact aattctctaa caacaacgaa	10380
gtcagtlacct ttagcaaat gtggcttgat tggagataag tctaaaagag agtcggatgc	10440
aaagttacat ggtcttaagt tggcgtacaa ttgaagtctt ttacggattt ttagtaaac	10500
ttgttcagg ttaacactac cggtacccca ttaggacca cccacagcac ctaacaaaac	10560
ggcatcagcc ttcttggagg cttccagcgc ctcatctgga agtggaaac ctgtagcatc	10620
gatagcagca ccaccaatta aatgatitct gaaatcgaac ttgacattgg aacgaacatc	10680
agaaatagct ttaagaacct taatggcttc ggctgtgatt tcttgacaa cgtggtcacc	10740
tggcaaacg acgatcttct taggggcaga cattacaatg gtatatcctt gaaatatata	10800
taaaaaaaa aaaaaaaaa aaaaaaaaa atgcagcttc tcaatgatat tcgaatacgc	10860
tttgaggaga tacagcctaa tatccgacaa actgttttac agatttacga tcgtacttgt	10920
taccatcat tgaatttga acatccgaac ctgggagttt tccctgaaac agatagtata	10980
tttgaacctg tataataata tatagtctag cgttttacgg aagacaatgt atgtatttcg	11040
gttcttggag aaactattgc atctattgca taggtaatct tgcacgtcgc atccccggtt	11100
cattttctgc gtttccatct tgcacttcaa tagcatatct ttgttaacga agcatctgtg	11160
cttcattttg tagaacaaaa atgcaacgcg agagcgctaa ttttcaaac aaagaatctg	11220
agctgcattt ttacagaaca gaaatgcaac gcgaaagcgc tatittacca acgaagaatc	11280
tgtgtctcat ttttgtaaaa caaaaatgca acgcgagagc gctaattttt caaacaaga	11340
atctgagctg catTTTTTt gaacagaaat gcaacgcgag agcgctattt taccaacaaa	11400
gaatctatac ttctTTTTt ttctacaaaa atgcatcccg agagcgctat ttttctaaca	11460
aagcatctta gattactttt tttctctttt gtgcgctcta taatgcagtc tcttgataac	11520
tttttgact gtaggtccgt taaggttaga agaaggctac ttgtgtctt atttctctt	11580

ccataaaaa agcctgactc cacttcccgc gtttactgat tactagcgaa gctgcgggtg 11640

cattttttca agataaaggc atccccgatt atattctata ccgatgtgga ttgcgcatac 11700

tttgtgaaca gaaagtgata gcgttgatga ttcttcatig gtcagaaaat tatgaacggt 11760

ttcttctatt ttgtctctat atactacgta taggaaatgt ttacattttc gtattgtttt 11820

cgattcactc tatgaatagt tcttactaca atttttttgt ctaaagagta atactagaga 11880

taaacataaa aaatgtagag gtcgagtta gatgcaagtt caaggagcga aagggtgatg 11940

ggtaggttat atagggatat agcacagaga tatatagcaa agagatactt ttgagcaatg 12000

tttgtggaag cggatattgc aatattttag tagctcgta cagtcgggtg cgtttttgg 12060

tttttgaaag tgcgtcttca gagcgctttt ggttttcaaa agcgctctga agttcctata 12120

ctttctagag aataggaact tcggaatagg aacttcaaag cgtttccgaa aacgagcgct 12180

tccgaaaatg caacgcgagc tgcgcacata cagctcactg ttcacgtcgc acctatatct 12240

gcgtgttgcc tgtatatata tatacatgag aagaacggca tagtgcgtgt ttatgcttaa 12300

atgcgtactt atatgcgtct atttatgtag gatgaaagg agtctagtag ctctgtgat 12360

attatcccat tccatgcggg gtatcgtag ctctctcag cactaccctt tagctgttct 12420

atatgctgcc actcctcaat tggattagtc tcaccttca atgctatcat ttcctttgat 12480

attggatcat atgcatagta ccgagaaact agaggatc 12518

<210> 774

<211> 1029

<212> DNA

<213> Artificial Sequence

<220><223> Anaerostipes caccae KARI-ABIG9 variant

<400> 774

atggaagaat gtaagatggc taagatttac taccaagaag actgtaactt gtccttgttg 60

gatggtgaaga ctatgccgt tatcggttac ggttctcaag gtcacgtca tgcctgaat 120

gctaaggaat ccggttgtaa cgttatcatt ggtttatagc aaggtgcgaa ggagtggaaa 180

agagctgaag aacaaggttt cgaagtctac accgtgctg aagctgctaa gaaggctgac 240

atcattatga tcttgatcaa cgatgaaaag caggctacca tgtacaaaa cgacatcgaa 300

ccaaacttgg aagccggtaa catgttgatg ttcgctcacg gtttcaacat ccatttcggt 360

tgtattgttc caccaaagga cgttgatgtc actatgatcg ctcaaaggg tccaggtcac 420

accgttagat ccgaatacga agaaggtaaa ggtgtcccat gcttggttgc tgtcgaacaa 480

gacgctactg gcaaggcttt ggatatggct ttggcctacg ctttagccat cgggtgtgct 540

agagccggtg tcttggaac taccttcaga accgaaactg aaaccgactt gttcggtgaa 600
caagctgttt tatgtgggtg tgtctgcgct ttgatgcagg ccggttttga aaccttggtt 660

gaagccggtt acgaccecaag aaacgcttac ttcgaatgta tccacgaaat gaagttgatc 720
gttgacttga tctaccaatc tggtttctcc ggtatgcgtt actctatctc caacactgct 780
gaatacgggtg actacattac cggtccaaag atcattactg aagataccaa gaaggctatg 840
aagaagattt tgtctgacat tcaagatggt acctttgcca aggacttctt ggttgacatg 900
tctgatgctg gttcccaggt ccacttcaag gctatgagaa agttggcctc cgaacacca 960
gctgaagttg tcggtgaaga aattagatcc ttgtactcct ggtccgacga agacaagttg 1020
attaacaac 1029

<210> 775
<211> 31
<212> DNA
<213> Artificial Sequence
<220><223> Primer pOLE1-NgoMI
<400> 775
cgtggccggc cgtgcaaaaa ctaactccga g 31
<210> 776
<211> 34
<212> DNA
<213> Artificial Sequence
<220><223> Primer pOLE1-PmII
<400> 776
gtactcacgt gtttgttgta atgttttagt gctg 34
<210> 777
<211> 1029
<212> DNA
<213> Artificial Sequence
<220><223> K9G9 ilvC gene
<400> 777
atggaagaat gtaagatggc taagatttac taccaagaag actgtaactt gtccttgttg 60
gatggtgaaga ctatcgccgt tatcggttac ggttctcaag gtcacgtca tgccctgaat 120
gctaaggaat ccggttgtaa cgttatcatt ggtttatacg aaggtgcgaa ggagtggaaa 180

agagctgaag aacaaggttt cgaagtctac accgctgctg aagctgctaa gaaggctgac 240
 atcattatga tcttgatcaa cgatgaaaag caggctacca tgtacaaaaa cgacatcgaa 300
 ccaaacttgg aagccggtaa catgttgatg ttcgctcacg gtttcaacat ccatttcggt 360
 tgtattgttc caccaaagga cgttgatgtc actatgatcg ctccaaaggg tccaggtcac 420
 accgttagat ccgaatacga agaaggtaaa gggtgcccat gcttggttgc tgtcgaacaa 480

 gagcctactg gcaaggcttt ggatatggct ttggcctacg ctttagccat cgggtggtgct 540
 agagccgggt tcttggaac taccttcaga accgaaactg aaaccgactt gttcggtgaa 600
 caagctgttt tatgtgggtg tgtctgcgct ttgatgcagg ccggttttga aaccttggtt 660
 gaagccggtt acgacceag aaacgcttac ttcgaatgta tccacgaaat gaagttgatc 720
 gttgacttga tctaccaatc tggtttctcc ggtatgcgtt actctatctc caacactgct 780
 gaatacgggt actacattac cgggtccaaag atcattactg aagatacga gaaggctatg 840
 aagaagattt tgctgacat tcaagatggt acctttgcca aggacttctt ggttgacatg 900

 tctgatgctg gttcccaggt ccacttcaag gctatgagaa agttggcctc cgaacaccca 960
 gctgaagttg tcggtgaaga aattagatcc ttgtactcct ggtccgacga agacaagttg 1020
 attaacaac 1029

 <210> 778
 <211> 9613
 <212> DNA
 <213> Artificial Sequence
 <220><223> pLH701 plasmid containing K9G9 variant
 <400> 778

 ggccgcacct ggtaaaacct ctagtggagt agtagatgta atcaatgaag cggaagccaa 60
 aagaccagag tagaggccta tagaagaaac tgcgatacct tttgtgatgg ctaacaaaac 120

 agacatcttt ttatatgttt ttacttctgt atatcgtgaa gtagtaagtg ataagcgaat 180
 ttggctaaga acgttgtaag tgaacaaggg acctcttttg cctttcaaaa aaggattaaa 240
 tggagttaat cattgagatt tagttttcgt tagattctgt atccctaaat aactccctta 300
 cccgacggga aggcacaaaa gacttgaata atagcaaacg gccagtagcc aagaccaaat 360
 aatactagag ttaactgatg gtcttaaaca ggcattacgt ggtgaactcc aagaccaata 420
 tacaaaatat cgataagtta ttcttgccca ccaatttaag gagcctacat caggacagta 480
 gtaccattcc tcagagaaga ggtatacata acaagaaaat cgcgtgaaca cttatataa 540

 cttagcccgt tattgagcta aaaaaccttg caaaatttcc tatgaataag aatacttcag 600

acgtgataaa aatttacttt ctaactcttc tcacgtcgcc cctatctgtt ctccgctct 660
accgtgagaa ataaagcatc gactacggca gttcgctgtc actgaactaa aacaataagg 720
ctagtccgaa tgatgaactt gcttgcgtgc aaacttctga gttgccgctg atgtgacact 780
gtgacaataa attcaaaccg gttatagcgg tctcctccgg taccggttct gccacctcca 840
atagagctca gtaggagica gaacctctgc ggtagcgtgc agtgactcat ccgcgtttcg 900
taagtgtgac gcgtgcacat ttgccccgtt cccgctcacc ttgcagcagg cggaaatatt 960

catcacgctg taggacgcaa aaaaaaata attaatcgta caagaatctt ggaaaaaaa 1020
ttgaaaaatt ttgtataaaa gggatgacct aacttgactc aatggctttt acaccagta 1080
ttttcccttt ccttgtttgt tacaattata gaagcaagac aaaaacatat agacaaccta 1140
ttcctaggag ttatatTTTT ttaccctacc agcaatataa gtaaaaaact gtttaaacag 1200
tatggaagaa tgtaagatgg ctaagattta ctaccaagaa gactgtaact tgccttgtt 1260
ggatggttaag actatcgccg ttatcggtta cggttctcaa ggtcacgctc atgccctgaa 1320
tgctaaggaa tccggttgta acgttatcat tggtttatac gaaggcgcga aggagtggaa 1380

aagagctgaa gaacaaggtt tcgaagtcta caccgctgct gaagctgcta agaagctga 1440
catcattatg atcttgatca acgatgaaaa gcaggctacc atgtacaaaa acgacatcga 1500
accaaacttg gaagccggtta acatgttgat gttcgctcac ggtttcaaca tccatttcgg 1560
ttgtattgtt ccaccaaagg acgttgatgt cactatgac gctccaaagg gtccaggta 1620
caccgttaga tccgaatcac aagaaggtaa aggtgtccca tgcttggttg ctgtcgaaca 1680
agacgctact ggcaaggtt tggatatggc ttggcctac gctttagcca tcggtggtgc 1740
tagagccggt gtcttgaaa ctaccttcag aaccgaaact gaaaccgact tgttcggtga 1800

acaagctgtt ttatgtggtg gtgtctgcgc ttgatgcag gccggttttg aaaccttggt 1860
tgaagccggt tacgaccaa gaaacgctta cttcgaatgt atccacgaaa tgaagttgat 1920
cgttgacttg atctaccaat ctggtttctc cggtatgcgt tactctatct ccaacactgc 1980
tgaatacggg gactacatta ccggtcctaaa gatcattact gaagatacca agaaggctat 2040
gaagaagatt ttgtctgaca ttcaagatgg tacctttgcc aaggacttct tggttgacat 2100
gtctgatgct gggtcccagg tccacttcaa ggctatgaga aagttggcct ccgaacaccc 2160
agctgaagtt gtcggtgaag aaattagatc ctgtactcc tgggtccgacg aagacaagtt 2220

gattaacaac tgaggccctg caggccagag gaaaaataa tcaagtgtg gaaacttttt 2280
ctcttggaat ttttgcaaca tcaagtcata gtcaattgaa ttgacccaat ttcacattta 2340
agattttttt ttttcatcc gacatacatc tgtacactag gaagccctgt ttttctgaag 2400
cagcttcaaa tatatatatt ttttacatat ttattatgat tcaatgaaca atctaattaa 2460

atcgaaaaca agaaccgaaa cgcgaaataa taatttatit agatgggtgac aagtgtataa	2520
gtcctcatcg ggacagctac gatttctctt tcggtttttg ctgagctact ggttgctgtg	2580
acgcagcggc attagcgcgg cgttatgagc taccctcgtg gcctgaaaga tggcgggaat	2640
aaagcggaac taaaaattac tgactgagcc atattgaggt caatttgtca actcgtcaag	2700
tcacgtttgg tggacggccc ctttccaacg aatcgtatat actaacatgc gcgcgcttcc	2760
tatatacaca tatacatata tatatatata tatatgtgtg cgtgtatgtg tacacctgta	2820
tttaatttcc ttactcgcgg gtttttcttt ttctcaatt cttggcttcc tctttctcga	2880
gcggaccgga tcttccgcgg tgccggcaga tctatttaaa tggcgcgcgg acgtcaggtg	2940
gcacttttcc gggaaatgtg cgcggaaccc ctatttgttt atttttctaa atacattcaa	3000
atatgtatcc gctcatgaga caataacctt gataaatgct tcaataatat tgaanaagga	3060
agagtatgag tattcaacat ttccgtgtcg cctttattcc cttttttgcg gcattttgcc	3120
ttcctgtttt tgctcaccca gaaacgctgg tgaaagtaaa agatgctgaa gatcagttgg	3180
gtgcacgagt gggttacatc gaactggatc tcaacagcgg taagatcctt gagagttttc	3240
gccccgaaga acgttttcca atgatgagca cttttaaagt tctgctatgt ggcgcggtat	3300
tatcccgat tgacgccggg caagagcaac tcggtcgccg catacactat tctcagaatg	3360
acttggttga gtactcacca gtcacagaaa agcatcttac ggatggcatg acagtaagag	3420
aattatgcag tgctgccata accatgagtg ataacttgc ggccaactta cttctgacaa	3480
cgatcgagg accgaaggag ctaaccgtt tttgcacaa catgggggat catgtaactc	3540
gccttgatcg ttgggaaccg gagctgaatg aagccatacc aaacgacgag cgtgacacca	3600
cgatgcctgt agcaatggca acaacgttgc gaaaactatt aactggcgaa ctacttactc	3660
tagcttccc gcaacaatta atagactgga tggaggcgga taaagttgca ggaccacttc	3720
tgcgctcggc ctttccggct ggctggttta ttgctgataa atctggagcc ggtgagcgtg	3780
ggtctcgcgg tatcattgca gcactggggc cagatggtaa gccctcccgt atcgtagtta	3840
tctacacgac ggggagtcag gcaactatgg atgaacgaaa tagacagatc gctgagatag	3900
gtgcctcact gattaagcat tggtaactgt cagaccaagt ttactcatat atactttaga	3960
ttgatttaaa acttcatttt taatttaaaa ggatctaggt gaagatcctt ttgataatc	4020
tcatgaccaa aatcccttaa cgtgagtttt cgttccactg agcgtcagac cccgtagaaa	4080
agatcaaagg atcttcttga gatccttttt ttctgcgcgt aatctgctgc ttgcaaaaa	4140
aaaaaccacc gctaccagcg gtggtttgtt tgccggatca agagctacca actcttttcc	4200
cgaaggtaac tggcttcagc agagcgcaga taccaatac tgttcttcta gtgtagccgt	4260

agttaggcca ccacttcaag aactctgtag caccgcctac atacctcgct ctgctaatec	4320
tgttaccagt ggctgctgcc agtggcgata agtcgtgtct taccgggttg gactcaagac	4380
gatatgtacc ggataaggcg cagcggtcgg gctgaacggg gggttcgtgc acacagccca	4440
gcttggagcg aacgacctac accgaactga gataacctaca gcgtgagcta tgagaaagcg	4500
ccacgcttcc cgaaggagga aaggcggaca ggtatccggt aagcggcagg gtcggaacag	4560
gagagcgcac gagggagctt ccagggggaa acgcctggta tctttatagt cctgtcgggt	4620
ttcgccacct ctgacttgag cgctgatttt tgtgatgtc gtcagggggg cggagcctat	4680
ggaaaaacgc cagcaacgcg gcctttttac ggttcttggc cttttgctgg cttttgtc	4740
acatgttctt tctgcgtta tccctgatt ctgtggataa ccgtattacc gccttgagt	4800
gagctgatac cgctcgccgc agccgaacga ccgagcgcag cgagtcagt agcgaggaag	4860
cggaagagcg cccaatacgc aaaccgcctc tccccgcggt ttggccgatt cattaatgca	4920
gctggcacga caggtttccc gactggaaag cgggcagtga gcgcaacga attaatgtga	4980
gttagctcac tcattagcca cccaggtt tacactttat gcttccggt cgtatgtgt	5040
gtggaattgt gagcggataa caatttcaca caggaaacag ctatgacct gattacgcca	5100
agctttttct ttccaatttt tttttttcgc tcattataaa aatcattacg accgagattc	5160
ccgggtaata actgataaa ttaaattgaa gctctaattt gtgagtttag tatacatgca	5220
tttacttata atacagtttt ttagttttgc tggccgcac ttcitcaaata tgcttcccag	5280
cctgcttttc tgtaacgttc accctctacc ttagcatccc ttccttttgc aaatagtcct	5340
cttccaacaa taataatgtc agatcctgta gagaccacat catccacgt tctatactgt	5400
tgaccaatg cgctccctt gtcactaaa cccacaccgg gtgtcataat caaccaatcg	5460
taaccttcat ctcttcacc catgtctctt tgagcaataa agccgataac aaaatctttg	5520
tcgtctctcg caatgtcaac agtaccctta gtatattctc cagtagatag ggagcccttg	5580
catgacaatt ctgctaacat caaaaggcct ctaggttctt ttgttacttc ttctgccgcc	5640
tgcttcaaac cgtaacaat acctgggccc accacaccgt gtgcattcgt aatgtctgcc	5700
cattctgcta ttctgtatac acccgagag tactgcaatt tgactgtatt accaatgtca	5760
gcaaattttc tgccttcgaa gagtaaaaaa ttgtacttgg cggataatgc ctttagcggc	5820
ttaactgtgc cctccatgga aaaatcagtc aagatatcca catgtgtttt tagtaaaaa	5880
atgttgggac ctaatgttc aactaactcc agtaattcct tgggtgtacg aacatccaat	5940
gaagcacaca agttgtttg cttttcgtgc atgatattaa atagcttggc agcaacagga	6000

ctaggatgag tagcagcacg ttccttatat gtagctttcg acatgattta tcttcgtttc 6060

ctgcaggttt ttgttctgtg cagttgggtt aagaatactg ggcaatttca tgtttcttca 6120

acactacata tgcgtatata taccaatcta agtctgtgct ccttccttcg ttcttccttc 6180

tgttcggaga ttaccgaatc aaaaaaattt caaggaaacc gaaatcaaaa aaaagaataa 6240

aaaaaaaaatg atgaattgaa aagcttgcat gcctgcaggt cgactctagt atactccgtc 6300

tactgtacga tacacttccg ctacaggctc tgccttttaa cgaggcctta ccactctttt 6360

gttactctat tgatccagct cagcaaaggc agtgtgatct aagattctat cttcgcgatg 6420

tagtaaaact agctagaccg agaaagagac tagaaatgca aaaggcactt ctacaatggc 6480

tgccatcatt attatccgat gtgacgctgc attttttttt tttttttttt tttttttttt 6540

tttttttttt tttttttttt ttgtacaaa tatcataaaa aaagagaatc tttttaagca 6600

aggattttct taacttcttc ggcgacagca tcaccgactt cgggtggtact gttggaacca 6660

cctaaatcac cagttctgat acctgcatcc aaaacctttt taactgcac ttcaatggct 6720

ttacttctt caggcaagtt caatgacaat ttcaacatca ttgcagcaga caagatagtg 6780

gcgatagggt tgaccttatt ctttggcaaa tctggagcgg aacctggca tggttcgtac 6840

aaaccaaatg cgggtgttctt gtctggcaaa gaggccaagg acgcagatgg caacaaaccc 6900

aaggagcctg ggataacgga ggcttcacgc gagatgatat caccaaact gtgtctggtg 6960

attataatac catttaggtg gggtgggttc ttaactagga tcatggcggc agaataatc 7020

aattgatgtt gaactttcaa ttaggggaat tcgttcttga tggtttcctc cacagttttt 7080

ctccataatc ttgaagagc caaaacatta gctttatcca aggaccaaact aggcaatggt 7140

ggctcatgtt gtagggccat gaaagcggc attcttga ttccttgac ttcctggaacg 7200

gtgtattgtt cactatccca agcgacacca tcaccatcgt cttcctttct cttaccaag 7260

taaatactc cactaatc tctaacaaca acgaagtcag tacctttagc aaattgtggc 7320

ttgattggag ataagtctaa aagagagtcg gatgcaaagt tacatggtct taagtggcg 7380

tacaattgaa gttctttacg gatttttagt aaacctgtt caggctaac actaccgta 7440

ccccatttag gaccaccac agcacctaac aaaacggcat cagccttctt ggaggttcc 7500

agcgctcat ctggaagtgg aacacctgta gcatcgatag cagcaccacc aattaaatga 7560

ttttcgaat cgaacttgac attggaacga acatcagaaa tagctttaag aaccttaatg 7620

gcttcggctg tgatttcttg accaacgtgg tcacctggca aaacgacgat cttcttaggg 7680

gcagacatta caatggtata tccttgaaat atatatataa aaaaaaaaaa aaaaaaaaaa 7740

aaaaaatgca gcttctcaat gatattcgaa tacgcttga ggagatacag ctaatatcc 7800

gacaaactgt ttacagatt tacgatcgt cttgttacc atcattgaat ttgaacatc 7860

cgaacctggg agttttccct gaaacagata gtatatattga acctgtataa taatatatag 7920
tctagcgctt tacggaagac aatgtatgta ttccggttcc tggagaaact attgcatcta 7980
ttgcataggt aatcttgcac gtgcgcatccc cggttcattt tctgcgttcc catcttgcac 8040
ttcaatagca tatctttgtt aacgaagcat ctgtgcttca tttttagaa caaaaatgca 8100

acgcgagagc gctaattttt caaacaaga atctgagctg catttttaca gaacagaaat 8160
gcaacgcgaa agcgctatit taccaacgaa gaatctgtgc ttcatitttg taaaacaaaa 8220
atgcaacgcg agagcgctaa tttttcaaac aaagaatctg agctgcattt ttacagaaca 8280
gaaatgcaac gcgagagcgc tattttacca acaagaatc tatacttctt tttgtttcta 8340
caaaaatgca tcccgagagc gctatitttt taacaaagca tcttagatta ctttttttct 8400
cctttgtgcg ctctataatg cagtctcttg ataacttttt gcactgtagg tccgttaagg 8460
ttagaagaag gctactttgg tgtctatitt ctcttcata aaaaaagcct gactccactt 8520

cccgcgttta ctgattacta gcgaagctgc ggggtgcattt tttcaagata aaggcatccc 8580
cgattatatt ctataccgat gtggattgcg catactttgt gaacagaaag tgatagcggt 8640
gatgattctt cattggctag aaaattatga acggtttctt ctatittgtc tctatatact 8700
acgtatagga aatgtttaca ttttcgtatt gttttcgatt cactctatga atagtcttta 8760
ctacaatttt tttgtctaaa gagtaatact agagataaac ataaaaaatg tagaggctga 8820
gtttagatgc aagtccaagg agcgaaaggt ggatgggtag gttatatagg gatatagcac 8880
agagatatat agcaaagaga tacttttgag caatgtttgt ggaagcggtta ttcgcaatat 8940

tttagtagct cgttacagtc cgggtcggtt ttggtttttt gaaagtgcgt cticagagcg 9000
cttttggttt tcaaaagcgc tctgaagttc ctatactttc tagagaatag gaacttcgga 9060
ataggaactt caaagcgtt ccgaaaacga gcgcttccga aatgcaacg cgagctgcgc 9120
acatacagct cactgttcac gtgcgacctt tatctgcgtg ttgcctgtat atatatatac 9180
atgagaagaa cggcatagtg cgtgtttatg cttaaatgcg tacttatatg cgtctattta 9240
ttaggatga aaggtagtct agtacctct gtgatattat cccattccat gcgggggtatc 9300
glatgcttcc ttcagcacta ccttttagct gttctatatg ctgccactcc tcaattggat 9360

tagtctcacc ctccaatgct atcatttcct ttgatattgg atcatatgca tagtaccgag 9420
aaactagagg atctccatt accgacattt gggcgctata cgtgcatatg ttcatgtatg 9480
tatctgtatt taaaacactt ttgtattatt ttctctcata tatgtgtata ggttttatag 9540
gatgatataa ttattacttc accacccttt atttcaggct gatatcttag ccttgttact 9600
agtcaccggt ggc 9613

<210> 779

<211> 28

<212> DNA

<213> Artificial Sequence

<220><223> N1090 primer

<400> 779

aatcccgatg tatgggtttg gttgccag 28

<210> 780

<211> 7589

<212> DNA

<213> Artificial Sequence

<220><223> Als integration plasmid

<400> 780

tatttgtatc gaggtgtcta gtcttctatt acactaatgc agtttcaggg ttttggaac 60

cacactgttt aaacagtgtt ccttaataca ggatacctct tttttttcc ttggttccac 120

taattcatcg gtttttttt tggaagacat cttttccaac gaaaagaata tacatatcgt 180

ttaagagaaa ttctccaaat ttgtaaagaa gcggacccag acttaaggcc gcccgcaaat 240

taaagccttc gagcgtccca aaaccttctc aagcaagggt ttcagtataa tgttacatgc 300

gtacacgcgt ctgtacagaa aaaaaagaaa aatttgaaat ataataacg ttcttaatac 360

taacataact ataaaaaat aaataggac ctagacttca ggttgtctaa ctcttcctt 420

ttcgggttaga gcggatgtgg ggggagggcg tgaatgtaag cgtgacataa ctaattacat 480

gattaattaa ctagagagct ttcgttttca tgagttcccc gaattcttcc ggaagcttgt 540

cacttgctaa attaatgtta tcaactgtagt caaccgggac atcgatgatg acaggacctt 600

cagcgttcat gccttgacgc agaacatctg ccagctggtc tggtgattct acgcgcaagc 660

cagttgctcc gaagctttcc gcatatttca cgatatcgat atttccgaaa tcgaccgcag 720

atgtacggtt atatttttcc aattgctgga atgcaacat gtcatatgtg ctgtcgttcc 780

atacaatgtg tacaattggt gcttttagtc gaactgctgt ctctaattcc attgctgaga 840

ataagaaacc gccgtcacca gagacagaaa ccacttttcc tcccggtttc accaatgaag 900

cgccgattgc ccaaggaagc gcaacgccga gtgtttgcat accgttactg atcattaatg 960

ttaacggctc gtagctgcgg aaataacgtg acatccaaat ggctgtcgaa ccgatatcgc 1020

aagttactgt aacatgatca tcgactgcat tacgcaactc ttttaacgatt tcaagagggt 1080

gcgctctgtc tgatttccaa tctgcaggca cctgctcacc ttcatgcata tattgtttta 1140

aatcagaaag gattttctgc tcacgctctg caaattccac tttcacagca tcgtgttcga	1200
tatgattgat cgtggacgga atgtcaccga tcaattcaag atcaggctgg taagcatgat	1260
caatgtcagc gataatctcg tctaaatgga taattgtccg gtctccattg atattccaga	1320
atttcggatc atattcaatc gggtcatagc cgatcgtcag aacaacatct gctgtctcta	1380
gcagtaaatc gccaggctgg ttgcggaaca aaccgatacg gccaaaatat tgatcctcta	1440
aatctctaga aagggtaccg gcagcttgat atgtttcaac aaatggaagc tgaacctttt	1500
tcaaaagctt gcgaaccgt ttaattgctt ccggtcttcc gcctttcatg ccgacaaaaa	1560
cgacaggaag ttttgcgttt tggatTTTTG ctatggccgc actgattgca tcctctgctg	1620
caggaccgag ttttggcgt gcaacagcac gcacgttttt cgtatttttg acttcattca	1680
caacatcttg cgaaaagctc acaaaagcgg cccagcctg ccctgctgac gctatcctaa	1740
atgcatttgt aacagcttcc ggtatatTTT ttacatcttg aacttctaca ctgtattttg	1800
taatcggtg gaatagccc gcattatcca aagattgatg tgtccgtttt aaacgatctg	1860
cacggatcac gtttccagca agcgcaacga cagggtctcc ttcagtgttc gctgtcagca	1920
ggcctgttgc caagttagag gcaaccggtc ctgatgtgac taacacgact cccggttttc	1980
cagttaaacg gccgactgct tgggccatga atgctgcgtt ttgttcgtgc cgggcaacga	2040
taatttcagg tctttatct tgtaaagcgt caaataccgc atcaattttt gcacctggaa	2100
tgccaaatac atgtgtgaca ctttgcctca ctaagcaatc aacaacaagc tccgccctc	2160
tgtttttcac aagggtattt tgttcttttg ttgcttttgt caacatctc agctctagat	2220
ttgaatatgt attacttgg tttggttata tatgacaaaa gaaaaagaag aacagaagaa	2280
taacgcaagg aagaacaata actgaaattg atagagaagt attatgtctt tgtcttttta	2340
taataaatca agtcagaaa tccgttagac aacatgaggg ataaaattta acgtgggcga	2400
agaagaagga aaaaagtttt tgtgagggcg taattgaagc gatctgttga ttgtagattt	2460
ttttttttg aggagtcaaa gtcagaagag aacagacaaa tggatttaac catccaatac	2520
ttttttggag caacgctaag ctcatgcttt tccattgggt acgtgctcag ttgttagata	2580
tggaaagaga ggatgctcac ggcagcgtga ctccaattga gcccgaaga gaggatgcca	2640
cgttttcccg acggctgcta gaatggaaaa aggaaaaata gaagaatccc attcctatca	2700
ttatttacgt aatgaccac acatttttga gatTTTcaac tattacgtat tacgataatc	2760
ctgctgtcat tatcattatt atctatatcg acgtatgcaa cgtatgtgaa gccaaagtagg	2820
caattattta gtactgtcag tattgttatt catttcagat cttaaacct accaggccaa	2880

ttcaacagac tgcggcaac ttcttgctcg gtctttccat ggtaagtgac agtgcagtaa	2940
taatatgaac caattttatt ttcggtacat aaaaatgctt ataaaacttt aactaataat	3000
tagagattaa atcgcaaacg gccgactcta gaggatcccc caccttggtt aactcgttgt	3060
atcatcactg gataacttcg tatagcatac attatacgaa gttatctagg gattcataac	3120
cattttctca atcgaattac acagaacaca ccgtacaaac ctctctatca taactactta	3180
atagtccac acgtactcgt ctaaatacac atcatcgtcc tacaagttca tcaaagtgtt	3240
ggacagacaa ctataccage atggatctct tgtatcgggt cttttctccc gctctctcgc	3300
aataacaatg aacactgggt caatcatagc ctacacaggt gaacagagta gcgtttatac	3360
agggtttata cggtgattcc tacggcaaaa atttttcatt tctaaaaaa aaaagaaaaa	3420
tttttcttc caacgctaga aggaaaagaa aaatctaatt aaattgattt ggtgattttc	3480
tgagagtcc ctttttcata tatcgaattt tgaatataaa aggagatcga aaaaattttt	3540
ctattcaatc tgttttctgg ttttatttga tagttttttt gtgtattatt attatggatt	3600
agtactgggt tatatgggtt tttctgtata acttcttttt attttagttt gtttaacttt	3660
attttgagtt acattatagt tcctaactg caagagaagt aacattaaaa ctcgagatgg	3720
glaaggaaaa gactcacgtt tcgaggccgc gattaaattc caacatggat gctgatttat	3780
atgggtataa atgggctcgc gataatgtcg ggcaatcagg tgcgacaatc tatcgattgt	3840
atgggaagcc cgatgcgcca gatttgtttc tgaacatgg caaaggtagc gttgccaatg	3900
atgttacaga tgagatggtc agactaaact ggctgacgga atttatgcct cttccgacca	3960
tcaagcattt tatccgtact cctgatgatg catggttact caccactgcg atccccggca	4020
aaacagcatt ccaggtatta gaagaatac ctgattcagg tgaaaatatt gttgatgcgc	4080
tggcagtgtt cctgcgccgg ttgcattcga ttctgtttg taattgtcct tttaacagcg	4140
atcgcgtatt tcgtctcgtc caggcgcaat cacgaatgaa taacggtttg gttgatgcga	4200
gtgattttga tgacgagcgt aatggctggc ctgttgaaca agtctggaaa gaaatgcata	4260
agcttttgcc attctcaccg gattcagtcg tcaactatgg tgattttctc cttgataacc	4320
ttatttttga cgaggggaaa ttaatagggt gtattgatgt tggacgagtc ggaatcgcag	4380
accgatacca ggatcttgcc atcctatgga actgcctcgg tgagttttct cttcattac	4440
agaaacggct ttttcaaaaa tatgttattg ataactctga tatgaataaa ttgcagtttc	4500
atttgatgct cgatgagttt ttctaagttt aacttgatac tactagattt tttctcttca	4560
tttataaaat ttttggttat aattgaagct ttagaagtat gaaaaaatcc ttttttttca	4620
ttctttgcaa ccaaaataag aagcttcttt tattcattga aatgatgaat ataaacctaa	4680
caaaagaaaa agactcgaat atcaaacatt aaaaaaaaaat aaaagagggt atctgttttc	4740

ccatttagtt ggagtttgca ttttctaata gatagaactc tcaattaatg tggatttagt	4800
ttctctgttc gataacttcg tatagcatac attatacgaa gttatctgaa cattagaata	4860
cgtaatccgc aatgcggggc cgcttaatta atctagagtc gacctgcagg catgcaagct	4920
tggcgtaatc atggtcatag ctgtttccctg tgtgaaattg ttatccgctc acaattccac	4980
acaacatacg agccggaagc ataaagtgt aagcctgggg tgcctaata gaagagctaac	5040
tcacattaat tgcgttgccg tcaactgccc ctttccagtc gggaaacctg tcgtgccagc	5100
tgcattaatg aatcgcccaa cgcgcgggga gaggcgggtt gcgtattggg cgctcttccg	5160
cttctcgtct cactgactcg ctgcgctcgg tcgttcggct gcggcgagcg gtatcagctc	5220
actcaaaggc ggtaatacgg ttatccacag aatcagggga taacgcagga aagaacatgt	5280
gagcaaaagg ccagcaaaag gccaggaacc gtaaaaaggc cgcggtgctg gcgtttttcc	5340
ataggtccg cccccctgac gagcatcaca aaaatcgacg ctcaagtcag aggtggcgaa	5400
acccgacagg actataaaga taccaggcgt tccccctgg aagctccctc gtgcgctctc	5460
ctgttccgac cctgccgctt accggatacc tgcgcgctt tctcccttcg ggaagcgtgg	5520
cgctttctca tagctcagcg tgtaggtatc tcagttcggg gtaggtcgtt cgctccaagc	5580
tgggctgtgt gcacgaacct cccgttcagc ccgaccctg cgccttatcc ggtaactatc	5640
gtcttgagtc caaccggta agacacgact tatcgccact ggagcagacc actggttaaca	5700
ggattagcag agcgaggtat gtaggcgggt ctacagagtt cttgaagtgg tggcctaact	5760
acggctacac tagaaggaca gtatttggtg tctgcgctct gctgaagcca gttaccttcg	5820
gaaaaagagt tggtagctct tgatccggca aacaaaccac cgctggtagc ggtggttttt	5880
ttgtttgcaa gcagcagatt acgcgcagaa aaaaaggatc tcaagaagat cctttgatct	5940
tttctacggg gtctgacgct cagtggaacg aaaactcacg ttaagggtatt ttggtcatga	6000
gattatcaaa aaggatcttc acctagatcc ttttaaatga aaaatgaagt tttaaatcaa	6060
tctaaagtat atatgagtaa acttggtctg acagttacca atgcttaatc agtgaggcac	6120
ctatctcagc gatctgtcta ttctgttcac ccatagttgc ctgactcccc gtcgtgtaga	6180
taactacgat acgggagggc ttaccatctg gccccagtgc tgcaatgata ccgcgagacc	6240
cacgctcacc ggctccagat ttatcagcaa taaaccagcc agccggaagg gccgagcgca	6300
gaagtggtec tgcaacttta tccgcctcca tccagtctat taattgttgc cgggaagcta	6360
gagtaagtag ttcgccagtt aatagtttgc gcaacgttgt tgccattgct acaggcatcg	6420
tgggtgcacg ctctcgtttt ggtatggctt cattcagctc cggttcccaa cgatcaaggc	6480
gagttacatg atcccccatg ttgtgcaaaa aagcggttag ctcttcgggt cctccgatcg	6540
ttgtcagaag taagtgggcc gcagtggtat cactcatggt tatggcagca ctgcataatt	6600

ctcttactgt catgccaatcc gtaagatgct tttctgtgac tggtagtac tcaaccaagt 6660

cattctgaga atagtgtatg cggcgaccga gttgctcttg cccggcgtca atacgggata 6720

ataccgcgcc acatagcaga actttaaaag tgetcatcat tggaaaacgt tcttcggggc 6780

gaaaactctc aaggatctta ccgctgttga gatccagttc gatgtaaccc actcgtgcac 6840

ccaactgac ttcagcatct tttactttca ccagcgtttc tgggtgagca aaaacaggaa 6900

ggcaaaatgc cgcaaaaaag ggaataaggc cgacacggaa atgttgaata ctcatactct 6960

tcctttttca atattattga agcatttacc aggggttattg tctcatgagc ggatacatat 7020

ttgaatgtat ttagaaaaat aaacaaatag gggttccgcg cacatttccc cgaaaagtgc 7080

cacctgacgt ctaagaaacc attattatca tgacattaac ctataaaaat aggcgtatca 7140

cgaggccctt tcgtctcgcg cgtttcgggtg atgacgggtga aaacctctga cacatgcagc 7200

tcccggagac ggtcacagct tgtctgtaag cggatgccgg gagcagacaa gcccgtcagg 7260

gcgcgtcagc ggggtgttggc ggggtgtcggg gctggcttaa ctatgcggca tcagagcaga 7320

ttgtactgag agtgcacat atgcgggtgtg aaataccgca cagatgcgta aggagaaaat 7380

accgcatcag gcgccattcg ccattcaggc tgcgcaactg ttgggaaggc cgatcggtgc 7440

gggcctcttc gctattacgc cagctggcga aagggggatg tgctgcaagg cgattaagtt 7500

gggtaacgcc agggttttcc cagtcacgac gttgtaaaac gacggccagt gaattcgagc 7560

tcggtacccg gggatccggc gcgccgttt 7589

<210> 781

<211> 338

<212> PRT

<213> Artificial Sequence

<220><223> k9_Jarvis_SH

<400> 781

Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn Leu Ser Leu Leu Asp

1 5 10 15

Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser Gln Gly His Ala His

20 25 30

Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val Ile Ile Gly Leu Phe

35 40 45

Glu Gly Ala Glu Glu Trp Lys Arg Ala Glu Glu Gln Gly Ile Glu Val

50 55 60

Tyr Thr Ala Ala Glu Ala Ala Lys Lys Ala Asp Ile Ile Met Ile Leu
 65 70 75 80
 Ile Pro Asp Glu Lys Gln Ala Thr Met Tyr Lys Asn Asp Ile Glu Pro
 85 90 95

 Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala His Gly Phe Asn Ile
 100 105 110
 His Phe Gly Cys Ile Val Pro Pro Lys Asp Val Asp Val Thr Met Ile
 115 120 125
 Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser Glu Tyr Glu Glu Gly
 130 135 140
 Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln Asp Ala Thr Gly Lys
 145 150 155 160

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

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

 Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly Ser Gln Val His Phe
 290 295 300
 Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro Ala Glu Val Val Gly

305 310 315 320
 Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp Glu Asp Lys Leu Ile
 325 330 335
 Asn Asn

<210> 782

<211> 338

<212> PRT

<213> Artificial Sequence

<220><223> K9_Kelly

<400

> 782

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

Tyr Thr Ala Ala Glu Ala Ala Lys Lys Ala Asp Ile Ile Met Ile Leu
 65 70 75 80
 Ile Pro Asp Glu Lys Gln Ala Thr Met Tyr Lys Asn Asp Ile Glu Pro
 85 90 95
 Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala His Gly Phe Asn Ile
 100 105 110
 His Phe Gly Cys Ile Val Pro Pro Lys Asp Val Asp Val Thr Met Ile
 115 120 125

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

165 170 175
Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu Thr Glu Thr Asp Leu
180 185 190

Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val Cys Ala Leu Met Gln
195 200 205

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

Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile Val Asp Leu Ile Tyr
225 230 235 240

Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile Ser Asn Thr Ala Glu
245 250 255

Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile Thr Glu Asp Thr Lys
260 265 270

Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln Asp Gly Thr Phe Ala
275 280 285

Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly Ser Gln Val His Phe
290 295 300

Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro Ala Glu Val Val Gly
305 310 315 320

Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp Glu Asp Lys Leu Ile
325 330 335

Asn Asn

<210> 783

<211> 338

<212> PRT

<213> Artificial Sequence

<220><223> K9_Norman_SH

<400> 783

Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn Leu Ser Leu Leu Asp
1 5 10 15

Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser Gln Gly His Ala His

20	25	30	
Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val Ile Ile Gly Leu Phe			
35	40	45	
Glu Cys Ala Glu Glu Trp Lys Arg Ala Glu Glu Gln Gly Phe Glu Val			
50	55	60	
Tyr Thr Ala Ala Glu Ala Ala Lys Lys Ala Asp Ile Ile Met Ile Leu			
65	70	75	80
Ile Pro Asp Glu Lys Gln Ala Thr Met Tyr Lys Asn Asp Ile Glu Pro			
85	90	95	
Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala His Gly Phe Asn Ile			
100	105	110	
His Phe Gly Cys Ile Val Pro Pro Lys Asp Val Asp Val Thr Met Ile			
115	120	125	
Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser Glu Tyr Glu Glu Gly			
130	135	140	
Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln Asp Ala Thr Gly Lys			
145	150	155	160
Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala Ile Gly Gly Ala Arg			
165	170	175	
Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu Thr Glu Thr Asp Leu			
180	185	190	
Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val Cys Ala Leu Met Gln			
195	200	205	
Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr Asp Pro Arg Asn Ala			
210	215	220	
Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile Val Asp Leu Ile Tyr			
225	230	235	240
Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile Ser Asn Thr Ala Glu			
245	250	255	
Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile Thr Glu Asp Thr Lys			
260	265	270	

Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln Asp Gly Thr Phe Ala
 275 280 285

Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly Ser Gln Val His Phe
 290 295 300

Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro Ala Glu Val Val Gly
 305 310 315 320

Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp Glu Asp Lys Leu Ile
 325 330 335

Asn Asn

<210> 784

<211> 24

<212> DNA

<213> Artificial Sequence

<220><223> N1237 primer

<400> 784

gatgatgcta ttgtgtgcag aggg 24

<210> 785

<211> 27

<212> DNA

<213> Artificial Sequence

<220><223> N1238 primer

<400> 785

ccttactgaa gtatttagtt atcatgg 27

<210> 786

<211> 24

<212> DNA

<213> Artificial Sequence

<220><223> Primer N1240

<400> 786

ggatgtatgg gctaaatgta cggg 24

<210> 787

<211> 24

<212> DNA

<213> Artificial Sequence

<220><223> N1241 Primer

<400> 787

catccagtgt cgaaaacgag ctcg

24

<210> 788

<211> 343

<212> PRT

<213> Artificial Sequence

<220><223> Anaerostipes caccae KARI- K9D3 variant

<400> 788

Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn

1 5 10 15

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

20 25 30

Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val

35 40 45

Ile Ile Gly Leu Tyr Glu Gly Ala Lys Asp Trp Lys Arg Ala Glu Glu

50 55 60

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

65 70 75 80

Ile Ile Met Ile Leu Ile Asn Asp Glu Lys Gln Ala Thr Met Tyr Lys

85 90 95

Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala

100 105 110

His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val

115 120 125

Asp Val Thr Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser

130 135 140

Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln

145 150 155 160

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

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

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

<210> 789

<211> 448

<212> DNA

<213> Artificial Sequence

<220><223> Saccharomyces cerevisiae CUP1 promoter

<400> 789

cccattaccg acatttgggc gctatacgtg catatgttca tgtatgtatc tgtattttaa 60

acacttttgt attatttttc ctcatatatg tgtataggtt tatacggatg atttaattat 120

tacttcacca cccittatit caggctgata tcttagcctt gttactagt agaaaaagac	180
atttttgctg tcagtcactg tcaagagatt cttttgctgg catttcttct agaagcaaaa	240
agagcgatgc gtcttttccg ctgaaccgtt ccagcaaaaa agactaccaa cgcaatatgg	300
attgtcagaa tcatataaaa gagaagcaaa taactccttg tcttgitatca attgcattat	360
aatatcttct tgttagtga atatcatata gaagtcacg aaatagatat taagaaaaac	420
aaactgtaca atcaatcaat caatcate	448
<210> 790	
<211> 1713	
<212> DNA	
<213> Artificial Sequence	
<220><223> Bacillus subtilis AlsS	
<400> 790	
ttgacaaaag caacaaaaga acaaaaatcc cttgtgaaaa acagaggggc ggagcttggt	60
gttgattgct tagtggagca aggtgtcaca catgtatttg gcattccagg tgcaaaaatt	120
gatgcggtat ttgacgcttt acaagataaa ggacctgaaa ttatcgttgc ccggcacgaa	180
caaaacgcag cattcatggc ccaagcagtc ggccgtttaa ctggaaaacc gggagtcgtg	240
ttagtcacat caggaccggg tgcctctaac ttggcaacag gcctgctgac agcgaacact	300
gaaggagacc ctgtcgttgc gcttctgga aacgtgatcc gtgcagatcg tttaaaacgg	360
acacatcaat ctttgataa tgcggcgcta ttccagccga ttacaaaata cagtgtagaa	420
gttcaagatg taaaaaatat accggaagct gttacaaatg catttaggat agcgtcagca	480
gggcaggctg gggccgcttt tgtgagcttt ccgaagatg ttgtgaatga agtcacaaat	540
acgaaaaacg tgcgtgctgt tgcagcgcca aaactcggtc ctgcagcaga tgatgcaatc	600
agtgcggcca tagcaaaaat ccaaacagca aaacttcctg tcgttttggg cgcatgaaa	660
ggcgggaagac cggaagcaat taaagcgtt cgcaagcttt tgaaaaaggt tcagcttcca	720
tttgttgaaa calatcaagc tgccggtacc ctttctagag atttagagga tcaatatit	780
ggccgtatcg gtttgttccg caaccagcct ggcgatttac tgctagagca ggcagatgtt	840
gttctgacga tcggctatga ccgattgaa tatgatccga aattctggaa tatcaatgga	900
gaccggacaa ttatccattt agacgagatt atcgtgaca ttgatcatgc ttaccagcct	960
gatcttgaat tgatcggiga cattccgtcc acgatcaatc atatcgaaca cgatgctgtg	1020
aaagtggaat ttgcagagcg tgagcagaaa atcctttctg atttaaaaca atatatgcat	1080
gaaggtgagc aggtgcctgc agattggaaa tcagacagag cgcaccctct tgaaatcgtt	1140

aaagagttgc gtaatgcagt cgatgatcat gttacagtaa cttgcgatat cggttcgcac 1200
gccatttggg tgtcacgtta tttccgcagc tacgagccgt taacattaat gatcagtaac 1260
ggatatgaaa cactcggcgt tgcgcttcct tgggcaatcg gcgcttcatt ggtgaaaccg 1320
ggagaaaaag tggtttctgt ctctggtgac ggcggtttct tattctcagc aatggaatta 1380
gagacagcag ttcgactaaa agcaccaatt gtacacattg tatggaacga cagcacatat 1440
gacatggttg cattccagca attgaaaaaa tataaccgta catctgcggt cgatttcgga 1500

aatatcgata tcgtgaaata tgcggaaagc ttcggagcaa ctggcttgcg cgtagaatca 1560
ccagaccagc tggcagatgt tctgcgtcaa ggcatagaac ctgaaggtcc tgtcatcatc 1620
gatgtcccgg ttgactacag tgataacatt aatttagcaa gtgacaagct tccgaaagaa 1680
ttcggggaac tcatgaaaac gaaagctctc tag 1713

<210> 791

<211> 571

<212> PRT

<213> Artificial Sequence

<220><223> Bacillus subtilis

<400> 791

Met Leu Thr Lys Ala Thr Lys Glu Gln Lys Ser Leu Val Lys Asn Arg

1 5 10 15

Gly Ala Glu Leu Val Val Asp Cys Leu Val Glu Gln Gly Val Thr His

20 25 30

Val Phe Gly Ile Pro Gly Ala Lys Ile Asp Ala Val Phe Asp Ala Leu

35 40 45

Gln Asp Lys Gly Pro Glu Ile Ile Val Ala Arg His Glu Gln Asn Ala

50 55 60

Ala Phe Met Ala Gln Ala Val Gly Arg Leu Thr Gly Lys Pro Gly Val

65 70 75 80

Val Leu Val Thr Ser Gly Pro Gly Ala Ser Asn Leu Ala Thr Gly Leu

85 90 95

Leu Thr Ala Asn Thr Glu Gly Asp Pro Val Val Ala Leu Ala Gly Asn

100 105 110

Val Ile Arg Ala Asp Arg Leu Lys Arg Thr His Gln Ser Leu Asp Asn

115 120 125

Ala Ala Leu Phe Gln Pro Ile Thr Lys Tyr Ser Val Glu Val Gln Asp
130 135 140

Val Lys Asn Ile Pro Glu Ala Val Thr Asn Ala Phe Arg Ile Ala Ser
145 150 155 160

Ala Gly Gln Ala Gly Ala Ala Phe Val Ser Phe Pro Gln Asp Val Val
165 170 175

Asn Glu Val Thr Asn Thr Lys Asn Val Arg Ala Val Ala Ala Pro Lys
180 185 190

Leu Gly Pro Ala Ala Asp Asp Ala Ile Ser Ala Ala Ile Ala Lys Ile
195 200 205

Gln Thr Ala Lys Leu Pro Val Val Leu Val Gly Met Lys Gly Gly Arg
210 215 220

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

Pro Phe Val Glu Thr Tyr Gln Ala Ala Gly Thr Leu Ser Arg Asp Leu
245 250 255

Glu Asp Gln Tyr Phe Gly Arg Ile Gly Leu Phe Arg Asn Gln Pro Gly
260 265 270

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

Pro Ile Glu Tyr Asp Pro Lys Phe Trp Asn Ile Asn Gly Asp Arg Thr
290 295 300

Ile Ile His Leu Asp Glu Ile Ile Ala Asp Ile Asp His Ala Tyr Gln
305 310 315 320

Pro Asp Leu Glu Leu Ile Gly Asp Ile Pro Ser Thr Ile Asn His Ile
325 330 335

Glu His Asp Ala Val Lys Val Glu Phe Ala Glu Arg Glu Gln Lys Ile
340 345 350

Leu Ser Asp Leu Lys Gln Tyr Met His Glu Gly Glu Gln Val Pro Ala
355 360 365

Asp Trp Lys Ser Asp Arg Ala His Pro Leu Glu Ile Val Lys Glu Leu

370 375 380
 Arg Asn Ala Val Asp Asp His Val Thr Val Thr Cys Asp Ile Gly Ser
 385 390 395 400

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

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

 Leu Arg Gln Gly Met Asn Ala Glu Gly Pro Val Ile Ile Asp Val Pro
 530 535 540
 Val Asp Tyr Ser Asp Asn Ile Asn Leu Ala Ser Asp Lys Leu Pro Lys
 545 550 555 560
 Glu Phe Gly Glu Leu Met Lys Thr Lys Ala Leu
 565 570

 <210> 792
 <211> 250
 <212> DNA
 <213> Artificial Sequence
 <220><223> Saccharomyces cerevisiae CYC1 terminator 2
 <400> 792

ccgcaaatta aagccttcga gcgtcccaaa accttctcaa gcaaggtttt cagtataatg 60

ttacatgcgt acacgcgtct gtacagaaaa aaaagaaaaa ttgaaatat aaataacgtt 120

cttaatacta acataactat aaaaaataa ataggacct agacttcagg ttgtctaact 180

ccttcctttt cggttagagc ggatgtgggg ggaggcgctg aatgtaagcg tgacataact 240

aattacatga 250

<210> 793

<211> 1181

<212> DNA

<213> Artificial Sequence

<220><223> Saccharomyces cerevisiae ILV5 promoter

<400> 793

taaaacctct agtggagtag tagatgtaat caatgaagcg gaagccaaaa gaccagagta 60

gaggcctata gaagaaactg cgataccttt tgtgatggct aaacaaacag acatcttttt 120

atatgttttt acttctgtat atcgtgaagt agtaagtgat aagcgaattt ggctaagaac 180

gttgtaagtg aacaaggac ctcttttgcc ttcaaaaaa ggattaaatg gagttaatca 240

ttgagattta gttttcgta gattctgtat ccctaaataa ctccttacc cgacgggaag 300

gcacaaaaga ctigaataat agcaaacggc cagtagccaa gaccaataa tactagagtt 360

aactgatggt cttaaacagg cattacgtgg tgaactcaa gaccaatata caaatatcg 420

ataagttatt cttgccacc aatttaagga gcctacatca ggacagtagt accattcctc 480

agagaagagg tatacataac aagaaaatcg cgtgaacacc ttatataact tagcccgtta 540

ttgagctaaa aaaccttgca aaatttccta tgaataagaa tacttcagac gtgataaaaa 600

tttactttct aactcttctc acgtgcccc tatctgttct tccgtctac cgtgagaaat 660

aaagcatcga gtacggcagt tcgtgtcac tgaactaaa caataaggct agttcgaatg 720

atgaacttgc ttgtgtcaa acttctgagt tgccgtgat gtgacactgt gacaataaat 780

tcaaaccggt tatagcggtc tcctccggta ccggttctgc cacctccaat agagtcagt 840

aggagtcaga acctctgcgg tggctgtcag tgactcatcc gcgtttcgta agttgtgcgc 900

gtgcacattt cgcccgttcc cgctcatctt gcagcaggcg gaaattttca tcacgtgta 960

ggacgcaaaa aaaaaataat taatcgtaca agaactttgg aaaaaaatt gaaaaatttt 1020

gtataaaagg gatgacctaa cttgactcaa tggtttttac acccagtatt ttccctttcc 1080

ttgtttgtta caattataga agcaagacaa aaacatatag acaacctatt cctaggagtt 1140

atattttttt accctaccag caatataagt aaaaaactag t 1181

<210> 794

<211> 759

<212> DNA

<213> Artificial Sequence

<220><223> *Saccharomyces cerevisiae* ILV5 terminator

<400>

794

```
ggccctgcag gcctatcaag tgctggaaac tttttctctt ggaatttttg caacatcaag      60
tcatagtcaa ttgaattgac ccaatttcac atttaagatt tttttttttt catccgacat      120
acatctgtac actaggaagc cctgtttttc tgaagcagct tcaaatatat atatttttta      180
catatttatt atgattcaat gaacaatcta attaaatcga aaacaagaac cgaaacgcga      240
ataaataatt tatttagatg gtgacaagtg tataagtcct catcgggaca gctacgattt      300
ctctttcggg tttggctgag ctactgggtg ctgtgacgca gcggcattag cgcggcgtta      360
tgagctacc tcgtggcctg aaagatggcg ggaataaagc ggaactaaaa attactgact      420
```

```
gagccatatt gaggtcaatt tgtcaactcg tcaagtcacg tttggtggac ggcccctttc      480
caacgaatcg tatatactaa catgcgcgcg ctctctatat acacatatac atatatatat      540
atatatatat gtgtgcgtgt atgtgtacac ctgtatttaa tttccttact cgcggggtttt      600
tcttttttct caattcttgg ctctctcttt ctcgagtata taatttttca ggtaaaattt      660
agtacgatag taaaatactt ctgcaactcg tcacatatac gtgtacataa tgtctgaacc      720
agctcaaaag aaacaaaagg ttgctaacaa ctctctaga                                759
```

<210> 795

<211> 643

<212> DNA

<213> Artificial Sequence

<220><223> *Saccharomyces cerevisiae* FBA1 promoter

<400> 795

```
gaaatgaata acaatactga cagtactaaa taattgccta cttggcttca catacgttgc      60
atacgtcgat atagataata atgataatga cagcaggatt atcgtaatat gtaatagttg      120
aaaatctcaa aaatgtgtgg gtcattacgt aaataatgat aggaatggga ttctttctatt      180
tttctttttt ccattctagc agccgtcggg aaaacgtggc atcctctctt tcgggctcaa      240
ttggagtcaac gctgccgtga gcatcctctc ttccatatac taacaactga gcacgtaacc      300
aatggaaaag catgagctta gcgttgctcc aaaaaagtat tggatgggta ataccatttg      360
```

tctgttctct tctgactttg actcctcaaa aaaaaaaaaa ctacaatcaa cagatcgctt 420
 caattacgcc ctcaaaaaa cttttttcct tcttcttcgc ccacgttaaa ttttatccct 480
 catgttgctt aacggatttc tgcacttgat ttattataaa aagacaaaga cataatactt 540
 ctctatcaat ttcagttatt gttcttcctt gcgttattct tctgttcttc tttttctttt 600
 gtcatatata accataacca agtaatacat attcaaatct aga 643

<210> 796

<211> 1188

<212> DNA

<213> Artificial Sequence

<220><223> Saccharomyces cerevisiae KARI ILV5

<400> 796

atgttgagaa ctcaagccgc cagattgac tgcaactccc gtgtcatcac tgctaagaga 60
 acctttgctt tggccaccgc tgctgctgct tacagcagac cagctgcccg tttcgtaag 120
 ccaatgatca ctaccctggg ttgaagcaa atcaacttcg gtggtactgt tgaaaccgtc 180
 tacgaaagag ctgactggcc aagagaaaag ttgttgact acttcaagaa cgacactttt 240
 gctttgatcg gttacggttc ccaaggttac ggtcaaggtt tgaacttgag agacaacggt 300
 ttgaacgtta tcatlgtgtt ccgtaagat ggtgttctt ggaaggctgc catcgaagac 360
 gggtgggttc caggcaagaa cttgttcact gttgaagatg ctatcaagag aggtagttag 420

gttatgaact tgttgtccga tgccgtcaa tcagaaacct ggctgtctat caagccattg 480
 ttgaccaagg gtaagacttt gtacttctcc cacggtttct cccagttctt caaggacttg 540
 actcagttg aaccacaaa ggacttagat gttatcttgg ttgctccaaa gggttccggt 600
 agaactgtca gatctttgtt caaggaaggt cgtggtatta actcttctta cgccgtctgg 660
 aacgatgtca ccggtaaagg tcacgaaaag gcccaagctt tggccgttgc cattggttcc 720
 ggttacgttt accaaaccac ttctgaaaga gaagtcaact ctgacttgta cggtgaaaga 780
 gggtgtttta tgggtgggtat ccacggtatg ttcttggtc aatacgacgt cttagagaaa 840

aacggtcact ccccatctga agctttcaac gaaaccgtcg aagaagctac ccaatctcta 900
 taccattga tcgtaagta cggtatggat tacatgtacg atgcttggtc caccaccgcc 960
 agaagagggt ctttgactg gtaccaatc ttcaagaatg ctttgaagcc tgttttccaa 1020
 gacttgtagc aatctacaa gaacggtacc gaaaccaaga gatctttgga attcaactct 1080
 caacctgact acagagaaaa gctagaaaag gaattagaca ccatcagaaa catggaaatc 1140
 tggaagggtg gtaaggaagt cagaaagttg agaccagaaa accaataa 1188

<210> 797

<211> 395

<212> PRT

<213> Artificial Sequence

<220><223> *Saccharomyces cerevisiae* KARI coding region

<400> 797

Met Leu Arg Thr Gln Ala Ala Arg Leu Ile Cys Asn Ser Arg Val Ile

1 5 10 15

Thr Ala Lys Arg Thr Phe Ala Leu Ala Thr Arg Ala Ala Ala Tyr Ser

20 25 30

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

35 40 45

Lys Gln Ile Asn Phe Gly Gly Thr Val Glu Thr Val Tyr Glu Arg Ala

50 55 60

Asp Trp Pro Arg Glu Lys Leu Leu Asp Tyr Phe Lys Asn Asp Thr Phe

65 70 75 80

Ala Leu Ile Gly Tyr Gly Ser Gln Gly Tyr Gly Gln Gly Leu Asn Leu

85 90 95

Arg Asp Asn Gly Leu Asn Val Ile Ile Gly Val Arg Lys Asp Gly Ala

100 105 110

Ser Trp Lys Ala Ala Ile Glu Asp Gly Trp Val Pro Gly Lys Asn Leu

115 120 125

Phe Thr Val Glu Asp Ala Ile Lys Arg Gly Ser Tyr Val Met Asn Leu

130 135 140

Leu Ser Asp Ala Ala Gln Ser Glu Thr Trp Pro Ala Ile Lys Pro Leu

145 150 155 160

Leu Thr Lys Gly Lys Thr Leu Tyr Phe Ser His Gly Phe Ser Pro Val

165 170 175

Phe Lys Asp Leu Thr His Val Glu Pro Pro Lys Asp Leu Asp Val Ile

180 185 190

Leu Val Ala Pro Lys Gly Ser Gly Arg Thr Val Arg Ser Leu Phe Lys

195 200 205

Glu Gly Arg Gly Ile Asn Ser Ser Tyr Ala Val Trp Asn Asp Val Thr
 210 215 220
 Gly Lys Ala His Glu Lys Ala Gln Ala Leu Ala Val Ala Ile Gly Ser
 225 230 235 240
 Gly Tyr Val Tyr Gln Thr Thr Phe Glu Arg Glu Val Asn Ser Asp Leu
 245 250 255
 Tyr Gly Glu Arg Gly Cys Leu Met Gly Gly Ile His Gly Met Phe Leu
 260 265 270
 Ala Gln Tyr Asp Val Leu Arg Glu Asn Gly His Ser Pro Ser Glu Ala
 275 280 285
 Phe Asn Glu Thr Val Glu Glu Ala Thr Gln Ser Leu Tyr Pro Leu Ile
 290 295 300
 Gly Lys Tyr Gly Met Asp Tyr Met Tyr Asp Ala Cys Ser Thr Thr Ala
 305 310 315 320
 Arg Arg Gly Ala Leu Asp Trp Tyr Pro Ile Phe Lys Asn Ala Leu Lys
 325 330 335
 Pro Val Phe Gln Asp Leu Tyr Glu Ser Thr Lys Asn Gly Thr Glu Thr
 340 345 350
 Lys Arg Ser Leu Glu Phe Asn Ser Gln Pro Asp Tyr Arg Glu Lys Leu
 355 360 365
 Glu Lys Glu Leu Asp Thr Ile Arg Asn Met Glu Ile Trp Lys Val Gly

370 375 380
 Lys Glu Val Arg Lys Leu Arg Pro Glu Asn Gln
 385 390 395

<210> 798

<211> 244

<212> DNA

<213> Artificial Sequence

<220><223> Saccharomyces cerevisiae CYC1 terminator

<400> 798

attaaagcct tcgagcgccc caaaaccttc tcaagcaagg ttttcagtat aatgttaccat 60

gcgtacacgc gtctgtacag aaaaaaaga aaaatttgaa atataaataa cgttcttaat 120

actaacataa ctataaaaa ataaataggg acctagactt caggttgtct aactccttcc 180

ttttcggtta gagcggatgt ggggggaggg cgtgaatgta agcgtgacat aactaattac 240

atga 244

<210> 799

<211> 1017

<212> DNA

<213> Artificial Sequence

<220><223> Pf5.IlvC-JEA1 variant

<400> 799

atgaaagttt tctacgataa agactgcgac ctgtcgatca tccaaggtaa gaaagttgcc 60

atcatcggtt tcggttccca gggccacgct caagcactca acctgaagga ttccggcgta 120

gacgtgactg ttggcctgcc taaaggcttt gctgatgtag ccaaggctga agcccacggc 180

tttaaagtga ccgacgttgc tgcagccgtt gccggtgccg acttgggtcat gatcctgatt 240

ccggacgagt tccagtccca gctgtacaag aacgaaatcg agccgaacat caagaagggc 300

gccactctgg ccttctccca cggcttcgcg atccactaca accaggttgt gcctcgtgcc 360

gacctcgacg tgatcatgat cgcgccgaag gctccaggcc acaccgtacg ttccgagttc 420

gtcaaggcgc gaggtatttc tgacctgacg gcgatctacc aggacgtttc cggcaacgcc 480

aagaacgtcg cctgtccta cgccgcaggc gtgggcggcg gccgtaccgg catcatcgaa 540

accacettca aggacgagac tgaaaccgac ctgttcggtg agcaggctgt tctgtgtggc 600

ggtaccgtcg agctgggtcaa agccgggttc gaaaccctgg ttgaagctgg ctacgtcca 660

gaaatggcct acttcgagtg cctgcacgaa ctgaagctga tcgttgacct catgtacgaa 720

ggcggtatcg ccaacatgaa ctactcgatc tccaacaacg ctgaatacgg cgagtacgtg 780

actggtccag aagtcacaa cgccgaatcc cgtcaggcca tgcgcaatgc tctgaagcgc 840

atccaggacg gcgaatacgc gaagatgttc atcagcgaag gcgctaccgg ctaccatcg 900

atgaccgcca agcgtcgtaa caacgtgct caccgtatcg aaatcatcgg cgagcaactg 960

cgctcgatga tgccttgat cggtgccaac aaaatcgtcg acaaagcaa gaactaa 1017

<210> 800

<211> 338

<212> PRT

<213> Artificial Sequence

<220><223> Pf5.IlvC-JEA1 variant

<400> 800

```

Met Lys Val Phe Tyr Asp Lys Asp Cys Asp Leu Ser Ile Ile Gln Gly
1           5           10           15
Lys Lys Val Ala Ile Ile Gly Phe Gly Ser Gln Gly His Ala Gln Ala
           20           25           30
Leu Asn Leu Lys Asp Ser Gly Val Asp Val Thr Val Gly Leu Pro Lys
           35           40           45
Gly Phe Ala Asp Val Ala Lys Ala Glu Ala His Gly Phe Lys Val Thr

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

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

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

```

Gly Gly Ile Ala Asn Met Asn Tyr Ser Ile Ser Asn Asn Ala Glu Tyr

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

305 310 315 320
Arg Ser Met Met Pro Trp Ile Gly Ala Asn Lys Ile Val Asp Lys Ala
325 330 335
Lys Asn

<210> 801

<211> 1644

<212> DNA

<213> Artificial Sequence

<220><223> codon optimized sequence

<400> 801

atgtatacag taggtgacta tctgttggac agattacacg aattaggtat agaagaaata 60
ttcggagtac caggtgacta caatttgcaa ttctagatc aaattatttc acacaaagat 120
atgaaatggg tgggaaatgc taatgagtta aatgcctcct atatggccga cgggtacgca 180

agaacgaaaa aggtgcggc attcttgact acatttgggtg ttggcgaatt atccgcagtt 240
aatggccttag cgggctccta tgctgagaac ctgcctgttg ttgagatcgt gggatctcct 300
acctcgaaag tgcagaacga aggtaagttt gttcaccata cgttggctga tggatgatttc 360
aagcacttta tgaagatgca cgaaccggtt actgctgcca ggactttatt gacagccgag 420
aatgcaactg ttgaaattga tagagtgttg tctgccttac taaaggaaag aaagccggtt 480
tacatcaatt tacctgtaga tgtagctgcc gctaaggctg aaaaaccatc ctgcctctt 540
aagaaggaaa attccacgtc gaatacatct gatcaagaga ttctgaacaa aatacaggaa 600

agtctgaaga atgccaagaa accaattgta atcacaggcc atgaaattat atcgttcggc 660

ctagagaaga ctgttactca gtttatttca aagactaagt tacctattac tactttgaac 720
 ttgtgtaaat catctgttga tgaagcattg ccttcatttt tggggattta caacggtact 780
 ctgtcagagc caaacttgaa ggaatttgtg gaatctgctg attttattct tatgttgggt 840
 gtaaagctta ccgattctag tacgggtgca ttactcacc atcttaatga aaataaaatg 900
 atttccttga atatcgatga aggtaaaatt ttcaacgaaa gaatccaaaa tticgacttc 960
 gaatccctga tatcatctct tcttgacttg tccgaaattg aataaaagg caagtacata 1020

gataaaaagc aagaagattt tgtaccttct aacgcgctgt tgtcacaaga tagactgtgg 1080
 caagctgtcg aaaatttgac ccaaagtaat gagacgatcg tggctgaaca aggcacttct 1140
 ttcttcggtg cctcatctat atttctgaaa tcgaaatcac attttattgg tcaacccttg 1200
 tggggatcta taggatacac ttccccgca gctctaggca gccaaattgc agataaagaa 1260
 tctagacatt tattgtttat cggagatgga tcattgcaac tgactgtcca agaattagga 1320
 ctagccatta gagagaagat aaaccaatc tgctttatca ttaataacga tggttacacg 1380
 gttgagaggg aaattcatgg tccgaaccag agttataatg acattcctat gtggaattac 1440

tcaaaactgc cagaaagttt cggggcaacg gaagacagag ttgtgtccaa aattgtgaga 1500
 acagaaaatg aattcgtatc cgtgatgaaa gaagctcaag cagatccaaa taggatgtat 1560
 tggatagaac ttattctagc aaaggagggt gcacctaaag ttttgaaaaa gatgggtaag 1620
 ttatttgcag aacaaaacaa gaggc 1644

<210> 802

<211> 1125

<212> DNA

<213> Artificial Sequence

<220><223> codon optimized sequence

<400> 802

atgtcaacag ccggtaaagt tattaagtgt aaagcggcag ttttgtggga agagaaaaag 60
 ccgtttagca tagaagaagt agaagtagcg ccacaaaag cacacgaggt tagaatcaag 120

atggttgcca ccggaatctg tagatccgac gaccatgtgg tgagtggcac tctagttact 180
 cctttgccag taatcgcgagg acacagggt gccggaatcg ttgaatccat aggtgaaggt 240
 gttaccactg ttcgtcctgg tgataaagt atccactgt tcaactctca atgtggtgtaag 300
 ttagagtct gcaaacatcc tgagggtaat ttctgcctta aaaatgattt gtctatgcct 360
 agaggtacta tgcaggatgg tacaagcaga ttacatgca gagggaaacc tataccat 420
 ttcttggta ctctacatt ttccaatac acagtgggtg acgagatac tgctgctaaa 480

atcgatgcag cttcaccact ggaaaaagt ttgcttgatag ggtgcggatt ttccaccggt 540

tacggttccg cagttaaagt tgcaaagggt acacagggtt cgacttgtgc agtattcgggt 600

ttaggaggag taggactaag cgttattatg ggggtgtaaag ctgcaggcgc agcaggaggt 660

ataggtgtag acatcaataa ggacaaattt gcaaaagcta aggaggtcgg ggctactgaa 720

tgtgttaacc ctcaagatta taagaaacca atacaagaag tccttactga aatgtcaaac 780

ggaggagttg atttctcttt tgaagttata ggccgtcttg atactatggt aactgcgttg 840

tcctgctgtc aagaggcata tggagtcagt gtgatcgtag gtgttcctcc tgattcacia 900

aatttgctga tgaatcctat gctgttgcta agcggtcgta catggaaggg agctatatatt 960

ggcggtttta agagcaagga tagtgttcca aaacttggtt cgcactttat ggcgaagaag 1020

tttgctcttg atcctttaat tacacatgta ttgccattcg agaaaatcaa tgaagggttt 1080

gatttggtta gaagtgggtga atctattcgt acaattttaa ctttt 1125

<210> 803

<211> 375

<212> PRT

<213> Artificial Sequence

<220><223> Equus caballus

<400> 803

Met Ser Thr Ala Gly Lys Val Ile Lys Cys Lys Ala Ala Val Leu Trp

1 5 10 15

Glu Glu Lys Lys Pro Phe Ser Ile Glu Glu Val Glu Val Ala Pro Pro

20 25 30

Lys Ala His Glu Val Arg Ile Lys Met Val Ala Thr Gly Ile Cys Arg

35 40 45

Ser Asp Asp His Val Val Ser Gly Thr Leu Val Thr Pro Leu Pro Val

50 55 60

Ile Ala Gly His Glu Ala Ala Gly Ile Val Glu Ser Ile Gly Glu Gly

65 70 75 80

Val Thr Thr Val Arg Pro Gly Asp Lys Val Ile Pro Leu Phe Thr Pro

85 90 95

Gln Cys Gly Lys Cys Arg Val Cys Lys His Pro Glu Gly Asn Phe Cys

100 105 110

Leu Lys Asn Asp Leu Ser Met Pro Arg Gly Thr Met Gln Asp Gly Thr
 115 120 125
 Ser Arg Phe Thr Cys Arg Gly Lys Pro Ile His His Phe Leu Gly Thr
 130 135 140
 Ser Thr Phe Ser Gln Tyr Thr Val Val Asp Glu Ile Ser Val Ala Lys
 145 150 155 160
 Ile Asp Ala Ala Ser Pro Leu Glu Lys Val Cys Leu Ile Gly Cys Gly
 165 170 175
 Phe Ser Thr Gly Tyr Gly Ser Ala Val Lys Val Ala Lys Val Thr Gln
 180 185 190
 Gly Ser Thr Cys Ala Val Phe Gly Leu Gly Gly Val Gly Leu Ser Val
 195 200 205
 Ile Met Gly Cys Lys Ala Ala Gly Ala Ala Arg Ile Ile Gly Val Asp
 210 215 220
 Ile Asn Lys Asp Lys Phe Ala Lys Ala Lys Glu Val Gly Ala Thr Glu
 225 230 235 240
 Cys Val Asn Pro Gln Asp Tyr Lys Lys Pro Ile Gln Glu Val Leu Thr
 245 250 255
 Glu Met Ser Asn Gly Gly Val Asp Phe Ser Phe Glu Val Ile Gly Arg
 260 265 270
 Leu Asp Thr Met Val Thr Ala Leu Ser Cys Cys Gln Glu Ala Tyr Gly
 275 280 285
 Val Ser Val Ile Val Gly Val Pro Pro Asp Ser Gln Asn Leu Ser Met
 290 295 300
 Asn Pro Met Leu Leu Leu Ser Gly Arg Thr Trp Lys Gly Ala Ile Phe
 305 310 315 320
 Gly Gly Phe Lys Ser Lys Asp Ser Val Pro Lys Leu Val Ala Asp Phe
 325 330 335
 Met Ala Lys Lys Phe Ala Leu Asp Pro Leu Ile Thr His Val Leu Pro
 340 345 350
 Phe Glu Lys Ile Asn Glu Gly Phe Asp Leu Leu Arg Ser Gly Glu Ser

355 360 365
 Ile Arg Thr Ile Leu Thr Phe
 370 375
 <210> 804
 <211> 548
 <212> PRT
 <213> Artificial Sequence
 <220><223> Lactococcus lactis
 <400> 804
 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 His Lys Asp Met Lys Trp Val 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 Val
 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 Glu 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 Val
 130 135 140
 Glu Ile Asp Arg Val Leu Ser Ala 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
 Ser Leu Pro Leu Lys Lys Glu Asn Ser Thr Ser Asn Thr Ser Asp Gln

180 185 190
 Glu Ile Leu Asn Lys Ile Gln Glu Ser Leu Lys Asn Ala Lys Lys Pro
 195 200 205
 Ile Val Ile Thr Gly His Glu Ile Ile Ser Phe Gly Leu Glu Lys Thr

 210 215 220
 Val Thr Gln Phe Ile Ser Lys Thr Lys Leu Pro Ile Thr Thr Leu Asn
 225 230 235 240
 Phe Gly Lys Ser Ser Val Asp Glu Ala Leu Pro Ser Phe Leu Gly Ile
 245 250 255
 Tyr Asn Gly Thr Leu Ser Glu Pro Asn Leu Lys Glu 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 Asn Glu Asn Lys Met Ile Ser Leu Asn
 290 295 300
 Ile Asp Glu Gly Lys Ile Phe Asn Glu Arg Ile Gln Asn Phe Asp Phe
 305 310 315 320
 Glu Ser Leu Ile Ser Ser Leu Leu Asp Leu Ser Glu Ile Glu Tyr Lys
 325 330 335
 Gly Lys Tyr Ile Asp Lys Lys Gln Glu Asp Phe Val Pro Ser Asn Ala

 340 345 350
 Leu Leu Ser Gln Asp Arg Leu Trp Gln Ala Val Glu Asn 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 Ser Ile Phe Leu Lys Ser Lys Ser His 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 Ala Ile Arg Glu Lys Ile 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 Asn Gln Ser Tyr Asn Asp Ile Pro Met Trp Asn Tyr

465 470 475 480

Ser Lys Leu Pro Glu Ser 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 Pro Asn Arg Met Tyr Trp Ile Glu Leu Ile Leu Ala Lys

515 520 525

Glu Gly Ala Pro Lys Val Leu Lys Lys Met Gly Lys Leu Phe Ala Glu

530 535 540

Gln Asn Lys Ser

545

<210> 805

<211> 1023

<212> DNA

<213> Artificial Sequence

<220><223> Saccharomyces cerevisiae TDH3 promoter

<400> 805

caccgcggtg gggcgcgccc tattttcgag gaccttgta ccttgagccc aagagagcca 60

agattttaat tttcctaaga cttgatgcaa attcccaaag ctaataacat gcaagacacg 120

tacggtaag aagacatatt tgacctctta acaggttcag acgcgactgc ctcacagta 180

agaccggtg aaaagaactt acctgaaaaa aacgaatata tactagcggt gaatgttagc 240

gtcaacaaca agaagtttaa tgacgcggag gccaggcaa aaagattcct tgattacgta 300

aggaggttag aatcattttg aataaaaaac acgctttttc agttcgagtt tatcattatc 360

aatactgcca tttcaaagaa tacgtaaata attaatagta gtgattttcc taactttatt 420

tagtcaaaaa attagccttt taattctgct gtaaccgta catgcccaaa atagggggcg 480

ggttacacag aatatataac atcgtagggtg tctgggtgaa cagttttatt ctggcatcca 540

ctaaatataa tggagcccg cttttaagct ggcatccaga aaaaaaaga atcccagcac 600

caaaatattg ttttcttcac caaccatcag ttcataggtc cattctctta gcgcaactac 660

agagaacagg ggcacaaaca ggcaaaaaac gggcacaacc tcaatggagt gatgcaacct 720

gcctggagta aatgatgaca caaggcaatt gaccacgca tgtatctatc tcattttctt 780

acaccttcta ttaccttctg ctctctctga ttggaaaaa gctgaaaaa aaggttgaaa 840

ccagttccct gaaattatc cctacttga ctaataagta tataaagacg gtaggtattg 900

attgtaattc tgtaaacta tttcttaaac ttcttaaatt ctactttat agttagtctt 960

tttttaggtt taaaacacc aagaacttag ttccaataa acacacataa actagtaaac 1020

aaa 1023

<210> 806

<211> 21

<212> DNA

<213> Artificial Sequence

<220><223> Primer OT1068

<400> 806

caaaagctga gctccaccgc g 21

<210> 807

<211> 44

<212> DNA

<213> Artificial Sequence

<220><223> Primer OT1067

<400> 807

gtttactagt ttatgtgtgt ttattcgaaa ctaagttctt ggtg 44

<210> 808

<211> 9333

<212> DNA

<213> Artificial Sequence

<220><223> pLH467

<400> 808

tcgcgcgttt cggtgatgac ggtgaaaacc tctgacacat gcagctcccg gagacggtca 60

cagcttgtct gtaagcggat gccgggagca gacaagcccg tcagggcgcg tcagcgggtg 120

ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180

accataccac agcttttcaa ttcaattcat cttttttttt ttattctttt ttttgatttc 240

ggtttctttg aaatTTTTTT gattcggtaa tctccgaaca gaaggaagaa cgaaggaagg 300
 agcacagact tagattggtg tatatacgca tatgtagtgt tgaagaaaca tgaattgcc 360
 cagtattctt aacccaactg cacagaacaa aaacctgcag gaaacgaaga taaatcatgt 420
 cgaaagctac atataaggaa cgtgctgcta ctcatcctag tcctgttgct gccaaagctat 480

 ttaatatcat gcacgaaaag caaacaact tgtgtgcttc attggatgtt cgtaccacca 540
 aggaattact ggagttagtt gaagcattag gtcccaaat ttgtttacta aaaacacatg 600
 tggatatctt gactgatttt tccatggagg gcacagttaa gccgctaaag gcattatccg 660
 ccaagtacaa ttttttactc ttccaagaca gaaaatttgc tgacattggt aatacagtca 720
 aattgcagta ctctgcggtt gtatacagaa tagcagaatg ggacagacatt acgaatgcac 780
 acggtgtggt gggcccaggt attgttagcg gtttgaagca ggcggcagaa gaagtaacaa 840
 aggaacctag aggccttttg atgttagcag aattgtcatg caagggtccc ctatctactg 900

 gagaatatac taagggtact gttgacattg cgaagagcga caaagatttt gttatcggct 960
 ttattgctca aagagacatg ggtggaagag atgaaggta cgattggttg attatgacac 1020
 ccggtgtggg tttagatgac aaggagacg cattgggtca acagtataga accgtggatg 1080
 atgtggtctc tacaggatct gacattatta ttgttgaag aggactattt gcaaaggga 1140
 gggatgctaa ggtagagggt gaacgttaca gaaaagcagg ctgggaagca tatttgagaa 1200
 gatgcggcca gcaaaactaa aaaactgtat tataagtaaa tgcatgtata ctaaactcac 1260
 aaattagagc ttcaatttaa ttatatcagt tattacccta tgcggtgtga aataccgcac 1320

 agatgcgtaa ggagaaaata ccgcatcagg aaattgtaaa cgtaaatatt ttgttaaaat 1380
 tcgcgttaaa tttttgttaa atcagctcat tttttaacca ataggccgaa atcgcaaaa 1440
 tccctataa atcaaaagaa tagaccgaga tagggttgag tgttgttcca gtttgaaca 1500
 agagtccact attaaagaac gtggactcca acgtcaaagg gcgaaaaacc gtctatcagg 1560
 gcgatggccc actacgtgaa ccatcacctt aatcaagttt tttgggtcgc aggtgccgta 1620
 aagcactaaa tcggaacctt aaaggagacc ccgatttag agcttgacgg ggaagccgg 1680
 cgaacgtggc gagaaaggaa gggaagaaag cgaaaggagc gggcgctagg gcgctggcaa 1740

 gtgtagcgtt cacgtgcgc gtaaccacca caccgccgc gcttaatgcg ccgtacagg 1800
 gcgctgcgc ccattcgcca ttacggctgc gcaactgttg ggaagggcga tcggtgcggg 1860
 cctcttcgtt attacgccag ctggcgaaag ggggatgtgc tgcaaggcga ttaagtggg 1920
 taacgccagg gttttccag tcacgacgtt gtaaacgac ggccagttag gcgctgaat 1980
 acgactcact ataggcgaa ttgggtaccg gggccccct cgaggtcgac tggccattaa 2040
 tctttcccat attagatttc gccaaagcat gaaagttaa gaaaggtctt tagacgaatt 2100

acccttcatt tcicaaactg gcgtcaagg atcctggat ggtttatcg tttatttct 2160

ggttcttata gcatcgtttt ggacttctct gtteccatta ggcggttcag gagccagcgc 2220

agaatcattc ttgaaggat acttatcctt tccaattttg attgtctgtt acgttggaca 2280

taaactgtat actagaaatt ggactttgat ggtgaaacta gaagatatgg atcttgatac 2340

cggcagaaaa caagtagatt tgactcttcg tagggaagaa atgaggattg agcgagaaac 2400

attagcaaaa agatccttcg taacaagatt ttacatttc tgggtttgaa gggaaagata 2460

tgagctatac agcgggaattt ccatatcact cagattttgt tatctaattt tttccttccc 2520

acgtccgcgg gaatctgtgt atattactgc atctagatat atgttatctt atcttggcgc 2580

gtacatttaa tttcaacgt attctataag aaattgcggg agtttttttc atgtagatga 2640

tactgactgc acgcaaatat aggcatgatt tataggcatg atttgatggc tgtaccgata 2700

ggaacgctaa gagtaacttc agaatcgtaa tctggcgga aaaaattcat ttgtaaactt 2760

taaaaaaaa agccaatatc cccaaaatta ttaagagcgc ctccattatt aactaaaatt 2820

tcactcagca tccacaatgt atcaggtatc tactacagat attacatgtg gcgaaaaaga 2880

caagaacaat gcaatagcgc atcaagaaaa aacacaaagc tttcaatcaa tgaatcgaaa 2940

atgtcattaa aatagtatat aaattgaaac taagtcataa agctataaaa agaaaattta 3000

tttaaatgca agatttaaag taaattcacg gccctgcagg cctcagctct tgtttgttc 3060

tgcaaataac ttacccatct tttcaaaac tttaggtgca ccttctttg ctagaataag 3120

ttctatcaa tacatcctat ttggatctgc ttgagcttct ttcatcacgg atacgaattc 3180

attttctgtt ctcaaatit tggacacaac tctgtcttc gttgccccga aactttctgg 3240

cagttttgag taattccaca taggaatgtc attataactc tggttcggac catgaatttc 3300

ccttcaacc gtgtaacct cgttattaat gataaagcag attgggttta tcttctctct 3360

aatggctagt cctaattctt ggacagtcag ttgcaatgat ccatctccga taaacaataa 3420

atgtctagat tctttatctg caatttggct gcctagagct gcggggaaag tgtatcctat 3480

agatccccac aagggttgac caataaaatg tgatttcgat ttcagaaata tagatgaggc 3540

accgaagaaa gaagtgcctt gttcagccac gatcgtctca ttactttggg tcaaattttc 3600

gacagcttg cactgtctat cttgtgacaa cagcgcgtta gaaggtacaa aatcttcttg 3660

ctttttatct atgtacttg ctttatattc aatttcggac aagtcaagaa gagatgatat 3720

cagggttcg aagtcgaaat ttggattct ttcgttgaaa attttacctt catcgatatt 3780

caaggaaatc atttatttt cattaagatg gtgagtaaat gcaccgtac tagaatcggc 3840

aagctttaca cccaacataa gaataaaaac agcagattcc acaaattcct tcaagtttgg	3900
ctctgacaga gtaccgttgt aaatccccaa aaatgagggc aatgcttcat caacagatga	3960
tttaccaaag ttcaaagtag taataggtaa cttagtcttt gaaataaact gagtaacagt	4020
cttctctagg ccgaacgata taatttcatg gcctgtgatt acaattgggt tcttggcatt	4080
cttcagactt tctgtatatt tgttcagaat ctcttgatca gatgtattcg acgtggaatt	4140
ttccttctta agaggcaagg atggtttttc agccttagcg gcagctacat ctacaggtaa	4200
attgatgtaa accggctttc ttctcttttag taaggcagac aacactctat caatttcaac	4260
agttgcattc tcggctgtca ataaagtcct ggcagcagta accggttcgt gcattctcat	4320
aaagtgcttg aaatcacat cagccaacgt atggatgaaca aacttacctt cgttctgcac	4380
tttcgaggta ggagatccca cgatctcaac aacaggcagg ttctcagcat aggagcccgc	4440
taagccatta actgcggata attcgccaac accaaatgta gtcaagaatg ccgcagcctt	4500
tttcgttctt gcgtaccgt cgccatata ggaggcattt aactcattag catttcccac	4560
ccatttcata tctttgtgtg aaataatttg atctagaaat tgcaaattgt agtcacctgg	4620
tactccgaat atttcttcta tacctaattc gtgtaatctg tccaacagat agtcacctac	4680
tgtatacatt ttgtttacta gtttatgtgt gtttattcga aactaagttc ttggtgtttt	4740
aaaactaaaa aaaagactaa ctataaaagt agaatttaag aagttaaga aatagattta	4800
cagaattaca atcaatacct accgtcttta tatacttatt agtcaagtag gggaataatt	4860
tcagggaact ggittcaacc ttttttttca gctttttcca aatcagagag agcagaaggt	4920
aatagaaggt gtaagaaaat gagatagata catgcgtggg tcaattgcct tgtgtcatca	4980
tttactccag gcaggttgca tcactccatt gaggttgtgc ccgttttttg cctgtttgtg	5040
cccctgttct ctgtagtgc gctaagagaa tggacctatg aactgatggt tggatgaaga	5100
aacaatattt tgggtgtggg attctttttt ttcttgatg ccagcttaaa aagcgggctc	5160
cattatattt agtggatgcc aggaataaac tgttcacca gacacctacg atgttatata	5220
ttctgtgtaa cccgcccctt attttgggca tgtacgggtt acagcagaat taaaaggcta	5280
attttttgac taaataaagt taggaaaac actactatta attatttacg tattctttga	5340
aatggcagta ttgataatga taaactcgaa ctgaaaaagc gtgtttttta ttcaaatga	5400
ttctaactcc cttacgtaat caaggaatct tttgccttg gcctccgct cattaaactt	5460
cttgttgttg acgtaacat tcaacgctag tatatatctg tttttttcag gtaagttctt	5520
ttcaacgggt ctactgatg aggcagtcgc gtctgaacct gttaagaggt caaatatgtc	5580
ttcttgaccg tacgtgtctt gcattgtatt agctttggga atttgcacat agtcatagga	5640
aaatttaaat ctggctctc ttgggtctca ggtgacaagg tcctcgaaaa tagggcgcg	5700

cccaccgcgg tggagctcca gcttttgttc cctttagtga gggttaattg cgcgcttggc 5760
 gtaatcatgg tcatagctgt ttctgtgtg aaattgttat ccgctcacia ttccacacia 5820
 cataggagcc ggaagcataa agtgtaaagc ctgggggtgcc taatgagtga ggtaactcac 5880
 attaattgcg ttgcgctcac tgcccgcctt ccagtcggga aacctgtcgt gccagctgca 5940

 ttaatgaatc ggccaacgcg cggggagagg cggtttgcgt attgggcgct ctccgcttc 6000
 ctgctcact gactcgctgc gctcggtcgt tcggctgcgg cgagcggat cagctcactc 6060
 aaaggcggta atacggttat ccacagaatc aggggataac gcaggaaaga acatgtgagc 6120
 aaaaggccag caaaaggcca ggaaccgtaa aaaggccgcg ttgctggcgt ttttccatag 6180
 gctccgccc cctgacgagc atcacaaaaa tcgacgctca agtcagaggt ggcgaaaccc 6240
 gacaggacta taaagatacc aggcgtttcc ccttggaagc tccctcgtgc gctctcctgt 6300
 tccgacctg ccgcttaccg gatacctgtc cgcctttctc ccttcgggaa gcgtggcgt 6360

 ttctcatagc tcacgctgta ggtatctcag ttccgtgtag gtcgttcgct ccaagctggg 6420
 ctgtgtgcac gaacccccg ttccagccga ccgctgcgc ttatccgta actatcgtct 6480
 tgagtccaac ccgtaagac acgacttacc gccactggca gcagccactg gtaacaggat 6540
 tagcagagcg aggtatgtag gcggtgctac agagtctctg aagtgggtgc ctaactacgg 6600
 ctacactaga aggacagtat ttggtatctg cgtctcgtg aagccagtta ccttcggaaa 6660
 aagagtgggt agctcttgat ccggcaaaaa aaccaccgct ggtagcgggt gttttttgt 6720
 ttgcaagcag cagattaccg gcagaaaaaa aggatctcaa gaagatcctt tgatctttc 6780

 tacggggtct gacgctcagt ggaacgaaaa ctacgctaa gggattttgg tcatgagatt 6840
 atcaaaaagg atcttcacct agatcctttt aaattaaaaa tgaagtttta aatcaatcta 6900
 aagtatatat gagliaaactt ggtctgacag ttaccaatgc ttaatcagt aggcacctat 6960
 ctccagcgtc tctctatttc gttcatccat agttgcctga ctcccgctg ttagataac 7020
 tacgatcagg gagggcttac catctggccc cagtgtgca atgataccg gagaccacg 7080
 ctaccggct ccagatttat cagcaataaa ccagccagcc ggaaggccg agcgcagaag 7140
 tggctctgca actttatccg cctccatcca gtctattaat tgttgccggg aagctagagt 7200

 aagtagttcg ccagttaata gtttgcgcaa cgttgttgc attgtctacg gcatcgttgt 7260
 gtacgctcg tcgttttgta tggcttcatt cagctccggt tccaacgat caaggcgagt 7320
 tacatgatcc ccatgttgt gcaaaaaagc ggttagctcc ttcggtcctc cgtcgttgt 7380
 cagaagtaag ttggccgag tgttatcact catggttatg gcagcactgc ataattctct 7440
 tactgtcatg ccatccgtaa gatcttttc tgtgactggg gagtactcaa ccaagtcatt 7500
 ctgagaatag tglatgcggc gaccgagttg ctcttgccc gcgtcaatc ggataatac 7560

cgcgccacat agcagaactt taaaagtgt catcattgga aaacgttctt cggggcgaaa 7620

actctcaagg atcttaccgc tgttgagatc cagttcgatg taaccactc gtgcacccaa 7680

ctgatcttca gcatctttta ctttcaccag cgtttctggg tgagcaaaaa caggaaggca 7740

aatgccgca aaaaaggga taaggcgac acggaaatgt tgaatactca tactcttctt 7800

ttttcaatat tattgaagca tttatcaggg ttattgtctc atgagcggat acatatttga 7860

atgtatttag aaaaataaac aaataggggt tccgcgcaca tttccccgaa aagtgccacc 7920

tgaacgaagc atctgtgctt cattttgtag acaaaaaatg caacgcgaga gcgctaattt 7980

ttcaacaaa gaatctgagc tgcattttta cagaacagaa atgcaacgcg aaagcgctat 8040

ttaccaacg aagaatctgt gcttcatttt tgtaaaacaa aaatgcaacg cgagagcgct 8100

aatttttcaa acaaagaatc tgagctgcat ttttacagaa cagaaatgca acgcgagagc 8160

gctattttac caacaaagaa tctatacttc tttttgttc tacaaaaatg catcccgaga 8220

gcgctatttt tctaacaag catcttagat tacttttttt ctctttgtg cgctctataa 8280

tgcagtctct tgataacttt ttgcaactga ggtccgttaa ggtagaaga aggctacttt 8340

ggtgtctatt ttctcttcca taaaaaagc ctgactccac ttcgccggtt tactgattac 8400

tagcgaagct gcgggtgcat ttttcaaga taaaggcatc cccgattata ttctataccg 8460

atgtggattg cgcatacttt gtgaacagaa agtgatagcg ttgatgattc ttcattggtc 8520

agaaaattat gaacggtttc ttctattttg tctctatata ctacgtatag gaaatgttta 8580

cattttcgta ttgttttca ttactctat gaatagtctt tactacaatt tttttgtcta 8640

aagagtaata ctagagataa acataaaaaa tftagaggtc gagtttagat gcaagttcaa 8700

ggagcgaaag gtggatgggt aggttatata gggatatagc acagagatat atagcaaaga 8760

gatacttttg agcaatgttt gtggaagcgg tattcgcaat attttagtag ctcgttacag 8820

tccggtgcgt ttttggtttt ttgaaagtgc gtcttcagag cgcttttgggt ttcaaaaagc 8880

gctctgaagt tcctatactt tctagagaat aggaacttcg gaataggaac ttcaaagcgt 8940

ttccgaaaac gagcgcttcc gaaaatgcaa cgcgagctgc gcacatacag ctcactgttc 9000

acgtcgcacc tatatctgcg tgttgccctgt atatatatat acatgagaag aacggcatag 9060

tgcgtgttta tgcttaaatg cgtacttata tgcgtctatt tatgtaggat gaaaggtagt 9120

ctagtacctc ctgtgatatt atcccattcc atgcggggta tcgtatgctt ccttcagcac 9180

taccctttag ctgttctata tgctgccact cctcaattgg attagtctca tccttcaatg 9240

ctatcatttc ctttgatatt ggatcatact aagaaacat tattatcatg acattaacct 9300

ataaaaaatag gcgtatcacg aggccctttc gtc 9333

<210> 809

<211> 8994

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic construct pRS425::GPM-sadB

<400> 809

ctagtcttag agcgccgcc accgcggtgg agctccagct tttgttcct ttagtgaggg 60

ttaattgcgc gcttggcgta atcatggtca tagctgtttc ctgtgtgaaa ttgttatccg 120

ctcacaattc cacacaacat aggagccgga agcataaagt gtaaagcctg gggtgcctaa 180

tgagtgaggt aactcacatt aattgcgttg cgtcactgc ccgctttcca gtcgggaaac 240

ctgtcgtgcc agctgcatta atgaatcggc caacgcgcgg ggagaggcgg ttgctgatt 300

gggcgctctt ccgcttcctc gctcactgac tcgtcgcgt cggctgttcg gctgcggcga 360

gcggtatcag ctactcaaa ggccgtaata cgtttatcca cagaatcagg ggataacgca 420

ggaaagaaca tgtgagcaaa aggccagcaa aaggccagga accgtaaaaa ggccgcgttg 480

ctggcgtttt tccataggt ccgccccct gacgagcatc acaaaaatcg acgtcaagt 540

cagaggtggc gaaacccgac aggactataa agataaccagg cgtttccccc tggaagctcc 600

ctcgtgcgt ctctgttcc gacctgccg cttaccgat accgtccgc ctttctccct 660

tcgggaagcg tggcgctttc tcatagctca cgtgtaggt atctcagttc ggtgtaggtc 720

gttcgtcca agctgggctg tgtgcacgaa cccccgttc agcccgaccg ctgcgcctta 780

tccgtaact atcgtcttga gtccaaccg gtaagacacg acttatcgcc actggcagca 840

gccactggtg acaggattag cagagcgagg tatgtaggcg gtgctacaga gttcttgaag 900

tgggtgccta actacggcta cactagaagg acagtatttg gtatctgcgc tctgctgaag 960

ccagttacct tcggaaaaag agttggtagc tcttgatccg gcaaacaaac caccgtggt 1020

agcggtggtt tttttgttg caagcagcag attacgcgca gaaaaaaagg atctcaagaa 1080

gacctttga tcttttctac ggggtctgac gctcagtga acgaaaactc acgttaaggg 1140

attttggtca tgagattatc aaaaaggatc ttcacctaga tctttttaa ttaaaaatga 1200

agttttaaat caatctaaag tataatatgag taaacttggc ctgacagtta ccaatgctta 1260

atcagtgagg cacctatctc agcgatctgt ctatttcgtt catccatagt tgcctgactc 1320

cccgctgtgt agataactac gatacgggag ggttaccat ctggccccag tgctgcaatg 1380

ataccgcgag acccacgtc accggctcca gatttatcag caataaacca gccagccgga 1440

agggccgagc gcagaagtgg tcctgcaact ttatccgcct ccatccagtc tattaattgt 1500

tgccgggaag ctagagtaag tagttcgcca gttaatagtt tgcgcaacgt tgttgccatt 1560

gctacaggca tcgtgggtgc acgctcgctg tttggatatg cttcattcag ctccggttcc 1620

caacgatcaa ggcgagttac atgatcccc atgttgtgca aaaaagcggg tagctccttc 1680

ggctcctcga tcgttgtcag aagtaagttg gccgcagtgt tatcactcat ggttatggca 1740

gcactgcata attctcttac tgtcatgcca tccgtaagat gcttttctgt gactggtgag 1800

tactcaacca agtcattctg agaatagtgt atgcggcgac cgagttgctc ttgcccggcg 1860

tcaatacggg ataataccgc gccacatagc agaactttaa aagtgtcat cattggaaaa 1920

cgttcttcgg ggcgaaaact ctcaaggatc tiaccgctgt tgagatccag ttcgatgtaa 1980

cccactcgtg cacccaactg atcttcagca tcttttactt tcaccagcgt ttctgggtga 2040

gcaaaaaacag gaagcaaaa tgccgcaaaa aagggaataa gggcgacacg gaaatgttga 2100

atactcatac tcttcctttt tcaatattat tgaagcattt atcaggggta ttgtctcatg 2160

agcggataca tatttgaatg tatttagaaa aataaacaaa taggggttcc gcgcacattt 2220

ccccgaaaag tgccacctga acgaagcatc tgtgttcat tttgtagaac aaaaatgcaa 2280

cgcgagagcg ctaatttttc aaacaagaa tctgagctgc atttttacag aacagaaatg 2340

caacgcgaaa gcgtatttt accaacgaag aatctgtgct tcatttttgt aaaacaaaaa 2400

tgcaacgcga gagcgctaatt ttttcaaaca aagaatctga gctgcatttt tacagaacag 2460

aatgcaacg cgagagcgct attttacaa caaagaatct atacttcttt ttgttctac 2520

aaaaatgcat cccgagagcg ctatttttct aacaaagcat cttagattac tttttttctc 2580

ctttgtgcgc tctataatgc agtctcttga taactttttg cactgtaggt ccgttaaggt 2640

tagaagaagg ctacttttgt gtctattttc tcttcataa aaaaagcctg actccacttc 2700

ccgcgtttac tgattactag cgaagctgcg ggtgcatttt ttcaagataa aggcaccccc 2760

gattatattc tataccgatg tggattgcgc atactttgtg aacagaaagt gatagcgttg 2820

atgattcttc attggtcaga aaattatgaa cggtttcttc tattttgtct ctatatacta 2880

cgatataggaa atgtttacat tttcgtattg ttttcgattc actctatgaa tagttcttac 2940

tacaattttt ttgtctaaag agtaatacta gagataaaca taaaaaatgt agaggtcgag 3000

tttagatgca agttcaagga gcgaaagggt gatgggtagg ttatataggg atatagcaca 3060

gagatatata gcaaagagat acttttgagc aatgtttgtg gaagcgggtat tcgcaatatt 3120

ttagtagctc gttacagtcc ggtgcgtttt tggttttttg aaagtgcgtc ttcagagcgc 3180

ttttggtttt caaaagcgct ctgaagttcc tatactttct agagaatagg aacttcggaa	3240
taggaacttc aaagcgttcc cgaaaacgag cgcttccgaa aatgcaacgc gagctgcgca	3300
catacagctc actgttcacg tcgcacctat atctgcgtgt tgcctgtata tatatataca	3360
tgagaagaac ggcatagttc gtgtttatgc ttaaatgcgt acttatatgc gtctatttat	3420
gtaggatgaa aggtagtcta gtacctctg tgatattatc ccattccatg cggggtatcg	3480
tatgttccct tcagcactac cctttagctg ttctatatgc tgccactcct caattggatt	3540
agtctcatcc ttcaatgcta tcatttcctt tgatattgga tcatactaag aaaccattat	3600
tatcatgaca ttaacctata aaaataggcg taccacgagg ccttttcgtc tcgcgcgttt	3660
cggatgatgac ggtgaaaacc tctgacacat gcagctcccg gagacgggtca cagcttgtct	3720
glaagcggat gccgggagca gacaagcccg tcaggcgcg tcagcgggtg ttggcgggtg	3780
tcggggctgg ctttaactat gcgcacatg gcagattgta ctgagagtgc accatatcga	3840
ctacgtcgta aggccgtttc tgacagagta aaattcttga gggaactttc accattatgg	3900
gaaatgcttc aagaaggat tgacttaaac tccatcaaat ggtcagggtca ttgagtgttt	3960
tttatttggt gtattttttt ttttttagag aaaatcctcc aatatcaaat taggaatcgt	4020
agtttcatga ttttctgtta cacctaactt tttgtgtggt gccctcctcc ttgtcaatat	4080
taatgttaaa gtgcaattct ttttccttat cacgttgagc cattagtatc aatttgctta	4140
cctgtattcc tttactatcc tcctttttct ccttcttgat aaatgtatgt agattgcgta	4200
tatagtttcg tctaccctat gaacatatcc cattttgtaa tttcgtgtcg tttctattat	4260
gaatttcatt tataaagttt atgtacaaat atcataaaaa aagagaatct ttttaagcaa	4320
ggattttctt aacttcttcg gcgacagcat caccgacttc ggtggtagtg ttggaaccac	4380
ctaaatcacc agttctgata cctgcatcca aaaccttttt aactgcatct tcaatggcct	4440
taccttcttc aggcaagttc aatgacaatt tcaacatcat tgcagcagac aagatagtgg	4500
cgatagggtc aaccttatcc tttggcaaat ctggagcaga accgtggcat ggttcgtaca	4560
aaccaaagtc ggtgttcttg tctggcaaag aggccaagga cgcagatggc aacaaacca	4620
aggaacctgg gataacggag gcttcatcgg agatgatatc accaaacatg ttgctggtga	4680
ttataatacc atttaggtgg gttgggttct taactaggat catggcggca gaatcaatca	4740
attgatgttg aaccttcaat gtagggaatt cgttcttgat ggtttcctcc acagtttttc	4800
tcataaatct tgaagaggcc aaaagattag ctttatccaa ggaccaaata ggcaatggtg	4860
gctcatgttg tagggccatg aaagcggcca ttcttgtgat tctttgact tctggaacgg	4920
tgtattgttc actatcccaa gcgacacat caccatcgtc ttctttctc ttaccaaagt	4980
aaatactcc cactaatct ctgacaaca cgaagtcagt accttagca aattgtggct	5040

tgattggaga taagtctaaa agagagtcgg atgcaaagtt acatgggtctt aagttggcgt	5100
acaattgaag ttctttacgg attttttagta aaccttggtc aggtctaaca ctaccgtac	5160
cccatttagg accagccaca gcaccttaaca aaacggcatc aaccttcttg gaggttcca	5220
gcgcctcatc tggaagtggg acacctgtag catcgatagc agcaccacca attaaatgat	5280
tttcgaaatc gaacttgaca ttggaacgaa catcagaaat agctttaaga accttaatgg	5340
cttcggctgt gatttcttga ccaacgtggt cacctggcaa aacgacgatc ttcttagggg	5400
cagacatagg ggcagacatt agaattggtat atccttgaaa tataatatata tattgctgaa	5460
atgtaaaagg taagaaaagt tagaaagtaa gacgattgct aaccacctat tggaaaaaac	5520
aataggtcct taaataatat tgtcaacttc aagtattgtg atgcaagcat ttagtcatga	5580
acgtttctct attctatatg aaaagccggt tccggcctct cacctttcct ttttctcca	5640
atttttcagt tgaaaaaggt atatgcgtca ggcgacctct gaaattaaca aaaaatttcc	5700
agtcacgaa ttigattctg tgcgatagcg cccctgtgtg ttctcgttat gttgaggaaa	5760
aaaataatgg ttgctaagag attcgaactc ttgcatctta cgatacctga gtattccac	5820
agttaactgc ggtcaagata tttcttgaat caggcgctt agaccgctcg gccaaacaac	5880
caattacttg ttgagaaata gagtataatt atcctataaa tataacgttt ttgaacacac	5940
atgaacaagg aagtacagga caattgattt tgaagagaat gtggattttg atgtaattgt	6000
tgggattcca tttttaataa ggcaataata ttaggtatgt ggatatacta gaagttctcc	6060
tcgaccgtcg atatgcggtg tgaaataccg cacagatgcg taaggagaaa ataccgcatc	6120
aggaaattgt aaacgttaat attttgtaa aattcgcgtt aaatttttgt taaatcagct	6180
cattttttaa ccaataggcc gaaatcggca aaatccctta taaatcaaaa gaatagaccg	6240
agatagggtt gagtgttgtt ccagtttga acaagagtcc actattaaag aacgtggact	6300
ccaacgtcaa agggcgaaaa accgtctatc agggcgatgg cccactacgt gaaccatcac	6360
cctaatacag ttttttgggg tcgaggtgcc gtaaagcact aaatcggaac ctaaaggga	6420
gccccgatt tagagcttga cggggaaagc cggcgaacgt ggcgagaaag gaagggaaga	6480
aagcgaaagg agcgggcgtt agggcgctgg caagtgtagc ggtcacgctg cgcgtaacca	6540
ccacaccgc cgcgttaat gcgccgtac agggcgctc gcgccattcg ccattcaggc	6600
tgcgcaactg ttgggaaggg cgatcggtgc gggcctcttc gctattacgc cagctggcga	6660
aagggggatg tgctgcaagg cgattaaagt gggtaacgcc agggttttcc cagtcacgac	6720
gttgtaaaac gacggccagt gagcgcgctt aatagactc actatagggc gaattgggta	6780
ccgggcccc cctcgaggtc gacgtatcg ataagcttga tatcgaattc ctgcagcccg	6840
ggggatccgc atgcttgcat ttagtcgtgc aatgtatgac ttttaagattt gtgagcagga	6900

agaaaaggga gaatcttcta acgataaacc cttgaaaaac tgggtagact acgctatgtt 6960

gagttgctac gcaggctgca caattacacg agaatgctcc cgcctaggat ttaaggctaa 7020

gggacgtgca atgcagacga cagatctaaa tgaccgtgtc ggtgaagtgt tcgccaaact 7080

tttcggttaa cacatgcagt gatgcacgcg cgatgggtgt aagttacata tatatatata 7140

tatatatata tagccatagt gatgtctaag taacctttat ggtatatttc ttaatgtgga 7200

aagatactag cgcgcgcacc cacacacaag cttcgtcttt tcttgaagaa aagaggaagc 7260

tcgctaaatg ggattccact ttccgttccc tgccagctga tggaaaaagg ttagtggaac 7320

gatgaagaat aaaaagagag atccactgag gtgaaatttc agctgacagc gagtttcatg 7380

atcgtgatga acaatggtaa cgagttgtgg ctgttgccag ggagggtggt tctcaacttt 7440

taatgtatgg ccaaatcgct acttgggttt gttatataac aaagaagaaa taatgaactg 7500

attctcttcc tcttcttctt cttttcttaa ttctgttgta attaccttcc ttigtatatt 7560

tttttgaat tattcttctt aataatccaa acaaacacac atattacaat agctagctga 7620

ggatgaaggc attagtttat catggggatc acaaaatttc gttagaagac aaacccaaac 7680

ccactctgca gaaaccaaca gacgttgtgg ttaggggtgt gaaaacaaca atttgcggta 7740

ctgacttggg aatatacaaa ggtaagaatc ctgaagtggc agatggcaga atcctgggtc 7800

atgagggcgt tggcgtcatt gaagaagtgg gcgaatccgt gacacaattc aaaaaggggg 7860

ataaagtttt aatctcctgc gttactagct gtggatcgtg tgattattgc aagaagcaac 7920

tgtattcaca ctgtagagac ggtggctgga ttttaggtta catgatcgac ggtgtccaag 7980

ccgaatacgt cagaatacca catgctgaca attcattgta taagatcccg caaactatcg 8040

atgatgaaat tgcagtacta ctgtccgata ttttacctac tggacatgaa attgggtgtc 8100

aatatggtaa cgttcaacca ggcatgctg tagcaattgt aggagcaggt cctgtttgaa 8160

tgtcagtttt gttaactgct caattttact cgcctagtac cattattgtt atcgacatgg 8220

acgaaaaccg tttaacaatta gcgaaggagc ttggggccac acacactatt aactccggtg 8280

ctgaaaatgt tgtcgaagct gtgcatcgta tagcagccga aggagtggat gtagcaatag 8340

aagctgttgg tataccgcga acctgggaca tctgtcagga aattgtaaaa cccggcgctc 8400

atattgccaa cgtgggagtt catggtgtta aggtggactt tgaaattcaa aagttgtgga 8460

ttaagaatct aaccatcacc actggttttg ttaacactaa tactaccca atgttgatga 8520

aggtagcctc tactgataaa ttgcctttta agaaaatgat tactcacagg tttaggttag 8580

ctgaaatcga acacgcatat caggttttct tgaatggcgc taaagaaaaa gctatgaaga 8640

ttattctatc taatgcaggt gccgcctaata taattaagag taagcgaatt tcttatgatt	8700
tatgatTTTT attattaaat aagttataaa aaaaataagt gtatacaaat tttaaagtga	8760
ctcttaggtt ttaaacgaa aattcttatt cttgagtaac tctttctgt aggtcaggtt	8820
gctttctcag gtatagcatg aggtcgctct tattgaccac acctctaccg gcatgccgag	8880
caaatgcctg caaatcgctc cccatttcac ccaattgtag atatgctaac tccagcaatg	8940
agttgatgaa tctcggtgtg tattttatgt cctcagagga caacacctgt ggta	8994

<210> 810

<211> 753

<212> DNA

<213> Artificial Sequence

<220><223> Saccharomyces cerevisiae GPM1 promoter

<400> 810

gcatgcttgc atttagtcgt gcaatgtatg actttaagat ttgtgagcag gaagaaaagg	60
gagaatcttc taacgataaa cccctgaaaa actgggtaga ctacgctatg ttgagttgct	120
acgcaggctg cacaattaca cgagaatgct cccgcctagg atttaaggct aagggacgtg	180
caatgcagac gacagatcta aatgaccgtg tcggtgaagt gttcgccaaa cttttcggtt	240
aacacatgca gtgatgcacg cgcgatgggt ctaagttaca tataatatata tatagccata	300
gtgatgtcta agtaaccttt atggtatatt tcttaatgtg gaaagatact agcgcgcgca	360

cccacacaca agcttcgtct tttcttgaag aaaagaggaa gctcgctaaa tgggattcca	420
ctttccgttc cctgccagct gatggaaaaa ggtagtgga acgatgaaga ataaaaagag	480
agatccactg aggtgaaatt tcagctgaca gcgagtttca tgatcgtgat gaacaatggt	540
aacgagttgt ggctgttgc agggaggggt gttctcaact tttaatgtat ggccaaatcg	600
ctacttgggt ttgttatata acaaagaaga aataatgaac tgattctctt cctccttctt	660
gtcctttctt aattctgttg taattacctt cctttgtaat ttttttgta attattcttc	720
ttaataatcc aaacaaacac acatattaca ata	753

<210> 811

<211> 1047

<212> DNA

<213> Artificial Sequence

<220><223> Achromobacter xylosoxidans sadB

<400> 811

atgaaagctc tggtttatca cggtgaccac aagatctcgc ttgaagacaa gcccaagccc	60

acccttcaaa agcccacgga tgtagtagta cgggttttga agaccacgat ctgcggcacg 120
gatctcggca tctacaaagg caagaatcca gaggtcgccg acgggcgcat cctgggccat 180
gaaggggtag gcgtcatcga ggaagtgggc gagagtgtca cgcagttcaa gaaaggcgac 240
aaggtcctga tticctgcgt cacttcttgc ggctcgtgcg actactgcaa gaagcagctt 300

tactccatt gccgcagcg cgggtggatc ctgggttaca tgatcgatgg cgtgcaggcc 360
gaatacgtcc gcacccgca tgccgacaac agcctctaca agatccccca gacaattgac 420
gacgaaatcg ccgtcctgct gagcgacatc ctccccaccg gccacgaaat cggcgtccag 480
tatgggaatg tccagccggg cgatcggtg gctattgtcg gcgcggggccc cgtcggcatg 540
tccgtactgt tgaccgcca gttctactcc cctcgacca tcatcgtgat cgacatggac 600
gagaatgcc tccagctcgc caaggagctc ggggcaacgc acaccatcaa ctccggcacg 660
gagaacgttg tcgaagccgt gcataggatt gcggcagagg gattcgatgt tgcgatcgag 720

gcggtgggca taccggcgac ttgggacatc tgccaggaga tcgtcaagcc cggcgcgcac 780
atcgccaacg tcggcgtgca tggcgtcaag gttgacttcg agattcagaa gctctggatc 840
aagaacctga cgatcaccac gggactggtg aacacgaaca cgacgcccac gctgatgaag 900
gtcgcctcga ccgacaagct tccgttgaag aagatgatta cccatcgctt cgagctggcc 960
gagatcgagc acgcctatca ggtattcttc aatggcgcca aggagaaggc gatgaagatc 1020
atcctctcga acgcaggcgc tgcctga 1047

<210> 812

<211> 348

<212> PRT

<213> Artificial Sequence

<220><223> Achromobacter xylosoxidans sadB

<400> 812

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

1 5 10 15

Lys Pro Lys Pro Thr Leu Gln Lys Pro Thr Asp Val Val Val Arg Val

20 25 30

Leu Lys Thr Thr Ile Cys Gly Thr Asp Leu Gly Ile Tyr Lys Gly Lys

35 40 45

Asn Pro Glu Val Ala Asp Gly Arg Ile Leu Gly His Glu Gly Val Gly

50

55

60

Val Ile Glu Glu Val Gly Glu Ser Val Thr Gln Phe Lys Lys Gly Asp
65 70 75 80
Lys Val Leu Ile Ser Cys Val Thr Ser Cys Gly Ser Cys Asp Tyr Cys
85 90 95
Lys Lys Gln Leu Tyr Ser His Cys Arg Asp Gly Gly Trp Ile Leu Gly
100 105 110
Tyr Met Ile Asp Gly Val Gln Ala Glu Tyr Val Arg Ile Pro His Ala
115 120 125
Asp Asn Ser Leu Tyr Lys Ile Pro Gln Thr Ile Asp Asp Glu Ile Ala
130 135 140
Val Leu Leu Ser Asp Ile Leu Pro Thr Gly His Glu Ile Gly Val Gln
145 150 155 160
Tyr Gly Asn Val Gln Pro Gly Asp Ala Val Ala Ile Val Gly Ala Gly
165 170 175
Pro Val Gly Met Ser Val Leu Leu Thr Ala Gln Phe Tyr Ser Pro Ser
180 185 190
Thr Ile Ile Val Ile Asp Met Asp Glu Asn Arg Leu Gln Leu Ala Lys
195 200 205
Glu Leu Gly Ala Thr His Thr Ile Asn Ser Gly Thr Glu Asn Val Val
210 215 220
Glu Ala Val His Arg Ile Ala Ala Glu Gly Val Asp Val Ala Ile Glu
225 230 235 240
Ala Val Gly Ile Pro Ala Thr Trp Asp Ile Cys Gln Glu Ile Val Lys
245 250 255
Pro Gly Ala His Ile Ala Asn Val Gly Val His Gly Val Lys Val Asp
260 265 270
Phe Glu Ile Gln Lys Leu Trp Ile Lys Asn Leu Thr Ile Thr Thr Gly
275 280 285
Leu Val Asn Thr Asn Thr Thr Pro Met Leu Met Lys Val Ala Ser Thr
290 295 300
Asp Lys Leu Pro Leu Lys Lys Met Ile Thr His Arg Phe Glu Leu Ala

305	310	315	320
Glu Ile Glu His Ala Tyr Gln Val Phe Leu Asn Gly Ala Lys Glu Lys			
325	330	335	
Ala Met Lys Ile Ile Leu Ser Asn Ala Gly Ala Ala			
340	345		
<210>	813		
<211>	39		
<212>	DNA		
<213>	Artificial Sequence		
<220><223>	Primer OT1074		
<400>	813		
cacacatatt acaatagcta gctgaggatg aaagctctg			39
<210>	814		
<211>	39		
<212>			
>	DNA		
<213>	Artificial Sequence		
<220><223>	Primer OT1075		
<400>	814		
cagagctttc atcctcagct agctattgta atatgtgtg			39
<210>	815		
<211>	9075		
<212>	DNA		
<213>	Artificial Sequence		
<220><223>	pLH435		
<400>	815		
ctagtcttag agcggccgcc accgcggtgg agctccagct tttgttcctt ttagtgaggg			60
ttaattgcgc gcttggcgta atcatggtca tagctgtttc ctgtgtgaaa ttgttatccg			120
ctcacaattc cacacaacat aggagccgga agcataaagt gtaaagcctg gggcgcctaa			180
tgagtgaggt aactcacatt aattgcgttg cgctcactgc ccgctttcca gtcgggaaac			240
ctgtcgtgcc agctgcatta atgaatcggc caacgcgcgg ggagaggcgg ttgcgtatt			300
gggcgctctt ccgcttcctc gctcactgac tcgctgcgct cggctgttcg gctgcggcga			360
gcggtatcag ctactcaaa ggccgtaata cggttatcca cagaatcagg ggataacgca			420
ggaaagaaca tgtgagcaaa aggccagcaa aaggccagga accgtaaaaa ggccgcgttg			480

ctggcgTTTT tccataggct ccgccccct gacgagcatc acaaaaatcg acgctcaagt	540
cagaggtggc gaaacccgac aggactataa agataccagg cgtttcccc tggaagctcc	600
ctcgtgcgct ctctgttcc gacctgccg cttaccggat acctgtccgc ctttctccct	660
tcgggaagcg tggcgctttc tcatagctca cgctgtaggt atctcagttc ggtgtaggtc	720
gttcgctcca agctgggctg tgtgcacgaa cccccgttc agcccgaccg ctgcgcctta	780
tccggttaact atcgtcttga gtccaaccg gtaagacacg acttatcgcc actggcagca	840
gccactggta acaggattag cagagcgagg tatgtaggcg gtgctacaga gttcttgaag	900
tgggtgccta actacggcta cactagaagg acagtatttg gtatctgcgc tctgctgaag	960
ccagttacct tcggaaaaag agttggtagc tcttgatccg gcaaaaaac caccgtggt	1020
agcggtggtt ttttgtttg caagcagcag attacgcgca gaaaaaagg atctcaagaa	1080
gatactttga tctttctac ggggtctgac gctcagtga acgaaaactc acgttaaggg	1140
attttggta tgagattatc aaaaaggatc ttcacctaga tctttttaa ttaaaatga	1200
agttttaa ataatctaaag tataatgag taaacttgg ctgacagtta ccaatgctta	1260
atcagtgagg cacctatctc agcgatctgt ctatttcgtt catccatagt tgcctgactc	1320
cccgtcgtgt agataactac gatacgggag ggcttaccat ctggccccag tgctgcaatg	1380
ataccgcgag acccacgctc accggctcca gatttatcag caataaacca gccagccgga	1440
agggccgagc gcagaagtgg tcctgcaact ttatccgcct ccatccagtc tattaattgt	1500
tgccgggaag ctagagtaag tagttcgcca gttaatagtt tgcgcaacgt tgttgccatt	1560
gtacagga tcgtgggtgc acgctcgtcg tttggatgg cttcattcag ctccggttcc	1620
caacgatcaa ggcgagttac atgatcccc atgttgtgca aaaaagcgg tagctccttc	1680
ggtcctccga tcgttgtcag aagtaagttg gccgcagtgt tactactcat ggttatggca	1740
gcactgcata attctcttac tgtcatgcca tccgtaagat gcttttctgt gactggtgag	1800
tactcaacca agtcattctg agaatagtgt atgcggcgac cgagttgctc ttgccggcg	1860
tcaatacggg ataataccgc gccacatagc agaacttta aagtgtcat cattggaaaa	1920
cgttcttcgg ggcgaaaact ctcaaggatc ttaccgtgt tgagatccag ttcgatgtaa	1980
cccactcgtg cacccaactg atcttcagca tcttttactt tcaccagcgt ttctgggtga	2040
gcaaaaacag gaaggcaaaa tgccgcaaaa aagggaataa gggcgacacg gaaatgttga	2100
atactcatac tcttctttt tcaatattat tgaagcattt atcagggtta ttgtctcatg	2160
agcggataca tatttgaatg tatttagaaa aataaaciaa taggggttcc gcgcacattt	2220
ccccgaaaag tgccacctga acgaagcatc tgtgttcat tttgtagaac aaaaatgcaa	2280

cgcgagagcg ctaatttttc aaacaaagaa tctgagctgc atttttacag aacagaaatg 2340

caacgcgaaa gcgctatttt accaacgaag aatctgtgct tcatttttgt aaaacaaaaa 2400

tgcaacgcga gagcgctaata ttttcaaaca aagaatctga gctgcatttt tacagaacag 2460

aaatgcaacg cgagagcgct attttaccac caaagaatct atacttcttt ttgtttctac 2520

aaaaatgcat cccgagagcg ctatttttct aacaaagcat cttagattac tttttttctc 2580

ctttgtgcgc tctataatgc agtctcttga taactttttg cactgtaggt ccgttaaggt 2640

tagaagaagg ctacttttgt gtctattttc tcttcataa aaaaagcctg actccacttc 2700

ccgcgtttac tgattactag cgaagctgcg ggtgcatttt ttcaagataa aggcattccc 2760

gattatattc tataccgatg tggattgcgc atactttgtg aacagaaagt gatagcgttg 2820

atgattcttc attggtcaga aaattatgaa cggttttcttc tattttgtct ctatatacta 2880

cgtataggaa atgtttacat tttcgtattg ttttcgattc actctatgaa tagttcttac 2940

tacaattttt ttgtctaaag agtaatacta gagataaaca taaaaaatgt agaggtcgag 3000

tttagatgca agttcaagga gcgaaagggt gatgggtagg ttatataggg atatagcaca 3060

gagatatata gcaaagagat acttttgagc aatgtttgtg gaagcgggat tcgcaatatt 3120

ttagtagctc gttacagtcc ggtgcgtttt tggttttttg aaagtgcgtc ttcagagcgc 3180

ttttggtttt caaaagcgtc ctgaagttcc tatactttct agagaatagg aacttcggaa 3240

taggaacttc aaagcgtttc cgaaaacgag cgcttccgaa aatgcaacgc gagctgcgca 3300

catacagctc actgttcacg tcgcacctat atctgcgtgt tgcctgtata tatatatata 3360

tgagaagaac ggcatagtgc gtgtttatgc ttaaatgcgt acttatatgc gtctatttat 3420

gtaggatgaa aggtagtcta gtacctctg tgatattatc ccattccatg cgggggatcg 3480

tatgttctct tcagcactac cctttagctg ttctatatgc tgccactcct caattggatt 3540

agtctcatcc ttcaatgcta tcatttcttt tgatattgga tcatactaag aaaccattat 3600

tatcatgaca ttaacctata aaaataggcg tatcacgagg ccttttcgtc tcgcgcgttt 3660

cggatgatgac ggtgaaaacc tctgacacat gcagctcccg gagacggtca cagcttgtct 3720

gtaagcggat gccgggagca gacaagcccg tcagggcgcg tcagcgggtg ttggcgggtg 3780

tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc accatatcga 3840

ctacgtcgta aggccgtttc tgacagagta aaattcttga gggaactttc accattatgg 3900

gaaatgcttc aagaaggat tgacttaaac tccatcaaat ggtcagggtca ttgagtgttt 3960

tttatttgtt gtattttttt ttttttagag aaaatcctcc aatatcaaat taggaatcgt 4020

agtttcatga ttttctgtta cacctaactt ttgtgtgggt gccctcctcc ttgtcaatat 4080
 taatgttaaa gtgcaattct ttttccttat caggttgagc cattagtatc aatttgctta 4140
 cctgtattcc ttactatcc tcttttttct ctttcttgat aaatgtatgt agattgcgta 4200
 tatagtttcg tctaccctat gaacatattc cattttgtaa ttctgtgtcg ttctattat 4260
 gaatttcatt tataaagttt atgtacaaat atcataaaaa aagagaatct ttttaagcaa 4320
 ggattttctt aacttcttcg gcgacagcat caccgacttc ggtggtagtg ttggaaccac 4380
 ctaaatacacc agttctgata cctgcatcca aaaccttttt aactgcatct tcaatggcct 4440

 taccttcttc aggcaagttc aatgacaatt tcaacatcat tgcagcagac aagatagtgg 4500
 cgatagggtc aaccttattc ttggcgaat ctggagcaga accgtggcat ggttcgtaca 4560
 aaccaaagtc ggtgttcttg tctggcaaag aggccaagga cgcagatggc aacaaacca 4620
 aggaacctgg gataacggag gcttcatcgg agatgatatc accaaacatg ttgctggtga 4680
 ttataatacc atttaggtgg gttgggttct taactaggat catggcggca gaatcaatca 4740
 attgatgttg aaccttcaat gtagggaatt cgttcttgat ggtttcctcc acagttttc 4800
 tccataatct tgaagaggcc aaaagattag ctttatccaa ggaccaaata ggcaatggtg 4860

 gctcatgttg tagggccatg aaagcggcca ttcttggat tcttgcact tctggaacgg 4920
 tgtattgttc actatccaa gcgacacat caccatcgtc ttctttctc ttaccaaagt 4980
 aaatactcc cactaattct ctgacaacaa cgaagtcagt accttagca aattgtggct 5040
 tgattggaga taagtctaaa agagagtcgg atgcaaagtt acatggtctt aagttggcgt 5100
 acaattgaag ttctttacgg atttttagta aacctgttc aggtctaaca ctaccgtac 5160
 cccatttagg accagccaca gcacetaaca aaacggcatc aaccttcttg gaggttcca 5220
 gcgcctcatc tggaagtggg acacctgtag catcgatagc agcaccacca attaatgat 5280

 ttctgaaatc gaacttgaca ttggaacgaa catcagaaat agctttaaga acctaatgg 5340
 cttcggtgtg gatctctga ccaacgtggc cacctggcaa aacgacgatc ttcttagggg 5400
 cagacatagg ggcagacatt agaatggtat atccttgaaa tatatatata tattgtgaa 5460
 atgtaaaagg taagaaaagt tagaaagtaa gacgattgct aaccacctat tggaaaaaac 5520
 aataggtcct taaataatat tgtcaacttc aagtattgtg atgcaagcat ttagtcatga 5580
 acgttctct attctatatg aaaagccggt tccggcctct cactttcct ttttctcca 5640
 attttccagt tgaaaaaggt atatgctca ggcgacctct gaaattaaca aaaaatttc 5700

 agtcatcgaa ttgattctg tgcgatagcg ccctgtgtg ttctcgttat gttgaggaaa 5760
 aaaataatgg ttgctaagag attcgaactc ttgcatctta cgatacctga gtattccac 5820
 agttaactgc ggtcaagata tttcttgaat caggcgctt agaccgctcg gccaaacaac 5880

caattacttg ttgagaaata gagtataatt atcctataaa tataacgttt ttgaacacac	5940
atgaacaagg aagtacagga caattgattt tgaagagaat gtggattttg atgtaattgt	6000
tgggattcca tttttaataa ggcaataata ttaggtatgt ggatatacta gaagttctcc	6060
tcgaccgtcg atatcggtg tgaaataaccg cacagatgcg taaggagaaa ataccgcatc	6120
aggaaattgt aaacgttaat attttgtaa aattcgcgtt aaatttttgt taaatcagct	6180
cattttttaa ccaataggcc gaaatcggca aaatccctta taaatcaaaa gaatagaccg	6240
agatagggtt gagtgttgtt ccagtttga acaagagtcc actattaaag aacgtggact	6300
ccaacgtcaa agggcgaaaa accgtctatc agggcgatgg cccactacgt gaaccatcac	6360
cctaatacaag ttttttggg tcgaggtgcc gtaaagcact aaatcggaac cctaaaggga	6420
gccccgatt tagagcttga cggggaaaagc cggcgaacgt ggcgagaaaag gaagggaaga	6480
aagcgaaagg agcgggcgct agggcgctgg caagtgtagc ggtcacgctg cgcgtaacca	6540
ccacacccgc cgcgcttaat gcgcgctac agggcgctc gcgccattcg ccattcaggc	6600
tgcgcaactg ttgggaaggc cgatcggtgc gggcctcttc gctattacgc cagctggcga	6660
aagggggatg tgctgcaagg cgattaagtt gggtaacgcc agggttttcc cagtcacgac	6720
gttgtaaaac gacggccagt gagcgcgctt aatagactc actatagggc gaattgggta	6780
ccgggcccc cctcgaggtc gacggtatcg ataagcttga tatcgaattc ctgcagcccg	6840
ggggatccgc atgcttgcat ttagtcgtgc aatgtatgac ttttaagattt gtgagcagga	6900
agaaaaggga gaatcttcta acgataaacc cttgaaaaac tgggtagact acgctatgtt	6960
gagttgctac gcaggctgca caattacacg agaatgctcc cgcctaggat ttaaggctaa	7020
gggacgtgca atgcagacga cagatctaaa tgaccgtgct ggtgaagtgt tcgccaaact	7080
tttcggttaa cacatgcagt gatgcacgcg cgatggtgct aagttacata tatatatata	7140
tatatatata tagccatagt gatgtctaag taacctttat ggtatatttc ttaatgtgga	7200
aagatactag cgcgcgcacc cacacacaag cttcgtcttt tcttgaagaa aagaggaagc	7260
tcgctaaatg ggattccact ttccgttccc tgccagctga tggaaaaagg ttagtggaac	7320
gatgaagaat aaaaagagag atccactgag gtgaaatttc agctgacagc gagtttcatg	7380
atcgtgatga acaatggtaa cgagttgtgg ctgttgccag ggagggtggt tctcaacttt	7440
taatgtatgg ccaaatcgct acttgggttt gttatataac aaagaagaaa taatgaactg	7500
attctcttcc tccttcttgt cttttcttaa ttctgttgta attaccttcc tttgtaattt	7560
tttttgtaat tattcttctt aataatcaa acaaacacac atattacaat agctagctga	7620
ggatgtcaac agccggtaaa gttattaagt gtaaagcggc agttttgtgg gaagagaaaa	7680
agccgtttag catagaagaa gtagaagtag cgccaccaa agcacacgag gtagaatca	7740

agatggttgc caccggaatc tgtagatccg acgacatgt ggtgagtggc actctagtta 7800

ctcctttgcc agtaatcgcg ggacacgagg ctgccggaat cgttgaatcc ataggtgaag 7860

gtgttaccac tgttcgtcct ggtgataaag tgatcccact gttcactcct caatgttgta 7920

agtgtagagt ctgcaaacat cctgagggta atttctgcct taaaaatgat ttgtctatgc 7980

ctagaggtag tatgcaggat ggtacaagca gatttacctg cagagggaaa cctatacacc 8040

atttccttgg tacttctaca ttttcccaat acacagtggg ggacgagata tctgtcgcta 8100

aaatcgatgc agcttcacca ctggaaaaag ttgtcttgat aggggtcgga ttttcaccg 8160

gttacggttc cgcagttaaa gttgcaaagg ttacacaggg ttcgacttgt gcagtattcg 8220

gtttaggagg agtaggacta agcgttattt tggggtgtaa agctgcaggc gcagcgagga 8280

ttataggtgt agacatcaat aaggacaaat ttgcaaaagc taaggaggtc ggggctactg 8340

aatgtgttaa cctcaagat tataagaaac caatacaaga agtccttact gaaatgtcaa 8400

acggtggagt tgatttctct ttgtaagtta taggccgtct tgatactatg gtaactgcgt 8460

tgtcctgctg tcaagaggca tatggagtca gtgtgatcgt aggtgttcct cctgattcac 8520

aaaatttgc gatgaatcct atgctgttgc taagcggctg tacatggaag ggagctatat 8580

ttggcggttt taagagcaag gatagtgttc caaaacttgt tgccgacttt atggcgaaga 8640

agtttgcctt tgatccttta attacacatg tatlgccatt cgagaaaatc aatgaagggt 8700

ttgatttgtt aagaagtggg gaatctattc gtacaatttt aactttttga ttaattaaga 8760

gtaagcgaat ttcttatgat ttatgatttt tattattaaa taagttataa aaaaaataag 8820

tgtatacaaa ttttaaagtg actcttaggt tttaaaacga aaattcttat tcttgagtaa 8880

ctctttcctg taggtcaggt tgctttctca ggtatagcat gaggtcgctc ttattgacca 8940

cacctctacc ggcatgccga gcaaatgcct gcaaatcgct cccatttca cccaattgta 9000

gatatgctaa ctccagcaat gagttgatga atctcggtgt gtattttatg tcctcagagg 9060

acaacacctg tggta 9075

<210> 816

<211> 11367

<212> DNA

<213> Artificial Sequence

<220><223> pLH441

<400> 816

tcgcgcgttt cggatgatgac ggtgaaaacc tctgacacat gcagctcccg gagacggtea 60

cagcttgtct gtaagcggat gccgggagca gacaagcccg tcagggcgcg tcagcgggtg 120

ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc	180
accatagcca tctcatgaa aactgtgtaa cataataacc gaagtgtcga aaaggtggca	240
ccttgtccaa ttgaacacgc tcgatgaaaa aaataagata tatataaggt taagtaaagc	300
gtctgttaga aaggaagttt ttccittttc ttgctctctt gtcttttcat ctactatttc	360
cttcgtgtaa tacagggtcg tcagatacat agatacaatt ctattacccc catccataca	420
atgccatctc atttcgatac tgttcaacta cagccgggcc aagagaaccc tggtgacaat	480
gtcacagat ccagagctgt accaatttac gccaccactt cttatgtttt cgaaaactct	540
aagcatggtt cgcaattgtt tggcttagaa gtccaggtt acgtctatc ccgtttccaa	600
aaccaacca gtaatgtttt ggaagaaaga attgtgctt tagaaggtgg tgctgctgct	660
ttggctgttt cctccggtca agccgctcaa acccttgcca tccaaggttt ggcacacact	720
ggtgacaaca tcgtttccac ttcttactta tacggtggta cttataacca gttcaaaatc	780
tcgttcaaaa gatttggat cgaggctaga ttgttgaag gtgacaatcc agaagaattc	840
gaaaaggctt ttgatgaaag aaccaaggct gtttatttgg aaaccattgg taatccaaag	900
tacaatgttc cggattttga aaaaattgtt gcaattgctc acaaacacgg tattecagtt	960
gtcgttgaca acacatttgg tgccggtggt tacttctgtc agccaattaa atacggtgct	1020
gatattgtaa cacattctgc taccaaatgg attggtggtc atggtactac tatcggtggt	1080
attattgttg actctggtaa gtteccatgg aaggactacc cagaaaagtt ccctcaattc	1140
tctcaacctg ccgaaggata tcacgtact atctacaatg aagcctacgg taacttggca	1200
tacatcgttc atgttagaac tgaactatta agagatttgg gtccattgat gaaccatttt	1260
gcctctttct tgctactaca aggtgttgaa acattatctt tgagagctga aagacacggt	1320
gaaaatgcat tgaagttagc caaatggtta gaacaatccc catacgtatc ttgggtttca	1380
taccctggtt tagcatctca ttctcatcat gaaaatgcta agaagtatct atctaacggt	1440
ttcgttggtg tcttatcttt cgggtgtaaaa gacttacaa atgccgcaa ggaaactgac	1500
ccattcaaac ttcttggtgc tcaagttgtt gacaatttaa agcttgctc taacttggcc	1560
aatgttgggtg atgccaagac cttagtcatt gctccatact tcactacca caaacaatta	1620
aatgacaaag aaaagttggc atctgggtgt accaaggact taattcgtgt ctctgttgg	1680
atcgaattta ttgatgacat tattgcagac ttccagcaat cttttgaaac tgttttcgt	1740
ggccaaaaac catgagtgtg cgtaatgagt tgtaaaatta tgtataaacc tactttctct	1800
cacaagttaa gcggtgtgaa ataccgcaca gatgcgtaag gagaaaatac cgcacagga	1860
aattgtaaac gttaatatit tgtaaaat cgcgttaaat ttttgtaaa tcagctcatt	1920

ttttaaccaa taggccgaaa tcggcaaaat cccttataaa tcaaaagaat agaccgagat	1980
agggttgagt gttgttccag tttggaacaa gagtccacta ttaaagaacg tggactccaa	2040
cgtcaaaggg cgaaaaaccg tctatcaggg cgatggccca ctacgtgaac catcaccta	2100
atcaagtttt ttggggtcga ggtgccgtaa agcactaaat cggaacccta aaggagagccc	2160
ccgatttaga gcttgacggg gaaagccggc gaacgtggcg agaaaggaag ggaagaaagc	2220
gaaaggagcg ggcgctaggg cgctggcaag tgtagcggtc acgtcgcg taaccaccac	2280
acccgccg cgctaatgcgc cgctacaggg cgcgtcgcg cattcgccat tcaggctgcg	2340
caactgttgg gaagggcgat cgggtcgggc ctcttcgcta ttacgccagc tggcgaaagg	2400
gggatgtgct gcaaggcgat taagtgggt aacgccaggg ttttcccagt cacgacgttg	2460
taaaacgacg gccagtgagc gcgcgtaata cgactcacta tagggcgaat tgggtaccgg	2520
gccccccctc gaggtcgacg gtatcgataa gcttgatc gaattcctgc agccccgggg	2580
atccactagt tctagagcgg ccgctctaga actagtacca caggtgttgt cctctgagga	2640
cataaaatac acaccgagat tcatcaactc attgctggag ttagcatac tacaattggg	2700
tgaaatgggg agcgatttgc aggcatttgc tcggcatgcc ggtagagggt tggtaataa	2760
gagcgacctc atgctatacc tgagaaagca acctgacctc caggaaagag ttactcaaga	2820
ataagaatth tcgttttaaa acctaagagt cactttaaaa tttgtataca cttatthttt	2880
ttataactta ttttaataata aaaatcataa atcataagaa attcgcttac tcttaattaa	2940
tcaaaaagtt aaaattgtac gaatagattc accacttctt aacaaatcaa acccttcatt	3000
gattttctcg aatggcaata catgtgtaat taaaggatca agagcaaact tcttcgcat	3060
aaagtccgca acaagttttg gaacactatc ctgtctctta aaaccgcaa atatagctcc	3120
cttccatgta cgaccgctta gcaacagcat aggattcatc gacaaattht gtgaatcagg	3180
aggaacacct acgatcacac tgactccata tgcctcttga cagcaggaca acgcagttac	3240
catagtatca agacggccta taacttcaaa agagaaatca actccaccgt ttgacatttc	3300
agtaaggact tcttgatttg gtttcttata atcttgaggg ttaacacatt cagtagcccc	3360
gacctctta gcttttgcaa atttgtcctt attgatgtct acacctataa tcttcgtgc	3420
gcctgcagct ttacacccca taataacgct tagtcctact cctcctaacc cgaatactgc	3480
acaagtcgaa cctgtgttaa cctttgcaac ttttaactgc gaaccgtaac cgggtgaaaa	3540
tccgcacct atcaagcaaa ctttttccag tggatgaagc gcatcgattt tagcgacaga	3600
tatctcgtcc accactgtgt attgggaaaa tgtagaagta ccaaggaaat ggtgtatagg	3660

tttccctctg catgtaaacc tgcttgtagc atcctgcata gtacctctag gcatagacaa	3720
atcattttta aggcagaaat taccctcagg atgtttgcag actctacact taccacattg	3780
aggagtgaac agtgggatca cttttacc accaggaaca gtggtaacac cttcacctat	3840
ggattcaacg attccggcag cctcgtgtcc cgcgattact ggcaaaggag taactagagt	3900
gccactcacc acatggctgt cggatctaca gattccgggtg gcaaccatct tgattctaac	3960
ctcgtgtgct tttgggtggc ctacttctac ttcttctatg ctaaaccggt ttttctcttc	4020
ccacaaaact gccgctttac acttaataac ttaccgggt gttgacatcc tcagctagct	4080
attgtaatat gtgtgtttgt ttggattatt aagaagaata attacaaaa aaattacaaa	4140
ggaaggtaat tacaacagaa ttaagaaagg acaagaagga ggaagagaat cagttcatta	4200
tttcttcttt gttatataac aaaccaagt agcgatttgg ccatcatta aaagttgaga	4260
accacctcc ctggcaacag ccacaactcg ttaccattgt tcatcacgat catgaaactc	4320
gctgtcagct gaaatttcac ctacgtggat ctctcttttt attcttcac gtccactaa	4380
cctttttcca tcagctggca gggaacggaa agtgggaatcc catttagcga gcttctctt	4440
ttcttcaaga aaagacgaag cttgtgtgtg ggtgcgcgcg ctagtatctt tccacattaa	4500
gaaatatacc ataaaggta cttagacatc actatggcta tatatatata tatatatata	4560
tatgtaactt agcaccatcg cgcgtgcac actgcatgtg ttaaccgaaa agtttggcga	4620
acacttcacc gacacggta tttagatctg tcgtctgcat tgcacgtccc ttagccttaa	4680
atcctaggcg ggagcattct cgtgtaattg tgcagcctgc gtagcaactc aacatagcgt	4740
agtctacca gtttttcaag ggtttatctg tagaagattc tcccttttct tctgtctac	4800
aaatcttaaa gtcataact gcacgactaa atgcaagcat gcggatcccc cgggctgcag	4860
gaattcgata tcaagcttat cgataccgtc gactggccat taatctttcc catattagat	4920
ttcgccaagc catgaaagt caagaaagg ctttagacga attacccttc atttctcaaa	4980
ctggcgtcaa gggatcctgg tatggtttta tcgttttatt tctggttctt atagcatcgt	5040
tttgacttc tctgttccca ttagcgggt caggagccag cgcagaatca ttctttgaag	5100
gatacttacc ctttcaatt ttgattgtct gttacgttgg acataaactg tatactagaa	5160
attggacttt gatggtgaaa ctagaagata tggatcttga taccggcaga aaacaagtag	5220
atttgactct tcgtagggaa gaaatgagga ttgagcgaga aacattagca aaaagatcct	5280
tcgtaacaag atttttcat ttctggtgtt gaagggaag atatgagcta tacagcggaa	5340
tttccatacc actcagattt tggtatctaa ttttttctt cccacgtccg cgggaatctg	5400
tgatatttac tgcatctaga tatatgttat cttatcttgg cgcgtacatt taattttcaa	5460
cgtattctat aagaaattgc gggagttttt ttcattgata tgatactgac tgcacgcaaa	5520

tataggcatg atttataggc atgatttgat ggctgtaccg ataggaacgc taagagtaac	5580
ttcagaatcg ttatcctggc ggaaaaaatt catttgtaaa ctttaaaaaa aaaagccaat	5640
atccccaaaa ttattaagag cgccctccatt attaaactaaa atttcactca gcatccacaa	5700
tgtatcaggt atctactaca gatattacat gtggcgaaaa agacaagaac aatgcaatag	5760
cgcacaaaga aaaaacacaa agctttcaat caatgaatcg aaaatgtcat taaaatagta	5820
tataaatgta aactaagta taaagctata aaaagaaaat ttattttaat gcaagattta	5880
aagtaaattc acggccctgc aggcctcagc tcttgttttg ttctgcaaat aacttaccca	5940
tctttttcaa aacttttagt gcacctctct ttgctagaat aagtctctat caatacatcc	6000
tatttggatc tgcttgagct tctttcatca cggatacgaa ttcatTTTTt gtctcacia	6060
ttttggacac aactctgtct tccgttgccc cgaaactttc tggcagtttt gagtaattcc	6120
acataggaat gtcattataa ctctggttcg gaccatgaat ttccctctca accgtgtaac	6180
catcgttatt aatgataaag cagattgggt ttatcttctc tctaattggct agtcctaatt	6240
cttggacagt cagttgcaat gatccatctc cgataaacia taaatgtcta gattctttat	6300
ctgcaatttg gctgcctaga gctgcgggga aagtgtatcc tatagatccc cacaagggtt	6360
gaccaataaa atgtgatttc gatttcagaa atatagatga ggcaccgaag aaagaagtgc	6420
cttgctcagc cagatcgtc tcattacttt gggtaaaatt ttgacagct tgccacagtc	6480
tatcttgta caacagcgcg ttagaaggta caaaatcttc ttgcttttta tctatgtact	6540
tgcttttata ttcaatttcg gacaagtcaa gaagagatga tatcagggat tcgaagtcga	6600
aattttggat tctttcgttg aaaattttac ctccatcgat attcaaggaa atcattttat	6660
tttcattaag atggtgagta aatgcacccg tactagaatc ggtaagcttt acaccaaca	6720
taagaataaa atcagcagat tccacaaatt cttcaagtt tggctctgac agagtaccgt	6780
tgtaaatccc caaaaatgag ggcaatgctt catcaacaga tgatttacca aagttcaaag	6840
tagtaatagg taacttagtc ttgaaataa actgagtaac agtcttctct aggccgaacg	6900
atataatttc atggcctgtg attacaattg gtttcttggc attcttcaga ctttctgtg	6960
ttttgttcag aatctcttga tcagatgtat tcgacgtgga atttcccttc ttaagaggca	7020
aggatggttt ttcagcctta gcggcagcta catctacagg taaattgatg taaaccggct	7080
ttctttctct tagtaaggca gacaacactc tatcaatttc aacagttgca ttctcggtg	7140
tcaataaagt cctggcagca gtaaccggtt cgtgcatctt cataaagtgc ttgaaatcac	7200
catcagccaa cgtatggatg acaaaacttac ctctgttctg cactttcgag gtaggagatc	7260
ccacgatctc aacaacaggc aggtttctcag cataggagcc cgctaagcca ttaactgcgg	7320
ataattcgcc aacaccaaat gtagtcaaga atgccgcagc ctttttcgtt ctgctgacc	7380

cgtcggccat ataggaggca tttaactcat tagcatttcc caccatttc atatctttgt	7440
gtgaaataat ttgatctaga aattgcaaat ttagtcacc tggctactcg aatatttctt	7500
ctatacctaa ttcgtgtaat ctgtccaaca gatagtcacc tactgtatac attttgttta	7560
ctagtttatg tgtgtttatt cgaaactaag ttcttgggtg tttaaaacta aaaaaagac	7620
taactataaa agtagaattt aagaagttaa agaaatagat ttacagaatt acaatcaata	7680
cctaccgtct ttatatactt attagtcaag taggggaata atttcaggga actggtttca	7740
accttttttt tcagcttttt ccaaatcaga gagagcagaa ggtaatagaa ggtgtaagaa	7800
aatgagatag atacatgcgt gggtaattg ccttgtgtca tcatttactc caggcagggt	7860
gcatcactcc atigaggttg tgcccgtttt ttgctgttt gtgcccctgt tctctgtagt	7920
tgcgctaaga gaatggacct atgaactgat ggttgggtgaa gaaaacaata ttttgggtgt	7980
gggattcttt tttttctgg atgccagctt aaaaagcggg ctccattata tttagtggat	8040
gccaggaata aactgttcac ccagacacct acgatgttat atattctgtg taaccgccc	8100
cctattttgg gcatgtacgg gttacagcag aattaaaagg ctaatttttt gactaaataa	8160
agttaggaaa atcactacta ttaattatit acgtattctt tgaaatggca gtattgataa	8220
tgataaactc gaactgaaaa agcgtgtttt ttattcaaaa tgattctaac tcccttacgt	8280
aatcaaggaa tctttttgcc ttggcctccg cgtcattaaa cttcttgttg ttgacgctaa	8340
cattcaacgc tagtatatat tcgttttttt caggtaagtt ctttcaacg ggtcttactg	8400
atgaggcagt cgcgtctgaa cctgttaaga ggtcaaatat gtcttcttga ccgtacgtgt	8460
cttgcagtgt attagctttg ggaatttgca tcaagtcata ggaaaattta aatcttggct	8520
ctcttgggct caaggtgaca aggtcctcga aaataggcg cgccccaccg cgggtggagct	8580
cagcttttgt tccctttagt gaggtttaat tgcgcgttg gcgtaatcat ggtcatagct	8640
gtttctgtg tgaaattgtt atccgctcac aattccacac aacatacgag ccggaagcat	8700
aaagtgtaaa gcctgggggtg cctaattgagt gagctaactc acattaattg cgttgcgctc	8760
actgcccgtt ttccagtcgg gaaacctgtc gtgccagctg cattaatgaa tcggccaacg	8820
cgcggggaga ggcgggtttgc gtattggcg ctcttccgtt tcctcgctca ctgactcgct	8880
gcgctcggtc gttcggctgc ggcgagcgg atcagctcac tcaaaggcgg taatacggtt	8940
atccacagaa tcaggggata acgcaggaaa gaacatgtga gcaaaaggcc agcaaaaggc	9000
caggaaccgt aaaaaggccg cgttgcctggc gtttttccat aggtctcgcc cccctgacga	9060
gcatcacaaa aatcgacgct caagtcagag gtggcgaaac ccgacaggac tataaagata	9120

ccaggcggtt ccccttgga gctccctcgt gcgctctcct gttccgaccc tgccgcttac	9180
cggatacctg tccgcctttc tcccttcggg aagcgtggcg ctttctcata gctcacgctg	9240
taggtatctc agttcgggtg aggtcgttcg ctccaagctg ggctgtgtgc acgaaccccc	9300
cgttcagccc gaccgctgcg ccttatccgg taactatcgt cttgagtcca acccggttaag	9360
acacgactta tcgccactgg cagcagccac tggtaacagg attagcagag cgaggtatgt	9420
aggcgggtgt acagagttct tgaagtgggt gcctaactac ggctacacta gaaggacagt	9480
atttggatc tcgctctcgc tgaagccagt taccttcgga aaaagagttg gtagctcttg	9540
atccggcaaa caaaccaccg ctggtagcgg tggttttttt gtttcaagc agcagattac	9600
gcgcagaaaa aaaggatctc aagaagatcc tttgatcttt tctacggggt ctgacgtca	9660
gtggaacgaa aactcacgtt aagggtttt ggtcatgaga ttatcaaaaa ggatcttcac	9720
ctagatcctt ttaaattaaa aatgaagttt taaatcaatc taaagtatat atgagtaaac	9780
ttggtctgac agttaccaat gcttaatcag tgaggcacct atctcagcga tctgtctatt	9840
tcgttcaccc atagtgcct gactccccgt cgtgtagata actacgatac gggagggtt	9900
accatctggc ccagtgctg caatgatacc gcgagacca cgctcaccgg ctccagattt	9960
atcagcaata aaccagccag ccggaagggc cgagcgcaga agtggctctg caactttatc	10020
cgctccatc cagttctatta attgttgccg ggaagctaga gtaagtagtt cgccagttaa	10080
tagtttgccg aacgttgttg ccattgctac aggcatcgtg gtgtcacgt cgtcgtttgg	10140
tatggcttca ttcagctccg gttcccaacg atcaaggcga gttacatgat ccccatgtt	10200
gtgcaaaaaa gcggttagct ccttcgggtc tccgatcgtt gtcagaagta agttggccgc	10260
agtgttatca ctcatggta tggcagcact gcataattct cttactgtca tgccatccgt	10320
aagatgcttt tctgtgactg gtgagtactc aaccaagtca ttctgagaat agtgtatgcg	10380
gcgaccgagt tgctcttgcc cggcgtcaat acgggataat accgcgccac atagcagaac	10440
tttaaaagtg ctcatcattg gaaaacgttc ttcggggcga aaactctcaa ggatcttacc	10500
gctgttgaga tccagttcga tgtaaccac tcgtgcaccc aactgatctt cagcatcttt	10560
tactttcacc agcgtttctg ggtgagcaaa aacaggaagg caaaatgccg caaaaaaggg	10620
aataagggcg acacggaaat gttgaatact catactcttc ctttttcaat attattgaag	10680
catttatcag ggttattgtc tcatgagcgg atacatattt gaatgtattt agaaaaataa	10740
acaaataggg gtccgcgca catttccccg aaaagtcca cctgggtcct tttcatcacg	10800
tgctataaaa ataattataa tttaaatttt ttaataataa tataaaatt aaaaatagaa	10860
agtaaaaaaa gaaattaaag aaaaaatagt tttgttttc cgaagatgta aaagactcta	10920
gggggatcgc caacaaatc taccttttat cttgctcttc ctgctctcag gtattaatgc	10980

cgaattgttt caicttgtct gtgtagaaga ccacacacga aaatcctgtg attttacatt 11040
 ttacttatcg ttaatcgaat gtatatctat ttaatctgct tttcttgtct aataaatata 11100
 tatgtaaagt acgctttttg ttgaaatttt ttaaaccctt gtttattttt ttttcttcat 11160
 tccgtaactc ttctacctic tttatttact ttctaaaatc caaatacaaa acataaaaat 11220

aaataaacac agagtaaatt cccaaattat tccatcatta aaagatacga ggcgcgtgta 11280
 agttacaggc aagcgatccg tcctaagaaa ccattattat catgacatta acctataaaa 11340
 ataggcgtat cacgaggccc tttcgtc 11367

<210> 817
 <211> 9491
 <212> DNA
 <213> Artificial Sequence
 <220><223> Synthetic construct pRS423 FBA ilvD(Strep)
 <400> 817

tcgcgcgttt cggatgatgac ggtgaaaacc tctgacacat gcagctcccg gagacggtca 60
 cagcttgtct gtaagcggat gccgggagca gacaagcccg tcagggcgcg tcagcgggtg 120

ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180
 accataaatt cccgttttaa gagcttgggtg agcgctagga gtcactgcca ggtatcgttt 240
 gaacacggca ttagtcaggg aagtcataac acagtccttt cccgcaattt tctttttcta 300
 ttactcttgg cctcctctag tacactctat atttttttat gcctcggtaa tgattttcat 360
 tttttttttt cccctagcgg atgactcttt ttttttctta gcgattggca ttatcacata 420
 atgaattata cattatataa agtaatgtga tttcttcgaa gaatatacta aaaaatgagc 480
 aggcaagata aacgaaggca aagatgacag agcagaaaag cctagtaaag cgtattacaa 540

atgaaaccaa gattcagatt gcgatctctt taaagggtgg tcccctagcg atagagcact 600
 cgatcttccc agaaaaagag gcagaagcag tagcagaaca ggccacacaa tcgcaagtga 660
 ttaacgtcca cacaggata gggtttctgg accatatgat acatgctctg gccaagcatt 720
 ccggctggtc gctaactggt gagtgacattg gtgacttaca catagacgac catcacacca 780
 ctgaagactg cgggattgct ctcggtcaag cttttaaaga ggccctactg gcgcgtggag 840
 taaaaagggt tggatcagga tttgcgcctt tggatgaggc actttccaga gcggtggtag 900
 atctttcgaa caggccgtac gcagttgtcg aacttggttt gcaaaggag aaagtaggag 960

atctctcttg cgagatgac cgcattttt ttgaaagctt tgcagaggt agcagaatta 1020
 ccctccacgt tgattgtctg cgaggcaaga atgatcatca ccgtagttag agtgcgttca 1080

aggctcttgc ggttgcata agagaagcca cctcgcccaa tggtaaccaac gatgttcctt	1140
ccaccaaagg tgttcttatg tagtgacacc gattatttaa agctgcagca tacgatatat	1200
atacatgtgt atatatgtat acctatgaat gtcagtaagt atgtatacga acagtatgat	1260
actgaagatg acaaggtaat gcatcattct atacgtgtca ttctgaacga ggcgcgcttt	1320
ccttttttct ttttgccttt tctttttttt tctcttgaac tcgacggatc tatgcggtgt	1380
gaaataccgc acagatgcgt aaggagaaaa taccgcatca ggaaattgta aacgttaata	1440
ttttgttaaa attcgcgtta aatttttgtt aaatcagctc attttttaac caataggccg	1500
aaatcggaac aatcccttat aaatcaaaag aatagaccga gatagggttg agtgttgttc	1560
cagtttgga caagagtcca ctattaaaga acgtggactc caacgtcaaa gggcgaaaaa	1620
ccgtctatca gggcgatggc ccactacgtg aaccatcacc ctaatcaagt tttttggggt	1680
cgaggtgccg taaagcacta aatcggaacc ctaaaggag ccccgattt agagcttgac	1740
ggggaaagcc gggaacgtg gcgagaaagg aagggaagaa agcgaaagga gcgggcgcta	1800
gggcgctggc aagtgtagcg gtcacgtgc gcgtaaccac cacacccgcc gcgcttaatg	1860
cgccgtaca gggcgcgtcg gcgcattcgc cattcaggct gcgcaactgt tgggaagggc	1920
gatcggctgc ggctcttcg ctattacgcc agctggcgaa agggggatgt gctgcaaggc	1980
gattaagttg ggtaacgcca gggttttccc agtcacgacg ttgtaaaacg acggccagtg	2040
agcgcgcgta atacgactca ctataggcg aattgggtac cgggcccc ctcgaggtcg	2100
acggcgcgcc actggtagag agcgactttg tatgccccaa ttgcgaaacc cgcgatatcc	2160
ttctcgattc tttagtacc gaccaggaca aggaaaagga ggtcgaaacg tttttgaaga	2220
aacaagagga actacacgga agctctaaag atggcaacca gccagaaact aagaaaatga	2280
agttgatgga tccaactggc accgctggct tgaacaacaa taccagcctt ccaacttctg	2340
taaataacgg cggtagccca gtgccaccag taccgttacc tttcggtata cctcctttcc	2400
ccatgtttcc aatgcccttc atgcctcaa cggtactat cacaatcct catcaagctg	2460
acgcaagccc taagaaatga ataacaatac tgacagtact aaataattgc ctacttggct	2520
tcacatacgt tgcatagctc gatatagata ataataataa tgacagcagg attatcgtaa	2580
tacgtaatag ttgaaaatct caaaaatgtg tgggtcatta cgtaataat gataggaatg	2640
ggattcttct atttttctt ttccattct agcagccgtc gggaaaacgt ggcacctct	2700
ctttcgggct caattggagt cacgtgccg tgagcatcct cttttccat atctaacaac	2760
tgagcacgta accaatggaa aagcatgagc ttagcgttgc tccaaaaaag tattggatgg	2820
ttaataccat ttgtctgttc tttctgact ttgactctc aaaaaaaaaa aatctacaat	2880
caacagatcg cttaattac gccctacaa aaacttttt cttcttctt cgcccacgtt	2940

aaatattatc cctcatgttg tctaacggat ttctgcactt gatttattat aaaaagacaa	3000
agacataata ctctctatc aatttcagtt attgttcttc cttgcgttat tcttctgttc	3060
ttctttttct tttgtcatat ataaccataa ccaagtaata catattcaaa ctagtatgac	3120
tgacaaaaaa actcttaaag acttaagaaa tcgtagtctt gtttacgatt caatggttaa	3180
atcacctaata cgtgctatgt tgcgtgcaac tggatatgcaa gatgaagact ttgaaaaacc	3240
tatcgtcggg gtcatttcaa cttgggctga aaacacacct tgtaatatcc acttacatga	3300
ctttggtaaa ctagccaaag tcgggtgtta ggaagctggg gcttggccag ttcagttcgg	3360
aacaatcacg gtttctgatg gaatcgccat gggaacccaa ggaatgcgtt tctccttgac	3420
atctcgtgat attattgcag attctattga agcagccatg ggaggtcata atgcggatgc	3480
ttttgtagcc attggcgggt gtgataaaaa catgcccggt tctgttatcg ctatggctaa	3540
catggatata ccagccattt ttgcttacgg cggaacaatt gcacctggta atttagacgg	3600
caaagatata gatttagtct ctgtctttga aggtgtcggc cattggaacc acggcgatat	3660
gaccaaagaa gaagttaaag ctttggatg taatgcttgt cccggtcctg gaggtgcgg	3720
tggtatgtat actgctaaca caatggcgac agctattgaa gttttgggac ttagccttcc	3780
gggttcatct tctcaccggc ctgaatccgc agaaaagaaa gcagatattg aagaagctgg	3840
tcgcgtggt gtcaaaatgc tcgaaatggg cttaaaacct tctgacattt taacgcgtga	3900
agcttttgaa gatgctatta ctgtaactat ggctctggga ggttcaacca actcaaccct	3960
tcacctctta gctattgcc atgctgctaa tgtggaattg acacttgatg atttcaatac	4020
tttccaagaa aaagtccctc atttggctga ttgaaacct tctggtcaat atgtattcca	4080
agacctttac aaggctcggag gggtagcagc agttatgaaa tatctccta aaaatggctt	4140
ccttcattgt gaccgtatca cttgtactgg caaaacagtc gctgaaaatt tgaagcttt	4200
tgatgattta acacctggc aaaaggttat tatgccgctt gaaaatccta aacgtgaaga	4260
tggctcgcctc attattctcc atggtaactt ggctccagac ggtgccgttg ccaaagtttc	4320
tgggtgtaaaa gtgcgtcgtc atgtcgggtc tgctaaggtc ttttaattctg aagaagaagc	4380
cattgaagct gtcttgaatg atgatattgt tgatgggat gttgtgtcg tacgttttgt	4440
aggaccaaag ggcggctcctg gtatgcctga aatgctttcc ctttcatcaa tgattgttg	4500
taaagggcaa ggtgaaaaag ttgcccttct gacagatggc cgcttctcag gtggtactta	4560
tggctctgtc gtgggtcata tcgtcctga agcacaagat ggcgggtcaa tcgcctacct	4620
gcaaacagga gacatagtca ctattgacca agacactaag gaattacact ttgatattc	4680

cgatgaagag ttaaaacatc gtcaagagac cattgaattg ccaccgctct attcacgcgg	4740
tatccttggg aaatatgctc acatcgtttc gtctgcttct aggggagccg taacagactt	4800
ttggaagcct gaagaaactg gcaaaaaatg ttgtcctggg tgctgtggtt aagcgccgc	4860
gttaattcaa attaatgat atagtttttt aatgagtatt gaatctgttt agaaataatg	4920
gaatattatt tttatttatt tatttatatt attggtcggc tcttttcttc tgaaggtaaa	4980
tgacaaaatg atatgaagga aataatgatt tctaaaattt tacaacgtaa gatattttta	5040
caaaagccta gctcatcttt tgtcatgcac tattttactc acgcttgaaa ttaacggcca	5100
gtccactgcg gagtcatttc aaagtcaccc taatcgatct atcgtttttg atagctcatt	5160
ttggagttcg cgattgtctt ctgttattca caactgtttt aatttttatt tcattctgga	5220
actcttcgag ttctttgtaa agtctttcat agtagcttac tttatctcc aacatattta	5280
acttcatgtc aatttcggct cttaaaattt ccacatcatc aagttcaaca tcacttttta	5340
acttgaattt attctctagc tcttccaacc aagcctcatt gctccttgat ttactggtga	5400
aaagtatac actttgcgcg caatccaggc caaaactttc ctgcaaagaa ttcaccaatt	5460
tctcgacatc atagtacaat ttgttttggt ctcccatcac aatttaatat acctgatgga	5520
ttcttatgaa gcgctgggta atggacgtgt cactctactt cgcctttttc cctactcctt	5580
ttagtacgga agacaatgct aataaataag agggtaataa taatattatt aatcggcaaa	5640
aaagattaaa cgccaagcgt ttaattatca gaaagcaaac gtcgtaccaa tccttgaatg	5700
cttccaatt gtatattaag agtcatcaca gcaacatatt ctgtgtatta aattaattat	5760
tattgatatt tgatattgta taaaaaacc aaatatgtat aaaaaaagtg aataaaaaat	5820
accaagtatg gagaaatata ttagaagtct atacgttaaa ccaccgcggt ggagctccag	5880
cttttgttcc ctttagtgag ggttaattgc gcgcttggcg taatcatggt catagctgtt	5940
tcctgtgtga aattgttacc cgctcacaat tccacacaac ataggagccg gaagcataaa	6000
gtgtaaagcc tggggtgcct aatgagttag gtaactcaca ttaattgcgt tgcgtcact	6060
gcccgttttc cagtccggaa acctgtctgt ccagctgcat taatgaatcg gccaacgcgc	6120
ggggagaggc ggtttgcgta ttgggcgctc ttccgcttcc tcgctcactg actcgctgcg	6180
ctcggtcgtt cggtgcggc gageggtatc agctcactca aaggcggtaa tacggttacc	6240
cacagaatca ggggataacg caggaaagaa catgtgagca aaaggccagc aaaaggccag	6300
gaaccgtaaa aaggccgcgt tgctggcggt ttccatagg ctccgcccc ctgacgagca	6360
tcacaaaaat cgacgtcaa gtcagagggt gcgaaaccgc acaggactat aaagatacca	6420

ggcgtttccc cctggaagct cctcgtgcg ctctcctgtt cggacctgc cgcttaccgg 6480
 atacctgtcc gcctttctcc cttegggaag cgtggcgctt tctcatagct cacgctgtag 6540
 gtatctcagt tcggtgtagg tcgttcgctc caagctgggc tgtgtgcacg aacccccgt 6600
 tcagcccgac cgctgcgcct tatccggtaa ctatcgtctt gagtccaacc cggtaaagaca 6660
 cgacttatcg ccaactggcag cagccactgg taacaggatt agcagagcga ggtatgtagg 6720
 cgggtgctaca gatttcttga agtgggtggc taactacggc tacactagaa ggacagtatt 6780
 tggatatctgc gcctgtctga agccagttac ctteggaaaa agagtggta gctcttgatc 6840

 cggcaaaaa accaccgtg gtagcgggtg tttttttgtt tgcaagcagc agattacgcg 6900
 cagaaaaaaa ggatctcaag aagatccttt gatcttttct acgggggtctg acgctcagtg 6960
 gaacgaaaac tcacgttaag ggatttttgg catgagatta tcaaaaagga tcttcaccta 7020
 gatcctttta aattaaaaat gaagttttta atcaatctaa agtatatatg agtaaacttg 7080
 gtctgacagt taccaatgct taatcagtga ggcacctatc tcagcgatct gcttatttcg 7140
 ttcattcata gttgcctgac tccccgtcgt gtagataact acgatacggg agggcttacc 7200
 atctggcccc agtgcctgca tgataccgcg agaccacgc tcaccggctc cagatttacc 7260

 agcaataaac cagccagccg gaaggccga ggcagaaagt ggtcctgcaa ctttatccgc 7320
 ctccatccag tctattaatt gttgccggga agctagagta agtagttcgc cagttaatag 7380
 tttgcgcaac gttgttgcca ttgctacagg catcgtggtg tcacgctcgt cgtttggtat 7440
 ggcttcattc agtccgggtt cccaacgac aaggcgagtt acatgatccc ccatgttgtg 7500
 caaaaaagcg gttagctcct tcggctcctc gatcgttgtc agaagtaagt tggccgcagt 7560
 gttatcactc atggttatgg cagcactgca taattctctt actgtcatgc catccgtaag 7620
 atgcttttct gtgactggtg agtactcaac caagtcattc tgagaatagt gtatgcggcg 7680

 accgagtgtc tcttgccccg cgtcaatacg ggataatacc gcgccacata gcagaacttt 7740
 aaaagtgtc atcattggaa aacgttcttc ggggcgaaaa ctctcaagga tcttaccgct 7800
 gttgagatcc agttcgatg aaccactcg tgcaccaac tgatcttcag catcttttac 7860
 tttcaccagc gtttctgggt gagcaaaaac aggaaggcaa aatgccgcaa aaaagggaat 7920
 aaggcgaca cggaatgtt gaatactcat actcttcctt tttcaatatt attgaagcat 7980
 ttatcagggt tattgtctca tgagcggata catatttgaa tgtatttaga aaaataaaca 8040
 aatagggtt ccgcgcacat tccccgaaa agtgcacct gaacgaagca tctgtgcttc 8100

 atttttaga acaaaaatgc aacgcgagag cgctaatttt tcaacaaaag aatctgagct 8160
 gcatttttac agaacagaaa tgcaacgcga aagcgtatt ttaccaacga agaactctgtg 8220
 cttcatTTTT gtaaaacaaa aatgaacgc gagagcgcta atttttcaaa caaagaatct 8280

gagctgcatt ttacagaac agaaatgcaa cgcgagagcg ctattttacc aacaaagaat 8340
ctatacttct tttttgttct acaaaaatgc atcccagagag cgctattttt ctaacaaagc 8400
atcttagatt acttttttct tcttttgtgc gctctataat gcagctctctt gataactttt 8460
tgcactgtag gtccgttaag gttagaagaa ggctactttg gtgtctattt tctcttccat 8520

aaaaaaagcc tgactccact tcccgcgttt actgattact agcgaagctg cgggtgcatt 8580
ttttcaagat aaaggcatcc ccgattatat tctataccga tgtggattgc gcatactttg 8640
tgaacagaaa gtgatagcgt tgatgattct tcatttgtca gaaaattatg aacggtttct 8700
tctattttgt ctctatatac tacgtatagg aaatgtttac attttcgtat tgttttcgat 8760
tcactctatg aatagttctt actacaattt ttttgtctaa agagtaatac tagagataaa 8820
cataaaaaat gtagaggctg agtttagatg caagttcaag gagcgaaagg tggatgggta 8880
ggttatatag ggatatagca cagagatata tagcaaagag atacttttga gcaatgtttg 8940

tggaagcggg attcgcaata ttttagtagc tcgttacagt ccggtgcgtt tttggttttt 9000
tgaaagtgcg tcttcagagc gcttttgggt ttcaaaagcg ctctgaagtt cctatacttt 9060
ctagagaata ggaacttcgg aataggaact tcaaagcgtt tccgaaaacg agcgttccg 9120
aaaatgcaac gcgagctgcg cacatacagc tcactgttca cgtcgacact atatctgcgt 9180
gttgccgtga tatatatata catgagaaga acggcatagt gcgtgtttat gcttaaatgc 9240
gtacttatat gcgtctattt atgtaggatg aaaggtagtc tagtacctcc tgtgatatta 9300
tcccattcca tgcgggggtat cgtatgcttc ctccagcact acccttttagc tgttctatat 9360

gtgcccactc ctcaattgga ttagtctcat cttcaatgc tatcatttcc ttigtatattg 9420
gatcatctaa gaaaccatta ttatcatgac attaacctat aaaaataggc gtatcacgag 9480
gccctttcgt c 9491

<210> 818
<211> 1000
<212> DNA
<213> Artificial Sequence
<220><223> FBA1 terminator nt 4861 to 5860
<400> 818

gttaattcaa attaatgat atagtttttt aatgagtatt gaatctgttt agaaataatg 60
gaatattatt tttatttatt tatttatatt attggtcggc tcttttcttc tgaaggtaaa 120

tgacaaaatg atatgaagga aataatgatt tctaaaattt tacaacgtaa gatattttta 180
caaaagccta gctcatcttt tgtcatgcac tattttactc acgcttgaaa ttaacggcca 240

gtccactgcg gagtcatttc aaagtcaccc taatcgatct atcgtttttg atagtcatt 300

ttggagttcg cgattgtctt ctgttattca caactgtttt aatttttatt tcattctgga 360

actcttcgag ttctttgtaa agtctttcat agtagcttac ttatcctcc aacatattta 420

acttcatgtc aatttcggct cttaaatttt ccacatcatc aagttcaaca tcattcttta 480

acttgaattt attctctagc tcttccaacc aagcctcatt gctccttgat ttactggatga 540

aaagtatac actttgcgcg caatccaggt caaaactttc ctgcaaagaa ttcaccaatt 600

tctcgacatc atagtacaat ttgttttgtt ctcccatcac aatttaatat acctgatgga 660

ttcttatgaa gcgctgggta atggacgtgt cactctactt cgcctttttc cctactcctt 720

ttagtacgga agacaatgct aataaataag agggtataaa taatattatt aatcggcaaa 780

aaagattaaa cgccaagcgt ttaattatca gaaagcaaac gtcgtaccaa tccttgaatg 840

cttccaatt gtatattaag agtcacacaca gcaacatatt cttgttatta aattaattat 900

tattgatttt tgatattgta taaaaaacc aaatatgtat aaaaaaagtg aataaaaaat 960

accaagtatg gagaaatata ttagaagtct atacgttaaa 1000

<210> 819

<211> 1713

<212> DNA

<213> Artificial Sequence

<220><223> Streptococcus mutans ilvD coding region

<400> 819

atgactgaca aaaaaactct taaagactta agaaatcgta gttctgttta cgattcaatg 60

gttaaatcac ctaatcgtgc tatgttgcgt gcaactggta tgcaagatga agactttgaa 120

aaacctatcg tcggtgtcat ttcaacttgg gctgaaaaca caccttgtaa tatccactta 180

catgactttg gtaaaactagc caaagtcggt gttaaggaag ctggtgcttg gccagttcag 240

ttcggaacaa tcacggtttc tgatggaatc gccatgggaa cccaaggaat gcgtttctcc 300

ttgacatctc gtgatattat tgcagattct attgaagcag ccatgggagg tcataatgcg 360

gatgcttttg tagccattgg cggttgtgat aaaaacatgc ccggttctgt tatcgctatg 420

gctaacaatg atatccagc catttttgct tacggcggaa caattgcacc tggttaattta 480

gacggcaaag atatcgattt agtctctgtc ttgaagggtg tcggccattg gaaccacggc 540

gatatgacca aagaagaagt taaagctttg gaatgtaatg cttgtcccg tcttgaggc 600

tgcggtggta tgtatactgc taacacaatg gcgacagcta ttgaagtttt gggacttagc 660

cttccgggtt catcttctca cccggtgaa tccgcagaaa agaaagcaga tattgaagaa 720

gctggtcgcg ctgttgtcaa aatgctcgaa atgggcttaa aaccttctga cattttaacg 780
 cgtgaagctt ttgaagatgc tattactgta actatggctc tgggaggttc aaccaactca 840
 acccttcacc tcttagctat tgcccatgct gctaagtgg aattgacact tgatgatttc 900
 aatactttcc aagaaaaagt tcctcatttg gctgatttga aaccttctgg tcaatatgta 960
 ttccaagacc ttacaaggt cggaggggta ccagcagtta tgaaatatct ccttaaaaat 1020
 ggcttccttc atggtgaccg taccacttgt actggcaaaa cagtcgctga aaatttgaag 1080

gcttttgatg atttaacacc tggcmetaag gttattatgc cgcttgmeta tcctaacgt 1140
 gaagatggtc cgctcattat tctccatggt aacttggctc cagacgggtc cgttgmeta 1200
 gtttctgggt metaagtgcg tcgtcatgtc ggtcctgcta aggtctttaa ttctgaagaa 1260
 gaagccattg aagctgtctt gaatgatgat attgttgatg gtgatgttgt tgcgtacgt 1320
 ttgttaggac caaaggcggt tcctgggtatg cctgaaatgc ttccctttc atcaatgatt 1380
 gtggtaaaag ggcaaggatga aaaagttgcc cttctgacag atggccgctt ctcaggtggt 1440
 acttatggtc ttgtcgtggg tcatatcgct cctgaagcac aagatggcgg tccaatcgcc 1500

tacctgmeta caggagacat agtcactatt gaccaagaca ctaaggaatt acactttgat 1560
 atctccgatg aagagttmeta acatcgtaaa gagaccattg aattgccacc gctctattca 1620
 cgcggtatcc ttggtmeta tgctcacatc gtttctctg cttctagggg agccgtaaca 1680
 gacttttggg agcctgaaga aactggmeta aaa 1713

<210> 820

<211> 571

<212> PRT

<213> Artificial Sequence

<220><223> Streptococcus mutans ilvD

<400> 820

Met Thr Asp Lys Lys Thr Leu Lys Asp Leu Arg Asn Arg Ser Ser Val

1 5 10 15

Tyr Asp Ser Met Val Lys Ser Pro Asn Arg Ala Met Leu Arg Ala Thr

20 25 30

Gly Met Gln Asp Glu Asp Phe Glu Lys Pro Ile Val Gly Val Ile Ser

35 40 45

Thr Trp Ala Glu Asn Thr Pro Cys Asn Ile His Leu His Asp Phe Gly

50 55 60

Lys Leu Ala Lys Val Gly Val Lys Glu Ala Gly Ala Trp Pro Val Gln

65	70	75	80
Phe Gly Thr Ile Thr Val Ser Asp Gly Ile Ala Met Gly Thr Gln Gly			
	85	90	95
Met Arg Phe Ser Leu Thr Ser Arg Asp Ile Ile Ala Asp Ser Ile Glu			
	100	105	110
Ala Ala Met Gly Gly His Asn Ala Asp Ala Phe Val Ala Ile Gly Gly			
	115	120	125
Cys Asp Lys Asn Met Pro Gly Ser Val Ile Ala Met Ala Asn Met Asp			
	130	135	140
Ile Pro Ala Ile Phe Ala Tyr Gly Gly Thr Ile Ala Pro Gly Asn Leu			
145	150	155	160
Asp Gly Lys Asp Ile Asp Leu Val Ser Val Phe Glu Gly Val Gly His			
	165	170	175
Trp Asn His Gly Asp Met Thr Lys Glu Glu Val Lys Ala Leu Glu Cys			
	180	185	190
Asn Ala Cys Pro Gly Pro Gly Gly Cys Gly Gly Met Tyr Thr Ala Asn			
	195	200	205
Thr Met Ala Thr Ala Ile Glu Val Leu Gly Leu Ser Leu Pro Gly Ser			
	210	215	220
Ser Ser His Pro Ala Glu Ser Ala Glu Lys Lys Ala Asp Ile Glu Glu			
225	230	235	240
Ala Gly Arg Ala Val Val Lys Met Leu Glu Met Gly Leu Lys Pro Ser			
	245	250	255
Asp Ile Leu Thr Arg Glu Ala Phe Glu Asp Ala Ile Thr Val Thr Met			
	260	265	270
Ala Leu Gly Gly Ser Thr Asn Ser Thr Leu His Leu Leu Ala Ile Ala			
	275	280	285
His Ala Ala Asn Val Glu Leu Thr Leu Asp Asp Phe Asn Thr Phe Gln			
	290	295	300
Glu Lys Val Pro His Leu Ala Asp Leu Lys Pro Ser Gly Gln Tyr Val			
305	310	315	320

Phe Gln Asp Leu Tyr Lys Val Gly Gly Val Pro Ala Val Met Lys Tyr
325 330 335

Leu Leu Lys Asn Gly Phe Leu His Gly Asp Arg Ile Thr Cys Thr Gly
340 345 350

Lys Thr Val Ala Glu Asn Leu Lys Ala Phe Asp Asp Leu Thr Pro Gly
355 360 365

Gln Lys Val Ile Met Pro Leu Glu Asn Pro Lys Arg Glu Asp Gly Pro
370 375 380

Leu Ile Ile Leu His Gly Asn Leu Ala Pro Asp Gly Ala Val Ala Lys
385 390 395 400

Val Ser Gly Val Lys Val Arg Arg His Val Gly Pro Ala Lys Val Phe
405 410 415

Asn Ser Glu Glu Glu Ala Ile Glu Ala Val Leu Asn Asp Asp Ile Val
420 425 430

Asp Gly Asp Val Val Val Val Arg Phe Val Gly Pro Lys Gly Gly Pro
435 440 445

Gly Met Pro Glu Met Leu Ser Leu Ser Ser Met Ile Val Gly Lys Gly
450 455 460

Gln Gly Glu Lys Val Ala Leu Leu Thr Asp Gly Arg Phe Ser Gly Gly
465 470 475 480

Thr Tyr Gly Leu Val Val Gly His Ile Ala Pro Glu Ala Gln Asp Gly
485 490 495

Gly Pro Ile Ala Tyr Leu Gln Thr Gly Asp Ile Val Thr Ile Asp Gln
500 505 510

Asp Thr Lys Glu Leu His Phe Asp Ile Ser Asp Glu Glu Leu Lys His
515 520 525

Arg Gln Glu Thr Ile Glu Leu Pro Pro Leu Tyr Ser Arg Gly Ile Leu
530 535 540

Gly Lys Tyr Ala His Ile Val Ser Ser Ala Ser Arg Gly Ala Val Thr
545 550 555 560

Asp Phe Trp Lys Pro Glu Glu Thr Gly Lys Lys

565

570

<210> 821

<211> 2145

<212> DNA

<213> Artificial Sequence

<220><223> GPM-sadB-ADHt segment

<400> 821

gcatgcttgc atttagtcgt gcaatgtatg actttaagat ttgtgagcag gaagaaaagg 60

gagaatcttc taacgataaa cccttgaaaa actgggtaga ctacgctatg ttgagttgct 120

acgcaggctg cacaattaca cgagaatgct cccgcctagg atttaaggct aagggacgtg 180

caatgcagac gacagatcta aatgaccgtg tcggtgaagt gttcgccaaa cttttcggtt 240

aacacatgca gtgatgcacg cgcgatgggt ctaagttaca tatatatata tatatatata 300

tatagccata gtgatgtcta agtaaccttt atggtatatt tcttaatgtg gaaagatact 360

agcgcgcgca cccacacaca agcttcgtct tttcttgaag aaaagaggaa gctcgctaaa 420

tgggattcca ctttccgttc cctgccagct gatggaaaaa ggtagtgga acgatgaaga 480

ataaaaagag agatccactg aggtgaaatt tcagctgaca gcgagtttca tgatcgtgat 540

gaacaatggt aacgagttgt ggctgttgcc agggagggtg gttctcaact tttaatgtat 600

ggccaaatcg ctacttgggt ttgttatata acaaagaaga aataatgaac tgattctctt 660

cctccttctt gtcctttctt aattctgttg taattacctt cctttgtaat ttttttgtta 720

attattcttc ttaataatcc aaacaaacac acatattaca atagctagct gaggatgaag 780

gcattagttt atcatgggga tcacaaaatt tcgttagaag acaaaccaaa acccactctg 840

cagaaacca cagacgttgt ggtaggggtg ttgaaaacaa caatttgcgg tactgacttg 900

ggaatataca aaggtaagaa tcctgaagtg gcagatggca gaatcctggg tcatgagggc 960

gttggcgtea ttgaagaagt gggcgaatcc gtgacacaat tcaaaaaggg ggataaagtt 1020

ttaatctcct gcgttactag ctgtggatcg tgtgattatt gcaagaagca actgtattca 1080

caactgtagag acggtggctg gattttaggt tacatgatcg acggtgtcca agccgaatac 1140

gtcagaatac cacatgtcga caattcattg tataagatcc cgcaaactat cgatgatgaa 1200

attgcagtac tactgtccga tattttacct actggacatg aaattgggtg tcaatatggt 1260

aacgttcaac caggcgatgc ttagcaatt gtaggagcag gtcctgttgg aatgtcagtt 1320

ttgttaactg ctcaatttta ctgcctagt accattattg ttatcgacat ggacgaaaac 1380

cgtttacaat tagcgaagga gcttggggcc acacacacta ttaactccgg tactgaaat 1440

gttgtcgaag ctgtgcatcg tatagcagcc gaaggagtgg atgtagcaat agaagctgtt 1500
 ggtataccgc caacctggga catctgtcag gaaattgtaa aaccggcgcc tcatattgcc 1560
 aacgtgggag ttcatggtgt taaggtggac ttgaaattc aaaagtgttg gattaagaat 1620
 ctaaccatca ccactggttt ggtaaacact aatactaccc caatgttgat gaaggtagcc 1680
 tctactgata aattgccttt aaagaaaatg attactcaca ggtttgagtt agctgaaatc 1740

gaacacgcat atcaggtttt ctggaatggc gctaaagaaa aagctatgaa gattattcta 1800
 tctaattgcag gtgccgccta attaatgaag agtaagcgaa tttcttatga tttatgattt 1860
 ttattattaa ataagttata aaaaaataa gtgtatacaa attttaaaagt gactcttagg 1920
 ttttaaaacg aaaattctta ttcttgagta actctttcct gtaggtcagg ttgctttctc 1980
 aggtatagca tgaggctcgt cttattgacc acacctctac cggcatgccg agcaaatgcc 2040
 tgcaaatcgc tccccatttc acccaattgt agatatgcta actccagcaa tgagttgatg 2100
 aatctcgggtg tgiattttat gtcctcagag gacaacacct gtggt 2145

<210> 822

<211> 4280

<212> DNA

<213> Artificial Sequence

<220><223> pUC19-URA3r

<400> 822

ggggatcctc tagagtcgac ctgcaggcat gcaagcttgg cgtaatcatg gtcatactg 60
 ttctctgtgt gaaattgtta tccgctcaca attccacaca acatacgagc cggaagcata 120
 aagtgtaaag cctgggggtgc ctaatgagt agctaactca cattaattgc gttgcgctca 180
 ctgcccgttt tccagtcggg aaacctgtcg tgccagctgc attaatgaat cggccaacgc 240
 gcggggagag gcggtttgcg tattgggcgc tcttccgctt cctcgctcac tgactcgctg 300
 cgctcggctg ttcggctcgc gcgagcggta tcagctcact caaaggcggg aatacggtta 360

tccacagaat caggggataa cgcaggaaaag aacatgtgag caaaaggcca gaaaaggcc 420
 aggaaccgta aaaaggccgc gttgctggcg tttttccata ggctccgccc cctgacgag 480
 catcacaaaa atcgacgtc aagtcagagg tggcgaaacc cgacaggact ataaagatac 540
 caggcgtttc cccctggaag ctccctctgt cgtctcctg ttccgacct gccgcttacc 600
 ggatactgt ccgcctttct ccttcggga agcgtggcg tttctcatag ctacgctgt 660
 aggtatctca gttcgggtga ggtcgttcgc tccaagctgg gctgtgtgca cgaaccccc 720
 gttcagcccg accgctgcgc cttatccggt aactatcgtc ttgagtcaa cccgtaaga 780

cacgacttat cgccactggc agcagccact ggtaacagga ttagcagagc gaggtatgta	840
ggcggtgcta cagagttctt gaagtgggtg cctaactacg gctacactag aaggacagta	900
tttgggtatct gcgctctgct gaagccagtt accttcggaa aaagagttgg tagctcttga	960
tccggcaaac aaaccaccgc tggtagcggg ggtttttttg tttgcaagca gcagattacg	1020
cgcagaaaaa aaggatctca agaagatcct ttgatctttt ctacggggtc tgacgctcag	1080
tggaacgaaa actcacgtta agggattttg gtcatgagat tatcaaaaag gatcttcacc	1140
tagatccttt taaattaaaa atgaagtttt aaatcaatct aaagtatata tgagtaaact	1200
tggtctgaca gttaccaatg cttaatcagt gaggcaccta tctcagcgat ctgtctattt	1260
cgttcatcca tagttgcttg actccccgtc gtgtagataa ctacgatacg ggagggctta	1320
ccatctggcc ccagtgtgc aatgataccg cgagaccac gctcaccggc tccagattta	1380
tcagcaataa accagccagc cggaagggcc gagcgagaa gtggtcctgc aactttatcc	1440
gcctccatcc agictattaa ttgttgccgg gaagctagag taagtagttc gccagttaat	1500
agtttgcca acgttgttc cattgtaca ggcatcgtgg tgtcacgctc gtcgtttggg	1560
atggcttcat tcagctccgg tteccaacga tcaaggcgag ttacatgatc ccccatgttg	1620
tgcaaaaaag cggttagctc cttcggctct ccgatcgttg tcagaagtaa gttggccgca	1680
gtgttatcac tcatggttat ggcagcactg cataattctc ttactgtcat gccatccgta	1740
agatgctttt ctgtgactgg tgagtactca accaagtcac tctgagaata gtgtatgcgg	1800
cgaccgagtt gctcttggcc ggcgtcaata cgggataata ccgcgccaca tagcagaact	1860
ttaaaagtgc tcatcattgg aaaacgttct tcggggcgaa aactctcaag gatcttaccg	1920
ctgttagat ccagttcgat gtaacccact cgtgcacca actgatcttc agcatctttt	1980
actttacca gcgtttctgg gtgagcaaaa acaggaaggc aaaatgccgc aaaaaggga	2040
ataaggcgca cacgaaatg ttgaatactc atactcttcc ttttcaata ttattgaagc	2100
atttatcagg gttattgtct catgagcgga tacatatgtg aatgtattta gaaaaataaa	2160
caaatagggg ttccgcgcac atttccccga aaagtgccac ctgacgtcta agaaaccatt	2220
attatcatga cattaacctt taaaaatagg cgtatcacga ggccctttcg tctcgcgct	2280
ttcggtgatg acggtgaaaa cctctgacac atgcagctcc cggagacggg cacagcttgt	2340
ctgtaagcgg atgccgggag cagacaagcc cgtcagggcg cgtcagcggg tgttggcggg	2400
tgtcggggct ggcttaacta tgcggcatca gagcagattg tactgagagt gcaccatatg	2460
cggtgtgaaa taccgcacag atgcgtaagg agaaaatacc gcatcaggcg ccattcgcca	2520
ttcaggctgc gcaactgttg ggaaggcgga tcgggtgcggg cctcttcgct attacgccag	2580
ctggcgaaag ggggatgtgc tgcaaggcga ttaagtggg taacgccagg gttttcccag	2640

tcacgacgtt gtaaacgac ggccagtga ttcgagctcg gtacccccgg ctctgagaca	2700
gtagtaggtt agtcatcgct ctaccgacgc gcaggaaaag aaagaagcat tgcggattac	2760
gtatttctaat gttcagcccc cggaacgcca gcaaatcacc acctatgcgc atgatactga	2820
gtcttgtaca cgctgggctt ccagtgact gagagtgcac cataccacag cttttcaatt	2880
caattcatca tttttttttt attctttttt ttgatttcgg tttctttgaa atttttttga	2940
ttcggtaate tccgaacaga aggaagaacg aaggaaggag cacagactta gatttggtata	3000
tatacgcata tgiagtgttg aagaacatg aaattgcca gtattcttaa cccaactgca	3060
cagaacaaaa acctgcagga aacgaagata aatcatgtcg aaagctacat ataaggaacg	3120
tgctgtact catcctagtc ctgttgctgc caagctatit aatatcatgc acgaaaagca	3180
aacaaacttg tigtcttcat tggatgttcg taccaccaag gaattactgg agttagtga	3240
agcattaggt cccaaaattt gtttactaaa aacacatgtg gatatcttga ctgatttttc	3300
catggagggc acagttaagc cgctaaaggc attatccgcc aagtacaatt ttttactctt	3360
cgaagacaga aaatttgctg acattggtaa tacagtcaaa ttgcagtact ctgcgggtgt	3420
atacagaata gcagaatggg cagacattac gaatgcacac ggtgtggtgg gccaggtat	3480
tgttagcggg ttgaagcagg cggcagaaga agtaacaaag gaacctagag gccttttgat	3540
gttagcagaa ttgtcatgca agggctccct atctactgga gaataacta aggggtactgt	3600
tgacattgcg aagagcgaca aagattttgt tatcggtttt attgctcaaa gagacatggg	3660
tggaagagat gaaggttacg attggttgat tatgacaccc ggtgtgggtt tagatgacaa	3720
gggagacgca ttgggtcaac agtatagaac cgtggatgat gtggtctcta caggatctga	3780
cattattatt gtiggaagag gactatttgc aaagggaagg gatgctaagg tagagggtga	3840
acgttacaga aaagcaggct gggaagcata ttgagaaga tgcggccagc aaaactaaaa	3900
aactgtatta taagtaaatg catgtatact aaactcaca attagagctt caatttaatt	3960
atatcagtta ttacctatg cgggtgtgaaa taccgcacag atgcgtaagg agaaaatacc	4020
gcatcaggaa attgtaaacg ttaatatit tttaaaattc gcgttaaatt tttgttaaat	4080
cagctcattt tttaaccaat aggcgaaat cggcaaaatc ttcagccgc ggaacgccag	4140
caaatcacca ccatgcgca tgatactgag tctgttacac gctgggcttc cagtgatgat	4200
acaacagatt agccaaggatg agcacggatg tctaaattag aattacgttt taatatcttt	4260
ttttccatat ctagggctag	4280
<210> 823	
<211> 30	
<212> DNA	

<213> Artificial Sequence

<220><223> Primer 114117-11A

<400> 823

gcatgcttgc atttagtcgt gcaatgtatg 30

<210> 824

<211> 54

<212> DNA

<213> Artificial Sequence

<220><223> Primer 114117-11B

<400> 824

gaacattaga atacgtaatc cgcaatgcac tagtaccaca ggtgttgtcc tctg 54

<210> 825

<211> 54

<212> DNA

<213> Artificial Sequence

<220><223> Primer 114117-11C

<400> 825

cagaggacaa cacctgtggc actagtgcac tgcggattac gtattctaata gttc 54

<210> 826

<211> 28

<212> DNA

<213> Artificial Sequence

<220><223> Primer 114117-11D

<400> 826

caccttggtc aactcgttgt atcatcac 28

<210> 827

<211> 100

<212> DNA

<213> Artificial Sequence

<220><223> Primer 114117-13A

<400> 827

ttttaagccg aatgagtgac agaaaaagcc cacaacttat caagtgatat tgaacaaagg 60

gcgaaacttc gcatgcttgc atttagtcgt gcaatgtatg 100

<210> 828

<211> 98

<212> DNA

<213> Artificial Sequence

<220><223> Primer 114117-13B

<400> 828

cccaattggt aaatattcaa caagagacgc gcagtacgta acatgcgaat tgcgtaattc 60

acggcgataa caccttggct aactcgttgt atcatcac 98

<210> 829

<211> 29

<212> DNA

<213> Artificial Sequence

<220><223> Primer 112590-34G

<400> 829

caaaagccca tgtcccacac caaaggatg 29

<210> 830

<211> 26

<212> DNA

<213> Artificial Sequence

<220><223> Primer 112590-34H

<400> 830

caccatcgcg cgtgcatcac tgcattg 26

<210> 831

<211> 28

<212> DNA

<213> Artificial Sequence

<220><223> Primer 112590-34F

<400> 831

tcggtttttg caatatgacc tgtgggcc 28

<210> 832

<211> 22

<212> DNA

<213> Artificial Sequence

<220><223> Primer 112590-49E

<400> 832

gagaagatgc ggccagcaaa ac

22

<210> 833

<211> 2745

<212> DNA

<213> Artificial Sequence

<220><223> pdc1:: PPDC1-ilvD-FBA1t-URA3r

<400> 833

atgactgaca aaaaaactct taaagactta agaaatcgta gttctgttta cgattcaatg 60
gttaaatcac ctaatcgtgc tatgttgcgt gcaactggta tgcaagatga agactttgaa 120
aaacctatcg tcggtgtcat ttcaacttgg gctgaaaaca caccttgtaa tatccactta 180

catgactttg gtaaaactagc caaagtcggt gttaaggaag ctggtgcttg gccagttcag 240
ttcggaacaa tcacggtttc tgatggaatc gccatgggaa cccaaggaat gcgtttctcc 300
ttgacatctc gtgatattat tgcagattct attgaagcag ccatgggagg tcataatgcg 360
gatgcttttg tagccattgg cggttgtgat aaaaacatgc ccggttctgt tatcgctatg 420
gctaacaatgg atatccagc catttttgct tacggcggaa caattgcacc tggtaattta 480
gacggcaaag atatcgattt agtctctgtc ttgaagggtg tcggccattg gaaccacggc 540
gatatgacca aagaagaagt taaagctttg gaatgtaatg cttgtcccgg tcctggaggc 600

tgcggtggta tgtatactgc taacacaatg gcgacagcta ttgaagtttt gggacttagc 660
cttccgggtt catctttctca cccggtgaa tccgcagaaa agaaagcaga tattgaagaa 720
gctggtcgcg ctgttgtcaa aatgctcgaa atgggcttaa aaccttctga cattttaacg 780
cgtgaagctt ttgaagatgc tattactgia actatggctc tgggagggtc aaccaactca 840
acccttcacc tcttagctat tgcccatgct gctaattgtg aattgacact tgatgatttc 900
aatactttcc aagaaaaagt tcttcatttg gctgatttga aaccttctgg tcaatatgta 960
ttccaagacc ttacaagggt cggaggggta ccagcagtta tgaatatct ccttaaaaat 1020

ggcttccttc atggtgaccg taccattgt actggcaaaa cagtcgctga aaatttgaag 1080
gcttttgatg atttaacacc tggtaaaaag gttattatgc cgcttgaaaa tcctaaacgt 1140
gaagatggtc cgctcattat tctccatggt aacttggtc cagacggtgc cgttgccaaa 1200
gtttctggtg taaaagtgcg tcgtcatgtc ggtcctgcta aggtctttaa ttctgaagaa 1260
gaagccattg aagctgtctt gaatgatgat attgttgatg gtgatgttgt tgtcgtacgt 1320
ttttaggac caaaggcggt tcctggtatg cctgaaatgc tttcccttc atcaatgatt 1380

gttggtaaag ggcaaggatga aaaagttgcc cttctgacag atggccgctt ctacagtggt 1440

acttatggtc ttgtcgtggg tcatatcgct cctgaagcac aagatggcgg tccaatcgcc 1500

tacctgcaaa caggagacat agtcactatt gaccaagaca ctaaggaatt acactttgat 1560

atctccgatg aagagttaaa acatcgtcaa gagaccattg aattgccacc gctctattca 1620

cgcggtatcc ttggtaaata tgctcacatc gtttcgtctg cttctagggg agccgtaaca 1680

gacttttggg agcctgaaga aactggcaaa aaatgttgtc ctggttgctg tggttaagcg 1740

gccgcgttaa ttcaaatata ttgatatagi tttttaatga gtattgaatc tgtttagaaa 1800

taatggaata ttatTTTTat ttattttttt atattattgg tcggctcttt tcttctgaag 1860

gtcaatgaca aaatgatatg aaggaaataa tgatttctaa aattttacaa cgtaagatat 1920

ttttacaaaa gcctagctca tcttttgtca tgcactatTT tactcacgct tgaaattaac 1980

ggccagtcca ctgcggagtc atttcaaagt cactctaate gatctatcgt ttttgatagc 2040

tcatttttga gticgcgatt gtcttctgtt attcacaact gttttaattt ttatttcatt 2100

ctggaactct tcgagttctt tgtaaagtct ttcatagtag cttactttat cctccaacat 2160

atttaacttc atgtcaattt cggtctttaa attttccaca tcatcaagtt caacatcacc 2220

ttttaacttg aattttattt ctagctcttc caaccaagcc tcattgctcc ttgatttact 2280

ggtgaaaagt gatacacttt gcgcgcaatc cagggtcaaaa ctttcttgca aagaattcac 2340

caattttctg acatcatagt acaatttggt ttgttctccc atcacaattt aatatacctg 2400

atggattctt atgaagcgtc gggtaattga cgtgtcactc tacttcgcct ttttccctac 2460

tccttttagt acggaagaca atgctaataa ataagagggt aataataata ttattaatcg 2520

gcaaaaaaga ttaaagccca agcgtttaat tatcagaaag caaacgtcgt accaatcctt 2580

gaatgcttcc caattgtata ttaagagtca tcacagcaac atattcttgt tattaaatta 2640

attattattg atttttgata ttgtataaaa aaaccaataa tgtataaaaa aagtgaataa 2700

aaaataccaa gtatggagaa atatattaga agtctatagc ttaaa 2745

<210> 834

<211> 27

<212> DNA

<213> Artificial Sequence

<220><223> Primer 114117-36D

<400> 834

gacttttggg agcctgaaga aactggc 27

<210> 835

<211> 338

<212> PRT

<213> Artificial Sequence

<220><223> K9_Ophelia_SH

<400> 835

Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn Leu Ser Leu Leu Asp

1 5 10 15

Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser Gln Gly His Ala His

20 25 30

Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val Ile Ile Gly Leu Phe

35 40 45

Glu Gly Ala Glu Glu Trp Lys Arg Ala Glu Glu Gln Gly Phe Glu Val

50 55 60

Tyr Thr Ala Ala Glu Ala Ala Lys Lys Ala Asp Ile Ile Met Ile Leu

65 70 75 80

Ile Pro Asp Glu Lys Gln Ala Thr Met Tyr Lys Asn Asp Ile Glu Pro

85 90 95

Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala His Gly Phe Asn Ile

100 105 110

His Phe Gly Cys Ile Val Pro Pro Lys Asp Val Asp Val Thr Met Ile

115 120 125

Ala Ser Lys Gly Pro Gly His Thr Val Arg Ser Glu Tyr Glu Glu Gly

130 135 140

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

145 150 155 160

Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala Ile Gly Gly Ala Arg

165 170 175

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

180 185 190

Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val Cys Ala Leu Met Gln

195 200 205

Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr Asp Pro Arg Asn Ala

210 215 220
Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile Val Asp Leu Ile Tyr
225 230 235 240
Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile Ser Asn Thr Ala Glu
245 250 255
Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile Thr Glu Asp Thr Lys
260 265 270
Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln Asp Gly Thr Phe Ala

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

<210> 836

<211> 26

<212>

> DNA

<213> Artificial Sequence

<220><223> 112590-30F

<400> 836

ccaggccaat tcaacagact gtcggc

26

<210> 837

<211> 2347

<212> DNA

<213> Artificial Sequence

<220><223> URA3r2 template DNA

<400> 837

gcattgcgga ttacgtattc taatgttcag gtgctggaag aagagctgct taaccgccgc

60

gcccaggggtg aagatccacg ctactttacc ctgcgtcgtc tggatttcgg cggctgtcgt	120
ctttcgtcgg caacgccggt tgatgaagcc tgggacggtc cgctctcctt aaacggtaaa	180
cgtatcgcca cctcttatcc tcacctgtc aagcgttate tgcaccagaa aggcattctt	240
tttaaatcct gcttactgaa cggttctgtt gaagtcgccc cgcgtgccgg actggcggat	300
gcgatttgcg atctggtttc caccggtgcc acgttggaag ctaacggcct gcgcgaagtc	360
gaagttatct atcgctcgaa agcctgcctg attcaacgcg atggcgaaat ggaagaatcc	420
aaacagcaac tgatcgacaa actgctgacc cgtattcagg gtgtgatcca ggccgcgcaa	480
tcaaaataca tcatgatgca cgcaccgacc gaacgtctgg atgaagtcac ggtacctact	540
gagagtgcac cataccacag cttttcaatt caattcatca tttttttttt attctttttt	600
ttgatttcgg ttcttttgaa atttttttga ttccgtaatc tccgaacaga aggaagaacg	660
aaggaaggag cacagactta gattgggtata tatacgcata tgtagtgttg aagaaacatg	720
aaattgccca gtattcttaa cccaactgca cagaacaaaa acctcagga aacgaagata	780
aatcatgtcg aaagctacat ataaggaacg tgctgtact catcctagtc ctgttgctgc	840
caagctatctt aatatcatgc acgaaaagca aacaaacttg tgtgcttcat tggatgttcg	900
taccaccaag gaattactgg agttagttag agcattaggt cccaaaattt gtttactaaa	960
aacacatgtg gatattctga ctgatttttc catggagggc acagttaagc cgctaaaggc	1020
attatccgcc aagtacaatt ttttactctt cgaagacaga aaatttgctg acattggtaa	1080
tacagtcaaa ttgcagtact ctgcgggtgt atacagaata gcagaatggg cagacattac	1140
gaatgcacac ggtgtgggtg gccaggtat ttttagcggg ttgaagcagg cggcagaaga	1200
agtaacaaag gaacctagag gccttttgat gtttagcagaa ttgtcatgca agggctccct	1260
atctactgga gaatatacta aggggtactgt tgacattgcg aagagcgaca aagattttgt	1320
tatcggtctt attgctcaaa gagacatggg tggaagagat gaaggttacg attggttgat	1380
tatgacaccc ggtgtgggtt tagatgacaa gggagacgca ttgggtcaac agtatagaac	1440
cgtggatgat gtggtctcta caggatctga cattattatt gttggaagag gactatttgc	1500
aaaggaagg gatgctaagg tagagggtga acgttacaga aaagcaggct gggaagcata	1560
tttgagaaga tgcggccagc aaaactaaaa aactgtatta taagtaaatg catgtatact	1620
aaactcacia attagagctt caatttaatt atatcagtta ttaccctatg cgggtgtgaaa	1680
taccgcacag atgcgtaagg agaaaatacc gcatcaggaa attgtaaacg ttaatatatt	1740
gttaaaatc gcgttaaat tttgttaaat cagctcattt ttttaaccaat aggccgaaat	1800
cggcaaaatc tctagatgac tggaagaaga gctgcttaac cgccgcgccc aggggtgaaga	1860

tccacgctac tttaccctgc gtcgtctgga tttcggcggc tgtcgtcttt cgctggcaac 1920

gccggttgat gaagcctggg acggtccgct ctctttaaac ggtaaacgta tcgccacctc 1980

ttatctcac ctgctcaagc gttatctcga ccagaaaggc atctctttta aatcctgctt 2040

actgaacggt tctgttgaag tcgccccgcg tgccggactg gcggatgcga ttgctgatct 2100

ggtttccacc ggtgccacgc tggaagctaa cggcctgcgc gaagtcgaag ttatctatcg 2160

ctcgaaagcc tgcctgattc aacgcgatgg cgaatggaa gaatccaaac agcaactgat 2220

cgacaaactg ctgaccgta ttcagggtgt gatccaggcg cgcgaatcaa aatacatcat 2280

gatgcacgca ccgaccgaac gtctggatga agtcatccag tgatgataca acgagtttagc 2340

caaggtg 2347

<210> 838

<211> 80

<212> DNA

<213> Artificial Sequence

<220><223> Primer 114117-45A

<400> 838

cttcgaagaa tataactaaa aatgagcagg caagataaac gaaggcaaag gcattgcgga 60

ttacgtattc taatgttcag 80

<210> 839

<211> 81

<212> DNA

<213> Artificial Sequence

<220><223> Primer 114117-45B

<400> 839

tatacacatg tatatatatc gtatgctgca gctttaaata atcggtgtca caccttggt 60

aactcgttgt atcatcactg g 81

<210> 840

<211> 90

<212> DNA

<213> Artificial Sequence

<220><223> Primer 384

<400> 840

atggttcatt taggtccaaa aaaaccacaa gccagaaagg gttccatggc cgatgtgcc 60

gcattgcgga ttacgtattc taatgttcag	90
<210> 841	
<211> 91	
<212> DNA	
<213> Artificial Sequence	
<220><223> Primer 385	
<400> 841	
ttaagcacgc atgataccaa cggacttacc ttcagcaatt cttttttggg ccaaagcagc	60
caccttggct aactcgttgt atcatcactg g	91
<210> 842	
<211> 24	
<212> DNA	
<213> Artificial Sequence	
<220><223> Primer N869	
<400> 842	
ctaggatgag tagcagcacg ttcc	24
<210> 843	
<211> 26	
<212> DNA	
<213> Artificial Sequence	
<220><223> Primer N871	
<400> 843	
ccaattccgt gatgtctctt tgttgc	26
<210> 844	
<211> 20	
<212> DNA	
<213> Artificial Sequence	
<220	
><223> Primer N946	
<400> 844	
gtgaacgagt tcacaaccgc	20
<210> 845	
<211> 22	
<212> DNA	

<213> Artificial Sequence

<220><223> Primer N947

<400> 845

gttcgttcca gaattatcac gc 22

<210> 846

<211> 26

<212> DNA

<213> Artificial Sequence

<220><223> Primer PDC5:KanMXF

<400> 846

gacttgaata atgcagcggc gcttgc 26

<210> 847

<211> 30

<212>

> DNA

<213> Artificial Sequence

<220><223> Primer PDC5::KanMXR

<400> 847

ccaccctctt caattagcta agatcatagc 30

<210> 848

<211> 90

<212> DNA

<213> Artificial Sequence

<220><223> Primer T001

<400> 848

atgcctacct tgtatactga tatcgaaatc ccacaattga aaatctcttt aaagcaaccg 60

gcattgcgga ttacgtattc taatgttcag 90

<210> 849

<211> 97

<212> DNA

<213> Artificial Sequence

<220><223> T002 primer

<400> 849

ttagttgtcc aaagagagat ttatgtgaac tgcttttggt tgaagatagg tatcaacacc 60

actttgcacc ttggctaact cgttgatca tcactgg 97

<210> 850

<211> 29

<212> DNA

<213> Artificial Sequence

<220><223> T003 primer

<400> 850

gaacgttaca gaaaagcagg ctgggaagc 29

<210> 851

<211> 25

<212> DNA

<213> Artificial Sequence

<220><223> T004 primer

<400> 851

gacctagtaa caacaattgt tccgc 25

<210> 852

<211> 27

<212> DNA

<213> Artificial Sequence

<220><223> T005 primer

<400> 852

ccgtgctacc tcttaatgtg tcacaag 27

<210> 853

<211> 338

<212> PRT

<213> Artificial Sequence

<220><223> K9_Pat_SH

<400> 853

Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn Leu Ser Leu Leu Asp

1 5 10 15

Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser Gln Gly His Ala His

20 25 30

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

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

 Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala His Gly Phe Asn Ile
 100 105 110
 His Phe Gly Cys Ile Val Pro Pro Lys Asp Val Asp Val Thr Met Ile
 115 120 125
 Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser Glu Tyr Glu Glu Gly
 130 135 140
 Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln Asp Ala Thr Gly Lys
 145 150 155 160

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

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

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

<210> 854

<211> 338

<212> PRT

<213> Artificial Sequence

<220><223> K9_Quetnin_SH

<400> 854

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

Tyr Thr Ala Ala Glu Ala Ala Lys Lys Ala Asp Ile Ile Met Ile Leu
 65 70 75 80
 Ile Pro Asp Glu Lys Gln Ala Thr Ile Tyr Lys Asn Asp Ile Glu Pro
 85 90 95
 Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala His Gly Phe Asn Ile
 100 105 110
 His Phe Gly Cys Ile Val Pro Pro Lys Asp Val Asp Val Thr Met Ile
 115 120 125

Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser Glu Tyr Glu Glu Gly
 130 135 140

Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln Asp Ala Thr Gly Lys
 145 150 155 160
 Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala Ile Gly Gly Ala Arg
 165 170 175
 Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu Thr Glu Thr Asp Leu
 180 185 190

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

Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile Thr Glu Asp Thr Lys
 260 265 270
 Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln Asp Gly Thr Phe Ala
 275 280 285
 Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly Ser Gln Val His Phe
 290 295 300
 Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro Ala Glu Val Val Gly
 305 310 315 320

Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp Glu Asp Lys Leu Ile
 325 330 335

Asn Asn

<210> 855

<211> 338

<212> PRT

<213> Artificial Sequence

<220><223> K9_Ralph_SH

<400> 855

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

245 250 255
 Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile Thr Glu Asp Thr Lys
 260 265 270
 Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln Asp Gly Thr Phe Ala
 275 280 285

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

<210> 856

<211> 338

<212> PRT

<213> Artificial Sequence

<220><223> K9_Ursala_SH

<400> 856

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

 Tyr Thr Ala Ala Glu Ala Ala Lys Lys Ala Asp Ile Ile Met Ile Leu
 65 70 75 80
 Ile Pro Asp Glu Lys Gln Ala Thr Met Tyr Lys Asn Asp Ile Glu Pro
 85 90 95
 Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala His Gly Phe Asn Ile

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

<210> 857

<211> 338

<212> PRT

<213> Artificial Sequence

<220><223> K9_Watson_SH

<400> 857

Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn Leu Ser Leu Leu Asp

1 5 10 15

Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser Gln Gly His Ala His

20 25 30

Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val Ile Ile Gly Leu Phe

35 40 45

Glu Gly Ala Glu Glu Cys Lys Arg Ala Glu Glu Gln Gly Phe Glu Val

50 55 60

Tyr Thr Ala Ala Glu Ala Ala Lys Lys Ala Asp Ile Ile Met Ile Leu

65 70 75 80

Ile Pro Asp Glu Lys Gln Ala Thr Met Tyr Lys Asn Asp Ile Glu Pro

85 90 95

Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala His Gly Phe Asn Ile

100 105 110

His Phe Gly Cys Ile Val Pro Pro Lys Asp Val Asp Val Thr Met Ile

115 120 125

Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser Glu Tyr Glu Glu Gly

130 135 140

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

145 150 155 160

Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala Ile Gly Gly Ala Arg

165 170 175

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

180 185 190

Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val Cys Ala Leu Met Gln

195 200 205

Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr Asp Pro Arg Asn Ala

210 215 220

Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile Val Asp Leu Ile Tyr

225 230 235 240

Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile Ser Asn Thr Ala Glu

245 250 255

Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile Thr Glu Asp Thr Lys

260 265 270

Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln Asp Gly Thr Phe Ala

275 280 285

Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly Ser Gln Val His Phe

290 295 300

Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro Ala Glu Val Val Gly

305 310 315 320

Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp Glu Asp Lys Leu Ile

325 330 335

Asn Asn

<210> 858

<211> 1017

<212> DNA

<213> Artificial Sequence

<220><223> K9_Watson_SH

<400> 858

atggctaaga ttactacca agaagactgt aactgtcct tgttggatgg taagactatc	60
gccgttatcg gttacgggtc tcaaggcac gctcatgcc tgaatgctaa ggaatccggt	120
tgtaacgtta tcattggttt attcgaaggt gcggaggagt gcaaaagagc tgaagaacaa	180
ggcttcgaag tctacaccgc tgctgaagct gctaagaagg ctgacatcat tatgatcttg	240
atcccagatg aaaagcaggc taccatgtac aaaaacgaca tcgaacaaa cttggaagcc	300
ggtaacatgt tgatgttcgc tcacggtttc aacatccatt tcggttgtat tgttccacca	360
aaggacgttg atgtcactat gatcgctcca aagggtccag gtcacaccgt tagatccgaa	420

tacgaagaag gtaaagggtg cccatgcttg gttgctgtcg aacaagacgc tactggcaag 480
gcttttgata tggctttggc ctacgcttta gccatcggtg gtgctagagc cgggtgtcttg 540
gaaactacct tcagaaccga aactgaaacc gacttggtcg gtgaacaagc tgttttatgt 600
ggtgggtgtct gcgctttgat gcaggccggt tttgaaacct tggttgaagc cggttacgac 660
ccaagaaacg cttacttcga atgtatccac gaaatgaagt tgatcgttga cttgatctac 720
caatctggtt tctccggtat gcgttactct atctccaaca ctgctgaata cggtgactac 780
attaccggtc caaagatcat tactgaagat accaagaagg ctatgaagaa gattttgtct 840

gacattcaag atggtacctt tgccaaggac ttcttggttg acatgtctga tgctggttcc 900
caggtccact tcaaggctat gagaaagttg gcctccgaac acccagctga agttgtcggc 960
gaagaaatta gatccttgta ctctgggtcc gacgaagaca agttgattaa caactga 1017

<210> 859

<211> 338

<212> PRT

<213> Artificial Sequence

<220><223> K9_Xavier_SH

<400> 859

Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn Leu Ser Leu Leu Asp

1 5 10 15

Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser Gln Gly His Ala His

20 25 30

Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val Ile Ile Gly Leu Phe

35 40 45

Glu Gly Ala Glu Glu Trp Lys Arg Ala Glu Glu Gln Gly Phe Glu Val

50 55 60

Tyr Thr Ala Ala Glu Ala Ala Lys Lys Ala Asp Ile Ile Met Ile Leu

65 70 75 80

Ile Pro Asp Glu Lys Gln Ala Ala Met Tyr Lys Asn Asp Ile Glu Pro

85 90 95

Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala His Gly Phe Asn Ile

100 105 110

His Phe Gly Cys Ile Val Pro Pro Lys Asp Val Asp Val Thr Met Ile

115 120 125

Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser Glu Tyr Glu Glu Gly

130 135 140

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

145 150 155 160

Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala Ile Gly Gly Ala Arg

165 170 175

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

180 185 190

Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val Cys Ala Leu Met Gln

195 200 205

Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr Asp Pro Arg Asn Ala

210 215 220

Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile Val Asp Leu Ile Tyr

225 230 235 240

Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile Ser Asn Thr Ala Glu

245 250 255

Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile Thr Glu Asp Thr Lys

260 265 270

Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln Asp Gly Thr Phe Ala

275 280 285

Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly Ser Gln Val His Phe

290 295 300

Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro Ala Glu Val Val Gly

305 310 315 320

Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp Glu Asp Lys Leu Ile

325 330 335

Asn Asn

<210> 860

<211> 1017

<212> DNA

<213> Artificial Sequence

<220><223> k9_Zeke_SH

<400> 860

atggctaaga ttactacca agaagactgt aacttgicct tgttggatgg taagactatc	60
gccgttatcg gttacggttc tcaaggtcac gctcatgccc tgaatgctaa ggaatccggt	120
tgtaacgtta tcattggttt attcgaaggt gcggaggagt ggaaaagagc tgaagaacaa	180
ggtttcgaag tctacaccgc tgctgaagct gctaagaagg ctgacatcat tatgatcttg	240
atcccagatg aaatgcagcg tgccatgtac aaaaacgaca tcgaaccaa cttggaagcc	300
ggtaacatgt tgatgttcgc tcacggtttc aacatccatt tcggttgtat tgttccacca	360

aaggacgttg atgtcactat gatcgctcca aagggtccag gtcacaccgt tagatccgaa	420
tacgaagaag gtaaaggtgt cccatgcttg gttgctgtcg aacaagacgc tactggcaag	480
gctttggata tggctttggc ctacgcttta gccatcggtg gtgctagagc cgggtgcttg	540
gaaactacct tcagaaccga aactgaaacc gacttggtcg gtgaacaagc tgttttatgt	600
ggtggtgtct gcgctttgat gcaggccggt ttgaaacct tggttgaagc cggttacgac	660
ccaagaaacg cttacttcga atgtatccac gaaatgaagt tgatcgttga ctgatctac	720
caatctgggt tctccggtat gcgttactct atctccaaca ctgctgaata cggtgactac	780

attaccggtc caaagatcat tactgaagat accaagaagg ctatgaagaa gatattgtct	840
gacattcaag atggtacctt tgccaaggac ttcttggttg acatgtctga tgctggttcc	900
cagggtccact tcaaggctat gagaaagtgt gcctccgaac acccagctga agttgtcggt	960
gaagaaatta gatccttgta ctcttggtcc gacgaagaca agttgattaa caactga	1017

<210> 861

<211> 338

<212> PRT

<213> Artificial Sequence

<220><223> K9_Zeke_SH

<400> 861

Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn Leu Ser Leu Leu Asp

1	5	10	15

Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser Gln Gly His Ala His

20	25	30

Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val Ile Ile Gly Leu Phe

35	40	45

Glu Gly Ala Glu Glu Trp Lys Arg Ala Glu Glu Gln Gly Phe Glu Val
50 55 60

Tyr Thr Ala Ala Glu Ala Ala Lys Lys Ala Asp Ile Ile Met Ile Leu
65 70 75 80

Ile Pro Asp Glu Met Gln Ala Ala Met Tyr Lys Asn Asp Ile Glu Pro
85 90 95

Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala His Gly Phe Asn Ile
100 105 110

His Phe Gly Cys Ile Val Pro Pro Lys Asp Val Asp Val Thr Met Ile
115 120 125

Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser Glu Tyr Glu Glu Gly
130 135 140

Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln Asp Ala Thr Gly Lys
145 150 155 160

Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala Ile Gly Gly Ala Arg
165 170 175

Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu Thr Glu Thr Asp Leu
180 185 190

Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val Cys Ala Leu Met Gln
195 200 205

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

Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile Val Asp Leu Ile Tyr
225 230 235 240

Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile Ser Asn Thr Ala Glu
245 250 255

Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile Thr Glu Asp Thr Lys
260 265 270

Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln Asp Gly Thr Phe Ala
275 280 285

Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly Ser Gln Val His Phe

290 295 300
 Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro Ala Glu Val Val Gly
 305 310 315 320
 Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp Glu Asp Lys Leu Ile
 325 330 335

Asn Asn

<210> 862

<211> 1017

<212> DNA

<213> Artificial Sequence

<220><223> K9_Annabel_SH

<400> 862

atggctaaga ttactacca agaagactgt aacttgctct tgttgatgg taagactatc 60
 gccgttatcg gttacggttc tcaaggcac gctcatgccc tgaatgctaa ggaatccggt 120
 tgtaacgtta tcattggttt attcgaaggt gcggaggagt ggaaaagagc tgaagaacaa 180
 ggtttcgaag tctacaccgc tgctgaagct gctaagaagg ctgacatcat tatgatcttg 240
 atcccagatg aaatgcaggc tatcatgtac aaaaacgaca tcgaaccaa cttggaagcc 300

 ggtaacatgt tgatgttcgc tcacggtttc aacatccatt tcggttgtat tgttccacca 360
 aaggacgttg atgtcactat gatcgctcca aagggtccag gtcacaccgt tagatccgaa 420
 tacgaagaag gtaaagggtg cccatgcttg gttgctgtcg aacaagacgc tactggcaag 480
 gcttttgata tggttttggc ctacgcttta gccatcgggtg gtgctagagc cgggtgtcttg 540
 gaaactacct tcagaaccga aactgaaacc gacttggtcg gtgaacaagc tgttttatgt 600
 ggtggtgtct gcgctttgat gcaggccggt ttgaaacct tgggtgaagc cggttacgac 660
 ccaagaaacg cttacttcga atgtatccac gaaatgaagt tgatcgttga cttgatctac 720

 caatctggtt tctccggtat gcgttactct atctccaaca ctgctgaata cggtgactac 780
 attaccggtc caaagatcat tactgaagat accaagaagg ctatgaagaa gattttgtct 840
 gacattcaag atggtacctt tgccaaggac ttcttggttg acatgtctga tgctggttcc 900
 cagggtccact tcaaggctat gagaaagttg gcctccgaac acccagctga agttgtcggc 960
 gaagaaatta gatccttgta ctctgggtcc gacgaagaca agttgattaa caactga 1017

<210> 863

<211> 338

<212> PRT

<213> Artificial Sequence

<220><223> K9_Annabel_SH

<400> 863

Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn Leu Ser Leu Leu Asp

1 5 10 15
Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser Gln Gly His Ala His

20 25 30
Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val Ile Ile Gly Leu Phe

35 40 45
Glu Gly Ala Glu Glu Trp Lys Arg Ala Glu Glu Gln Gly Phe Glu Val

50 55 60
Tyr Thr Ala Ala Glu Ala Ala Lys Lys Ala Asp Ile Ile Met Ile Leu

65 70 75 80
Ile Pro Asp Glu Met Gln Ala Ala Met Tyr Lys Asn Asp Ile Glu Pro

85 90 95
Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala His Gly Phe Asn Ile

100 105 110
His Phe Gly Cys Ile Val Pro Pro Lys Asp Val Asp Val Thr Met Ile

115 120 125
Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser Glu Tyr Glu Glu Gly

130 135 140
Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln Asp Ala Thr Gly Lys

145 150 155 160
Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala Ile Gly Gly Ala Arg

165 170 175
Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu Thr Glu Thr Asp Leu

180 185 190
Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val Cys Ala Leu Met Gln

195 200 205
Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr Asp Pro Arg Asn Ala

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

 260 265 270
 Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln Asp Gly Thr Phe Ala
 275 280 285
 Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly Ser Gln Val His Phe
 290 295 300
 Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro Ala Glu Val Val Gly
 305 310 315 320
 Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp Glu Asp Lys Leu Ile

 325 330 335
 Asn Asn

<210> 864

<211> 340

<212> PRT

<213> Artificial Sequence

<220><223> Lactococcus lactis KARI encoded

<400> 864

Met Ala Val Thr Met Tyr Tyr Glu Asp Asp Val Glu Val Ser Ala Leu
 1 5 10 15
 Ala Gly Lys Gln Ile Ala Val Ile Gly Tyr Gly Ser Gln Gly His Ala
 20 25 30
 His Ala Gln Asn Leu Arg Asp Ser Gly His Asn Val Ile Ile Gly Val

 35 40 45
 Arg His Gly Lys Ser Phe Asp Lys Ala Lys Glu Asp Gly Phe Glu Thr
 50 55 60
 Phe Glu Val Gly Glu Ala Val Ala Lys Ala Asp Val Ile Met Val Leu

65 70 75 80
 Ala Pro Asp Glu Leu Gln Gln Ser Ile Tyr Glu Glu Asp Ile Lys Pro
 85 90 95
 Asn Leu Lys Ala Gly Ser Ala Leu Gly Phe Ala His Gly Phe Asn Ile

 100 105 110
 His Phe Gly Tyr Ile Lys Val Pro Glu Asp Val Asp Val Phe Met Val
 115 120 125
 Ala Pro Lys Ala Pro Gly His Leu Val Arg Arg Thr Tyr Thr Glu Gly
 130 135 140
 Phe Gly Thr Pro Ala Leu Phe Val Ser His Gln Asn Ala Ser Gly His
 145 150 155 160
 Ala Arg Glu Ile Ala Met Asp Trp Ala Lys Gly Ile Gly Cys Ala Arg

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

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

 290 295 300
 Tyr Arg Glu Ala Ala Lys Asn Leu Glu Ile Glu Lys Ile Gly Ala Glu
 305 310 315 320

Leu Arg Gln Ala Met Pro Phe Thr Gln Ser Gly Asp Asp Asp Ala Phe

325

330

335

Lys Ile Tyr Gln

340

<210> 865

<211> 1074

<212> DNA

<213> Artificial sequence

<220><223> AOX1 from Pichia stipitis

<400> 865

atgctttctg tgcagactac aagagccgcc aagttgcaac taggtcaatt acctctgatt 60

gcctacaccg ccagaagtgg aagacttcac caccaattct actccaccgt tgctgaaaag 120

acagccaacc ctacgcaaaa cactcagat aaaactaata ttttgatat tagaaccaag 180

gtgtacgatg agactgatat aagaaaacat gacgacaatc agtttatcac ccatacttta 240

tttctctatc ctacgttttc gcaggaagat tgtttgaaag ttggttacga acatcgtcct 300

cctcggactt ttggtgacaa aatggctttc agaggcattg aacttgatcag aggttctttc 360

gactttgtca ccggttacaa gaagccaaaag acacaggctg atatacattc aggttttaaa 420

ggtaccagat acgagatgac agaaggtaaa tggttgacca gatgtatatt cttagaaagt 480

attgctggag ttccaggtgc cgttgcttcc ttcatacagac acttgcatc tttacgtttg 540

ttgaagagag acaaagcctg gatcgaaacc ttacttgatg aagcattcaa cgaaagaatg 600

catttactta ccttcatcaa gattggcaaa cctagtgtgt tcaccagaac aatcatctac 660

gtcggccaag gtgtattctg caacttgctt ttcctgttct acttgccaa ccctaagtat 720

tgtaacaggt ttgttggtta ccttgaagaa gaggtgttta gtacatacac tcatttcgtt 780

catgaattac aatcaggtaa gcttcccaag ttgagaaca tcaagattcc aaccattgca 840

tggcaatact ggccagagtt gaccgagaat tcgtccatgt tggacttgat tttagaatc 900

agagccgacg aagccaagca cagagaagtc aaccacacct tggccaattt agatcaaaga 960

aaggacagaa atccatttgg gttggcaatt ccggaatctca aggaaccgca acccgaaagt 1020

ggcttaaagg ttacgaaacc ccacggctgg gagaaggaag aattgaaatt gtaa 1074

<210> 866

<211> 357

<212> PRT

<213> Artificial Sequence

<220><223> AOX1 from *Pichia stipitis*

<400> 866

Met Leu Ser Val Gln Thr Thr Arg Ala Ala Lys Leu Gln Leu Gly Gln

1 5 10 15

Leu Pro Ser Ile Ala Tyr Thr Ala Arg Ser Gly Arg Leu His His Gln

20 25 30

Phe Tyr Ser Thr Val Ala Glu Lys Thr Ala Asn Pro Thr Pro Asn Thr

35 40 45

Ser Asp Lys Thr Asn Ile Phe Asp Ile Arg Thr Lys Val Tyr Asp Glu

50 55 60

Thr Asp Ile Arg Lys His Asp Asp Asn Gln Phe Ile Thr His Pro Leu

65 70 75 80

Phe Pro His Pro Thr Phe Ser Gln Glu Asp Cys Leu Lys Val Gly Tyr

85 90 95

Glu His Arg Pro Pro Arg Thr Phe Gly Asp Lys Met Ala Phe Arg Gly

100 105 110

Ile Glu Leu Val Arg Gly Ser Phe Asp Phe Val Thr Gly Tyr Lys Lys

115 120 125

Pro Lys Thr Gln Ala Asp Ile Asp Ser Gly Phe Lys Gly Thr Arg Tyr

130 135 140

Glu Met Thr Glu Gly Lys Trp Leu Thr Arg Cys Ile Phe Leu Glu Ser

145 150 155 160

Ile Ala Gly Val Pro Gly Ala Val Ala Ser Phe Ile Arg His Leu His

165 170 175

Ser Leu Arg Leu Leu Lys Arg Asp Lys Ala Trp Ile Glu Thr Leu Leu

180 185 190

Asp Glu Ala Phe Asn Glu Arg Met His Leu Leu Thr Phe Ile Lys Ile

195 200 205

Gly Lys Pro Ser Trp Phe Thr Arg Thr Ile Ile Tyr Val Gly Gln Gly

210 215 220

Val Phe Cys Asn Leu Phe Phe Leu Phe Tyr Leu Ala Asn Pro Lys Tyr

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

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

 340 345 350
 Glu Glu Leu Lys Leu

 355

<210> 867

<211> 1992

<212> DNA

<213> Artificial Sequence

<220><223> AOX2 from *Pichia stipitis*

<400> 867

atggccattc ctgaagaatt cgatattctt gtcctgggtg gtggatccag tggatcctgt	60
attgccggaa gattggccaa cttggaccac tccttgaaag ttggtcttat cgaggctggt	120
gagaacaatc ttaacaaccc atgggtctac ctccaggta tttaccaag aaacatgaag	180
ttggactcca aaactgcttc tttctacacc tccaaccctt ctcctcattt gaatggtaga	240
agagctattg tcccatgtgc caacatcttg ggtgggtggt cttcgatcaa cttcatgatg	300
tacaccagag gttccgcttc tgattacgat gactttgaag ctgagggatg gaagaccaag	360
gatttgcttc ctttgatgaa gaagactgag acttacaaa gagcttgcaa caacctgaa	420
attcacggtt ttgaaggccc aatcaagggt tctttcggta actacactta cccagtttgt	480
caagacttct tgagagcaac tgaatcccaa ggtattccat acgttgacga cttggaagac	540

ttggtgactg ctcatggcgc tgaacactgg ctgaaatgga tcaacagaga cactggcgtcgt 600
 cgttccgact ctgctcatgc cttcgttcat tctacgatga gaaaccacga caatctgtac 660

 ttgatctgca acaccaaagt tgacaagatt attgttgaag acggaagagc tgctgctgtc 720
 agaaccgttc caagtaaac tttgaacgca aagaagccaa ctcaacaggt ttatcgtgct 780
 agaacgcaaa tcgttttgc tttgtgtacc atctcttctc ctctggttct gcaaagatcc 840
 ggttttggtg acccaatcaa attgagagcc gctgggtgta agcctttggt caacttgcca 900
 ggtgttggaa gaaacttcca agaccactac tgcttcttct ctcttacag aattaagccc 960
 caatacagat ctttcgatga cttcgtacgt ggtgacgcta acattcaaaa gaaggtattc 1020
 gaccaatggt acgctaacgg tactgggtcca ttggccacca acggtattga agccggtgtc 1080

 aagattagac caactccaga agaattatct cagatggacg agtccttcca agagggttac 1140
 agagagtact tcgaagacaa accagacaag ccagttatgc actattccat cattgctggt 1200
 ttcttcgggtg accacaccaa gattccacct ggaaagtaca tgaccatgtt ccacttcttg 1260
 gagtacccat tctccagagg ttctatccac atcacctctc cagaccata cgcaactcca 1320
 gactttgacc caggtttcat gaacgatgaa agagacatgg ctctatggt ctggtcttac 1380
 aagaagtcca gagagactgc cagaaaaatg gaccactttg ctggtgaggt tacttcccac 1440
 caccctctgt tccatactc atccgaggcc agagcttacg agatggattt ggagacctcc 1500

 aacgcctacg gtggaccact gaacttgact gctggctctg ctacaggttc ttggactcag 1560
 cctttgaaga agcctgctgg aagaaacgaa ggacatgtta cttccaacca agtcgagctt 1620
 catccagaca ttgagtacga tgaggaggat gataaggcca ttgagaacta cattcgtgag 1680
 cacactgaga ccacatggca ctgtctggga acctgttcca ttggtccaag agagggttcc 1740
 aagatcgta aatgggggtg tgttttggat cacagatcta acgtttacgg agtcaagggc 1800
 ctgaagggtg gtgacttgc cgtctgtcca gacaatgttg gttgtaacac ctacaccacc 1860
 gctcttttga tcggtgaaaa gactgccacc ttggttgggtg aagacttagg atacacaggt 1920

 gaggccttag acatgactgt acctcagttc aagttgggca cttacagaa gactggtctt 1980
 gctagattct ag 1992

 <210> 868
 <211> 663
 <212> PRT
 <213> Artificial Sequence
 <220><223> AOX2 from Pichia stipitis
 <400> 868

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

245 250 255
 Val Tyr Arg Ala Arg Thr Gln Ile Val Leu Ser Cys Gly Thr Ile Ser
 260 265 270
 Ser Pro Leu Val Leu Gln Arg Ser Gly Phe Gly Asp Pro Ile Lys Leu
 275 280 285

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

 Thr Asn Gly Ile Glu Ala Gly Val Lys Ile Arg Pro Thr Pro Glu Glu
 355 360 365
 Leu Ser Gln Met Asp Glu Ser Phe Gln Glu Gly Tyr Arg Glu Tyr Phe
 370 375 380
 Glu Asp Lys Pro Asp Lys Pro Val Met His Tyr Ser Ile Ile Ala Gly
 385 390 395 400
 Phe Phe Gly Asp His Thr Lys Ile Pro Pro Gly Lys Tyr Met Thr Met
 405 410 415

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

 His Pro Leu Phe Pro Tyr Ser Ser Glu Ala Arg Ala Tyr Glu Met Asp
 485 490 495

Leu Glu Thr Ser Asn Ala Tyr Gly Gly Pro Leu Asn Leu Thr Ala Gly
 500 505 510
 Leu Ala His Gly Ser Trp Thr Gln Pro Leu Lys Lys Pro Ala Gly Arg
 515 520 525
 Asn Glu Gly His Val Thr Ser Asn Gln Val Glu Leu His Pro Asp Ile
 530 535 540

Glu Tyr Asp Glu Glu Asp Asp Lys Ala Ile Glu Asn Tyr Ile Arg Glu
 545 550 555 560
 His Thr Glu Thr Thr Trp His Cys Leu Gly Thr Cys Ser Ile Gly Pro
 565 570 575
 Arg Glu Gly Ser Lys Ile Val Lys Trp Gly Gly Val Leu Asp His Arg
 580 585 590
 Ser Asn Val Tyr Gly Val Lys Gly Leu Lys Val Gly Asp Leu Ser Val
 595 600 605

Cys Pro Asp Asn Val Gly Cys Asn Thr Tyr Thr Thr Ala Leu Leu Ile
 610 615 620
 Gly Glu Lys Thr Ala Thr Leu Val Gly Glu Asp Leu Gly Tyr Thr Gly
 625 630 635 640
 Glu Ala Leu Asp Met Thr Val Pro Gln Phe Lys Leu Gly Thr Tyr Glu
 645 650 655
 Lys Thr Gly Leu Ala Arg Phe
 660

<210> 869

<211> 1410

<212> DNA

<213> Artificial sequence

<220><223> YER081W Chr 5

<400> 869

atgacaagca ttgacattaa caacttaca aatacctttc aacaagctat gaatatgagc	60
ggctccccag gcgtgtttg tacttcacct acgcaatctt tcatgaatac cgttccacag	120
cgcttgaatg ctgtaaagca cccaaaaatt ttgaagcctt tctcaacggg tgatatgaag	180
atcttactat tagaaaacgt taatcaaact gctattacaa tcttcgaaga gcaaggttac	240

caagtcgaat tciataaatc ttcattgccc gaggaagagt tgatcgaaaa gatcaaggac 300
gttcatgcta ttggatcag atcaaagact agattaactt caaatgtctt acaacatgcg 360
aagaatctgg tttgtattgg ttgtttctgt atcggtacca accaagttga cttagactac 420

gctaccagca gaggtattgc tgttttcaac tcgcctttct ccaactcaag atcagtagca 480
gaattggtea tcgctgaat cattagttta gcaagacaac taggtgatag atctatcgaa 540
ttacataccg gtacatggaa taaggttgct gctagatgtt gggaggttaag aggaaaaact 600
cttggattta ttgggtacgg tcacattggt tcccaattat cagttcttgc agaagctatg 660
ggtttgcagt tgttgtacta cgatattgta actatcatgg ccttgggtac tgccagacaa 720
gtttctacat tagatgaatt gttgaataaa tctgattttg tgacactaca tgtaccagct 780
actcctgaaa ctgaaaaaat gttatctgcc ccacaatttg ctgctatgaa ggatggcgct 840

tatgttatta atgcttcaag aggtactgtc gtggacattc catctttgat ccaagccgtg 900
aaagccaaca aaattgcagg tgctgctttg gatgtttatc cacatgaacc agctaagaac 960
ggagaagggt catttaacga tgagctaaat agctggactt ctgaattagt ttcattacca 1020
aatatcatct tgacaccaca cattgggtggc tctaccgaag aagcccaaag ctcaatcggt 1080
attgaagtgg ctaccgcatt gtccaaatac atcaatgaag gtaactctgt cggttcagtc 1140
aacttcccag aagtggcatt gaaatcattg tcttacgacc aagagaacac tgtgcgtgtg 1200
ttatacattc accaaaatgt accaggtggt ttgaagaccg tcaatgatat ttatcgaac 1260

cataacatcg aaaagcaatt ttccgattca aatggtgaaa ttgcttattt aatggctgat 1320
atctcttctg ttgaccaaag cgatattaaa gatatttatg aacaactaaa tcaaacctct 1380
gctaagatct caattagatt gctatattaa 1410

<210> 870

<211> 469

<212> PRT

<213> Artificial Sequence

<220><223> YER081W Chr 5

<400> 870

Met Thr Ser Ile Asp Ile Asn Asn Leu Gln Asn Thr Phe Gln Gln Ala

1 5 10 15

Met Asn Met Ser Gly Ser Pro Gly Ala Val Cys Thr Ser Pro Thr Gln

20 25 30

Ser Phe Met Asn Thr Val Pro Gln Arg Leu Asn Ala Val Lys His Pro

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

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

 145 150 155 160
 Glu Leu Val Ile Ala Glu Ile Ile Ser Leu Ala Arg Gln Leu Gly Asp
 165 170 175
 Arg Ser Ile Glu Leu His Thr Gly Thr Trp Asn Lys Val Ala Ala Arg
 180 185 190
 Cys Trp Glu Val Arg Gly Lys Thr Leu Gly Ile Ile Gly Tyr Gly His
 195 200 205
 Ile Gly Ser Gln Leu Ser Val Leu Ala Glu Ala Met Gly Leu His Val

 210 215 220
 Leu Tyr Tyr Asp Ile Val Thr Ile Met Ala Leu Gly Thr Ala Arg Gln
 225 230 235 240
 Val Ser Thr Leu Asp Glu Leu Leu Asn Lys Ser Asp Phe Val Thr Leu
 245 250 255
 His Val Pro Ala Thr Pro Glu Thr Glu Lys Met Leu Ser Ala Pro Gln
 260 265 270
 Phe Ala Ala Met Lys Asp Gly Ala Tyr Val Ile Asn Ala Ser Arg Gly

 275 280 285

Thr Val Val Asp Ile Pro Ser Leu Ile Gln Ala Val Lys Ala Asn Lys

290 295 300

Ile Ala Gly Ala Ala Leu Asp Val Tyr Pro His Glu Pro Ala Lys Asn

305 310 315 320

Gly Glu Gly Ser Phe Asn Asp Glu Leu Asn Ser Trp Thr Ser Glu Leu

325 330 335

Val Ser Leu Pro Asn Ile Ile Leu Thr Pro His Ile Gly Gly Ser Thr

340 345 350

Glu Glu Ala Gln Ser Ser Ile Gly Ile Glu Val Ala Thr Ala Leu Ser

355 360 365

Lys Tyr Ile Asn Glu Gly Asn Ser Val Gly Ser Val Asn Phe Pro Glu

370 375 380

Val Ala Leu Lys Ser Leu Ser Tyr Asp Gln Glu Asn Thr Val Arg Val

385 390 395 400

Leu Tyr Ile His Gln Asn Val Pro Gly Val Leu Lys Thr Val Asn Asp

405 410 415

Ile Leu Ser Asn His Asn Ile Glu Lys Gln Phe Ser Asp Ser Asn Gly

420 425 430

Glu Ile Ala Tyr Leu Met Ala Asp Ile Ser Ser Val Asp Gln Ser Asp

435 440 445

Ile Lys Asp Ile Tyr Glu Gln Leu Asn Gln Thr Ser Ala Lys Ile Ser

450 455 460

Ile Arg Leu Leu Tyr

465

<210

> 871

<211> 1365

<212> DNA

<213> Artificial Sequence

<220><223> YOR375C Chr 15

<400> 871

atgtcagagc cagaatttca acaagcttac gaagaagttg tctcctcttt ggaagactct 60

actcttttgc aacaacaccc agaatacaga aaggttttgc caattgtttc tgttccagaa 120

agaatcatac aattcagagt cacctgggaa aatgacaagg gtgaacaaga agttgctcaa 180
 ggttacagag tgcaataataa ctccgccaag ggtccataca aggggtggtct acgtttccat 240
 ccttccgtga acttgtctat cttgaaattc ttgggtttcg aacaaatctt caagaactcc 300
 ttgaccggcc tagacatggg tggtaggtaaa ggtggtctat gtgtggactt gaagggaaga 360

tctaataacg aaatcagaag aatctgttat gctttcatga gagaattgag cagacacatt 420
 ggtcaagaca ctgacgtgcc agctggtgat atcggtgttg gtggtcgtga aattggttac 480
 ctgttcggtg cttacagatc atacaagaac tcttggaag gtgtcttaac cggttaagggt 540
 ttgaactggg gtggttcttt gatcagacca gaagccactg gttacggttt agtttactat 600
 actcaagcta tgatcgacta tgccacaaac ggtaaggaat ctttcgaagg taagcgcgtc 660
 accatctctg gtagtggtaa cgttgctcaa tacgtgcct tgaaggttat tgagctaggt 720
 ggtactgtcg tttccctatc tgactccaag ggttgatca tctctgaaac tggatcacc 780

tccgaacaag tcgtgatat ttccagtgt aaggtcaact tcaagtcctt ggaacaaatc 840
 gtcaacgaat actctacttt ctccgaaaac aaagtgaat acattgctgg tgctcgtcca 900
 tggaccacag tcaaaaagg cgacattgt ttgcatgtg ccacccaaaa tgaagtcagc 960
 ggtgaagaag ccaaggcctt ggttgctcaa ggtgtcaagt ttattgccga aggttccaac 1020
 atgggttcca ctccagaagc tattgccgtc ttgaaactg ctcgttcac cgccactgga 1080
 ccaagcgaag ctgtttgta cggtccacca aaggctgcta acttgggtgg tgttgctgtt 1140
 tctggtttag aaatggcaca aaactctcaa agaatcacat ggactagcga aagagttgac 1200

caagagttga agagaattat gatcaactgt ttcaatgaat gtatcgacta tgccaagaag 1260
 tacactaagg acggttaagg cttgccatct ttggtcaaag gtgctaatat cgcaagtttc 1320
 atcaaggtct ctgatgtat gtttgaccaa ggtgatgtat tttaa 1365

<210> 872

<211> 454

<212> PRT

<213> Artificial Sequence

<220><223> YOR375C Chr 15 reverse complement

<400> 872

Met Ser Glu Pro Glu Phe Gln Gln Ala Tyr Glu Glu Val Val Ser Ser

1 5 10 15

Leu Glu Asp Ser Thr Leu Phe Glu Gln His Pro Glu Tyr Arg Lys Val

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

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

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

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

<210> 873

<211> 343

<212> PRT

<213> Artificial Sequence

<220><223> 878 C1

<400> 873

Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn

1 5 10 15
 Leu Ser Leu Leu Asp Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser

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

 65 70 75 80
 Ile Ile Met Ile Leu Ile Asn Asp Glu Lys Gln Ala Thr Met Tyr Lys
 85 90 95
 Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala
 100 105 110
 His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val
 115 120 125
 Asp Val Thr Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser

 130 135 140
 Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln
 145 150 155 160
 Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala
 165 170 175
 Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu
 180 185 190
 Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val

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

 260 265 270

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

325 330 335
 Glu Asp Lys Leu Ile Asn Asn
 340

<210> 874

<211> 343

<212> PRT

<213> Artificial Sequence

<220><223> 879 A7

<400> 874

Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn
 1 5 10 15
 Leu Ser Leu Leu Asp Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser
 20 25 30
 Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val

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

100 105 110
 His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val
 115 120 125

Asp Val Thr Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser
 130 135 140
 Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln
 145 150 155 160
 Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala
 165 170 175
 Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu
 180 185 190
 Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val
 195 200 205
 Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr
 210 215 220
 Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile
 225 230 235 240
 Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile
 245 250 255
 Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile
 260 265 270
 Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln
 275 280 285
 Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly
 290 295 300
 Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro
 305 310 315 320
 Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
 325 330 335
 Glu Asp Lys Leu Ile Asn Asn
 340
 <210> 875
 <211> 343
 <212> PRT
 <213> Artificial Sequence

<220><223> 879 C2

<400> 875

Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn

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

65 70 75 80
Ile Ile Met Ile Leu Ile Asn Asp Glu Lys Gln Ala Thr Met Tyr Lys
 85 90 95
Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala
 100 105 110
His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val
 115 120 125
Asp Val Thr Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser

130 135 140
Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln
145 150 155 160
Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala
 165 170 175
Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu
 180 185 190
Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val

195 200 205
Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr
210 215 220
Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile

225 230 235 240
 Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile
 245 250 255
 Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile

 260 265 270
 Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln
 275 280 285
 Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly
 290 295 300
 Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro
 305 310 315 320
 Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp

 325 330 335
 Glu Asp Lys Leu Ile Asn Asn
 340
 <210> 876
 <211> 343
 <212> PRT
 <213> Artificial Sequence
 <220><223> 880 A11
 <400> 876
 Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn
 1 5 10 15
 Leu Ser Leu Leu Asp Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser
 20 25 30
 Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val

 35 40 45
 Ile Ile Gly Leu Tyr Glu Gly Ala Lys Glu Trp Lys Arg Ala Glu Glu
 50 55 60
 Gln Gly Phe Glu Val Tyr Thr Ala Ala Glu Ala Ala Lys Lys Ala Asp
 65 70 75 80
 Ile Ile Met Ile Leu Ile Asn Asp Glu Lys Gln Ala Thr Met Tyr Lys

85 90 95
 Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala

 100 105 110
 His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val
 115 120 125
 Asp Val Thr Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser
 130 135 140
 Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln
 145 150 155 160
 Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala

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

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

 290 295 300
 Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Gln
 305 310 315 320
 Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
 325 330 335

Glu Asp Lys Leu Ile Asn Asn

340

<210> 877

<211> 343

<212> PRT

<213> Artificial Sequence

<220><223> 880 B4

<400> 877

Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn

1 5 10 15
Leu Ser Leu Leu Asp Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser

20 25 30
Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val

35 40 45
Ile Ile Gly Leu Tyr Glu Gly Ala Glu Glu Trp Lys Arg Ala Glu Glu

50 55 60
Gln Gly Phe Glu Val Tyr Thr Ala Ala Glu Ala Ala Lys Lys Ala Asp

65 70 75 80
Ile Ile Met Ile Leu Ile Asn Asp Glu Lys Gln Ala Thr Met Tyr Lys

85 90 95
Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala

100 105 110
His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val

115 120 125
Asp Val Thr Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser

130 135 140
Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln

145 150 155 160
Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala

165 170 175
Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu

180 185 190

Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val

195 200 205

Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr

210 215 220

Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile

225 230 235 240

Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile

245 250 255

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

260 265 270

Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln

275 280 285

Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly

290 295 300

Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro

305 310 315 320

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

325 330 335

Glu Asp Lys Leu Ile Asn Asn

340

<210> 878

<211> 343

<212> PRT

<213> Artificial Sequence

<220><223> 880 D11

<400> 878

Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn

1 5 10 15

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

20 25 30

Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val

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

 100 105 110
 His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val
 115 120 125
 Asp Val Thr Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser
 130 135 140
 Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln
 145 150 155 160
 Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Val Leu Ala

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

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

Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly

290 295 300

Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro

305 310 315 320

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

325 330 335

Glu Asp Lys Leu Ile Asn Asn

340

<210> 879

<211> 343

<212> PRT

<213> Artificial Sequence

<220><223> 881 A2

<400> 879

Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn

1 5 10 15

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

20 25 30

Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val

35 40 45

Ile Ile Gly Leu Tyr Glu Gly Ala Lys Glu Trp Lys Arg Ala Glu Glu

50 55 60

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

65 70 75 80

Ile Ile Met Ile Leu Ile Asn Asp Glu Met Gln Ala Thr Met Tyr Lys

85 90 95

Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala

100 105 110

His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val

115 120 125

Asp Val Thr Met Val Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser

130 135 140
 Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln
 145 150 155 160
 Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala
 165 170 175
 Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu
 180 185 190
 Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val

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

260 265 270
 Thr Glu Asp Thr Lys Lys Ala Met Lys Thr Ile Leu Ser Asp Ile Gln
 275 280 285
 Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly
 290 295 300
 Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro
 305 310 315 320
 Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp

325 330 335
 Glu Asp Lys Leu Ile Asn Asn

340

<210> 880

<211> 343

<212> PRT

<213> Artificial Sequence

<220><223> 881 G3

<400> 880

Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn

1 5 10 15

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

20 25 30

Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val

35 40 45

Ile Ile Gly Leu Phe Glu Gly Ala Lys Glu Trp Lys Arg Ala Glu Glu

50 55 60

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

65 70 75 80

Ile Ile Met Ile Leu Ile Asn Asp Glu Lys Gln Ala Thr Met Tyr Lys

85 90 95

Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala

100 105 110

His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val

115 120 125

Asp Val Thr Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser

130 135 140

Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln

145 150 155 160

Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala

165 170 175

Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu

180 185 190

Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val

195 200 205

Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr

210 215 220

Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile

225 230 235 240

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

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

<210> 881

<211> 343

<212> PRT

<213> Artificial Sequence

<220><223> 881 G9

<400> 881

Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn

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

65 70 75 80
Ile Ile Met Ile Leu Ile Asn Asp Glu Glu Gln Ala Thr Met Tyr Lys
85 90 95

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

340

<210> 882

<211> 343

<212> PRT

<213> Artificial Sequence

<220><223> 882 B12

<400> 882

Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn

1 5 10 15

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

20 25 30

Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val

35 40 45

Ile Ile Gly Leu Tyr Glu Gly Ala Lys Glu Trp Lys Arg Ala Glu Glu

50 55 60

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

65 70 75 80

Ile Ile Met Ile Leu Ile Asn Asp Glu Lys Gln Ala Thr Met Tyr Lys

85 90 95

Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala

100 105 110

His Gly Phe Asn Ile Arg Phe Gly Cys Ile Val Pro Pro Lys Asp Val

115 120 125

Asp Val Thr Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser

130 135 140

Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln

145 150 155 160

Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala

165 170 175

Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu

180 185 190

Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val

195 200 205
 Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr
 210 215 220
 Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile

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

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

340

<210> 883

<211> 343

<212> PRT

<213> Artificial Sequence

<220><223> 882 C10

<400> 883

Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn

1 5 10 15
 Leu Ser Leu Leu Asp Gly Lys Thr Ile Ala Val Ile Gly Tyr Ser Ser
 20 25 30
 Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val
 35 40 45
 Ile Ile Gly Leu Tyr Glu Gly Ala Lys Glu Trp Lys Ser Ala Glu Glu

50 55 60
 Gln Gly Phe Glu Val Tyr Thr Ala Ala Glu Ala Ala Lys Lys Ala Asp

 65 70 75 80
 Ile Ile Met Ile Leu Ile Asn Asp Glu Lys Gln Ala Thr Met Tyr Lys
 85 90 95
 Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala
 100 105 110
 His Gly Phe Asn Ile His Phe Gly Tyr Ile Val Pro Pro Lys Asp Val
 115 120 125
 Asn Val Thr Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser

 130 135 140
 Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln
 145 150 155 160
 Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala
 165 170 175
 Ile Gly Gly Ala Arg Ala Asp Val Leu Glu Thr Thr Phe Arg Thr Glu
 180 185 190
 Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val

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

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

Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro
 305 310 315 320
 Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp

 325 330 335
 Glu Asp Lys Leu Ile Asn Asn
 340

<210> 884

<211> 24

<212> DNA

<213> Artificial Sequence

<220><223> K9_309I_111711r

<400> 884

ctttctcata gccttaatgt ggac 24

<210> 885

<211> 24

<212> DNA

<213> Artificial Sequence

<220><223> K9_309L_111711r

<400> 885

ctttctcata gcctttaagt ggac 24

<210> 886

<211> 24

<212> DNA

<213> Artificial Sequence

<220><223> K9_309Y_111711r

<400> 886

ctttctcata gccttatagt ggac 24

<210> 887

<211> 24

<212> DNA

<213> Artificial Sequence

<220><223> K9_309C_111711r

<400> 887

cttttcata gccttacagt ggac 24

<210> 888

<211> 24

<212> DNA

<213> Artificial Sequence

<220><223> K9_309W_111711r

<400> 888

cttttcata gccttcagt ggac 24

<210> 889

<211> 24

<212> DNA

<213> Artificial Sequence

<220><223> K9_309P_111711r

<400> 889

cttttcata gcctttgggt ggac 24

<210> 890

<211> 24

<212> DNA

<213> Artificial Sequence

<220><223> K9_309H_111711r

<400> 890

cttttcata gccttatggt ggac 24

<210> 891

<211> 24

<212> DNA

<213> Artificial Sequence

<220><223> K9_309Q_111711r

<400> 891

cttttcata gccttttgggt ggac 24

<210> 892

<211> 24

<212> DNA

<213> Artificial Sequence

<220><223> K9_309M_111711r
 <400> 892
 ctttctcata gccttcatgt ggac 24
 <210> 893
 <211> 24
 <212> DNA
 <213> Artificial Sequence
 <220><223> K9_309N_111711r
 <400> 893
 ctttctcata gccttattgt ggac 24
 <210> 894
 <211> 24
 <212> DNA
 <213> Artificial Sequence
 <220><223> K9_309V_111711r
 <400> 894
 ctttctcata gccttaacgt ggac 24

 <210> 895
 <211> 24
 <212> DNA
 <213> Artificial Sequence
 <220><223> K9_309A_111711r
 <400> 895
 ctttctcata gccttagcgt ggac 24
 <210> 896
 <211> 24
 <212> DNA
 <213> Artificial Sequence
 <220><223> K9_309D_111711r
 <400> 896
 ctttctcata gccttatcgt ggac 24
 <210> 897
 <211> 24
 <212> DNA

<213> Artificial Sequence
 <220><223> K9_309E_111711r
 <400> 897
 ctttctcata gccttttcgt ggac 24

<210> 898
 <211> 24
 <212> DNA
 <213> Artificial Sequence
 <220><223> K9_309G_111711r
 <400> 898
 ctttctcata gccttaccgt ggac 24

<210> 899
 <211> 24
 <212> DNA
 <213> Artificial Sequence
 <220><223> K9_309S_111711r
 <400> 899
 ctttctcata gccttagagt ggac 24

<210> 900
 <211> 24
 <212> DNA
 <213> Artificial Sequence
 <220><223> K9_309T_111711r
 <400> 900
 ctttctcata gccttagtgt ggac 24

<210> 901
 <211> 24
 <212> DNA
 <213> Artificial Sequence
 <220><223> K9_309R_111711r
 <400> 901
 ctttctcata gcctttctgt ggac 24

<210> 902

<211> 24

<212> DNA

<213> Artificial Sequence

<220><223> K9_309K_111711r

<400> 902

cttttcata gccttcttgt ggac 24

<210> 903

<211> 46

<212> DNA

<213> Artificial Sequence

<220><223> Kshort1

<400> 903

aaaccggttt aaacagtatg gctaagattt actaccaaga agactg 46

<210> 904

<211> 22

<212> DNA

<213> Artificial Sequence

<220><223> YGrev

<400> 904

ttctgtttta tcagaccgct tc 22

<210> 905

<211> 4978

<212> DNA

<213> Artificial Sequence

<220><223> pBAD-ps-JEA1

<400> 905

aagaaaccaa ttgtccatat tgcacagac attgccgtca ctgcgtcttt tactggctct 60

tctcgtaac caaacggta accccgctta ttaaaagcat tctgtaacaa agcgggacca 120

aagccatgac aaaaacgcgt aacaaaagtg tctataatca cggcagaaaa gtccacattg 180

attatttgca cggcgtcaca ctttgcctatg ccatagcatt tttatccata agattagcgg 240

atcctacctg acgtttttta tcgcaactct ctactgttcc tccataaccg ttttttgggc 300

taacaggagg aagttaaac agtatgaaag ttttctacga taaagactgc gacctgtcga 360

tcatccaagg taagaaagtt gccatcatcg gcttcggttc ccagggccac gctcaagcac 420

tcaacctgaa ggattccggc gtagacgtga ctgttggcct gcctaaaggc tttgctgatg	480
tagccaaggc tgaagccccc ggcttttaaag tgaccgacgt tgctgcagcc gttgccggtg	540
ccgacttggg catgatcctg attccggacg agttccagtc ccagctgtac aagaacgaaa	600
tcgagccgaa catcaagaag ggcgccactc tggccttctc ccacggcttc gcgatccact	660
acaaccagggt tgtgcctcgt gccgacctcg acgtgatcat gatcgcgccg aaggctccag	720
gccacaccgt acgttccgag ttctgcaagg gcggaggat tcttgacctg atcgcgatct	780
accaggacgt ttccggcaac gccaaagaac tgcacctgtc ctacgccga ggcgtgggcg	840
gcggccgtac cggcatcacc gaaaccacct tcaaggacga gactgaaacc gacctgttcg	900
gtgagcaggc tgttctgtgt ggcgttaccg tcgagctggg caaagccggt ttcgaaaccc	960
tggttgaagc tggctacgt ccagaaatgg cctacttcca gtgcctgcac gaactgaagc	1020
tgatcgttga cctcatgtac gaaggcggta tcgccaacat gaactactcg atctccaaca	1080
acgtgaata cggcgaglac gtgactggtc cagaagtcac caacgccga tcccgtcagg	1140
ccatcgcaa tgcctgaag cgcattccagg acggcgaata cgcgaagatg ttcattcagc	1200
aaggcgtac cggtaccaca tcgatgaccg ccaagcgtcg taacaacgt gtcacggta	1260
tcgaaatcat cggcgagcaa ctgcgtcga tgatgccttg gatcggtgcc aacaaaatcg	1320
tcgacaaagc caagaactaa ggccctgcag gccttggcgg atgagagaag attttcagcc	1380
tgatacagat taaatcagaa cgcagaagcg gtctgataaa acagaatttg cctggcggca	1440
gtagcgcggt ggtccacct gaccccatgc cgaactcaga agtgaaacgc ctagcgccg	1500
atggtagtgt ggggtctccc catgcgagag tagggaactg ccaggcatca aataaacga	1560
aaggctcagt cgaagactg ggcctttcgt tttatctgtt gttgtcggg gaacgtctc	1620
ctgagtagga caaatccgc gggagcggat ttgaacgttg cgaagcaacg gcccgagg	1680
tggcgggcag gacgcccgc ataaactgcc aggcataaaa ttaagcagaa ggccatcctg	1740
acggatggcc ttttgcgtt tctacaaact cttttgttta ttttctaaa tacattcaaa	1800
tatgtatccg ctcatgagac aataaccctg ataaatgctt caataatatt gaaaaaggaa	1860
gagtatgagt attcaacatt tccgtgtgc ccttattccc tttttgagg cattttgcct	1920
tcctgttttt gctcaccag aaacgtggt gaaagtaaaa gatgctgaag atcagttggg	1980
tgcacgagt ggttacatcg aactggatct caacagcggg aagatcctg agagttttcg	2040
ccccgaagaa cgttttccaa tgatgagcac ttttaaagtt ctgctatgtg gcgcggtatt	2100
atcccggtgt gacgccgggc aagagcaact cggtcgccgc atacactatt ctcagaatga	2160
cttgggttag tactcaccag tcacagaaaa gcactttacg gatggcatga cagtaagaga	2220

attatgcagt gctgccataa ccatgagtga taacactgcg gccaaacttac ttctgacaac 2280

gatcggagga ccgaaggagc taaccgcttt ttgacacaac atgggggatc atgtaactcg 2340

ccttgatcgt tgggaaccgg agctgaatga agccatacca aacgacgagc gtgacaccac 2400

gatgcctgta gcaatggcaa caacgttgcg caaactatta actggcgaac tacttactct 2460

agcttcccgg caacaattaa tagactggat ggaggcggat aaagtgtcag gaccacttct 2520

gcgctcggcc ctcccgctg gctggtttat tctgataaaa tctggagccg gtgagcgtgg 2580

gtctcgcggt atcattgcag cactggggcc agatggtaag ccctcccgtg tcgtagttag 2640

ctacacgacg gggagtcagg caactatgga tgaacgaaat agacagatcg ctgagatagg 2700

tgctcactg attaagcatt ggtaactgtc agaccaagtt tactcatata tactttagat 2760

tgatttaaaa cttcatTTTT aatttaaaag gatctagggtg aagatccttt ttgataatct 2820

catgacaaaa atcccttaac gtgagttttc gttccactga gcgtcagacc ccgtagaaaa 2880

gatcaaagga tcttcttgag atcctttttt tctgcgcgta atctgctgct tgcaaacaaa 2940

aaaaccaccg ctaccagcgg tggtttggtt gccggatcaa gagctaccaa ctctttttcc 3000

gaaggtaact ggcttcagca gagcgcagat accaaatact gtccttctag tgtagccgta 3060

gttaggccac cacttcaaga actctgtagc accgcctaca tacctcgctc tgctaactct 3120

gttaccagtg gctgctgcca gtggcgataa gtcgtgtctt accgggttgg actcaagacg 3180

atagttaccg gataaggcgc agcggtcggg ctgaacgggg ggttcgtgca cacagcccag 3240

cttgagcga acgacctaca ccgaactgag atacctacag cgtgagctat gagaaagcgc 3300

cacgttccc gaaggagaaa aggcggacag gtatccggtg agcggcaggg tcggaacagg 3360

agagcgcacg agggagcttc cagggggaaa cgcttggtat ctttatagtc ctgtcgggtt 3420

tcgccacctc tgacttgagc gtcgattttt gtgatgctcg tcaggggggc ggagcctatg 3480

gaaaaacgcc agcaacgcgg cttttttacg gttcctggcc ttttctggc cttttgctca 3540

catgttcttt cctgcgttat cccctgattc tgtggataac cgtattaccg cctttgagtg 3600

agctgatacc gctgccgca gccgaacgac cgagcgcagc gactcagtga gcgaggaagc 3660

ggaagagcgc ctgatgcggt attttctcct tacgcatctg tgcggtatct cacaccgcat 3720

atggtgcact ctcagtacaa tctgctctga tgccgcatag ttaagccagt atacactccg 3780

ctatcgctac gtgactgggt catggctgcg ccccacaccc cgccaacacc cgctgacgcg 3840

ccctgacggg ctgtctgct cccggcatcc gcttacagac aagctgtgac cgtctccggg 3900

agctgcatgt gtcagaggtt ttcaccgtca tcaccgaaac gcgcgaggca gcagatcaat 3960

tcgcgcgcga aggcgaagcg gcatgcataa tgtgcctgtc aaatggacga agcagggatt 4020
 ctgcaaaccc tatgtactc cgtcaagccg tcaattgtct gattcgttac caattatgac 4080
 aacttgacgg ctacatcatt cactttttct tcacaaccgg cacggaactc gctcgggctg 4140
 gccccggtgc attttttaaa taccgcgag aaatagagtt gatcgtcaaa accaaccattg 4200
 cgaccgacgg tggcgaatagg catccgggtg gtgctcaaaa gcagcttcgc ctggctgata 4260
 cgttggtcct cgcgccagct taagacgcta atccctaact gctggcggaa aagatgtgac 4320
 agacgcgacg gcgacaagca aacatgctgt gcgacgctgg cgatatcaaa attgctgtct 4380

gccaggtgat cgtgatgta ctgacaagcc tcgcgtaccc gattatccat cgggtgatgg 4440
 agcgactcgt taatcgcttc catgcgccgc agtaacaatt gctcaagcag atttatcgcc 4500
 agcagctccg aatagcggcc ttccccttgc ccggcgtaa tgatttgccc aaacaggctc 4560
 ctgaaatcgc gctggtgcgc ttcatccggg cgaaagaacc ccgtattggc aaatattgac 4620
 ggccagttaa gccattcatg ccagtaggcg cgcggacgaa agtaaaccca ctggtgatac 4680
 cattcgcgag cctccggatg acgaccgtag tgatgaatct ctccctggcg gaacagcaaa 4740
 atatcacccg gtcggcaaac aaattctcgt cctgatttt tcaccacccc ctgaccgcga 4800

atggtgagat tgagaatata acctttcatt ccagcgggtc ggtcgataaa aaaatcgaga 4860
 taaccgttgg cctcaatcgg cgtaaaccg gccaccagat gggcattaaa cgagtatccc 4920
 ggcagcaggg gatcattttg cgcttcagcc atacttttca tactcccgcc attcagag 4978

<210> 906

<211> 36

<212> DNA

<213> Artificial Sequence

<220><223> KPSH1

<400> 906

aaacagtatg gctaagattt actaccaaga agactg 36

<210> 907

<211> 1017

<212> DNA

<213> Artificial Sequence

<220><223> K9_Eliza_SH

<400> 907

atggctaaga ttactacca agaagactgt aactgtcct tgttgatgg taagactatc 60
 gccgttatcg gttacgggtc tcaaggcac gctcatgcc tgaatgctaa ggaatccggt 120

tgtaacgtta tcattggttt attcgaagat gcggaggaat ggaaaagagc tgaagaacaa 180
 ggtttcgaag tctacaccgc tgctgaagct gctaagaagg ctgacatcat tatgatcttg 240
 atcccagatg aaaagcaggc taccatgtac aaaaacgaca tcgaacaaaa cttggaagcc 300
 ggtaacatgt tgatgttcgc tcacggtttc aacatccatt tcggttgat tgttccacca 360
 aaggacgttg atgtcactat gatcgctcca aagggtccag gtcacaccgt tagatccgaa 420

tacgaagaag gtaaagggtg cccatgcttg gttgctgtcg aacaagacgc tactggcaag 480
 gcttttgata tggctttggc ctacgcttta gccatcgggtg gtgctagagc cgggtgtcttg 540
 gaaactacct tcagaaccga aactgaaacc gacttggttcg gtgaacaagc tgttttatgt 600
 ggtgggtgtc gcgctttgat gcaggccggt ttgaaacct tgggtgaagc cggttacgac 660
 ccaagaaacg ctacttcga atgtatccac gaaatgaagt tgatcgttga ctgatctac 720
 caatctggtt tctccggtat gcgttactct atctccaaca ctgctgaata cggtgactac 780
 attaccggtc caaagatcat tactgaagat accaagaagg ctatgaagaa gattttgtct 840

gacattcaag atggtaacct tgccaaggac ttcttggttg acatgtctga tgctggttcc 900
 cagggtccact tcaaggctat gagaaagttg gcctccgaac acccagctga agttgtcgg 960
 gaagaaatta gatccttgta ctcttggtcc gacgaagaca agttgattaa caactga 1017

<210> 908

<211> 1017

<212> DNA

<213> Artificial Sequence

<220><223> K9_Frank_SH

<400> 908

atggctaaga ttactacca agaagactgt aactgtcct tgttgatgg taagactatc 60
 gccgttatcg gtiacggttc tcaaggtcac gtcctgccc tgaatgctaa ggaatccggt 120
 tgtaacgtta tcattggttt attcgaaggt gcggaggagt ggaaaagagc tgaagaacaa 180

ggtttcgaag tctacaccgc tgctgaagct gctaagaagg ctgacatcat tatgatcttg 240
 atcccagatg aaaagctggc taccatgtac aaaaacgaca tcgaacaaaa cttggaagcc 300
 ggtaacatgt tgatgttcgc tcacggtttc aacatccatt tcggttgat tgttccacca 360
 aaggacgttg atgtcactat gatcgctcca aagggtccag gtcacaccgt tagatccgaa 420
 tacgaagaag gtaaagggtg cccatgcttg gttgctgtcg aacaagacgc tactggcaag 480
 gcttttgata tggctttggc ctacgcttta gccatcgggtg gtgctagagc cgggtgtcttg 540
 gaaactacct tcagaaccga aactgaaacc gacttggttcg gtgaacaagc tgttttatgt 600

ggtaggtgtct gcgctttgat gcaggccggt tttagaacct tggttgaagc cggttacgac 660
ccaagaaacg cttacttcga atgtatccac gaaatgaagt tgatcgttga cttgatctac 720
caatctggtt tctccggtat gcgttactct atctccaaca ctgctgaata cggtgactac 780
attaccggtc caaagatcat tactgaagat accaagaagg ctatgaagaa gattttgtct 840
gacattcaag atggtacctt tgccaaggac ttcttggttg acatgtctga tgctggttcc 900
cagggtccact tcaaggttat gagaaagttg gcctccgaac acccagctga agttgtcggc 960
gaagaaatta gatccttgta ctcttggtcc gacgaagaca agttgattaa caactga 1017

<210> 909

<211> 1017

<212> DNA

<213> Artificial Sequence

<220><223> K9_Grace_SH

<400> 909

atggctaaga ttactacca agaagactgt aacttgtcct tgttgatgg taagactatc 60
gccgttatcg gtiacggttc tcaaggtcac gctcatgccc tgaatgctaa ggaatccggt 120
tgtaacgtta tcattggttt attcgaaggt gcggaggagt ggaaaagagc tgaagaacaa 180
ggtttcgaag tctacaccgc tgctgaagct gctaagaagg ctgacatcat catgatcttg 240
atcccagatg aaaagcaggc taccatgtac aaaaacgaca tcgaacaaa cttggaagcc 300
ggtaacatgt tgatgttcgc tcacggtttc aacatccatt tcggttgtat tgttccacca 360

aaggacgttg atgtcactat gatcgctcca aagggtccag gtcacaccgt tagatccgaa 420
tacgaagaag gtaaagggtt cccatgcttg gttgctgtcg aacaagacgc tactggcaag 480
gcttttgata tggttttggc ctacgcttta gccatcgggtg gtgctagagc cgggtgtcttg 540
gaaactacct tcagaaccga aactgaaacc gacttggttcg gtgaacaagc tgttttatgt 600
ggtaggtgtct gcgctttgat gcaggccggt tttagaacct tggttgaagc cggttacgac 660
ccaagaaacg cttacttcga atgtatccac gaaatgaagt tgatcgttga cttgatctac 720
caatctggtt tctccggtat gcgttactct atctccaaca ctgctgaata cggtgactac 780

ataaccggtc caaagatcat tactgaagat accaagaagg ctatgaagaa gattttgtct 840
gacattcaag atggtacctt tgccaaggac ttcttggttg atatgtctga tgatggttcc 900
cagggtccact tcaaggttat gagaaagttg gcctccgaac acccagctga agttgtcggc 960
gaagaaatta gatccttgta ctcttggtcc gacgaagaca agttgattaa caactga 1017

<210> 910

<211> 1017

<212> DNA

<213> Artificial Sequence

<220><223> K9_Ingrid_SH

<400> 910

atggctaaga ttactacca agaagactgt aacttgctct tgttggatgg taagactatc	60
gccgttatcg gttacgggtc tcaaggtcac gctcatgccc tgaatgctaa ggaatccggt	120
tgtaacgtta tcattggttt attcgaaggt gcggaggagt ggaaaagagc tgaagaacaa	180
ggtttcgaag tctacaccgc tgctgaagct gctaagaagg ctgacatcat tatgatcttg	240
atcccagatg aaaagcaggc taccatttac aaaaacgaca tcgaacaaa ctgggaagcc	300
ggtaacatgt tgatgttcgc tcacggtttc aacatccatt tcggttgtat tgttccacca	360
aaggacgttg atgtcactat gatcgctcca aagggtccag gtcacaccgt tagatccgaa	420
tacgaagaag gtaaagggtg cccatgcttg gttgctgtcg aacaagacgc tactggcaag	480
gctttggata tggctttggc ctacgcttta gccatcggtg gtgctagagc cgggtgtcttg	540
gaaactacct tcagaaccga aactgaaacc gacttggtcg gtgaacaagc tgttttatgt	600
ggtggtgtct gcgctttgat gcaggccggt tttgaaacct tggttgaagc cggttacgac	660
ccaagaaacg cttacttcga atgtatccac gaaatgaagt tgatcgttga ctgatctac	720
caatctgggt tctccggat gcgttactct atctccaaca ctgctgaata cggtgactac	780
attaccggtc caaagatcat tactgaagat accaagaagg ctatgaagaa gattttgtct	840
gacattcaag atggtacctt tgccaaggac ttcttggttg acatgtctga tgctggttcc	900
cagatccact tcaaggetat gagaaagttg gcctccgaac acccagctga agttgtcgg	960
gaagaaatta gatccttgta ctctgggtcc gacgaagaca agttgattaa caactga	1017

<210> 911

<211> 1017

<212> DNA

<213> Artificial Sequence

<220><223> K9_Jarvis_SH

<400> 911

atggctaaga ttactacca agaagactgt aacttgctct tgttggatgg taagactatc	60
gccgttatcg gttacgggtc tcaaggtcac gctcatgccc tgaatgctaa ggaatccggt	120
tgtaacgtta tcattggttt attcgaaggt gcggaggagt ggaaaagagc tgaagaacaa	180
ggatatcgaag tctacaccgc tgctgaagct gctaagaagg ctgacatcat tatgatcttg	240

atcccagatg aaaagcaggc taccatgtac aaaaacgaca tcgaacaaa cttggaagcc 300

ggtaacatgt tgatgttcgc tcacggtttc aacatccatt tcggttgat tgttccacca 360

aaggacgttg atgtcactat gatcgctcca aagggtccag gtcacaccgt tagatccgaa 420

tacgaagaag gtaaagggtg cccatgcttg gttgctgtcg aacaagacgc tactggcaag 480

gctttggata tggctttggc ctacgcttta gccatcggtg gtgctagagc cgggtgtcttg 540

gaaactacct tcagaaccga aactgaaacc gacttggtcg gtgaacaagc tgttttatgt 600

ggtggtgtct gcgctttgat gcaggccggt tttgaaacct tggttgaagc cggttacgac 660

ccaagaaacg cttacttca atgtatccac gaaatgaagt tgatcgttga cttgatctac 720

caatctggtt tctccggtat gcgttactct atctccaaca ctgctgaata cggtgactac 780

attaccggtc caaagatcat tactgaagat accaagaagg ctatgaagaa gattttgtct 840

gacattcaag atggtacctt tgccaaggac ttcttggttg acatgtctga tgctggttcc 900

cagggtccact tcaaggctat gagaaagttg gcctccgaac acccagctga agttgtcggc 960

gaagaaatta gatccttgta ctctggtcc gacgaagaca agttgattaa caactga 1017

<210> 912

<211> 1017

<212> DNA

<213> Artificial Sequence

<220><223> K9_Kelly_SH

<400> 912

atggctaaga ttactacca agaagactgt aactgtcct tgttggatgg taagactatc 60

gccgttatcg gttacggttc tcaaggtcac gtcctatgcc tgaatgctaa ggaatccggt 120

tgtaacgtta tcattggttt attcgaaggt ggggaggagt ggaatagagc tgaagaacaa 180

ggtttcgaag tctacaccgc tgctgaagct gctaaaaagg ctgacatcat tatgatcttg 240

atcccagatg aaaagcaggc taccatgtac aaaaacgaca tcgaacaaa cttggaagcc 300

ggtaacatgt tgatgttcgc tcacggtttc aacatccatt tcggttgat tgttccacca 360

aaggacgttg atgtcactat gatcgctcca aagggtccag gtcacaccgt tagatccgaa 420

tacgaagaag gtaaagggtg cccatgcttg gttgctgtcg aacaagacgc tactggcaag 480

gctttggata tggctttggc ctacgcttta gccatcggtg gtgctagagc cgggtgtcttg 540

gaaactacct tcagaaccga aactgaaacc gacttggtcg gtgaacaagc tgttttatgt 600

ggtggtgtct gcgctttgat gcaggccggt tttgaaacct tggttgaagc cggttacgac 660

ccaagaaacg cttacttca atgtatccac gaaatgaagt tgatcgttga cttgatctac 720

caatctggtt tctccggtat gcgttactct atctccaaca ctgctgaata cggtgactac 780
attaccggtc caaagatcat tactgaagat accaagaagg ctatgaagaa gattttgtct 840
gacattcaag atggtacctt tgccaaggac ttcttggttg acatgtctga tgctggttcc 900

caggtccact tcaaggctat gagaaagttg gcctccgaac acccagctga agttgtcggc 960
gaagaaatta gatccttgta ctcttggtcc gacgaagaca agttgattaa caactga 1017

<210> 913

<211> 1017

<212> DNA

<213> Artificial Sequence

<220><223> 9_Norman_SH

<400> 913

atggctaaga ttactacca agaagactgt aacttgcct tgttgatgg taagactatc 60
gccgttatcg gttacggttc tcaaggcac gctcatgcc tgaatgctaa ggaatccggt 120
tgtaacgtta tcattggttt attcgaatgt gcggaggagt ggaaaagagc tgaagaacaa 180
ggtttcgaag tctacaccgc tgctgaagct gctaagaagg ctgacatcat tatgatcttg 240

atcccagatg aaaagcaggc taccatgtac aaaaacgaca tcgaacaaa ctggaagcc 300
ggtaacatgt tgatgttcgc tcacggtttc aacatccatt tcggttgat tgttccacca 360
aaggacgttg atgtcactat gatcgctcca aagggtccag gtcacaccgt tagatccgaa 420
tacgaagaag gtaaaggtgt cccatgcttg gttgctgtcg aacaagacgc tactggcaag 480
gctttggata tggctttggc ctacgcttta gccatcggcg gtgctagagc cgggtgtcttg 540
gaaactacct tcagaaccga aactgaaacc gacttggtcg gtgaacaagc tgttttatgt 600
ggtggtgtct gcgctttgat gcaggccggt ttgaaacct tggttgaagc cggttacgac 660

ccaagaaacg ctacttcga atgtatccac gaaatgaagt tgatcgttga ctgatctac 720
caatctggtt tctccggtat gcgttactct atctccaaca ctgctgaata cggtgactac 780
attaccggtc caaagatcat tactgaagat accaagaagg ctatgaagaa gattttgtct 840
gacattcaag atggtacctt tgccaaggac ttcttggttg acatgtctga tgctggttcc 900
caggtccact tcaaggctat gagaaagttg gcctccgaac acccagctga agttgtcggc 960
gaagaaatta gatccttgta ctcttggtcc gacgaagaca agttgattaa caactga 1017

<210> 914

<211> 1017

<212> DNA

<213> Artificial Sequence

<220><223> K9_Ophelia_SH

<400> 914

atggctaaga ttactacca agaagactgt aacttgctct tgttggatgg taagactatc	60
gccgttatcg gttacgggtc tcaagggtcac gctcatgccc tgaatgctaa ggaatccggt	120
tgtaacgtta tcattgggtt attcgaaggt gcggaggagt ggaaaagagc tgaagaacaa	180
ggtttcgaag tctacaccgc tgctgaagct gctaagaagg ctgacatcat tatgatcttg	240
atcccagatg aaaagcaggc taccatgtac aaaaacgaca tcgaacaaa cttggaagcc	300
ggtaacatgt tgatgttcgc tcacgggttc aacatccatt tcggttgat tgttccacca	360
aaggacgttg atgtcactat gatcgcttca aagggtccag gtcacaccgt tagatccgaa	420

tacgaagaag gtaaagggtg cccatgcttg gttgctgtcg aacaagacgc tactggcaag	480
gcttttgata tggctttggc ctacgcttta gccatcgggtg gtgctagagc cgggtgtcttg	540
gaaactacct tcagaaccga aactgaaacc gacttggttc gtgaacaagc tgttttatgt	600
ggtgggtgtc gcgctttgat gcaggccggt ttgaaacct tggttgaagc cggttacgac	660
ccaagaaacg ctacttcga atgtatccat gaaatgaagt tgatcgttga cttgatctac	720
caatctggtt tctccggtat gcgttactct atctccaaca ctgctgaata cggtgactac	780
attaccggtc caaagatcat tactgaagat accaagaagg ctatgaagaa gattttgtct	840

gacattcaag atggtacctt tgccaaggac ttcttggttg acatgtctga tgctggttcc	900
cagggtccact tcaaggttat gagaaagttg gcctccgaac acccagctga agttgtcgggt	960
gaagaaatta gatccttgta ctctgggtcc gacgaagaca agttgattaa caactga	1017

<210> 915

<211> 1017

<212> DNA

<213> Artificial Sequence

<220><223> K9_Pat_SH

<400> 915

atggctaaga ttactacca agaagactgt aacttgctct tgttggatgg taagactatc	60
gccgttatcg gttacgggtc tcaagggtcac gctcatgccc tgaatgctaa ggaatccggt	120
tgtaacgtta tcattgggtt attagaaggt gcggaggagt ggaaaagagc tgaagaacaa	180
ggtttcgaag tctacaccgc tgctgaagct gctaagaagg ctgacatcat tatgatcttg	240
atcccagatg aaaagcaggc taccatgtac aaaaacgaca tcgaacaaa cttggaagcc	300

ggtaacatgt tgatgttcgc tcacggtttc aacatccatt tcggttgat tgttcacca 360
aaggacgttg atgtcactat gatcgctcca aagggtccag gtcacaccgt tagatccgaa 420
tacgaagaag gtaaaggtgt cccatgcttg gttgctgtcg aacaagacgc tactggcaag 480
gctttggata tggctttggc ctacgcttta gccatcggtg gtgctagagc cgggtgtcttg 540
gaaactacct tcagaaccga aactgaaacc gacttggtcg gtgaacaagc tgttttatgt 600

ggtgggtgtct gcgctttgat gcaggccggt tttgaaacct tgggtgaagc cggttacgac 660
ccaagaaacg cttacttcga atgtatccac gaaatgaagt tgatcgttga cttgatctac 720
caatctggtt tctccggtat gcgttactct atctccaaca ctgctgaata cggtgactac 780
attaccggtc caaagatcat tactgaagat accaagaagg ctatgaagaa gattttgtct 840
gacattcaag atggtaacct tgccaaggac ttcttggttg acatgtctga tgctggttcc 900
caggtcact tcaaggctat gagaaagttg gcctccgaac acccagctga agttgtcgg 960
gaagaaatta gatccttgta ctccgtgtcc gacgaagaca agttgattaa caactga 1017

<210> 916

<211> 1017

<212> DNA

<213> Artificial Sequence

<220><223> K9_Quentin

<400> 916

atggctaaga ttactacca agaagactgt aacttgtcct tgttggatgg taagactatc 60
gccgttatcg gtiacggttc tcaaggtcac gtcctatgcc tgaatgctaa ggaatccggt 120
tgtaacgtta tcattggttt attcgaaggt gcggaggagt ggaaaagagc tgaagaacaa 180
ggtttcgaag tctacaccgc tgctgaagct gctaagaagg ctgacatcat tatgatcttg 240
atcccagatg aaaagcaggc taccatttac aaaaacgaca tcgaacaaaa cttggaagcc 300
ggtaacatgt tgatgttcgc tcacggtttc aacatccatt tcggttgat tgttcacca 360

aaggacgttg atgtcactat gatcgctcca aagggtccag gtcacaccgt tagatccgaa 420
tacgaagaag gtaaaggtgt cccatgcttg gttgctgtcg aacaagacgc tactggcaag 480
gctttggata tggctttggc ctacgcttta gccatcggtg gtgctagagc cgggtgtcttg 540
gaaactacct tcagaaccga aactgaaacc gacttggtcg gtgaacaagc tgttttatgt 600
ggtgggtgtct gcgctttgat gcaggccggt tttgaaacct tgggtgaagc cggttacgac 660
ccaagaaacg cttacttcga atgtatccac gaaatgaagt tgatcgttga cttgatctac 720
caatctggtt tctccggtat gcgttactct atctccaaca ctgctgaata cggtgactac 780

attaccggtc caaagatcat tactgaagat accaagaagg ctatgaagaa gattttgtct 840
gacattcaag atggtacctt tgccaaggac ttcttggttg acatgtctga tgctggttcc 900
caggtccact tcaaggctat gagaaagttg gcctccgaac acccagctga agttgtcggt 960
gaagaaatta gatccttgta ctcttggtcc gacgaagaca agttgattaa caactga 1017
<210> 917
<211> 1017
<212> DNA
<213> Artificial Sequence
<220><223> K9_Ralph_SH
<400> 917
atggctaaga ttactacca agaagactgt aacttgctct tgttggatgg taagactatc 60
gccgttatcg gttacggttc tcaaggteac gtcctatgcc tgaatgctaa ggaatccggt 120

tgtaacgtta tcattggttt attcgaaggt gcggaggagt ggaaaagagc tgaagaacaa 180
ggtttagaag tctacaccgc tgctgaagct gctaagaagg ctgacatcat tatgatcttg 240
atcccagatg aaaagcaggc taccatgtac aaaaacgaca tcgaaccaaa ctggaagcc 300
ggtaacatgt tgatgttcgc tcacggtttc aacatccatt tcggttgtat tgttccacca 360
aaggacgttg atgtcactat gatcgctcca aagggtccag gtcacaccgt tagatccgaa 420
tacgaagaag gtaaagggtg cccatgcttg gttgctgtcg aacaagacgc tactggcaag 480
gctttggata tggttttggc ctacgcttta gccatcgggtg gtgctagagc cgggtgtcttg 540

gaaactacct tcagaaccga aactgaaacc gacttggtcg gtgaacaagc tgttttatgt 600
gggtgtgtct gcgctttgat gcaggccggt tttgaaacct tgggtgaagc cggttacgac 660
ccaagaaacg cttacttcga atgtatccac gaaatgaagt tgatcgttga ctgatctac 720
caatctggtt tcctccggtat gcgttactct atctccaaca ctgctgaata cggtgactac 780
attaccggtc caaagatcat tactgaagat accaagaagg ctatgaagaa gattttgtct 840
gacattcaag atggtacctt tgccaaggac ttcttggttg acatgtctga tgctggttcc 900
caggtccact tcaaggctat gagaaagttg gcctccgaac acccagctga agttgtcggt 960

gaagaaatta gatccttgta ctcttggtcc gacgaagaca agttgattaa caactga 1017
<210> 918
<211> 1017
<212> DNA
<213> Artificial Sequence
<220><223> K9_Ursala

<400> 918

atggctaaga ttactacca agaagactgt aacttgctct tgttggatgg taagactatc	60
gccgttatcg gttacgggtc tcaaggtcac gctcatgccc tgaatgctaa ggaatccggt	120
tgtaacgtta tcattgggtt attcgaaggt gtggaggagt ggaaaagagc tgaagaacaa	180
ggtttcgaag tctacaccgc tgctgaagct gctaagaagg ctgacatcat tatgatcttg	240
atcccagatg aaaagcaggc taccatgtac aaaaacgaca tcgaacaaa ctiggaagcc	300
ggtaacatgt tgatgttcgc tcacggtttc aacatccatt tcggttgtat tgttccacca	360
aaggacgttg atgtcactat gatcgctcca aagggtccag gtcacaccgt tagatccgaa	420
tacgaagaag gtaaagggtg cccatgcttg gttgctgtcg aacaagacgc tactggcaag	480
gctttggata tggctttggc ctacgcttta gccatcggtg gtgctagagc cgggtgtcttg	540
gaaactacct tcagaaccga aactgaaacc gacttggtcg gtgaacaagc tgttttatgt	600
ggtaggtgtt gcgctttgat gcaggccggt ttgaaacct tgggtgaagc cggttacgac	660
ccaagaaacg cttacttcga atgtatccac gaaatgaagt tgatcgttga ctgatctac	720
caatctgggt tctccggtat gcgttactct atctccaaca ctgctgaata cggtgactac	780
attaccgggc caaagatcat tactgaagat accaagaagg ctatgaagaa gattttgtct	840
gacattcaag atggtacctt tgccaaggac ttcttggttg acatgtctga tgctggttcc	900
caggctccact tcaaggctat gagaaagttg gcctccgaac acccagctga agttgtcgg	960
gaagaaatta gatccttgta ctcttggtcc gacgaagaca agttgattaa caactga	1017

<210> 919

<211> 1017

<212> DNA

<213> Artificial Sequence

<220><223> K9_Xavier_SH

<400> 919

atggctaaga ttactacca agaagactgt aacttgctct tgttggatgg taagactatc	60
gccgttatcg gttacgggtc tcaaggtcac gctcatgccc tgaatgctaa ggaatccggt	120
tgtaacgtta tcattgggtt attcgaaggt gcggaggagt ggaaaagagc tgaagaacaa	180
ggtttcgaag tctacaccgc tgctgaagct gctaagaagg ctgacatcat tatgatcttg	240
atcccagatg aaaagcaggc tgccatgtac aaaaacgaca tcgaacaaa ctiggaagcc	300
ggtaacatgt tgatgttcgc tcacggtttc aacatccatt tcggttgtat tgttccacca	360
aaggacgttg atgtcactat gatcgctcca aagggtccag gtcacaccgt tagatccgaa	420

tacgaagaag gtaaagggtg cccatgcttg gttgctgtcg aacaagacgc tactggcaag 480

gcttttgata tggttttggc ctacgcttta gccatcggtg gtgctagagc cgggtgtcttg 540

gaaactacct tcagaaccga aactgaaacc gacttggtcg gtgaacaagc tgttttatgt 600

ggtggtgtct gcgctttgat gcaggccggt tttgaaacct tggttgaagc cggttacgac 660

ccaagaaacg cttacttcga atgtatccac gaaatgaagt tgatcgttga cttgatctac 720

caatctgggt tctccggtat gcgttactct atctccaaca ctgctgaata cggtgactac 780

attaccggtc caaagatcat tactgaagat accaagaagg ctatgaagaa gattttgtct 840

gacattcaag atggtacctt tgccaaggac ttcttggttg acatgtctga tgctggttcc 900

caggtccact tcaaggctat gagaaagttg gcctccgaac acccagctga agttgtcgg 960

gaagaaatta gatccttgta ctctgggtcc gacgaagaca agttgattaa caactga 1017

<210> 920

<211> 44

<212> DNA

<213> Artificial Sequence

<220><223> oMT93A

<400> 920

gatcccagat gaaatgcagg ctgccatgta caaaaacgac atcg 44

<210> 921

<211> 44

<212> DNA

<213> Artificial Sequence

<220><223> oMT93Arev

<400> 921

cgatgtcgtt tttgtacatg gcagcctgca tttcatctgg gatc 44

<210> 922

<211> 44

<212> DNA

<213> Artificial Sequence

<220><223> oMT93I

<400> 922

gatcccagat gaaatgcagg ctatcatgta caaaaacgac atcg 44

<210> 923

<211> 44
 <212> DNA
 <213> Artificial Sequence
 <220><223> oMT93Irev
 <400> 923
 cgatgctgtt tttgtacatg atagcctgca tttcatctgg gatc 44
 <210> 924
 <211> 22
 <212> DNA
 <213> Artificial Sequence
 <220><223> pHR81-F
 <400> 924
 acaccagta ttttcccttt cc 22
 <210> 925
 <211> 22
 <212> DNA
 <213> Artificial Sequence
 <220><223> pHR81- Rev
 <400> 925
 ctagtgtaca gatgtatgtc gg 22
 <210> 926
 <211> 570
 <212> PRT
 <213> Artificial Sequence
 <220><223> Lactococcus lactis DHAD
 <400> 926
 Met Glu Phe Lys Tyr Asn Gly Lys Val Glu Ser Ile Glu Leu Asn Lys
 1 5 10 15
 Tyr Ser Lys Thr Leu Thr Gln Asp Pro Thr Gln Pro Ala Thr Gln Ala
 20 25 30
 Met His Tyr Gly Ile Gly Phe Lys Asp Glu Asp Phe Lys Lys Ala Gln
 35 40 45
 Val Gly Ile Val Ser Met Asp Trp Asp Gly Asn Pro Cys Asn Met His

50	55	60	
Leu Gly Thr	Leu Gly Ser Lys Ile Lys Asn Ser Val	Asn Gln Thr Asp	
65	70	75	80
Gly Leu Ile Gly Leu Gln Phe His Thr	Ile Gly Val Ser Asp Gly Ile		
	85	90	95
Ala Asn Gly Lys Leu Gly Met Arg Tyr Ser Leu Val Ser Arg Glu Val			
	100	105	110
Ile Ala Asp Ser Ile Glu Thr Asn Ala Gly Ala Glu Tyr Tyr Asp Ala			
	115	120	125
Ile Val Ala Val Pro Gly Cys Asp Lys Asn Met Pro Gly Ser Ile Ile			
	130	135	140
Gly Met Ala Arg Leu Asn Arg Pro Ser Ile Met Val Tyr Gly Gly Thr			
145	150	155	160
Ile Glu His Gly Glu Tyr Lys Gly Glu Lys Leu Asn Ile Val Ser Ala			
	165	170	175
Phe Glu Ala Leu Gly Gln Lys Ile Thr Gly Asn Ile Ser Glu Glu Asp			
	180	185	190
Tyr His Gly Val Ile Cys Asn Ala Ile Pro Gly Gln Gly Ala Cys Gly			
	195	200	205
Gly Met Tyr Thr Ala Asn Thr Leu Ala Ser Ala Ile Glu Thr Leu Gly			
	210	215	220
Met Ser Leu Pro Tyr Ser Ala Ser Asn Pro Ala Val Ser Gln Glu Lys			
225	230	235	240
Glu Asp Glu Cys Asp Glu Ile Gly Leu Ala Ile Lys Asn Leu Leu Glu			
	245	250	255
Lys Asp Ile Lys Pro Ser Asp Ile Met Thr Lys Glu Ala Phe Glu Asn			
	260	265	270
Ala Ile Thr Ile Val Met Val Leu Gly Gly Ser Thr Asn Ala Val Leu			
	275	280	285
His Ile Ile Ala Met Ala Asn Ala Ile Gly Val Glu Ile Thr Gln Asp			
290	295	300	

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

 Cys Leu Thr Val Thr Gly Lys Thr Leu Ala Glu Asn Val Glu Thr Ala
 355 360 365
 Leu Asp Leu Asp Phe Asp Ser Gln Asp Ile Ile Arg Pro Leu Glu Asn
 370 375 380
 Pro Ile Lys Ala Thr Gly His Leu Gln Ile Leu Tyr Gly Asn Leu Ala
 385 390 395 400
 Glu Gly Gly Ser Val Ala Lys Ile Ser Gly Lys Glu Gly Glu Phe Phe
 405 410 415

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

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

 Pro Lys Ala Thr Arg Gly Val Leu Ala Lys Phe Ala Lys Leu Thr Arg

545 550 555 560
 Pro Ala Ser Glu Gly Cys Val Thr Asp Leu
 565 570
 <210> 927
 <211> 343
 <212> PRT
 <213> Artificial Sequence
 <220><223> K9C4
 <400> 927
 Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn
 1 5 10 15
 Leu Ser Leu Leu Asp Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser
 20 25 30
 Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val
 35 40 45
 Ile Ile Gly Leu Phe Glu Gly Ala Glu Glu Trp Lys Arg Ala Glu Glu
 50 55 60
 Gln Gly Phe Glu Val Tyr Thr Ala Ala Glu Ala Ala Lys Lys Ala Asp
 65 70 75 80
 Ile Ile Met Ile Leu Ile Pro Asp Glu Lys Gln Ala Thr Met Tyr Lys
 85 90 95
 Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala
 100 105 110
 His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val
 115 120 125
 Asp Val Thr Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser
 130 135 140
 Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln
 145 150 155 160
 Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala
 165 170 175
 Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu

180 185 190
 Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val
 195 200 205
 Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr

 210 215 220
 Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile
 225 230 235 240
 Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile
 245 250 255
 Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile
 260 265 270
 Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln

 275 280 285
 Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly
 290 295 300
 Ser Gln Val His Ala Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro
 305 310 315 320
 Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
 325 330 335
 Glu Asp Lys Leu Ile Asn Asn
 340

<210> 928

<211> 343

<212> PRT

<213> Artificial Sequence

<220><223> K9C8

<400> 928

Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn
 1 5 10 15
 Leu Ser Leu Leu Asp Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser
 20 25 30
 Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val

35 40 45
 Ile Ile Gly Leu Phe Glu Gly Ala Glu Glu Trp Lys Arg Ala Glu Glu

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

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

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

 245 250 255
 Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile
 260 265 270
 Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln
 275 280 285

Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly
290 295 300

Ser Gln Val His Thr Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro

305 310 315 320

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

325 330 335

Glu Asp Lys Leu Ile Asn Asn

340

<210> 929

<211> 4978

<212> DNA

<213> Artificial Sequence

<220><223> K9SB2_SH in pBAD.KARI

<400> 929

aagaaaccaa ttgtccatat tgcacagac attgccgtca ctgcgtcttt tactggctct 60

tctcgctaac caaacggta accccgctta ttaaaagcat tctgtaacaa agcgggacca 120

aagccatgac aaaaacgcgt aacaaaagt tctataatca cggcagaaaa gtccacattg 180

attatttgca cggcgtcaca ctttgctatg ccatagcatt tttatccata agattagcgg 240

atcctacctg acgtttttta tcgcaactct ctactgtttc tccatacccg ttttttgggc 300

taacaggagg aagttaaac agtatggcta agatttacta ccaagaagac tgtaacttgt 360

ccttggtgga tggtaagact atcgccgtta tcggttacgg ttctcaaggt cacgctcatg 420

ccctgaatgc taaggaatcc ggttgtaacg ttatcattgg tttattcgaa ggtgcggagg 480

agtggaaaag agctgaagaa caaggtttcg aagtctacac cgctgctgaa gctgctaaga 540

aggctgacat cattatgac ttgatcccag atgaaaagca ggctaccatg tacaaaaacg 600

acatcgaacc aaacttgga gccgtaaca tgttgatgtt cgctcacggt ttcaacatcc 660

atttcggttg tattgttcca ccaaaggacg ttgatgtcac tatgatcgct ccaaagggtc 720

caggtcacac cgtagatcc gaatacgaag aaggtaaagg tgtcccatgc ttggttgctg 780

tcgaacaaga cgtactggc aaggttttg atatggcttt ggctacgct ttagccatcg 840

gtggtgctag agccggtgtc ttggaaacta ccttcagaac cgaaactgaa accgacttgt 900

tcggtgaaca agctgtttta tgtggtggtg tctgcgcttt gatgcaggcc ggttttgaaa 960

ccttggttga agccggttac gaccaagaa acgcttactt cgaatgtatc cacgaaatga 1020

agttgatcgt tgacttgatc taccaatctg gtttctccgg tatgcgttac tctatctcca	1080
acactgctga atacggtgac tacattaccg gtccaaagat cattactgaa gataccaaga	1140
aggctatgaa gaagattttg tctgacattc aagatggtac ctttgccaag gacttcttgg	1200
ttgacatgtc tgatgctggg tcccagggtcc acttcaaggc tatgagaaaag ttggcctccg	1260
aacaccacgc tgaagttgtc ggtgaagaaa ttagatcctt gtactcctgg tccgacgaag	1320
acaagttgat taacaactga ggccctgcag gccttgccgg atgagagaag attttcagcc	1380
tgatacagat taaatcagaa cgcagaagcg gtctgataaa acagaatttg cctggcggca	1440
gtagcgcggg ggtcccactt gaccccatgc cgaactcaga agtgaaacgc cgtagcgcgg	1500
atggtagtgt ggggtctccc catgcgagag tagggaactg ccaggcatca aataaaacga	1560
aaggctcagt cgaagactg ggcccttcgt tttatctgtt gtttgcggg gaacgctctc	1620
ctgagtagga caaatccgcc gggagcggat ttgaacgttg cgaagcaacg gcccgagggg	1680
tgccgggcag gacgcccgcc ataaactgcc aggcatacaa ttaagcagaa ggccatcctg	1740
acggatggcc tttttgcgtt tctacaaact cttttgttta tttttctaaa tacattcaaa	1800
tatgtatccg ctcatgagac aataaccctg ataatgctt caataatatt gaaaaaggaa	1860
gagtatgagt attcaacatt tccgtgtcgc ccttattccc ttttttgcgg cattttgcct	1920
tcctgttttt gctcaccag aaacgctggg gaaagtaaaa gatgctgaag atcagttggg	1980
tgcacgagtg ggttacatcg aactggatct caacagcggg aagatccttg agagttttcg	2040
ccccgaagaa cgttttccaa tgatgagcac ttttaaagtt ctgctatgtg gcgcggtatt	2100
atcccgtgtt gacgccgggc aagagcaact cggtcgccgc atacactatt ctcagaatga	2160
cttggttaga tactcaccag tcacagaaaa gcattttacg gatggcatga cagtaagaga	2220
attatgcagt gctgccataa ccatgagtga taacactgcg gccaaattac ttctgacaac	2280
gatcggagga ccgaaggagc taaccgcttt ttgacacaac atgggggatc atgtaactcg	2340
ccttgatcgt tgggaaccgg agctgaatga agccatacca aacgacgagc gtgacaccac	2400
gatgcctgta gcaatggcaa caacgttgcg caaactatta actggcgaac tacttactct	2460
agcttcccgg caacaattaa tagactggat ggaggcggat aaagtgcag gaccacttct	2520
gcgctcggcc ctccggctg gctggtttat tgctgataaa tctggagccg gtgagcgtgg	2580
gtctcgggt atcattgcag cactggggcc agatggtaag cctcccgtg tcgtagtatt	2640
ctacacgacg gggagtcagg caactatgga tgaacgaaat agacagatcg ctgagatagg	2700
tgccctactg attaagcatt ggtaactgtc agaccaagtt tactcatata tactttagat	2760
tgatttaaaa cttcattttt aatttaaaag gatctagggt aagatccttt ttgataatct	2820
catgacaaaa atccctaac gtgagttttc gttccactga gcgtcagacc ccgtagaaaa	2880

gatcaaagga tcttcttgag atcctttttt tctgcgcgta atctgctgct tgcaaacaaa	2940
aaaaccaccg ctaccagcgg tggtttgttt gccggatcaa gagctaccaa ctctttttcc	3000
gaaggtaact ggcttcagca gagcgcagat accaaatact gtccttctag tntagccgta	3060
gttaggccac cacttcaaga actctgtagc accgcctaca tacctcgctc tgctaactct	3120
gttaccagtg gctgctgcca gtggcgataa gtcgtgtctt accgggttgg actcaagacg	3180
atagttaccg gataaggcgc agcggtcggg ctgaacgggg ggctcgtgca cacagcccag	3240
cttgagcga acgacctaca ccgaactgag atacctacag cgtgagctat gagaaagcgc	3300
cacgttcccc gaaggagaaa aggcggacag gtatccgta agcggcaggg tcggaacagg	3360
agagcgcacg agggagcttc caggggaaa cgcctggtat ctttatagtc ctgtcgggtt	3420
tcgccacctc tgacttgagc gtcgattttt gtgatgctcg tcaggggggc ggagcctatg	3480
gaaaaacgcc agcaacgcgg cttttttacg gttcctggcc ttttctggc cttttgctca	3540
catgttcttt ctgcgttat cccctgattc tgggataac cgtattaccg cttttgagt	3600
agctgatacc gctcgccga gccgaacgac cgagcgcagc gactcagtga gcgaggaagc	3660
ggaagagcgc ctgatcggtt attttctct tacgcatctg tgcggtatct cacaccgcat	3720
atggtgcaat ctgagtaca tctgctctga tgcgcatag ttaagccagt atacactccg	3780
ctatcgctac gtgactgggt catggctgcg ccccgacacc cgccaacacc cgctgacgcg	3840
ccctgacggg ctgtctgct cccggcatcc gcttacagac aagctgtgac cgtctccggg	3900
agctgcatgt gtcagaggtt ttcacgtca tcaccgaaac gcgcgaggca gcagatcaat	3960
tcgcgcgcga aggcgaagcg gcatgcataa tgtgcctgtc aaatggacga agcagggatt	4020
ctgcaaaccc tatgctactc cgtcaagccg tcaattgtct gattcgttac caattatgac	4080
aacttgacgg ctacatcatt cactttttct tcacaaccgg cacggaactc gctcgggctg	4140
gccccggtgc atttttttaa taccgcgag aaatagagtt gatcgtcaaa accaacattg	4200
cgaccgacgg tggcgtatgg catccgggtg gtgctcaaaa gcagcttcgc ctggctgata	4260
cgttggctct cgcgccagct taagacgcta atccctaact gctggcggaa aagatgtgac	4320
agacgcgacg gcgacaagca aacatgctgt gcgacgtgg cgatatcaaa attgctgtct	4380
gccaggtgat cgctgatgta ctgacaagcc tcgcgtaccc gattatccat cgggtgatgg	4440
agcgactcgt taatcgttc catgcgccg agtaacaatt gctcaagcag atttatgcc	4500
agcagctccg aatagcggc tccccctgc ccggcgtaa tgatttgccc aaacaggtcg	4560
ctgaaatgcg gctggtgctc ttcacccggg cgaaagaacc ccgtattggc aaatattgac	4620
ggccagttaa gccattcatg ccagtaggcg cgcggacgaa agtaaaccca ctggtgatac	4680

cattcgcgag cctccggatg acgaccgtag tgatgaatct ctctggcgg gaacagcaaa 4740

atatcacccg gtcggcaaac aaattctcgt cctgatTTTT tcaccacccc ctgaccgga 4800

atggtgagat tgagaatata acctttcatt ccagcgggc ggtcgataaa aaaatcgaga 4860

taaccgttgg cctcaatcgg cgttaaacc gccaccagat gggcattaaa cgagtatccc 4920

ggcagcagg gatcatTTTg cgcttcagcc atactTTTca tactcccc attcagag 4978

<210> 930

<211> 12319

<212> DNA

<213> Artificial Sequence

<220><223> pLH744

<400> 930

aggccagagg aaaataatat caagtgtg aaactTTTT tcttgaatt ttgcaacat 60

caagtcatag tcaattgaat tgaccaatt tcacattaa gatTTTTTT tttcatccg 120

acatacatct gtacactagg aagccctgtt tttctgaagc agcttcaaat atatatattt 180

ttacatatt tattatgatt caatgaacaa tctaattaaa tcgaaaacaa gaaccgaaac 240

gcgaataaat aatttattta gatggtgaca agtgataag tcctcatcgg gacagctacg 300

atttctcttt cggTTTTggc tgagctactg gttgctgtga cgcagcggca ttagcgcggc 360

gttatgagct accctcgtgg cctgaaagat ggcgggaata aagcggaact aaaaattact 420

gactgagcca tattgaggtc aatttgtcaa ctgtcaagt cacgtttggt ggacggcccc 480

ttccaacga atcgatatata ctaacatgcg cgcgttctct atatacatat atacatatat 540

atatatatat atatgtgtgc gtgtatgtgt acacctgtat ttaatttctt tactcgcggg 600

ttttctttt ttctcaattc ttggttctt ctttctcgag cggaccggat cctcgcaac 660

tccaaaatga gctatcaaaa acgatagatc gattaggatg actttgaaat gactccgcag 720

tggactggcc gttaatttca agcgtgagta aaatagtca tgacaaaaga tgagctaggc 780

ttttgtaaaa atatcttacg ttgtaaaatt ttagaaatca ttatttctt catatcattt 840

gtcattgac cttcagaaga aaagagccga ccaataatat aaataataa ataaaaataa 900

tattccatta ttctaaaca gattcaatac tcattaaaaa actatatcaa ttaatttgaa 960

ttaacgcggc cgcttaacca cagcaaccag gacaacattt ttgccagtt tcttcaggct 1020

tccaaaagtc gtgtacggct ccctagaag cagacgaaac gatgtgagca tatttaccaa 1080

ggataccgcg tgaatagagc ggtggcaatt caatggcttc ttgacgatgt tttaactctt 1140

catcgagat atcaaagtgt aattccttag tgtcttggtc aatagtgact atgtctctg 1200

tttgcaggta ggcgattgga ccgccatctt gtgcttcagg agcgatatga cccacgacaa	1260
gaccataagt accacctgag aagcggccat ctgtcagaag ggcaactttt tcaccttgcc	1320
ctttaccaac aatcattgat gaaagggaaa gcatttcagg cataccagga ccgccctttg	1380
gtcctacaaa acgtacgaca acaacatcac catcaacaat atcatcattc aagacagctt	1440
caatggcttc ttcttcagaa ttaaagacct tagcaggacc gacatgacga cgcactttta	1500
caccagaaac tttagcaacg gcaccgtctg gagccaagtt accatggaga ataatgagcg	1560
gaccatcttc acgttttaga ttttcaagcg gcataataac cttttgacca ggtgttaaat	1620
catcaaaagc cttcaaattt tcagcgactg ttttgccagt acaagtata cggtcaccat	1680
gaaggaagcc atttttaagg agatatattca taactgctgg taccctccg accttgtaaa	1740
ggtcttggaa tacatatga ccagaagggt tcaaatcagc caaatgagga actttttctt	1800
ggaaagtatt gaaatcatca agtgtcaatt ccacattagc agcatgggca atagctaaga	1860
ggtgaagggt tgagttaggt gaacctccca gagccatagt tacagtaata gcatcttcaa	1920
aagcttcacg cgttaaaatg tcagaagggt ttaagcccat ttcgagcatt ttgacaacag	1980
cgcgaccagc ttcttcaata tctgttttct tttctgcgga ttcagccggg tgagaagatg	2040
aaccgggaag gctaagtcac aaaacttcaa tagctgtcgc cattgtgtta gcagtataca	2100
taccaccgca gcttcagga ccgggacaag cattacattc caaagcttta acttcttctt	2160
tggatcatat gccgtgggtc caatggccga caccttcaa gacagagact aaatcgatat	2220
ctttgccgtc taaattacca ggtgcaattg ttccgccgta agcaaaaatg gctgggatat	2280
ccatgttagc catagcgata acagaaccgg gcatgttttt atcacaaccg ccaatggcta	2340
caaaagcatc cgcattatga cctcccatgg ctgcttcaat agaacttgca ataatacac	2400
gagatgtcaa ggagaaacgc attccttggg ttcccatggc gattccatca gaaaccgtga	2460
ttgttccgaa ctgaactggc caagcaccag ctctcttaac accgactttg gctagtttac	2520
caaagtcattg taagtggata ttacaagggtg tgttttcagc ccaagttgaa atgacaccga	2580
cgatagggtt ttcaaagtct tcatcttgca taccagttgc acgcaacata gcacgattag	2640
gtgatttaac cattgaatcg taaacagaac tacgatttct taagtcttta agagtttttt	2700
tgtcagtcac actcacgtgg aacttagatt agattgctat gctttctctc taacgagcaa	2760
gaagtaaaaa aagtgtaat agaacaagaa aaatgaaact gaagcttgag aaattgaaga	2820
ccgtttatta gcttaaatat caatgggagg tcatcgaaag agaaaaaat caagaaagaa	2880
actctcaaga aaaagaaacg tgataaaaat ttttattgcc tctctcgacg aagagaaaga	2940
aacgaggcgg tccttttttt cttttccaaa ctttttagtac gggtaattag cgacacccta	3000

gaggaagaaa gaggggaaat ttagtatgct gtgcttgggt gtcttgaagt ggtacggcga 3060

tgcgcgagat ccgagaaaaat ctggaagagt aaaaaggggg tagaagcgtt ttgaagctat 3120

gccggcagat ctattttaaata ggcgcgccga cgtcaggtgg cacttttcgg ggaaatgtgc 3180

gcggaacccc tatttgttta tttttctaaa tacattcaaa tatgtatccg ctcattgagac 3240

aataaccctg ataatgctt caataatatt gaaaaaggaa gagtatgagt attcaacatt 3300

tccgtgtgc cttattccc ttttttgcgg cttttgcct tctgttttt gctcaccag 3360

aaacgtggt gaaagtaaaa gatgctgaag atcagttggg tgcacgagt gggttacatcg 3420

aactggatct caacagcggg aagatccttg agagttttcg ccccgaagaa cgttttccaa 3480

tgatgagcac ttttaaagtt ctgctatgtg gcgcggtatt atcccgtatt gacgccgggc 3540

aagagcaact cgttcgccgc atacactatt ctcaaatga cttggttag tactcaccag 3600

tcacagaaaa gcatcttacg gatggcatga cagtaagaga attatgcagt gctgccataa 3660

ccatgagtga taacactcgc gccaaattac ttttgacaac gatcggagga ccgaaggagc 3720

taaccgcttt ttgcacaac atgggggatc atgtaactcg cttgatcgt tgggaaccgg 3780

agctgaatga agccatacca aacgacgagc gtgacaccac gatgcctgta gcaatggcaa 3840

caacgttcgc caaactatta actggcgaac tacttactct agcttcccgg caacaattaa 3900

tagactggat ggaggcggat aaagttgcag gaccacttct gcgctcggcc ctccggctg 3960

gctggtttat tgctgataaa tctggagccg gtgagcgtgg gtctcgggt atcattgcag 4020

cactggggcc agatggtaag ccctcccgta tcgtagtat ctacacgacg gggagtcagg 4080

caactatgga tgaacgaaat agacagatcg ctgagatagg tgctcactg attaagcatt 4140

ggtaactgtc agaccaagtt tactcatata tacttttagat tgatttaaaa ctcatTTTT 4200

aattttaaag gatctagggtg aagatccttt ttgataatct catgacaaa atccctaac 4260

gtgagttttc gttccactga gcgtcagacc ccgtagaaaa gatcaaagga tcttcttgag 4320

atcctttttt tctgcgcgta atctgctgct tgcaaacaaa aaaaccaccg ctaccagcgg 4380

tggtttgttt gccggatcaa gagctaccaa ctctttttcc gaaggtaact ggcttcagca 4440

gagcgcagat accaaatact gttcttctag tgtagccgta gttaggccac cacttcaaga 4500

actctgtagc accgcctaca tacctcgtc tgctaactct gttaccagt gctgctgcca 4560

gtggcgataa gtcgtgtctt accgggttgg actcaagacg atagttaccg gataaggcgc 4620

agcggtcggg ctgaacgggg gggttcgtgca cacagcccag cttggagcga acgacctaca 4680

ccgaactgag atacctacag cgtgagctat gagaaagcgc cacgttccc gaaggagaaa 4740

aggcggacag gtatccggtg agcggcaggg tcggaacagg agagcgcacg agggagcttc 4800
cagggggaaa cgcttggtat ctttatagtc ctgtcgggtt tgcacacctc tgacttgagc 4860
gtcgatTTTT gtgatgctcg tcaggggggc ggagcctatg gaaaaacgcc agcaacgcgg 4920
cctttttacg gticctggcc ttttgctggc cttttgctca catgttcttt cctgcgttat 4980
cccctgattc tgtggataac cgtattaccg cctttgagtg agctgatacc gctcgccgca 5040
gccgaacgac cgagcgcagc gagtcaagtga gcgaggaagc ggaagagcgc ccaatacgca 5100
aacccctctt ccccgcgctt tggccgattc attaatgcag ctggcacgac aggtttcccg 5160

actggaagc gggcagtgag cgcaacgcaa ttaatgtgag ttagctcact cattaggcac 5220
cccaggcttt acactttatg ctccggctc gtatgttggtg tggaaattgtg agcggataac 5280
aatttcacac aggaacacgc tatgaccatg attacgcaa gctttttctt tccaattttt 5340
tttttttcgt cattataaaa atcattacga ccgagattcc cgggtaataa ctgatataat 5400
taaattgaag ctctaatttg tgagtttagt atacatgcat ttacttataa tacagttttt 5460
tagttttgct ggccgcatct tctcaaatat gcttccagc ctgcttttct gtaacgttca 5520
ccctctacct tagcatccct tccctttgca aatagtcctc ttccaacaat aataatgtca 5580

gatcctgtag agaccacatc atccacggtt ctatactgtt gaccaaatgc gtctcccttg 5640
tcatctaaac ccacaccggg tgtcataatc aaccaatcgt aaccttcac tcttccacc 5700
atgtctcttt gagcaataaa gccgataaca aaatctttgt cgctcttcgc aatgtcaaca 5760
gtacccttag tatattctcc agtagatagg gagcccttgc atgacaattc tgctaacatc 5820
aaaaggcctc taggttcctt tgttacttct tctgccgctt gttcaaacc gctaacaata 5880
cctgggcccc ccacaccgtg tgcattcgta atgtctgccc attctgctat tctgtataca 5940
cccgagagt actgcaattt gactgtatta ccaatgtcag caaattttct gtcttcgaag 6000

agtaaaaaat tgtacttggc ggataatgcc tttagcggct taactgtgcc ctccatggaa 6060
aaatcagtca agatatccac atgtgttttt agtaaaaaa ttttgggacc taatgcttca 6120
actaactcca gtaattcctt ggtggtacga acatccaatg aagcacacaa gtttgtttgc 6180
ttttcgtgca tgatattaaa tagcttggca gcaacaggac taggatgagt agcagcacgt 6240
tccttatatg tagctttcga catgatttat ctccgtttcc tgcaggtttt tgttctgtgc 6300
agttgggtta agaatactgg gcaatttcac gtttcttcaa cactacatat gcgtatatat 6360
accaatctaa gtctgtgctc ctcccttcgt tcttcttctt gttcggagat taccgaatca 6420

aaaaaatctt aaggaaaccg aaatcaaaaa aaagaataaa aaaaaaatga tgaattgaaa 6480
agcttgcatg cctgcaggtc gactctagta tactccgtct actgtacgat acacttccgc 6540
tcaggctcct gtcttttaac gaggccttac cactcttttg ttactctatt gatccagctc 6600

agcaaaggca gtgtgatcta agattctatc ttcgcgatgt agtaaaacta gctagaccga 6660
gaaagagact agaaatgcaa aaggcacttc tacaatggct gccatcatta ttatccgatg 6720
tgacgtcgca tttttttttt tttttttttt tttttttttt tttttttttt 6780
ttgtacaaat atcataaaaa aagagaatct ttttaagcaa ggattttctt aactttcttcg 6840

gcgacagcat caccgacttc ggtggctactg ttggaaccac ctaaatacacc agttctgata 6900
cctgcatcca aaaccttttt aactgcatct tcaatggctt taccttcttc aggcaagttc 6960
aatgacaatt tcaacatcat tgcagcagac aagatagtgg cgatagggtt gaccttattc 7020
tttggcaaat ctggagcggg accatggcat ggttcgtaca aaccaaagtc ggtgttcttg 7080
tctggcaaag aggccaagga cgcagatggc aacaaacca aggagcctgg gataacggag 7140
gcttcatcgg agatgatatc accaaacatg ttgctggtga ttataatacc atttaggtgg 7200
gttgggttct taactaggat catggcggca gaatcaatca attgatgttg aactttcaat 7260

gtagggaatt cgttcttgat ggtttcctcc acagtttttc tccataatct tgaagaggcc 7320
aaaacattag ctttatccaa ggaccaaata ggcaatgggtg gctcatgttg tagggccatg 7380
aaagcggcca ttcttgtgat tctttgcact tctggaacgg tgtattgttc actatcccaa 7440
gcgacaccat caccatcgtc ttctttcttc ttaccaaagt aaatacctcc cactaattct 7500
ctaacaacaa cgaagtcagt accttttagca aattgtggct tgattggaga taagtctaaa 7560
agagagtcgg atgcaaagtt acatgggtctt aagttggcgt acaattgaag ttctttacgg 7620
atttttagta aaccttgctc aggtctaaca ctaccggtac ccattttagg accaccacaa 7680

gcacctaaac aaacggcatc agccttcttg gaggttcca gcgcctcatc tgggaagtgga 7740
acacctgtag catcgaatgc agcaccacca attaaatgat tttcgaaatc gaacttgaca 7800
ttggaacgaa catcagaaat agctttaaga accttaatgg cttcggtgtt gatttcttga 7860
ccaacgtggt cacctggcaa aacgacgatc ttcttagggg cagacattac aatggtatat 7920
ccttgaaata tatataaaaa aaaaaaaaaa aaaaaaaaaa aaaaatgcag cttctcaatg 7980
atattcgaat acgctttgag gagatacagc ctaatatccg acaaaactgtt ttacagattt 8040
acgatcgtac ttgttaccca tcattgaatt ttgaacatcc gaacctggga gttttccctg 8100

aaacagatag tatatttgaa cctgtataat aatataatgt ctagcgcttt acggaagaca 8160
atgtatgtat ttcggttctt ggagaaacta ttgcatctat tgcataggta atcttgacag 8220
tcgcatcccc ggttcatttt ctgcgtttcc atcttgcaat tcaatagcat atctttgtta 8280
acgaagcatc tgtgttcat tttgtagaac aaaaatgcaa cgcgagagcg ctaatttttc 8340
aaacaagaa tctgagctgc atttttacag aacagaaatg caacgcgaaa gcgctatttt 8400
accaacgaag aatctgtgct tcatttttgt aaaacaaaaa tgcaacgcga gagcgctaat 8460

ttttcaaaca aagaatctga gctgcatTTT tacagaacag aaatgcaacg cgagagcgct	8520
atTTTaccAA caaagaatct atacttcttt ttgtttctac aaaaatgcat cccgagagcg	8580
ctatTTTTct aacaaagcat cttagattac tttttttctc ctttgtgcgc tetataatgc	8640
agtctcttga taactTTTTg cactgtaggt ccgttaaggt tagaagaagg ctactttggt	8700
gtctatTTTc tcttccataa aaaaagcctg actccacttc ccgcgtttac tgattactag	8760
cgaagctgcg ggtgcatTTT ttcaagataa aggcattccc gattatattc tataccgatg	8820
tggattgcgc atactttgtg aacagaaagt gatagcgttg atgattcttc attggtcaga	8880
aaattatgaa cggtttcttc tattttgtct ctatatacta cgtataggaa atgtttacat	8940
tttcgtattg ttttcgattc actctatgaa tagttcttac tacaatTTTT ttgtctaaag	9000
agtaatacta gagataaaca taaaaaatgt agaggctcgag tttagatgca agttcaagga	9060
gcgaaagggtg gatgggtagg ttatataggg atatagcaca gagatatata gcaaagagat	9120
acttttgagc aatgtttgtg gaagcgggtat tcgcaatatt ttagtagctc gttacagtcc	9180
ggtgcgtttt tggTTTTtg aaagtgcgtc ttcagagcgc ttttggtttt caaaagcgct	9240
ctgaagtTcc tatactttct agagaatagg aacttcggaa taggaacttc aaagcgtttc	9300
cgaaaacgag cgcttccgaa aatgcaacgc gagctgcgca catacagctc actgttcacg	9360
tcgcacctat atctgcgtgt tgccgtgata tatatataca tgagaagaac ggcatagtgc	9420
gtgtttatgc ttaaattcgt acttatatgc gtctatttat gtaggatgaa aggtagtcta	9480
gtacctctg tgatattatc ccattccatg cggggtatcg tatgcttctc tcagcactac	9540
cctttagctg ttctatatgc tgccactcct caattggatt agtctcatcc ttcaatgcta	9600
tcatttctt tgatattgga tcatatgcat agtaccgaga aactagagga tctccatta	9660
ccgacatttg ggcgtatata gtgcatatgt tcatgtatgt atctgtattt aaaacacttt	9720
tgtattattt ttctcatat atgtgtatag gtttatacgg atgatttaat tattacttca	9780
ccacccttta tttcaggctg atatcttagc cttgttacta gattaatcat gtaattagtt	9840
atgtcacgct tacattcacg ccttcccccc acatccgctc taaccgaaaa ggaaggagtt	9900
agacaacctg aagtctaggt ccctatttat ttttttatag ttatgttagt attaagaacg	9960
ttatttataat ttcaaatttt tctttttttt ctgtacagac gcgtgtacgc atgtaacatt	10020
atactgaaaa cttgcttga gaaggttttg ggacgctcga aggcTTtaat ttgcgggcgg	10080
ccgcacctgg taaaacctct agtggagtag tagatgtaat caatgaagcg gaagccaaaa	10140
gaccagagta gaggcctata gaagaaactg cgataccttt tgtgatggct aaacaaacag	10200

acatcttttt atatgttttt acttctgtat atcgtgaagt agtaagtgat aagcgaattt	10260
ggctaagaac gttgtaagtg aacaagggac ctcttttgcc tttcaaaaaa ggattaaatg	10320
gagttaatca ttgagattta gttttcgta gattctgtat ccctaaataa ctcccttacc	10380
cgacgggaag gcacaaaaga cttgaataat agcaaacggc cagtagccaa gaccaaataa	10440
tactagagtt aactgatggt cttaaacagg cattacgtgg tgaactcaa gaccaatata	10500
caaaatcgcg ataagttatt cttgcccacc aatttaagga gcctacatca ggacagtagt	10560
accattcctc agagaagagg tatacataac aagaaaatcg cgtgaacacc ttatataact	10620
tagcccgtta ttgagctaaa aaaccttgca aaatttccta tgaataagaa tacttcagac	10680
gtgataaaaa ttfactttct aactcttctc acgctgcccc tatctgttct tccgctctac	10740
cgtgagaaat aaagcatcga gtacggcagt tcgctgtcac tgaactaaaa caataaggct	10800
agttcgaatg atgaacttgc ttgctgtcaa acttctgagt tgccgctgat gtgacactgt	10860
gacaataaat tcaaaccggt tatagcggtc tcctccggtt ccggttctgc cacctccaat	10920
agagctcagt aggagtcaga acctctgcgg tggtctgcag tgactcatcc gcgtttcgta	10980
agttgtgcgc gtgcacatct cgcccgttcc cgtctatctt gcagcaggcg gaaattttca	11040
tcacgctgta ggacgcaaaa aaaaaataat taatcgtaca agaactctgg aaaaaaatt	11100
gaaaaatctt gtataaaagg gatgacctaa cttgactcaa tggcttttac acccagtatt	11160
ttccctttcc ttgtttgtta caattataga agcaagacaa aaacatatag acaacctatt	11220
cctaggagtt atatTTTTTT accctaccag caatataagt aaaaaactgt ttaaacagta	11280
tggaagaatg taagatggct aagatttact accaagaaga ctgtaacttg tccttgttgg	11340
atggtaagac tatcgccgtt atcggttacg gttctcaagg tcacgctcat gccctgaatg	11400
ctaaggaatc cggttgtaac gttatcattg gtttattcga aggtgaggag gagtggaaaa	11460
gagctgaaga acaaggtttc gaagtctaca ccgctgctga agctgctaag aaggctgaca	11520
tcattatgat ctgatccca gatgaaaagc aggctaccat gtacaaaaac gacatcgaac	11580
caaaacttga agccggtaac atgttgatgt tcgctcacgg tttcaacatc ctttcggtt	11640
gtattgttcc accaaaggac gttgatgtca ctatgatcgc tccaaagggt ccaggtcaca	11700
ccgttagatc cgaatacga gaaggtaaag gtgtcccatg cttggttgct gtcgaacaag	11760
acgtactgg caaggctttg gatatggctt tggcctacgc tttagccatc ggtggtgcta	11820
gagccggtgt cttggaaact accttcagaa ccgaaactga aaccgacttg ttcggtgaac	11880
aagctgtttt atgigggtgt gtctgcgctt tgatgcaggc cggttttgaa acctgggttg	11940
aagccggtta cgaccaaga aacgttact tcgaatgtat ccacgaaatg aagttgatcg	12000
ttgacttgat ctaccaatct ggtttctccg gtatgcgtta ctctatctcc aacactgctg	12060

aatacggatga ctacattacc ggtccaaaga tcattactga agataccaag aaggctatga 12120
 agaagatttt gtctgacatt caagatggta cctttgcca ggacttcttg gttgacatgt 12180
 ctgatgctgg ttcccaggtc cacttcaagg ctatgagaaa gttggcctcc gaacacccag 12240
 ctgaagtgtt cggatgaagaa attagatcct tgtactcctg gtccgacgaa gacaagtga 12300

ttaacaactg aggcctgc 12319

<210> 931

<211> 12304

<212> DNA

<213> Artificial Sequence

<220><223> K9SB2_SH_DHAD

<400> 931

tccattacc gacatttggg cgctatacgt gcatatgttc atgtatgtat ctgtatttaa 60
 aacacttttg tattattttt cctcatatat gtgtataggt ttatacggat gatttaatta 120
 ttacttcacc accctttatt tcaggtgat atcttagcct tgttactaga ttaatcatgt 180
 aattagttaa gtcacgctta cattcacgcc ctccccccac atccgctcta accgaaaagg 240
 aaggagttag acaacctgaa gtctaggtcc ctatttattt ttttatagtt atgttagtat 300

taagaacgtt atttataatt caaatttttc ttttttttct gtacagacgc gtgtacgcat 360
 glaacattat actgaaaacc ttgcttgaga aggttttggg acgctcgaag gctttaattt 420
 gcggggcgcc gcacctggta aaacctctag tggagtagta gatgtaatca atgaagcgga 480
 agccaaaaga ccagagtaga ggcctataga agaaactgcg ataccttttg tgatggctaa 540
 acaaacagac atctttttat atgtttttac ttctgtatat cgtgaagtag taagtataa 600
 gcgaatttgg ctaagaacgt tgtaagttaa caaggacct cttttgcctt tcaaaaaagg 660
 attaatgga gttaatcatt gagatttagt tttcgtaga ttctgtatcc ctataaact 720

cccttaccg acgggaaggc acaaaagact tgaataatag caaacggcca gtagccaaga 780
 ccaaataata ctagagttaa ctgatgggtc taaacaggca ttacgtggtg aactccaaga 840
 ccaatataca aaatatcgat aagttattct tgcccaccaa ttaaggagc ctacatcagg 900
 acagtagtac cattctcag agaagaggta tacataacaa gaaaatcgcg tgaacacctt 960
 atataactta gccgttatt gagctaaaaa accttgcaaa atttcctatg aataagaata 1020
 cttcagacgt gataaaaatt tactttctaa ctcttctcac gctgcccta tctgttcttc 1080
 cgctctaccg tgagaaataa agcatcgagt acggcagttc gctgtcactg aactaaaaca 1140

ataaggctag ttcaatgat gaacttgctt gctgtcaaac ttctgagttg ccgctgatgt 1200

gacactgtga caataaatc aaaccggtta tagcggcttc ctccggtacc ggttctgcca 1260
 cctccaatag agctcagtag gagtcagaac ctctcggtg gctgtcagtg actcatccgc 1320
 gtttcgtaag ttgtgcgcgt gcacatttcg cccgttcccg ctcatcttgc agcaggcgga 1380
 aattttcatc acgtcttagg acgcaaaaaa aaaataatta atcgtacaag aatcttgaa 1440
 aaaaaattga aaaattttgt ataaaaggga tgacctaaact tgactcaatg gcttttacac 1500
 ccagtatfff ccttttctt gtttgttaca attatagaag caagacaaaa acatatagac 1560

 aacctattcc taggagtat atttttttac cctaccagca atataagtaa aaaactgttt 1620
 aaacagtatg gctaagattt actaccaaga agactgtaac ttgtccttgt tggatggtaa 1680
 gactatcgcc gttatcggtt acggttctca aggtcacgct catgccctga atgctaagga 1740
 atccggttgt aacgttatca ttggtttatt cgaagggtgcg gaggagtgga aaagagctga 1800
 agaacaaggt ttcaagctt acaccgctgc tgaagctgct aagaaggctg acatcattat 1860
 gatcttgatc ccagatgaaa agcaggctac catgtacaaa aacgacatcg aaccaaactt 1920
 ggaagccggt aacatgttga tgttcgctca cggtttcaac atccatttcg gttgtattgt 1980

 tccaccaaag gacgttgatg tcaatgatg cgtccaaag ggtccaggtc acaccgttag 2040
 atccgaatac gaagaaggta aagggtgccc atgcttgggt gctgtcgaac aagacgctac 2100
 tggcaaggct ttggatatgg ctttggccta cgttttagcc atcgggtggtg ctagagccgg 2160
 tgtcttgaa actacctca gaaccgaaac tgaaccgac ttgttcgggtg aacaagctgt 2220
 tttatgttgt ggtgtctgcg cttttagtca ggccggtttt gaaaccttgg ttgaagccgg 2280
 ttacgacca agaaacgctt acttcgaatg tatccacgaa atgaagtga tcgttgactt 2340
 gatctaccaa tctggtttct ccggtatgcg ttactctatc tccaacactg ctgaatacgg 2400

 tgactacatt accggtccaa agatcattac tgaagatacc aagaaggcta tgaagaagat 2460
 ttgtctgac attcaagatg gtaccttgc caaggacttc ttggttgaca tgtctgatgc 2520
 tggttccag gtccactca aggctatgag aaagtggcc tccgaacacc cagctgaagt 2580
 tgtcggtgaa gaaattagat cttgtactc ctggtccgac gaagacaagt tgattaacaa 2640
 ctgaggccct gcaggccaga ggaaaaaat atcaagtgtt ggaaactttt tctcttgaa 2700
 tttttgaa atcaagtc atgcaattga attgaccaa tttcacattt aagatttttt 2760
 tttttcatc cgacatacat ctgtacacta ggaagccctg tttttctgaa gcagcttcaa 2820

 atatatatat tttttacata tttattatga ttcaatgaac aatctaatta aatcgaaaac 2880
 aagaaccgaa acgcaataa ataatttatt tagatggtga caagtgtata agtcctcatc 2940
 gggacagcta cgatttctct ttccggtttt gctgagctac tggttgctgt gacgcagcgg 3000
 cattagcgcg gcgttatgag ctacctctgt ggctgaaag atggcgggaa taaagcggaa 3060

ctaaaaatta ctgactgagc catattgagg tcaatttgc aactcgtcaa gtcacgtttg	3120
gtggacggcc cttttcaac gaatcgata tactaacatg cgcgcgttc ctatatacac	3180
atatacatat atatatatat atatatgtgt gctgtatgt gtacacctgt atttaatttc	3240
cttactcgcg ggtttttctt ttttctcaat tcttggcttc ctccttctcg agcggaccgg	3300
atcctcgca actccaaaat gagctatcaa aaacgataga tcgattagga tgactttgaa	3360
atgactccgc agtggactgg ccgttaattt caagcgtgag taaaatagtg catgacaaaa	3420
gatgagctag gcttttgtaa aaatatctta cgttgtaaaa ttttagaaat cattatttcc	3480
ttcatatcat tttgtcattg accttcagaa gaaaagagcc gaccaataat ataaataaat	3540
aaataaaaat aatattccat tatttctaaa cagattcaat actcattaaa aaactatc	3600
aattaatttg aattaacgcg gccgttaac cacagcaacc aggacaacat ttttgccag	3660
tttcttcagg ctccaaaag tctgttacgg ctcctctaga agcagacgaa acgatgtgag	3720
catatttacc aaggataccg cgtgaataga gcggtggcaa ttcaatggtc tcttgacgat	3780
gttttaactc ttcacggag atatcaaagt gtaattcctt agtgtcttgg tcaatagtga	3840
ctatgtctcc tgtttgcagg taggcgattg gaccgccatc ttgtgcttca ggagcgatat	3900
gaccacgac aagaccataa gtaccacctg agaagcggcc atctgtcaga agggcaactt	3960
ttcaccttg cctttacca acaatcattg atgaaagga aagcatttca ggcataccag	4020
gaccgccctt tggctctaca aaacgtacga caacaacatc accatcaaca atatcatcat	4080
tcaagacagc ttcaatggct tcttcttcag aattaaagac cttagcagga ccgacatgac	4140
gagcacttt tacaccagaa actttggcaa cggcaccgtc tggagccaag ttaccatgga	4200
gaataatgag cggaccatct tcacgtttag gattttcaag cggcataata accttttgac	4260
caggtgttaa atcatcaaaa gccttcaaat ttccagcgac tgttttgcca gtacaagtga	4320
tacggtcacc atgaaggaag ccatttttaa ggagatattt cataactgct ggtaccctc	4380
cgaccttgta aaggtcttgg aatacatatt gaccagaagg tttcaatca gccaaatgag	4440
gaacttttc ttggaaagta ttgaaatcat caagtgtcaa ttccacatta gcagcatggg	4500
caatagctaa gaggtgaagg gttgagttgg ttgaacctcc cagagccata gttacagtaa	4560
tagcatcttc aaaagcttca cgcgttaaaa tgcagaagg ttttaagccc atttcgagca	4620
ttttgacaac agcgcgacca gcttcttcaa tatctgctt cttttctgcg gattcagccg	4680
ggtgagaaga tgaacccgga aggctaagtc caaaaacttc aatagctgtc gccatttgt	4740
tagcagtata cataccaccg cagcttcag gaccgggaca agcattacat tccaaagctt	4800
taacttttc tttggtcata tcgccgtggt tccaatggcc gacaccttca aagacagaga	4860

ctaaatcgat atctttgccg tctaaattac cagggtgcaat tgttccgccg taagcaaaaa	4920
tggctgggat atccatgtta gccatagcga taacagaacc ggcatgttt ttatcacaac	4980
cgccaatggc tacaaaagca tccgcattat gacctcccat ggctgcttca atagaatctg	5040
caataatatc acgagatgtc aaggagaaac gcattccttg ggttcccatg gcgattccat	5100
cagaaaccgt gattgttccg aactgaactg gccaagcacc agcttcttca acaccgactt	5160
tggctagttt accaaagtca tgtaagtgga tattacaagg tgtgttttca gcccaagttg	5220
aatgacacc gacgataggt ttttcaaagt ctctcatctg cataccagtt gcacgcaaca	5280
tagcacgatt aggtgattta accattgaat cgtaaacaga actacgattt ctttaagtctt	5340
taagagtttt ttgtcagtc atactcacgt ggaacttaga ttagattgct atgctttctc	5400
tctaacgagc aagaagtaaa aaaagttgta atagaacaag aaaaatgaaa ctgaagcttg	5460
agaaattgaa gaccgtttat tagcttaaat atcaatggga ggtcatcgaa agagaaaaaa	5520
atcaagaaag aaactctcaa gaaaaagaaa cgtgataaaa atttttattg cctctctcga	5580
cgaagagaaa gaaacgaggc ggtccctttt ttcttttcca aacctttagt acgggtaatt	5640
agcgacaccc tagaggaaga aagaggggaa atttagtatg ctgtgcttgg gtgtcttgaa	5700
gtggtacggc gatgcgcgga gtccgagaaa atctggaaga gtaaaaaggg gtagaagcg	5760
ttttgaagct atgccggcag atctatttta atggcgcgcc gacgtcaggt ggcacttttc	5820
ggggaaatgt gcgcggaacc cctatttgtt tatttttcta aatacattca aatatgtatc	5880
cgctcatgag acaataaccc tgataaatgc ttcaataata ttgaaaaagg aagagtatga	5940
gtattcaaca ttccgtgtc gcccttattc ctttttttgc ggcattttgc ctctctgttt	6000
ttgctcacc agaaacgtg gtgaaagtaa aagatgctga agatcagttg ggtgcacgag	6060
tgggttacat cgaactggat ctcaacagcg gtaagatcct tgagagtttt cgccccgaag	6120
aacgttttcc aatgatgagc actttttaaag ttctgctatg tggcgcggtt ttatcccgtt	6180
ttgacgccgg gcaagagcaa ctcggtcgcc gcatacacta ttctcagaat gacttggttg	6240
agtactcacc agtcacagaa aagcatctta cggatggcat gacagtaaga gaattatgca	6300
gtgctgcat aaccatgagt gataacactg cggccaactt acttctgaca acgatcggag	6360
gaccgaagga gctaaccgct tttttgcaca acatggggga tcatgtaact cgccttgatc	6420
gttgggaacc ggagctgaat gaagccatac caaacgacga gcgtgacacc acgatgcctg	6480
tagcaatggc aacaacgttg cgcaactat taactggcga actacttact ctagcttccc	6540
ggcaacaatt aatagactgg atggaggcgg ataaagtgc aggaccactt ctgcgctcgg	6600

cccttcgggc tggctgggtt attgctgata aatctggagc cggtagcgt gggctctcgcg 6660
 giatcattgc agcactgggg ccagatggta agccctcccg tatcgtagtt atctacacga 6720
 cggggagtcg ggcaactatg gatgaacgaa atagacagat cgctgagata ggtgcctcac 6780
 tgattaagca ttggtaactg tcagaccaag ttactcata tatacttttag attgatttaa 6840
 aacttcattt ttaattttaa aggatctagg tgaagatcct ttttgataat ctcatgacca 6900
 aaatccctta acgtgagttt tcgttccact gagcgtcaga ccccgtagaa aagatcaaag 6960
 gatcttcttg agatcctttt tttctgcgcg taatctgctg cttgcaaaca aaaaaaccac 7020

cgctaccagc ggtggtttgt ttgccggatc aagagctacc aactcttttt ccgaaggtaa 7080
 ctggcttcag cagagcgagc ataccaata ctgttcttct agttagcgcg tagttaggcc 7140
 accacttcaa gaactctgta gcaccgccta catacctcgc tctgctaata ctgttaccag 7200
 tggctgctgc cagtggcgat aagtctgtc ttaccgggtt ggactcaaga cgatagttac 7260
 cggataaggc gcagcggtcg ggctgaacgg ggggttcgtg cacacagccc agcttggagc 7320
 gaacgaccta caccgaactg agatacctac agcgtgagct atgagaaagc gccacgcttc 7380
 ccgaaggag aaaggcggac aggtatccgg taagcggcag ggtcggaaca ggagagcgca 7440

cgaggagct tccaggggga aacgcctggt atctttatag tcctgtcggg tttcgccacc 7500
 tctgacttga gcgtcgattt ttgtgatgct cgtcaggggg gcggagccta tggaaaaacg 7560
 ccagcaacgc ggctttttta cggttcctgg ccttttgctg gccttttgct cacatgttct 7620
 ttcctgcgtt atccctgat tctgttgata accgtattac cgcctttgag tgagctgata 7680
 ccgctcgccg cagccgaacg accgagcgca gcgagtcagt gagcgaggaa gcggaagagc 7740
 gcccaatacg caaacgcct ctccccgcgc gttggccgat tcattaatgc agctggcacg 7800
 acaggtttcc cgactggaaa gcgggcagtg agcgcaacgc aattaatgtg agttagctca 7860

ctcattagc accccaggct ttacacttta tgettccggc tcgtatgttg tgtggaattg 7920
 tgagcggata acaatttcac acaggaaaca gctatgacca tgattacgcc aagctttttc 7980
 tttccaattt ttttttttc gtcattataa aaatcattac gaccgagatt cccgggtaat 8040
 aactgatata attaaattga agctctaatt tgtgagtta gtatacatgc atttacttat 8100
 aatacagttt ttagttttg ctggccgcat ctctcaaat atgcttcca gcctgctttt 8160
 ctgtaacgtt caccctctac cttagcatcc ctccctttg caaatagtcc tcttccaaca 8220
 ataataatgt cagatcctgt agagaccaca tcattccagg ttctatactg ttgaccaat 8280

gcgtctccct tgtcatctaa acccacaccg ggtgtcataa tcaaccaatc gtaaccttca 8340
 tctcttccac ccatgtctct ttgagcaata aagccgataa caaatctttt gtcgctcttc 8400
 gcaatgtcaa caglacctt agtatattct ccagtagata gggagccctt gcatgacaat 8460

tctgctaaca tcaaaaggcc tctaggttcc ttgtttactt cttctgccgc ctgcttcaaa	8520
ccgctaacaa tacctgggcc caccacaccg tgtgcattcg taatgtctgc ccattctgct	8580
attctgtata caccgcaga gtactgcaat ttgactgtat taccaatgtc agcaaatttt	8640
ctgtcttcga agagtaaaaa attgtacttg gcggataatg cctttagcgg cttactgtg	8700
ccctccatgg aaaaatcagt caagatatcc acatgtgttt ttagtaacaa aattttggga	8760
cctaattgctt caactaactc cagtaattcc ttggtggtac gaacatccaa tgaagcacac	8820
aagtttgttt gcttttcgtg catgatatta aatagcttgg cagcaacagg actaggatga	8880
gtagcagcac gttccttata tgtagctttc gacatgattt atcttcgttt cctgcaggtt	8940
tttgttctgt gcagttgggt taagaatact gggcaatttc atgtttcttc aacactacat	9000
atgcgtatat ataccaatct aagtctgtgc tccttccttc gttcttcctt ctgttcggag	9060
attaccgaat caaaaaaatt tcaaggaaac cgaaatcaaa aaaaagaata aaaaaaaat	9120
gatgaattga aaagcttgca tgcctgcagg tcgactctag tatactccgt ctactgtacg	9180
atacacttcc gctcaggtec ttgtccttta acgaggcctt accactcttt tgttactcta	9240
ttgatccagc tcagcaaagg cagtgtgac taagattcta tcttcgcgat gtagtaaaac	9300
tagctagacc gagaaagaga ctagaatgc aaaaggcact tctacaatgg ctgccatcat	9360
tattatccga tgtgacgtg catttttttt tttttttttt tttttttttt tttttttttt	9420
tttttttttt ttttgtacaa atatcataaa aaaagagaat cttttaagc aaggattttc	9480
ttaacttctt cggcgacagc atcaccgact tcggtggtac tgttgaacc acctaaatca	9540
ccagttctga tacctgcac caaaaccttt ttaactgcat cttcaatggc ttacacttct	9600
tcaggcaagt tcaatgacaa tttcaacatc attgcagcag acaagatagt ggcataggg	9660
ttgaccttat tctttggcaa atctggagcg gaacctggc atggttcgta caaaccaaat	9720
gcggtgttct tgtctggcaa agaggccaag gacgcagatg gcaacaaacc caaggagcct	9780
gggataacgg aggcttcac ggagatgata tcaccaacaa tgttgctggt gattataata	9840
ccatttaggt gggttgggtt cttaactagg atcatggcgg cagaatcaat caattgatgt	9900
tgaactttca atgtagggaa ttcttcttg atggtttcct ccacagtttt tctccataat	9960
cttgaagagg caaaaacatt agctttatcc aaggacaaa taggcaatgg tggtcatgt	10020
tgtagggccaa tgaagcggc cattcttgtg attctttgca cttctggaac ggtgtattgt	10080
tcactatccc aagcgacacc atcaccatcg tcttccttcc tcttaccaaa gtaaatacct	10140
cccactaatt ctctaacaac aacgaagtca gtacctttag caaattgtgg cttgattgga	10200
gataagtcta aaagagagtc ggatgcaaag ttacatggtc ttaagttggc gtacaattga	10260
agttctttac ggatttttag taaacctgtt tcagggtctaa cactaccggt accccattta	10320

ggaccaccca cagcacctaa caaaacggca tcagccttct tggaggcttc cagcgctca	10380
tctggaagtg gaacacctgt agcatcgata gcagcaccac caattaaatg attttcgaaa	10440
tcgaacttga cattggaacg aacatcagaa atagctttta gaaccttaat ggcttcggct	10500
gtgatttctt gaccaacgtg gtcacctggc aaaacgacga tcttcttagg ggcagacatt	10560
acaatggtat atccttgaaa tatatataaa aaaaaaaaaa aaaaaaaaaa aaaaaaatgc	10620
agcttctcaa tgatattcga atacgctttg aggagatata gcctaataatc cgacaaactg	10680
ttttacagat ttacgatcgt acttggtacc catcattgaa ttttgaacat ccgaacctgg	10740
gagttttccc tgaaacagat agtatatttg aacctgtata ataatatata gtctagcgct	10800
ttacggaaga caatgtatgt atttcggctc ctggagaaac tattgcatct attgcatagg	10860
taatcttgca cgtcgcaccc ccggttcatt ttctgcgttt ccatcttgca cttcaatagc	10920
atatctttgt taacgaagca tctgtgcttc atttttaga acaaaaatgc aacgcgagag	10980
cgctaatttt tcaaacaaag aatctgagct gcatttttac agaacagaaa tgcaacgcga	11040
aagcgctatt ttaccaacga agaactctgt cttcattttt gtaaaacaaa aatgcaacgc	11100
gagagcgcta atttttcaaa caaagaatct gagctgcatt tttacagaac agaaatgcaa	11160
cgcgagagcg ctattttacc aacaagaat ctatacttct ttttgttctt acaaaaatgc	11220
atcccagagag cgctattttt ctaacaaagc atcttagatt acttttttc tcctttgtgc	11280
gtctataat gcagtctctt gataactttt tgcaactgtg gtccgttaag gttagaagaa	11340
ggctactttg gtgtctattt tctcttccat aaaaaaagcc tgactccact tcccgcgttt	11400
actgattact agcgaagctg cgggtgcatt ttttcaagat aaaggcatcc ccgattatat	11460
tctataccga tgtggattgc gcatactttg tgaacagaaa gtgatagcgt tgatgattct	11520
tcattggtea gaaaattatg aacggtttct tctattttgt ctctatatac tacgtatagg	11580
aatgttttac attttcgtat tgttttcgat tcactctatg aatagtctt actacaattt	11640
ttttgtctaa agagtaatac tagagataaa cataaaaaat gtagaggtcg agtttagatg	11700
caagtccaag gagcgaaagg tggatgggta ggttatatag ggatatagca cagagatata	11760
tagcaaagag atacttttga gcaatgtttg tggaagcggg attcgcaata ttttagtagc	11820
tcgttacagt ccggtgcgtt tttggttttt tgaaagtgcg tcttcagagc gcttttggtt	11880
ttcaaaagcg ctctgaagtt cctatacttt ctagagaata ggaacttcgg aataggaact	11940
tcaaagcgtt tccgaaaacg agcgcttccg aaaatgcaac gcgagctgcg cacatacagc	12000
tcactgttca cgtcgcacct atatctgcgt gttgcctgta tatatatata catgagaaga	12060

acggcatagt gcgtgtttat gcttaaatgc gtacttatat gcgtctatit atgtaggatg 12120
aaaggtagtc tagtacctcc tgtgatatta tccattcca tgcggggtat cgtatgcttc 12180
cttcagcaact acccttttagc tgttctatat gctgccactc ctcaattgga ttagtctcat 12240
ccttcaatgc tatcatttcc ttgatattg gatcatatgc atagtaccga gaaactagag 12300
gatc 12304
<210> 932
<211> 9612
<212> DNA
<213> Artificial Sequence
<220><223> K9SB2_SH_81
<400> 932
catcatcaca gtttaaacag tatggctaag atttactacc aagaagactg taacttgtec 60

ttgttggatg gtaagactat cgccgttatc gggtacggtt ctcaaggta cgctcatgcc 120
ctgaatgcta aggaatccgg ttgtaacgtt atcattgggt tattcgaagg tgcggaggag 180
tggaagagag ctgaagaaca aggtttcgaa gtctacaccg ctgctgaagc tgctaagaag 240
gctgacatca ttatgatctt gatccagat gaaaagcagg ctaccatgta caaaaacgac 300
atcgaaccaa acttgggaagc cggtaacatg ttgatgttcg ctacggttt caacatccat 360
ttcggttgta ttgtccacc aaaggacgtt gatgtcacta tgatcgctcc aaagggtcca 420
ggtcacaccg ttagatccga atacgaagaa ggtaaagggtg tccatgctt ggttgctgtc 480

gaacaagacg ctactggcaa ggctttggat atggctttgg cctacgcttt agccatcggt 540
gggtgtagag ccggtgtctt ggaaactacc ttcagaaccg aaactgaaac cgacttgctc 600
gggtgaacaag ctgttttatg tgggtgtgtc tgcgctttga tgcaggccgg ttttgaaacc 660
ttggttgaag ccggttacga cccaagaaac gcttacttcg aatgtatcca cgaaatgaag 720
ttgatcggtt acttgatcta ccaatctggt ttctccggtg tgcgttactc tatctccaac 780
actgctgaat acggtgacta cattaccggt ccaaagatca ttactgaaga taccaagaag 840
gctatgaaga agattttgtc tgacattcaa gatggtacct ttgccaagga ctcttggtt 900

gacatgtctg atgctggctc ccaggtccac ttcaaggcta tgagaaagt ggccctccga 960
caccagctg aagtgtcggt tgaagaaatt agatccttgt actcctggct cgacgaagac 1020
aagttgatta acaactgagg ccctgcaggc cagaggaaaa taatatcaag tgctggaaac 1080
ttttctctt ggaatttttg caacatcaag tcatagtcaa ttgaattgac ccaatttcac 1140
atttaagatt ttttttttt catccgacat acatctgtac actaggaagc cctgtttttc 1200
tgaagcagct tcaatatat atatttttta catatttatt atgattcaat gaacaatcta 1260

attaaatcga aaacaagaac cgaaacgcga ataaataatt tatttagatg gtgacaagtg 1320

tataagtcct catcgggaca gctacgattt ctctttcggg tttggctgag ctactggttg 1380

ctgtgacgca gcggcattag cgcggcgtta tgagctaccc tcgtggcctg aaagatggcg 1440

ggaataaagc ggaactaaaa attactgact gagccatatt gaggtcaatt tgtcaactcg 1500

tcaagtcacg tttgggtggc ggcccccttc caacgaatcg tataactaa catgcgcgcg 1560

cttcctatat acacataac atatatatat atatatatat gtgtgcgtgt atgtgtacac 1620

ctgtatttaa tttccttact cgcggttttt tcttttttct caattcttgg cttcctcttt 1680

ctcgagcggc ccggatcctc cgcggtgccg gcagatctat ttaaatggcg cgccgacgtc 1740

aggtggcact tttcggggaa atgtgcgcgg aacccttatt tgtttatttt tctaaataca 1800

ttcaaatatg tatccgctca tgagacaata accctgataa atgcttcaat aatattgaaa 1860

aaggaagagt atgagtattc aacatttcg tgctgccctt attccctttt ttgcggcatt 1920

ttgccttcct gtttttgctc acccagaaac gctgggtgaaa gtaaaagatg ctgaagatca 1980

gttgggtgca cgagtgggtt acatcgaact ggatctcaac agcggtaaga tccttgagag 2040

ttttcgcccc gaagaacgtt ttccaatgat gagcactttt aaagtctctgc tatgtggcgc 2100

ggtattatcc cgtattgacg ccgggcaaga gcaactcggg cgccgcatc actattctca 2160

gaatgacttg gttgagtact caccagtcac agaaaagcat cttacggatg gcatgacagt 2220

aagagaatta tgcagtgtcg ccataacatc gattgataac actgcggcca acttacttct 2280

gacaacgacg ggaggaccga aggagctaac cgcttttttg cacaacatgg gggatcatgt 2340

aactcgcctt gatcgttggg aaccggagct gaatgaagcc ataccaaacg acgagcgtga 2400

caccacgatg cctgtagcaa tggcaacaac gttgcgcaaa ctattaactg gcgaactact 2460

tactctagct tcccggaac aattaataga ctggatggag gcggataaag ttgcaggacc 2520

acttctgcgc tcggcccttc cggctggctg gtttattgct gataaatctg gagccggtga 2580

gcgtgggtct cgcggtatca ttgcagcact ggggccagat ggtaagccct cccgtatcgt 2640

agttatctac acgacgggga gtcaggcaac tatggatgaa cgaaatagac agatcgtga 2700

gataggtgcc tcactgatta agcattggta actgtcagac caagtttact catatatact 2760

ttagattgat ttaaaacttc atttttaatt taaaaggatc taggtgaaga tcctttttga 2820

taatctcatg accaaaaacc cttaacgtga gttttcgttc cactgagcgt cagaccccggt 2880

agaaaagatc aaaggatctt cttgagatcc tttttttctg cgcgtaatct gctgcttgca 2940

aacaaaaaaa ccaccgctac cagcgggtggg ttgtttgccg gatcaagagc taccaactct 3000

ttttccgaag gtaactggct tcagcagagc gcagatacca aatactgttc ttctagtgtg	3060
gccgtagtta ggccaccact tcaagaactc ttagcaccg cctacatacc tcgtctgtct	3120
aatcctgtta ccagtggctg ctgccagtgg cgataagtcg tgtcttaccg ggttggactc	3180
aagacgatag ttaccggata aggcgcagcg gtcgggctga acggggggtt cgtgcacaca	3240
gcccagcttg gagcgaacga cctacaccga actgagatac ctacagcgtg agctatgaga	3300
aagcgccacg cttcccgaag ggagaaaggc ggacaggtat ccggtaaagc gcagggtcgg	3360
aacaggagag cgcacgaggg agcttccagg gggaacgcc tggtatcttt atagtcctgt	3420
cggttttcgc cacctctgac ttgagcgtcg atttttgtga tgctcgtcag gggggcggag	3480
cctatggaaa aacgccagca acgcggcctt ttacggttc ctggcctttt gctggccttt	3540
tgctcacatg ttctttcctg cgttatcccc tgattctgtg gataaccgta ttaccgcctt	3600
tgagtgagct gataccgctc gccgcagccg aacgaccgag cgcagcaggt cagtgagcga	3660
ggaagcggaa gagcgcccaa tacgcaaacc gcctctcccc gcgcgttggc cgattcatta	3720
atgcagctgg cagcacaggt ttcccgactg gaaagcgggc agtgagcga acgcaattaa	3780
tgtgagttag ctactcatt aggcacccca ggctttacac tttatgcttc cggctcgtat	3840
gttgtgtgga attgtgagcg gataacaatt tcacacagga aacagctatg accatgatta	3900
cgccaagctt tttctttcca atttttttt ttctgtcatt ataaaaatca ttacgaccga	3960
gattccccgg taataactga tataattaaa ttgaagctct aatttgtgag tttagtatac	4020
atgcatttac ttataatata gtttttttagt ttgtctggcc gcatcttctc aaatatgctt	4080
cccagcctgc ttttctgtaa cgttcacctt ctaccttagc atcccttccc ttgcaaata	4140
gtcctcttcc aacaataata atgtcagatc ctgtagagac cacatcatcc acggttctat	4200
actgttgacc caatgcgtct ccttgtcat ctaaaccac accgggtgtc ataataacc	4260
aatcgtaac ttcatctctt ccacccatgt ctctttgagc aataaagccg ataacaaaat	4320
ctttgtcgt cticgcaatg tcaacagtac ccttagtata ttctccagta gataggagc	4380
ccttgcatga caattctgt aacatcaaaa ggctcttagg ttctttgtt acttctctg	4440
ccgctgctt caaacgcta acaatacctg ggcccaccac accgtgtgca ttcgtaatgt	4500
ctgccattc tgctattctg tataacccg cagagtactg caatttgact gtattacaa	4560
tgtcagcaaa ttttctgtct tcgaagagta aaaaattgta cttggcggat aatgccttta	4620
gcggttaac tgtgccctcc atggaaaaat cagtcaagat atccacatgt gtttttagta	4680
aacaaatttt gggaacctaat gcttcaacta actccagtaa ttccttggtg gtacgaacat	4740
ccaatgaagc acacaagttt gtttgctttt cgtgcatgat attaaatagc ttggcagcaa	4800
caggactagg atgagtagca gcacgttctt tatatgtagc tttcgacatg atttatctt	4860

gtttcctgca ggtttttgtt ctgtgcagtt gggttaagaa tactgggcaa tticatgttt 4920
cttcaacact acatatgcgt atataacca atctaagtct gtgctccttc cttcgttctt 4980
ccttctgttc ggagattacc gaatcaaaaa aatttcaagg aaaccgaaat caaaaaaaag 5040
aataaaaaaa aaatgatgaa ttgaaaagct tgcatgcctg caggtcgact ctagtatact 5100

ccgtctactg tacgatacac ttccgctcag gtccttgtcc tttaacgagg ctttaccact 5160
cttttgttac tctattgac cagctcagca aaggcagtgt gatctaagat tctatcttcg 5220
cgatgtagta aaactagcta gaccgagaaa gagactagaa atgcaaaagg cacttctaca 5280
atggctgcca tcattattat ccgatgtgac gctgcatttt tttttttttt tttttttttt 5340
tttttttttt tttttttttt ttttttttgt acaaatatca taaaaaaaga gaatcttttt 5400
aagcaaggat tttcttaact tcttcggcga cagcatcacc gacttcggtg gtactgttgg 5460
aaccacctaa atcaccagtt ctgatacctg catccaaaac ctttttaact gcattctcaa 5520

tggctttacc ttcttcaggc aagtccaatg acaatttcaa catcattgca gcagacaaga 5580
tagtggcgat agggttgacc ttattctttg gcaaatctgg agcggaaacca tggcatggtt 5640
cgtacaaacc aaatgcggtg ttcttgtctg gcaaagaggc caaggacgca gatggcaaca 5700
aacccaagga gcctgggata acggaggctt catcggagat gatacacca aacatgttgc 5760
tggtgattat aataccattt aggtgggttg ggttcttaac taggatcatg gcggcagaat 5820
caatcaattg atgttgaact ttcaatgtag ggaattcggt cttgatgggt tcctccacag 5880
ttttctcca taatcttgaa gaggccaaaa cattagcttt atccaaggac caaataggca 5940

atggctggctc atgtttagg gccatgaaag cgccattct tgtgattctt tgcattctg 6000
gaacggtgta ttgttacta tccaagcga caccatcacc atcgtcttcc tttctcttac 6060
caaagtaaat acctccact aattctctaa caacaacgaa gtcagtacct ttagcaaatt 6120
gtggcttgat tggagataag tctaaaagag agtcggatgc aaagttacat ggtcttaagt 6180
tggcgtacaa ttgaagtctt ttacggattt ttagtaaacc ttgttcaggt ctaacactac 6240
cggtaaccca tttaggacca ccacagcac ctaacaaaac ggcatcagcc ttcttgagg 6300
cttcacgcgc ctcatctgga agtggaaacac ctgtagcatc gatagcagca ccaccaatta 6360

aatgattttc gaaatcgaac ttgacattgg aacgaacatc agaaatagct ttaagaacct 6420
taatggcttc ggctgtgatt tcttgaccaa cgtggtcacc tggcaaaacg acgatcttct 6480
taggggcaga cattacaatg gtatatcctt gaaatatata taaaaaaaaa aaaaaaaaaa 6540
aaaaaaaaaa atgcagcttc tcaatgatat tcgaatacgc tttgaggaga tacagcctaa 6600
tatccgacaa actgttttac agatttacga tcgtacttgt tacccatcat tgaattttga 6660
acatccgaac ctgggagttt tccctgaaac agatagtata tttgaacctg tataataata 6720

tatagtctag cgctttacgg aagacaatgt atgtatttcg gttcctggag aaactattgc	6780
atctattgca taggtaatct tgcacgtcgc atccccgggt cattttctgc gtttccatct	6840
tgcacttcaa tagcatactt ttgttaacga agcatctgtg cttcattttg tagaacaaaa	6900
atgcaacgcg agagcgctaa tttttcaaac aaagaatctg agctgcattt ttacagaaca	6960
gaaatgcaac gcgaaagcgc tattttacca acgaagaatc tgtgcttcat ttttgtaaaa	7020
caaaaatgca acgcgagagc gctaattttt caaacaaaga atctgagctg catttttaca	7080
gaacagaaat gcaacgcgag agcgctattt taccaacaaa gaatctatac ttcttttttg	7140
ttctacaaaa atgcatcccg agagcgctat ttttctaaca aagcatctta gattactttt	7200
tttctccttt gtgcgcctta taatgcagtc tcttgataac ttttgcact gtaggtccgt	7260
taagggttaga agaaggctac ttgggtgtct attttctctt ccataaaaaa agcctgactc	7320
cacttccgcg gtttactgat tactagcgaa gctgcgggtg cattttttca agataaaggc	7380
atccccgatt atattctata ccgatgtgga ttgcgcatac tttgtgaaca gaaagtata	7440
gcgttgatga ttcttcattg gtcagaaaat tatgaacggt ttcttctatt ttgtctctat	7500
atactacgta taggaaatgt ttacattttc gtattgtttt cgattcactc tatgaatagt	7560
tcttactaca atttttttgt ctaaagagta atactagaga taaacataaa aaatgtagag	7620
gtcgagtta gatgcaagtt caaggagcga aaggtggatg ggtaggttat ataggatat	7680
agcacagaga tatatagcaa agagatactt ttgagcaatg tttgtggaag cggatttcgc	7740
aatattttag tagctcgta cagtcgggtg cgttttttgt ttttgaaag tgcgtcttca	7800
gagcgctttt ggttttcaaa agcgctctga agttcctata ctttctagag aataggaact	7860
tcggaatagg aacttcaaag cgtttccgaa aacgagcgct tccgaaaatg caacgcgagc	7920
tgcgcacata cagctcactg ttcacgtcgc acctatatct gcgtgttgcc tgtatatata	7980
tatacatgag aagaacggca tagtgcgtgt ttatgcttaa atgcgtactt atatgcgtct	8040
atttatgtag gatgaaaggt agtctagtag ctctgtgat attatcccat tccatgcggg	8100
gtatcgtatg cttccttcag cactaccctt tagctgttct atatgctgcc actcctcaat	8160
tggattagtc tcatccttca atgctatcat ttctttgat attggatcat atgcatagta	8220
ccgagaaact agaggatctc ccattaccga catttgggcg ctatacgtgc atatgttcat	8280
gtatgtatct gtatttaaaa cacttttgta ttatttttcc tcatatatgt gtataggttt	8340
atacggatga tttaattatt acttcaccac cctttatttc aggtgatat cttagccttg	8400
ttactagtca ccggtggcgg ccgcacctgg taaaacctct agtggagtag tagatgtaat	8460

caatgaagcg gaagccaaaa gaccagagta gaggcctata gaagaaactg cgataccttt 8520
 tgtgatggct aaacaaacag acatcttttt atatgttttt acttctgtat atcgtgaagt 8580
 agtaagtgat aagcgaatit ggctaagaac gttgtaagtg aacaaggac ctcttttgcc 8640
 tttaaaaaa ggattaaatg gagttaatca ttgagattta gttttcgta gattctgtat 8700
 ccctaataa ctcccttacc cgacgggaag gcacaaaaga cttgaataat agcaaacggc 8760
 cagtagccaa gaccaaataa tactagagt aactgatggt cttaaacagg cattacgtgg 8820
 tgaactcaa gaccaatata caaaatatcg ataagttatt cttgcccacc aatttaagga 8880

gcctacatca ggacagtagt accattcttc agagaagagg tatacataac aagaaaatcg 8940
 cgtgaacacc ttatataact tagcccgta ttgagctaaa aaaccttgca aaatttccta 9000
 tgaataagaa tacttcagac gtgataaaaa ttactttct aactcttctc acgtgcccc 9060
 tatctgttct tccgtctac cgtgagaaat aaagcatcga gtacggcagt tcgtgtcac 9120
 tgaactaaaa caataaggct agttcgaatg atgaacttgc ttgctgtcaa acttctgagt 9180
 tgccgtgat gtgacactgt gacaataaat tcaaaccggt tatagcggtc tcctccgta 9240
 ccggttctgc cacctccaat agagctcagt aggagtcaga acctctgcgg tggctgtcag 9300

tgactcatcc gcgtttcgta agttgtgcgc gtgcacattt cgcccggttc cgctcatctt 9360
 gcagcaggcg gaaattttca tcacgtgta ggacgcaaaa aaaaaataat taatcgtaca 9420
 agaactcttg aaaaaaattt gaaaaatttt gtataaaagg gatgacctaa ctgactcaa 9480
 tggcttttac acccagtatt ttccctttcc ttgtttgtta caattataga agcaagacaa 9540
 aaacatatag acaacctatt cctaggagt atattttttt accctaccag caatataagt 9600
 aaaaaactgt tt 9612

<210> 933

<211> 343

<212> PRT

<213> Artificial Sequence

<220><223> 882 C7

<400> 933

Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn

1 5 10 15

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

20 25 30

Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val

35 40 45

Ile Ile Gly Leu Tyr Gly Gly Ala Lys Glu Trp Lys Arg Ala Glu Glu
50 55 60

Gln Gly Phe Glu Val Tyr Thr Ala Ala Glu Ala Ala Lys Lys Ala Asp
65 70 75 80

Ile Ile Met Ile Leu Ile Asn Asp Glu Lys Gln Ala Thr Met Tyr Lys
85 90 95

Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala
100 105 110

His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val
115 120 125

Asp Val Thr Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser
130 135 140

Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln
145 150 155 160

Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala
165 170 175

Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu
180 185 190

Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val
195 200 205

Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr
210 215 220

Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile
225 230 235 240

Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile
245 250 255

Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile
260 265 270

Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln
275 280 285

Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly

290 295 300
 Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro
 305 310 315 320

Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
 325 330 335
 Glu Asp Lys Leu Ile Asn Asn
 340

<210> 934

<211> 343

<212> PRT

<213> Artificial Sequence

<220><223> 882 F9

<400> 934

Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn
 1 5 10 15
 Leu Ser Leu Leu Asp Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser
 20 25 30

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

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

145 150 155 160

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

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

 Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly
 290 295 300
 Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro
 305 310 315 320
 Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
 325 330 335
 Glu Asp Lys Leu Ile Asn Asn
 340
 <210> 935
 <211> 343
 <212> PRT
 <213> Artificial Sequence

 <220><223> 882 G6
 <400> 935

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

Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile
 Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln
 Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly
 Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro

Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
325 330 335
Glu Asp Lys Leu Ile Asn Asn
340

<210> 936

<211> 343

<212> PRT

<213> Artificial Sequence

<220><223> 882 G12

<400> 936

Met	Glu	Glu	Cys	Lys	Met	Ala	Lys	Ile	Tyr	Tyr	Gln	Glu	Asp	Cys	Asn
1				5				10					15		
Leu	Ser	Leu	Leu	Asp	Gly	Lys	Thr	Ile	Ala	Val	Ile	Gly	Tyr	Gly	Ser
			20					25					30		

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

Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala
 100 105 110
 His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val
 115 120 125
 Asp Val Thr Met Ile Ala Pro Lys Gly Pro Gly His Thr Leu Arg Ser
 130 135 140
 Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln
 145 150 155 160

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

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

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

<210> 937

<211> 343

<212> PRT

<213> Artificial Sequence

<220><223> 883 C4

<400> 937

Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn

1 5 10 15

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

20 25 30

Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val

35 40 45

Ile Ile Gly Leu Tyr Glu Gly Ala Lys Glu Trp Lys Arg Ala Glu Glu

50 55 60

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

65 70 75 80

Ile Ile Met Ile Leu Ile Asn Asp Glu Lys Gln Ala Thr Met Tyr Lys

85 90 95

Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala

100 105 110

His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val

115 120 125

Asp Val Thr Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser

130 135 140

Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln

145 150 155 160

Asp Ala Thr Gly Lys Ala Leu Asp Met Val Leu Ala Tyr Ala Leu Ala

165 170 175

Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu

180 185 190

Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val

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

 Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile
 260 265 270
 Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln
 275 280 285
 Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly
 290 295 300
 Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro
 305 310 315 320

 Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
 325 330 335
 Glu Asp Lys Leu Ile Asn Asn
 340
 <210> 938
 <211> 343
 <212> PRT
 <213> Artificial Sequence
 <220><223> 883 G9
 <400> 938
 Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn
 1 5 10 15
 Leu Ser Leu Leu Asp Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser
 20 25 30

 Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val
 35 40 45
 Ile Ile Gly Leu Tyr Glu Gly Ala Lys Glu Trp Lys Arg Ala Glu Glu

50	55	60	
Gln Gly Phe Glu Val Tyr Thr Ala Ala Glu Ala Ala Lys Lys Ala Asp			
65	70	75	80
Ile Ile Met Ile Leu Ile Asn Asp Glu Lys Gln Ala Thr Met Tyr Lys			
	85	90	95
Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala			
	100	105	110
His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val			
	115	120	125
Asp Val Thr Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser			
	130	135	140
Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln			
145	150	155	160
Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala			
	165	170	175
Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu			
	180	185	190
Thr Glu Thr Asp Met Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val			
	195	200	205
Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr			
	210	215	220
Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile			
225	230	235	240
Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile			
	245	250	255
Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile			
	260	265	270
Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln			
	275	280	285
Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly			
	290	295	300

Ser Gln Val His Phe Met Ala Met Arg Lys Leu Ala Ser Glu His Pro
 305 310 315 320
 Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
 325 330 335
 Glu Asp Lys Leu Ile Asn Asn
 340
 <210> 939
 <211> 571
 <212> PRT
 <213> Artificial Sequence

 <220><223> Streptococcus mutans DHAD
 <400> 939
 Met Thr Asp Lys Lys Thr Leu Lys Asp Leu Arg Asn Arg Ser Ser Val
 1 5 10 15
 Tyr Asp Ser Met Val Lys Ser Pro Asn Arg Ala Met Leu Arg Ala Thr
 20 25 30
 Gly Met Gln Asp Glu Asp Phe Glu Lys Pro Ile Val Gly Val Ile Ser
 35 40 45
 Thr Trp Ala Glu Asn Thr Pro Cys Asn Ile His Leu His Asp Phe Gly

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

 115 120 125
 Cys Asp Lys Asn Met Pro Gly Ser Val Ile Ala Met Ala Asn Met Asp
 130 135 140
 Ile Pro Ala Ile Phe Ala Tyr Gly Gly Thr Ile Ala Pro Gly Asn Leu
 145 150 155 160

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

405 410 415
 Asn Ser Glu Glu Glu Ala Ile Glu Ala Val Leu Asn Asp Asp Ile Val
 420 425 430
 Asp Gly Asp Val Val Val Val Arg Phe Val Gly Pro Lys Gly Gly Pro

 435 440 445
 Gly Met Pro Glu Met Leu Ser Leu Ser Ser Met Ile Val Gly Lys Gly
 450 455 460
 Gln Gly Glu Lys Val Ala Leu Leu Thr Asp Gly Arg Phe Ser Gly Gly
 465 470 475 480
 Thr Tyr Gly Leu Val Val Gly His Ile Ala Pro Glu Ala Gln Asp Gly
 485 490 495
 Gly Pro Ile Ala Tyr Leu Gln Thr Gly Asp Ile Val Thr Ile Asp Gln

 500 505 510
 Asp Thr Lys Glu Leu His Phe Asp Ile Ser Asp Glu Glu Leu Lys His
 515 520 525
 Arg Gln Glu Thr Ile Glu Leu Pro Pro Leu Tyr Ser Arg Gly Ile Leu
 530 535 540
 Gly Lys Tyr Ala His Ile Val Ser Ser Ala Ser Arg Gly Ala Val Thr
 545 550 555 560
 Asp Phe Trp Lys Pro Glu Glu Thr Gly Lys Lys
 565 570

<210> 940

<211> 1641

<212> DNA

<213> Artificial Sequence

<220><223> M. caseolyticus KAD

<400> 940

atgaaacaac gtatcgggca atacttgatc gatgccctac acgttaatgg tgtcgataag 60
 atctttggag tcccagggtga ttctacttta gccttttttg acgatatcat aagacatgac 120
 aacgtggaat ggggtgggaaa tactaatgag ttgaacgccg cttacgccgc tgatggttac 180
 gctagagtta atggattagc cgctgtatct accacttttg gggttggcga gttatctgct 240
 gtgaatggta ttgctggaag ttacgcagag cgtgttcctg taatcaaaat ctcaggcggc 300

ccttcacag ttgctcaaca agaggtaga tatgtccacc attcattggg tgaaggaatc 360

 ttgtattcat attcaaagat gtacgtcac ataaccgcaa caactacaat cttatccgtt 420
 gacaacgcag tcgacgaaat tgatagagt attcattgtg ctttgaagga aaagaggcca 480
 gtgcatattc atttgcctat tgacgtagcc ttaactgaga ttgaaatccc tcatgcacca 540
 aaagtttaca cacacgaatc ccagaacgtc gatgcttaca ttcaagctgt tgagaaaaag 600
 ttaatgtctg caaaacaacc agtaatcata gcaggatcatg aaatcaattc attcaagttg 660
 cacgaacaac tggaacagtt tgtcaatcag acaaacatcc ctgttgaca actttccttg 720
 ggtaagtctg ctttcaatga agagaatgaa cattaccttg gtatctacga tggcaaaatc 780

 gcaaaggaaa atgtgagaga gtacgtcac aatgtgatg tcatattgaa cataggtgcc 840
 aaactgactg attctgttac agctggattt tctacaagt tcgatacaaa caacataatc 900
 tacattaacc ataatgactt caaagctgaa gatgtgatgt ctgataatgt ttcactgatt 960
 gatcttgtga atggcctgaa ttctattgac tatagaaatg aaacacacta cccatcttat 1020
 caaagatctg atatgaaata cgaattgaat gacgcaccac ttacacaatc taactatttc 1080
 aaaatgatga acgtttttct agaaaaagat gacatcctac tagctgaaca aggtacatcc 1140
 tttttcggcg catatgactt atccctatac aagggaatc agtttatcgg tcagccttta 1200

 tgggggtcaa tagggtatc ttttccatct ttactaggaa gtcaactagc agacatgcat 1260
 aggagaaaca ttttgcctat aggcgatggg agtttacaac ttactgttca agccctaagt 1320
 acaatgatta gaaaggatat caaaccaatc attttcgtta tcaataacga cggttacacc 1380
 gtgaaagac ttatccacgg catggaagag ccatacaatg atatccaaat gtggaactac 1440
 aagcaattgc cagaagtatt tgggtgaaaa gatactgtaa aagttcatga tgctaaaacc 1500
 tccaacgaac tgaaaactgt aatggattct gttaaagcag acaaatgata catgcatttc 1560
 attgaagtgc atatggcagt agaggacgcc ccaagaagt tgattgatat agctaaagcc 1620

 tttagtgatg ctaacaagta a 1641

 <210> 941
 <211> 1647
 <212> DNA
 <213> Artificial Sequence
 <220><223> L. grayi KAD
 <400> 941
 atgtacaccg tcggccaata cttagtagac cgttagaag agatcgcat cgataaggtt 60
 tttgtgtcc cgggtgacta caacctgacc tttttggact acatccagaa ccacgaaggt 120

ctgagctggc aaggtaatac gaatgaactg aatgccgcgt acgcagctga tggctatgct 180
 cgtgaacgcg gtgttagcgc tttggtcacg accttcggcg ttggtgagct gtccgcaatc 240
 aatggcaccg caggtagctt cgcgagacaa gttccggtga ttcataatcgt gggcagcccc 300

accatgaatg ttcagagcaa caagaaactg gtcatcaca gccctgggtat gggcaacttt 360
 cacaacttca gcgagatggc gaaagaagtc accgccgcaa ccacgatgct gacggaagag 420
 aatgcggcgt cggagattga tcgtgttctg gaaaccgccc tgcctggagaa acgcccagtg 480
 tacatcaatc tgccgatcga cattgtctac aaggcgatcg tcaagccggc gaaagccctg 540
 caaaccgaga agagctctgg cgagcgtgag gcacaactgg cggagatcat tctgagccat 600
 ctggagaagg ctgcacagcc gatttgtatt gcgggtcacg agatcgcgcg cttccagatc 660
 cgtgagcgtt tcgagaattg gattaatcaa acgaaactgc cggtgaccaa tctggcctac 720

ggcaagggta gcttcaacga agaaaacgag catttcattg gtacctatta tctgcathtt 780
 agcgataaga acgtgctgga ctacgtggat aactccgact ttgtcctgca ctttgggtgtt 840
 aaaatcattg ataacagcac ctccagcttc tcccaaggct tcaaaaccga gaacaccctg 900
 actgcggcga acgatatcat tatgtgccg gacggtagca cgtattctgg tattagcctg 960
 aatggcctgc tggccgagct ggaaaaactg aatttcacgt ttgccgacac cgcagcaaag 1020
 caggcggagt tggcgggttt tgagccgcag gctgaaaccc cgttgaaaca ggaccgtttt 1080
 caccaggcgg tgatgaattt tctgcaagct gacgatgtcc tggttacgga acagggcacc 1140

tcttcttttg gcttgatgct ggcgcctctg aaaaagggta tgaacttgat ctgcgaaacg 1200
 ctgtggggta gcattggta cacgttgccg gcgatgattg gtagccaaat tgcggcacccg 1260
 gagcgtcgtc atacctgag cattggtgat ggtagctttc agctgactgc gcaggaaatg 1320
 agcaccattt tccgtgagaa actgaccca gtcattctca tcattaacaa tgatggctat 1380
 accgttgagc gtgcgatcca tggcgaagat gaaagctata acgacattcc gacgtggaac 1440
 ttgcaactgg tggcggaaac cttcgggtgtt gacgccgaaa ccgtcgacac tcacaatgtg 1500
 ttcacggaga ctgatttcgc caacaccctg gcggcaattg acgcgacgcc gcagaaagca 1560

cacgttgtgg aagttcacat ggaacaaatg gatatgccgg agagcctgcg ccagatcggt 1620
 ctggcactgt ccaagcagaa tagctaa 1647

<210> 942
 <211> 491
 <212> PRT
 <213> Artificial Sequence
 <220><223> E. coli KARI

<400> 942

Met Ala Asn Tyr Phe Asn Thr Leu Asn Leu Arg Gln Gln Leu Ala Gln
1 5 10 15

Leu Gly Lys Cys Arg Phe Met Gly Arg Asp Glu Phe Ala Asp Gly Ala
20 25 30

Ser Tyr Leu Gln Gly Lys Lys Val Val Ile Val Gly Cys Gly Ala Gln
35 40 45

Gly Leu Asn Gln Gly Leu Asn Met Arg Asp Ser Gly Leu Asp Ile Ser
50 55 60

Tyr Ala Leu Arg Lys Glu Ala Ile Ala Glu Lys Arg Ala Ser Trp Arg
65 70 75 80

Lys Ala Thr Glu Asn Gly Phe Lys Val Gly Thr Tyr Glu Glu Leu Ile
85 90 95

Pro Gln Ala Asp Leu Val Ile Asn Leu Thr Pro Asp Lys Gln His Ser
100 105 110

Asp Val Val Arg Thr Val Gln Pro Leu Met Lys Asp Gly Ala Ala Leu
115 120 125

Gly Tyr Ser His Gly Phe Asn Ile Val Glu Val Gly Glu Gln Ile Arg
130 135 140

Lys Asp Ile Thr Val Val Met Val Ala Pro Lys Cys Pro Gly Thr Glu
145 150 155 160

Val Arg Glu Glu Tyr Lys Arg Gly Phe Gly Val Pro Thr Leu Ile Ala
165 170 175

Val His Pro Glu Asn Asp Pro Lys Gly Glu Gly Met Ala Ile Ala Lys
180 185 190

Ala Trp Ala Ala Ala Thr Gly Gly His Arg Ala Gly Val Leu Glu Ser
195 200 205

Ser Phe Val Ala Glu Val Lys Ser Asp Leu Met Gly Glu Gln Thr Ile
210 215 220

Leu Cys Gly Met Leu Gln Ala Gly Ser Leu Leu Cys Phe Asp Lys Leu
225 230 235 240

Val Glu Glu Gly Thr Asp Pro Ala Tyr Ala Glu Lys Leu Ile Gln Phe
245 250 255

Gly Trp Glu Thr Ile Thr Glu Ala Leu Lys Gln Gly Gly Ile Thr Leu
260 265 270

Met Met Asp Arg Leu Ser Asn Pro Ala Lys Leu Arg Ala Tyr Ala Leu
275 280 285

Ser Glu Gln Leu Lys Glu Ile Met Ala Pro Leu Phe Gln Lys His Met
290 295 300

Asp Asp Ile Ile Ser Gly Glu Phe Ser Ser Gly Met Met Ala Asp Trp
305 310 315 320

Ala Asn Asp Asp Lys Lys Leu Leu Thr Trp Arg Glu Glu Thr Gly Lys
325 330 335

Thr Ala Phe Glu Thr Ala Pro Gln Tyr Glu Gly Lys Ile Gly Glu Gln
340 345 350

Glu Tyr Phe Asp Lys Gly Val Leu Met Ile Ala Met Val Lys Ala Gly
355 360 365

Val Glu Leu Ala Phe Glu Thr Met Val Asp Ser Gly Ile Ile Glu Glu
370 375 380

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

Ile Ala Arg Lys Arg Leu Tyr Glu Met Asn Val Val Ile Ser Asp Thr
405 410 415

Ala Glu Tyr Gly Asn Tyr Leu Phe Ser Tyr Ala Cys Val Pro Leu Leu
420 425 430

Lys Pro Phe Met Ala Glu Leu Gln Pro Gly Asp Leu Gly Lys Ala Ile
435 440 445

Pro Glu Gly Ala Val Asp Asn Gly Gln Leu Arg Asp Val Asn Glu Ala
450 455 460

Ile Arg Ser His Ala Ile Glu Gln Val Gly Lys Lys Leu Arg Gly Tyr
465 470 475 480

Met Thr Asp Met Lys Arg Ile Ala Val Ala Gly

485 490

<210> 943

<211> 338

<212> PRT

<213> Artificial Sequence

<220><223> Pseudomonas fluorescens KARI

<400> 943

Met Lys Val Phe Tyr Asp Lys Asp Cys Asp Leu Ser Ile Ile Gln Gly

1 5 10 15

Lys Lys Val Ala Ile Ile Gly Tyr Gly Ser Gln Gly His Ala Gln Ala

20 25 30

Cys Asn Leu Lys Asp Ser Gly Val Asp Val Thr Val Gly Leu Arg Lys

35 40 45

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

50 55 60

Asp Val Ala Ala Ala Val Ala Gly Ala Asp Leu Val Met Ile Leu Thr

65 70 75 80

Pro Asp Glu Phe Gln Ser Gln Leu Tyr Lys Asn Glu Ile Glu Pro Asn

85 90 95

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

100 105 110

Tyr Asn Gln Val Val Pro Arg Ala Asp Leu Asp Val Ile Met Ile Ala

115 120 125

Pro Lys Ala Pro Gly His Thr Val Arg Ser Glu Phe Val Lys Gly Gly

130 135 140

Gly Ile Pro Asp Leu Ile Ala Ile Tyr Gln Asp Ala Ser Gly Asn Ala

145 150 155 160

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

165 170 175

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

180 185 190

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

195 200 205
 Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr Ala Pro Glu Met Ala Tyr
 210 215 220
 Phe Glu Cys Leu His Glu Leu Lys Leu Ile Val Asp Leu Met Tyr Glu

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

 290 295 300
 Arg Arg Asn Asn Ala Ala His Gly Ile Glu Ile Ile Gly Glu Gln Leu
 305 310 315 320
 Arg Ser Met Met Pro Trp Ile Gly Ala Asn Lys Ile Val Asp Lys Ala
 325 330 335
 Lys Asn

<210> 944

<211> 348

<212> PRT

<213> Artificial Sequence

<220><223> Achromobacter xylosoxidans ADH

<400> 944

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

1 5 10 15
 Lys Pro Lys Pro Thr Leu Gln Lys Pro Thr Asp Val Val Val Arg Val
 20 25 30
 Leu Lys Thr Thr Ile Cys Gly Thr Asp Leu Gly Ile Tyr Lys Gly Lys
 35 40 45
 Asn Pro Glu Val Ala Asp Gly Arg Ile Leu Gly His Glu Gly Val Gly

50 55 60
 Val Ile Glu Glu Val Gly Glu Ser Val Thr Gln Phe Lys Lys Gly Asp

 65 70 75 80
 Lys Val Leu Ile Ser Cys Val Thr Ser Cys Gly Ser Cys Asp Tyr Cys
 85 90 95
 Lys Lys Gln Leu Tyr Ser His Cys Arg Asp Gly Gly Trp Ile Leu Gly
 100 105 110
 Tyr Met Ile Asp Gly Val Gln Ala Glu Tyr Val Arg Ile Pro His Ala
 115 120 125
 Asp Asn Ser Leu Tyr Lys Ile Pro Gln Thr Ile Asp Asp Glu Ile Ala

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

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

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

Asp Lys Leu Pro Leu Lys Lys Met Ile Thr His Arg Phe Glu Leu Ala
 305 310 315 320
 Glu Ile Glu His Ala Tyr Gln Val Phe Leu Asn Gly Ala Lys Glu Lys

325 330 335
 Ala Met Lys Ile Ile Leu Ser Asn Ala Gly Ala Ala
 340 345

<210> 945

<211> 347

<212> PRT

<213> Artificial Sequence

<220><223> B. indica ADH

<400> 945

Met Lys Ala Leu Val Tyr Arg Gly Pro Gly Gln Lys Leu Val Glu Glu
 1 5 10 15
 Arg Gln Lys Pro Glu Leu Lys Glu Pro Gly Asp Ala Ile Val Lys Val
 20 25 30

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

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

145 150 155 160

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

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

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

 <210> 946
 <211> 1032

 <212> DNA
 <213> Artificial Sequence
 <220><223> K9_Eliza
 <400> 946

atggaagaat gtaagatggc taagatttac taccaagaag actgtaactt gtccttggtg 60
gatggaaga ctatcgccgt tatcggttac ggttctcaag gtcacgtca tgccctgaat 120
gctaaggaat ccggttgtaa cgttatcatt ggtttattcg aagatgcgga ggaatggaaa 180
agagctgaag aacaaggttt cgaagtctac accgctgctg aagctgctaa gaaggctgac 240
atcattatga tcttgatccc agatgaaaag caggctacca tgtacaaaaa cgacatcgaa 300
ccaaacttgg aagccggtaa catgttgatg ttcgctcacg gtttcaacat ccatttcggt 360

tgtattgttc caccaaagga cgttgatgtc actatgatcg ctccaaaggg tccaggtcac 420
accgttagat ccgaatacga agaaggtaaa ggtgtcccat gcttggttgc tgtcgaacaa 480
gacgctactg gcaaggcttt ggatatggct ttggcctacg ctttagccat cgggtgtgct 540
agagccggtg tcttgaaac taccttcaga accgaaactg aaaccgactt gttcggtgaa 600
caagctgttt tatgtggtgg tgtctgcgct ttgatgcagg ccggttttga aaccttggtt 660
gaagccggtt acgaccaag aaacgcttac ttcgaatgta tccacgaaat gaagttgatc 720
gttgacttga tctaccaatc tggtttctcc ggtatgcgtt actctatctc caacactgct 780

gaatacgggtg actacattac cggtccaaag atcattactg aagataccaa gaaggtatg 840
aagaagattt tgctgacat tcaagatggt acctttgcca aggacttctt ggttgacatg 900
tctgatgctg gttcccaggt ccacttcaag gctatgagaa agttggcctc cgaacaccca 960
gctgaagtgg tcggtgaaga aattagatcc ttgtactcct ggtccgacga agacaagtgg 1020
attaacaact ga 1032

<210> 947

<211> 1032

<212> DNA

<213> Artificial Sequence

<220><223> K9_Frank

<400> 947

atggaagaat gtaagatggc taagatttac taccaagaag actgtaactt gtccttggtg 60
gatggaaga ctatcgccgt tatcggttac ggttctcaag gtcacgtca tgccctgaat 120
gctaaggaat ccggttgtaa cgttatcatt ggtttattcg aagtgcgga ggagtggaaa 180
agagctgaag aacaaggttt cgaagtctac accgctgctg aagctgctaa gaaggctgac 240
atcattatga tcttgatccc agatgaaaag ctggctacca tgtacaaaaa cgacatcgaa 300
ccaaacttgg aagccggtaa catgttgatg ttcgctcacg gtttcaacat ccatttcggt 360
tgtattgttc caccaaagga cgttgatgtc actatgatcg ctccaaaggg tccaggtcac 420

accgttagat ccgaatacga agaaggtaaa ggtgtcccat gcttggttgc tgtcgaacaa 480

gacgctactg gcaaggcttt ggatatggct ttggcctacg ctttagccat cgggtgtgct 540

agagccgggtg tcttggaac taccttcaga accgaaactg aaaccgactt gtccggtgaa 600

caagctgttt tatgtgggtg tgtctgcgct ttgatgcagg ccggttttga aaccttggtt 660

gaagccgggtt acgaccaag aaacgcttac ttcgaatgta tccacgaaat gaagttgatc 720

gttgacttga tctaccaatc tggtttctcc ggtatgcgtt actctatctc caaactgct 780

gaatacgggtg actacattac cgggtccaaag atcattactg aagatacga gaaggctatg 840

aagaagattt tgtctgacat tcaagatggt acctttgcca aggacttctt ggttgacatg 900

tctgatgctg gtccccagggt ccacttcaag gctatgagaa agttggcctc cgaacaccca 960

gctgaagttg tcggtgaaga aattagatcc ttgtactctt ggtccgacga agacaagttg 1020

attaacaact ga 1032

<210> 948

<211> 912

<212> DNA

<213> Artificial Sequence

<220><223> K9_Grace

<400> 948

gctaaggaat ccggttgtaa cgttatcatt ggtttattcg aagggtcgga ggagtggaaa 60

agagctgaag aacaaggttt cgaagtctac accgctgctg aagctgctaa gaaggctgac 120

atcatcatga tcttgatccc agatgaaaag caggctacca tgtacaaaaa cgacatcgaa 180

ccaaacttgg aagccggtaa catgttgatg ttcgctcagc gtttcaacat ccatttcggt 240

tgtattgttc caccaaagga cgttgatgtc actatgatcg ctccaaaggg tccaggtcac 300

accgttagat ccgaatacga agaaggtaaa ggtgtcccat gcttggttgc tgtcgaacaa 360

gacgctactg gcaaggcttt ggatatggct ttggcctacg ctttagccat cgggtgtgct 420

agagccgggtg tcttggaac taccttcaga accgaaactg aaaccgactt gtccggtgaa 480

caagctgttt tatgtgggtg tgtctgcgct ttgatgcagg ccggttttga aaccttggtt 540

gaagccgggtt acgaccaag aaacgcttac ttcgaatgta tccacgaaat gaagttgatc 600

gttgacttga tctaccaatc tggtttctcc ggtatgcgtt actctatctc caaactgct 660

gaatacgggtg actacataac cgggtccaaag atcattactg aagatacga gaaggctatg 720

aagaagattt tgtctgacat tcaagatggt acctttgcca aggacttctt ggttgatatg 780

tctgatgatg gtccccagggt ccacttcaag gctatgagaa agttggcctc cgaacaccca 840

gctgaagttg tcggtgaaga aattagatcc ttgtactcct ggtccgacga agacaagttg 900
attaacaact ga 912

<210> 949
<211> 1032
<212> DNA
<213> Artificial Sequence

<220><223> K9_Ingrid
<400> 949

atggaagaat gtaagatggc taagatttac taccaagaag actgtaactt gtccttggtg 60
gatggtgaaga ctatcgccgt tatcggttac ggttctcaag gtcacgtca tgccttgaat 120
gctaaggaat ccggttgtaa cgttatcatt ggtttattcg aaggtgcgga ggagtggaaa 180
agagctgaag aacaaggttt cgaagtctac accgctgctg aagctgctaa gaaggctgac 240
atcattatga tcttgatccc agatgaaaag caggctacca tttaaaaaa cgacatcgaa 300
ccaaacttgg aagccggtaa catgttgatg ttcgctcacg gtttcaacat ccatttcggt 360
tgtattgttc caccaaagga cgttgatgtc actatgatcg ctccaaaggg tccaggtcac 420

accgttagat ccgaatacga agaaggtaaa ggtgtcccat gcttggttgc tgtcgaacaa 480
gacgctactg gcaaggcttt ggatatggct ttggcctacg ctttagccat cgggtgtgct 540
agagccggtg tcttggaaac taccttcaga accgaaactg aaaccgactt gttcggtgaa 600
caagctgttt tatgtggtgg tgtctgcgct ttgatgcagg ccggttttga aaccttggtt 660
gaagccggtt acgaccaag aaacgcttac ttcgaatgta tccacgaaat gaagttgatc 720
gttgacttga tctaccaatc tggtttctcc ggtatgcgtt actctatctc caacactgct 780
gaatacgggtg actacattac cggtcctaaag atcattactg aagatacca gaaggctatg 840

aagaagattt tgtctgacat tcaagatggt acctttgcca aggacttctt ggttgacatg 900
tctgatgctg gttcccagat ccacttcaag gctatgagaa agttggcctc cgaacacca 960
gctgaagttg tcggtgaaga aattagatcc ttgtactcct ggtccgacga agacaagttg 1020
attaacaact ga 1032

<210> 950
<211> 1032
<212> DNA
<213> Artificial Sequence

<220><223> K9_Jarvis
<400> 950

atggaagaat gtaagatggc taagatttac taccaagaag actgtaactt gtccttggtg 60
gatggtaaga ctatcgccgt tateggttac ggttctcaag gtcacgtca tgcctgaat 120

gctaaggaat ccggttgtaa cgttatcatt ggtttattcg aaggtgcgga ggagtggaaa 180
agagctgaag aacaaggtat cgaagtctac accgctgctg aagctgctaa gaaggtgac 240
atcattatga tcttgatccc agatgaaaag caggctacca tgtacaaaaa cgacatcgaa 300
ccaaacttgg aagccggtaa catgttgatg ttcgctcacg gtttcaacat ccatttcggt 360
tgtattgttc caccaaagga cgttgatgtc actatgatcg ctccaaaggg tccaggtcac 420
accgttagat ccgaatacga agaaggtaaa ggtgtcccat gcttggttgc tgcgaacaa 480
gacgctactg gcaaggcttt ggatatggct ttggcctacg ctttagccat cgggtgtgct 540

agagccggtg tcttggaac taccttcaga accgaaactg aaaccgactt gttcggtgaa 600
caagctgttt tatgtggtgg tgtctgcgtt ttgatgcagg ccggttttga aaccttggtt 660
gaagccggtt acgacccaag aaacgcttac ttcgaatgta tccacgaaat gaagttgatc 720
gttgacttga tctaccaatc tggtttctcc ggtatgcgtt actctatctc caacactgct 780
gaatacgggt actacattac cgggtccaaag atcattactg aagatacga gaaggtatg 840
aagaagattt tgctgacat tcaagatggc acctttgcca aggacttctt ggttgacatg 900
tctgatgctg gttcccaggc ccacttcaag gctatgagaa agttggcctc cgaacacca 960

gctgaagttg tcggtgaaga aattagatcc ttgtactcct ggtccgacga agacaagttg 1020
attaacaact ga 1032

<210> 951
<211> 1032
<212> DNA
<213> Artificial Sequence
<220><223> K9_Kelly
<400> 951

atggaagaat gtaagatggc taagatttac taccaagaag actgtaactt gtccttggtg 60
gatggtaaga ctatcgccgt tateggttac ggttctcaag gtcacgtca tgcctgaat 120
gctaaggaat ccggttgtaa cgttatcatt ggtttattcg aaggtgggga ggagtggaat 180
agagctgaag aacaaggttt cgaagtctac accgctgctg aagctgctaa aaaggtgac 240

atcattatga tcttgatccc agatgaaaag caggctacca tgtacaaaaa cgacatcgaa 300
ccaaacttgg aagccggtaa catgttgatg ttcgctcacg gtttcaacat ccatttcggt 360
tgtattgttc caccaaagga cgttgatgtc actatgatcg ctccaaaggg tccaggtcac 420

accgtagat ccgaatacga agaaggtaaa ggtgtcccat gcttggttgc tgcgaacaa 480
gacgctactg gcaaggcttt ggatatggct ttggcctacg ctttagccat cgggtgtgct 540
agagccggtg tcttggaac taccttcaga accgaaactg aaaccgactt gttcggtgaa 600
caagctgttt taigtgggtg tgtctgcgct ttgatgcagg ccggttttga aaccttggtt 660

gaagccggtt acgaccaag aaacgcttac ttcgaatgta tccacgaaat gaagttgatc 720
gttgacttga tctaccaate tggtttctcc ggtatgcgtt actctatctc caacactgct 780
gaatacgggtg actacattac cgggtccaaag atcattactg aagatacaca gaaggctatg 840
aagaagattt tgtctgacat tcaagatggt acctttgcca aggacttctt ggttgacatg 900
tctgatgctg gttcccagggt ccacttcaag gctatgagaa agttggcctc cgaacaccca 960
gctgaagttg tcggtgaaga aattagatcc ttgtactcct ggtccgacga agacaagttg 1020
attaacaact ga 1032

<210> 952

<211> 1032

<212> DNA

<213> Artificial Sequence

<220><223> K9_Norman

<400> 952

atggaagaat gtaagatggc taagatttac taccaagaag actgtaactt gtccttgttg 60
gatggtgaaga ctatgccgt tatcggttac ggttctcaag gtcacgtca tgcctgaat 120
gctaaggaat ccggttgtaa cgttatcatt ggtttattcg aatgtgcgga ggagtggaaa 180
agagctgaag aacaagggtt cgaagtctac accgctgctg aagctgctaa gaaggctgac 240
atcattatga tcttgatccc agatgaaaag caggctacca tgtacaaaaa cgacatcgaa 300
ccaaacttgg aagccggtaa catgttgatg ttcgctcacg gtttcaacat ccatttcggt 360

tgtattgttc caccaaagga cgttgatgtc actatgatcg ctcaaagggt tccaggctac 420
accgtagat ccgaatacga agaaggtaaa ggtgtcccat gcttggttgc tgcgaacaa 480
gacgctactg gcaaggcttt ggatatggct ttggcctacg ctttagccat cgggtgtgct 540
agagccggtg tcttggaac taccttcaga accgaaactg aaaccgactt gttcggtgaa 600
caagctgttt taigtgggtg tgtctgcgct ttgatgcagg ccggttttga aaccttggtt 660
gaagccggtt acgaccaag aaacgcttac ttcgaatgta tccacgaaat gaagttgatc 720
gttgacttga tctaccaate tggtttctcc ggtatgcgtt actctatctc caacactgct 780

gaatacgggtg actacattac cgggtccaaag atcattactg aagatacaca gaaggctatg 840

aagaagattt tgctgacat tcaagatggt acctttgcca aggacttctt ggttgacatg 900

tctgatgctg gtccccaggt ccacttcaag gctatgagaa agttggcctc cgaacaccca 960

gctgaagttg tcggtgaaga aattagatcc ttgtactcct ggtccgacga agacaagttg 1020

attaacaact ga 1032

<210> 953

<211> 1032

<212> DNA

<213> Artificial Sequence

<220><223> K9_Ophelia

<400> 953

atggaagaat gtaagatggc taagatttac taccaagaag actgtaactt gtccttggtg 60

gatggaaga ctatcgccgt tatcggttac ggttctcaag gtcacgtca tgccctgaat 120

gctaaggaat ccggttgtaa cgttatcatt ggtttattcg aaggtgcgga ggagtggaaa 180

agagctgaag aacaagttt cgaagtctac accgctgctg aagctgctaa gaaggctgac 240

atcattatga tcttgatccc agatgaaaag caggctacca tgtacaaaaa cgacatcgaa 300

ccaaacttgg aagccggtaa catgttgatg ttgctcacg gtttcaacat ccatttcggt 360

tgtattgttc caccaaagga cgttgatgtc actatgatcg cttcaaaggg tccaggtcac 420

accgttagat ccgaatacga agaaggtaaa ggtgtcccat gcttgggtgc tgcgaacaa 480

gacgctactg gcaaggcttt ggatatggct ttggcctacg ctttagccat cggtggtgct 540

agagccggtg tcttggaaac taccttcaga accgaaactg aaaccgactt gttcggtgaa 600

caagctgttt tatgtggtgg tgtctgcgct ttgatgcagg ccggttttga aaccttggtt 660

gaagccggtt acgaccaag aaacgcttac ttcgaatgta tccatgaaat gaagttgatc 720

gttgacttga tctaccaatc tggtttctcc ggtatgcgtt actctatctc caacactgct 780

gaatacgggt actacattac cgggtccaaag atcattactg aagatacaa gaaggctatg 840

aagaagattt tgctgacat tcaagatggt acctttgcca aggacttctt ggttgacatg 900

tctgatgctg gtccccaggt ccacttcaag gctatgagaa agttggcctc cgaacaccca 960

gctgaagttg tcggtgaaga aattagatcc ttgtactcct ggtccgacga agacaagttg 1020

attaacaact ga 1032

<210> 954

<211> 1032

<212> DNA

<213> Artificial Sequence

<220><223> K9_Pat

<400> 954

atggaagaat gtaagatggc taagatttac taccaagaag actgtaactt gtccttggtg 60
gatggaaga ctatcgccgt tatcggttac ggttctcaag gtcacgtca tgccctgaat 120
gctaaggaat ccggttgtaa cgttatcatt ggtttattag aaggatcgga ggagtggaaa 180

agagctgaag aacaagggtt cgaagtctac accgtgctg aagctgctaa gaaggtgac 240
atcattatga tcttgatccc agatgaaaag caggctacca tgtacaaaa cgacatcgaa 300
ccaaacttgg aagccggtaa catgttgatg ttgctcagc gtttcaacat ccatttcggt 360
tgtattgttc caccaaagga cggtgatgac actatgatcg ctccaaaggg tccaggtcac 420
accgttagat ccgaatacga agaaggtaaa ggtgtcccat gcttggttgc tgtcgaacaa 480
gacgctactg gcaaggcttt ggatatggct ttggcctacg ctttagccat cgggtggtgct 540
agagccggtg tcttggaac taccttcaga accgaaactg aaaccgactt gttcggtgaa 600

caagctgttt tatgtggtgg tgtctgcgct ttgatgcagg ccggttttga aaccttggtt 660
gaagccggtt acgaccaag aaacgcttac ttgcaatgta tccacgaaat gaagttgatc 720
gttgacttga tctaccaatc tggtttctcc ggtatgcgtt actctatctc caacactgct 780
gaatacgggtg actacattac cgggtccaaag atcattactg aagatacga gaaggtatg 840
aagaagattt tgctgacat tcaagatggt accttgcca aggacttctt ggttgacatg 900
tctgatgctg gttcccaggc ccacttcaag gctatgagaa agttggcctc cgaacaccca 960
gctgaagttg tcggtgaaga aattagatcc ttgtactcct ggtccgacga agacaagttg 1020

attaacaact ga 1032

<210> 955

<211> 1032

<212> DNA

<213> Artificial Sequence

<220><223> K9_Quentin

<400> 955

atggaagaat gtaagatggc taagatttac taccaagaag actgtaactt gtccttggtg 60
gatggaaga ctatcgccgt tatcggttac ggttctcaag gtcacgtca tgccctgaat 120
gctaaggaat ccggttgtaa cgttatcatt ggtttattcg aaggatcgga ggagtggaaa 180
agagctgaag aacaagggtt cgaagtctac accgtgctg aagctgctaa gaaggtgac 240
atcattatga tcttgatccc agatgaaaag caggctacca ttacaaaaa cgacatcgaa 300

ccaaacttgg aagccggtaa catgttgatg ttcgctcacg gtttcaacat ccatttcggt 360
 tgtattgttc caccaaagga cgttgatgic actatgatcg ctccaaaggg tccaggtcac 420
 accgttagat ccgaatacga agaaggtaaa ggtgtcccat gcttggttgc tgcgaacaa 480
 gacgctactg gcaaggcttt ggatatggct ttggcctacg ctttagccat cgggtgtgct 540
 agagccggtg tcttggaac taccttcaga accgaaactg aaaccgactt gttcggtgaa 600
 caagctgttt tatgtggtgg tgtctgcgct ttgatgcagg ccggttttga aaccttggtt 660
 gaagccggtt acgaccaag aaacgcttac ttcgaatgta tccacgaaat gaagttgatc 720

gttgacttga tctaccaatc tggtttctcc ggtatgcgtt actctatctc caacactgct 780
 gaatacgggtg actacattac cgggtccaaag atcattactg aagatacga gaaggctatg 840
 aagaagattt tgtctgacat tcaagatggt acctttgcca aggacttctt ggttgacatg 900
 tctgatgctg gttcccggtt ccacttcaag gctatgagaa agttggcctc cgaacaccca 960
 gctgaagtgg tcggtgaaga aattagatcc ttgtactcct ggtccgacga agacaagtgg 1020
 attaacaact ga 1032

<210> 956

<211> 1032

<212> DNA

<213> Artificial Sequence

<220><223> K9_Ralph

<400> 956

atggaagaat gtaagatggc taagatttac taccaagaag actgttaactt gtccttggtg 60
 gatggtgaaga ctatcgccgt tatcggttac ggttctcaag gtcacgtca tgcctgaat 120
 gctaaggaat ccggttgtaa cgttatcatt ggtttattcg aaggtgcgga ggagtggaaa 180
 agagctgaag aacaaggctt agaagtctac accgctgctg aagctgctaa gaaggctgac 240
 atcattatga tcttgatccc agatgaaaag caggctacca tgtacaaaaa cgacatcgaa 300
 ccaaacttgg aagccggtaa catgttgatg ttcgctcacg gtttcaacat ccatttcggt 360
 tgtattgttc caccaaagga cgttgatgic actatgatcg ctccaaaggg tccaggtcac 420

accgttagat ccgaatacga agaaggtaaa ggtgtcccat gcttggttgc tgcgaacaa 480
 gacgctactg gcaaggcttt ggatatggct ttggcctacg ctttagccat cgggtgtgct 540
 agagccggtg tcttggaac taccttcaga accgaaactg aaaccgactt gttcggtgaa 600
 caagctgttt tatgtggtgg tgtctgcgct ttgatgcagg ccggttttga aaccttggtt 660
 gaagccggtt acgaccaag aaacgcttac ttcgaatgta tccacgaaat gaagttgatc 720
 gttgacttga tctaccaatc tggtttctcc ggtatgcgtt actctatctc caacactgct 780

gaatacgggtg actacattac cggtcctaaag atcattactg aagataccaa gaaggctatg 840

aagaagattt tgtctgacat tcaagatggt acctttgccca aggacttctt ggttgacatg 900

tctgatgctg gttcccagggt ccacttcaag gctatgagaa agttggcctc cgaacaccca 960

gctgaagttg tcggtgaaga aattagatcc ttgtactcct ggtccgacga agacaagttg 1020

attaacaact ga 1032

<210> 957

<211> 1032

<212> DNA

<213> Artificial Sequence

<220><223> K9_Sophia

<400> 957

atggaagaat gtaagatggc taatatattac taccaagaag actgtaactt gtccttggtg 60

gatggaaga ctatcgccgt tatcggttac ggttctcaag gtcacgctca tgcctgaat 120

gctaaggaat ccggttgtaa cgttatcatt ggtttattcg aaggtgcgga ggagtggaaa 180

agagctgaag aacaagggtt cgaagtctac accgctgctg aagctgctaa gaaggctgac 240

atcattatga tcttgatccc agatgaaatg caggctacca tgtacaaaaa cgacatcgaa 300

ccaaacttgg aagccggtaa catgttgatg ttcgctcacg gtttcaacat ccatttcggt 360

tgtattgttc caccaaagga cgttgatgtc actatgatcg ctccaaaggg tccaggtcac 420

atcgtagat ccgaatacga agaaggtaaa ggtgtcccat gcttggttgc tgtcgaacaa 480

gacgctactg gcaaggcttt ggatatggct ttggcctacg ctttagccat cgggtggtgct 540

agagccggtg tcttggaaac taccttcaga accgaaactg aaaccgactt gttcggtgaa 600

caagctgttt tatgtggtgg tgtctgcgct ttgatgcagg ccggttttga aaccttggtt 660

gaagccggtt acgaccaag aaacgcttac ttgcaatgta tccacgaaat gaagttgatc 720

gttgacttga tctaccaatc tggtttctcc ggtatgcgtt actctatctc caacactgct 780

gaatacgggtg actacattac cggtcctaaag atcattactg aagataccaa gaaggctatg 840

aagaagattt tgtctgacat tcaagatggt acctttgccca aggacttctt ggttgacatg 900

tctgatgctg gttcccagggt ccacttcaag gctatgagaa agttggcctc cgaacaccca 960

gctgaagttg tcggtgaaga aattagatcc ttgtactcct ggtccgacga agacaagttg 1020

attaacaact ga 1032

<210> 958

<211> 1032

<212> DNA

<213> Artificial Sequence

<220><223> K9_Tiberius

<400> 958

atggaagaat gtaagatggc taagatttac taccaagtag actgtaactt gtccttggtg	60
gatggtaaga ctattgccgt tatcggttac ggtttctcaag gtcacgtca tgcctgaat	120
gctaaggaat ccggttgtaa cgttatcatt ggtttattcg aaggtgcgga ggagtggaaa	180
agagctgaag aacaaggttt cgaagtctac accgctgctg aagccgctaa gaaggtgac	240
atcattatga tcttgatccc agatgaaaag caggctacca ttacaaaaa cgacatcgaa	300
ccaaacttgg aagccggtaa catgttgatg ttgctcacg gtttcaacat ccatttcggt	360
tgtattgttc caccaaagga cgttgatgtc actatgatcg ctccaaaggg tccaggtcac	420
atcgttagat ccgaatacga agaaggtaaa ggtgtcccat gcttggttgc tgcgaacaa	480
gacgctactg gcaaggcttt ggatatggct ttggcctacg ctttagccat cgggtgtgct	540
agagccggtg tcttggaaac taccttcaga accgaaactg aaaccgactt gttcggtgaa	600
caagctgttt tatgtgggtg tgtctgcgct ttgatgcagg ccggttttga aaccttggtt	660
gaagccggtt acgaccaag aaacgcttac ttcgaatgta tccacgaaat gaagttgatc	720
gttgacttga tctaccaate tggtttctcc ggtatgcgtt actctatctc caacactgct	780
gaatacgggt actacattac cgggtccaaag atcattactg aagatacaca gaaggctatg	840
aagaagattt tgtctgacat tcaagatggt acctttgcca aggacttctt ggttgacatg	900
tctgatgctg gttcccaggc ccacttcaag gctatgagaa agttggcctc cgaacaccca	960
gctgaagttg tcggtgaaga aattagatcc ttgtactcct ggtccgacga agacaagttg	1020
attaacaact ga	1032

<210> 959

<211> 1032

<212> DNA

<213> Artificial Sequence

<220><223> K9_Ursala

<400> 959

atggaagaat gtaagatggc taagatttac taccaagaag actgtaactt gtccttggtg	60
gatggtaaga ctatgccgt tatcggttac ggtttctcaag gtcacgtca tgcctgaat	120
gctaaggaat ccggttgtaa cgttatcatt ggtttattcg aaggtgtgga ggagtggaaa	180
agagctgaag aacaaggttt cgaagtctac accgctgctg aagctgctaa gaaggtgac	240

atcattatga tcttgatccc agatgaaaag caggctacca tgtacaaaaa cgacatcgaa 300
ccaaacttgg aagccggtaa catgttgatg ttcgctcacg gtttcaacat ccatttcggt 360

tgtattgttc caccaaagga cgttgatgtc actatgatcg ctccaaaggg tccaggtcac 420
accgttagat ccgaatacga agaaggtaaa ggtgtcccat gcttgggttc tgcgaacaa 480
gacgctactg gcaaggcttt ggatatggct ttggcctacg ctttagccat cgggtgtgct 540
agagccggtg tcttggaac taccttcaga accgaaactg aaaccgactt gttcggtgaa 600
caagctgttt tatgtgggtg tgtttgcgct ttgatgcagg ccggttttga aaccttggtt 660
gaagccggtt acgaccaag aaacgcttac ttcgaatgta tccacgaaat gaagttgatc 720
gttgacttga tctaccaatc tggtttctcc ggtatgcgtt actctatctc caacactgct 780

gaatacgggtg actacattac cgggtccaaag atcattactg aagatacga gaaggctatg 840
aagaagattt tgtctgacat tcaagatggt acctttgcca aggacttctt ggttgacatg 900
tctgatgctg gtccccaggt ccacttcaag gctatgagaa agttggcctc cgaacaccca 960
gctgaagttg tcggtgaaga aattagatcc ttgtactcct ggtccgacga agacaagttg 1020
attaacaact ga 1032

<210> 960
<211> 1032
<212> DNA
<213> Artificial Sequence
<220><223> K9_Victor
<400> 960

atggaagaat gtaagatggc taagatttac taccaagaag actgtaactt gtccttgttg 60

gatggtgaaga ctatgccgt tatcggttac ggtttctcaag gtcacgtca tgccctgaat 120
gctaaggaat ccggttgtaa cgttatcatt ggtttattcg aaggtgcgga ggagtggaaa 180
agagctgaag aacaaggctt cgaagtctac accgctgctg aagctgctaa gaaggctgac 240
atcattatgc tcttgatccc agatgaaaag caggctacca tgtacaaaaa cgacatcgaa 300
ccaaacttgg aagccggtaa catgttgatg ttcgctcacg gtttcaacat ccatttcggt 360
tgtattgttc caccaaagga cgttgatgtc actatgatcg ctccaaaggg tccaggtcac 420
accgttagat ccgaatacga agaaggtaaa ggtgtcccat gcttgggttc tgcgaacaa 480

gacgctactg gcaaggcttt ggatatggct ttggcctacg ctttagccat cgggtgtgct 540
agagccggtg tcttggaac taccttcaga accgaaactg aaaccgactt gttcggtgaa 600
caagctgttt tatgtgggtg tgtctgcgct ttgatgcagg ccggttttga aaccttggtt 660

gaagccggtt acgaccaag aaacgcttac ttcgaatgta tccacgaaat gaagttgac	720
gttgacttga tctaccaatc tggtttctcc ggtatgcgtt actctatctc caacactgct	780
gaatacgggtg actacattac cgggtccaaag atcattactg aagataccaa gaaggctatg	840
aagaagattt tgictgacat tcaagatggt acctttgcca aggacttctt ggttgacatg	900
tctgatgctg gttcccaggt ccacttcaag gctatgagaa agttggcctc cgaacaccca	960
gctgaagttg tcggtgaaga aattagatcc ttgtactcct ggtccgacga agacaagttg	1020
attaacaact ga	1032
<210> 961	
<211> 1032	
<212> DNA	
<213> Artificial Sequence	
<220><223> K9_Watson	
<400> 961	
atggaagaat gtaagatggc taagatttac taccaagaag actgtaactt gtccttggtg	60
gatggtaaga ctatcgccgt tatcggttac ggttctcaag gtcacgtca tgccctgaat	120
gctaaggaat ccggttgtaa cgttatcatt ggtttattcg aagggtcgga ggagtgcaaa	180
agagctgaag aacaaggctt cgaagtctac accgtgctg aagctgctaa gaaggctgac	240
atcattatga tcttgatccc agatgaaaag caggctacca tgtacaaaaa cgacatcgaa	300
ccaaacttgg aagccggtaa catgttgatg ttcgtcacg gtttcaacat ccatttcggt	360
tgtattgttc caccaaagga cgttgatgtc actatgatcg ctccaaaggg tccaggtcac	420
accgttagat ccgaatacga agaaggtaaa ggtgtcccat gcttgggtgc tgcgaacaa	480
gacgtactg gcaaggcttt ggatatggct ttggcctacg ctttagccat cgggtggtgct	540
agagccggtg tcttggaac taccttcaga accgaaactg aaaccgactt gttcggtgaa	600
caagctgttt tatgtggtgg tgtctgcgtt ttgatgcagg ccggttttga aaccttggtt	660
gaagccggtt acgaccaag aaacgcttac ttcgaatgta tccacgaaat gaagttgac	720
gttgacttga tctaccaatc tggtttctcc ggtatgcgtt actctatctc caacactgct	780
gaatacgggtg actacattac cgggtccaaag atcattactg aagataccaa gaaggctatg	840
aagaagattt tgictgacat tcaagatggt acctttgcca aggacttctt ggttgacatg	900
tctgatgctg gttcccaggt ccacttcaag gctatgagaa agttggcctc cgaacaccca	960
gctgaagttg tcggtgaaga aattagatcc ttgtactcct ggtccgacga agacaagttg	1020
attaacaact ga	1032

<210> 962

<211> 978

<212> DNA

<213> *Lactobacillus plantarum*

<400> 962

atggatttat ttgagtcatt agcacaaaaa attactggta aagatcaaac aattgttttc	60
cctgaaggaa ctgaaccccg aattgtcggg gcggcagcgc gattagctgc agacggcttg	120
gttaagccga ttgttttagg tgcaacggac aaagttcagg ctgtggctaa cgatttgaat	180
gcggatttaa caggcggtca agtccttgat cctgcgacat acccggtga agataagcaa	240
gcaatgcttg atgccctcgt tgaacggcgg aaaggtaga atacgccaga acaagcggct	300
aaaatgctgg aagatgaaaa ctactttggc acgatgctcg tttatatggg caaagcggat	360
gggatgggtt cagggtgcaat ccatccaact ggtgatacgg tacggccagc gttacaaatt	420
attaagacca agcccgggtc acaccgaatc tcgggtgcat ttatcatgca aaagggtgag	480
gaacgctacg tctttgctga ctgtgccatc aatattgac ccatgccga tacgttagcg	540
gaaattgcca ctgagagtgc ggctactgct aagggtctcg atattgacc gaaagttgcg	600
atgctcagct tctcaactaa gggttcggct aagggtgaaa tggtcactaa agtgcaagaa	660
gcaacggcca aggcgcaagc tgctgaaccg gaattggcta tcgatggtga acttcaattt	720
gacgcggcct tcgttgaaaa agttggtttg caaaaggctc ctggttccaa agtagctggt	780
catgccaatg tctttgtatt tccagagctt cagtctggta atattggcta taagattgcg	840
caacgatttg gtcattttga agcgggtggg cctgtcttgc aaggcctgaa caagccggtc	900
tccgacttgt cagtggtatg cagtgaagaa gacgtttata aggttgcgat tattacagca	960
gccaaggat tagcttaa	978

<210> 963

<211> 2367

<212> DNA

<213> *Lactobacillus plantarum*

<400> 963

atgacaacag attactcatc accagcatat ttgcaaaaag ttgataagta ctggcgtgct	60
gccaaactact tatcagttgg tcaactttat ttaaaagata atccactatt acaacggcca	120
ttgaaggcca gtgacgttaa ggttcatcca attggtcact gggggacgat tgccggtcaa	180
aactttatct atgctcatct taaccgggtc atcaacaagt acggtttgaa gatgtttctac	240
gttgaaggtc caggtcattg tgggtcaagt atggtttcaa actcttacct tgacgttact	300

tacaccgata tttatccaga aattacgcag gatgttgaag ggatgcaaaa gctcttcaag	360
caattctcat tcccaggtag ggttgcttcc catgcggcac ctgaaacacc cggttcaatc	420
cacgaaggtag gcaacttgg ttactcaatt tcacacgggg ttggggcaat tcttgacaat	480
cctgacgaaa tcgccgagggt tgttgttggg gatggggaat ccgaaacggg tccattagca	540
acttcatggc aatcaacgaa gtccattaac ccaatcaacg acggggctgt ttaccaatc	600
ttgaacttaa atggttttta gatttctaatt ccaacgattt ttggtcggac ttctgatgct	660
aagattaagg aatacttcga aagcatgaat tgggaaccaa tcttcgttga aggtgacgat	720
cctgaaaagg ttaccaccgc cttagctaag gccatggatg aagccgttga aaagatcaag	780
gcaatccaga agcatgctcg cgaataaac gatgcaacat tgccagtatg gccaatgatc	840
gtcttccgcg cacctaagggt ctggactggg ccgaagtcac gggacgggtga taagatcgaa	900
ggttcattcc gtgtcatca aattccgatt cctgttgatc aaaatgacat ggaacatgcg	960
gatgcttttag ttgattggct cgaatcata caaccaaag aactcttcaa tgaagatggc	1020
tctttgaagg atgataattaa agaaattatt cctactgggg acagtcggat ggctgctaac	1080
ccaatcacca atggtaggggt cgatccgaaa gccttgaact taccaaaact ccgtgattat	1140
gcggtcgata cgtccaaaga aggcgcgaat gttaagcaag atatgatcgt ttggtcagac	1200
tatttgcggg atgcatcaa gaaaaatcct gataacttcc ggttggtcgg acctgatgaa	1260
accatgtcta accgtttata tgggtgtcttc gaaaccacta atcgtcaatg gatggaagac	1320
attcatccag atagtacca atatgaagca ccagctggcc gggctcttaga tgctcagtta	1380
tctgaacacc aagctgaagg ttggttagaa ggttacgtct taactggacg tcatgggtta	1440
tttgccagtt atgaagcctt cctacgcgtt gtggactcaa tgttgacgca acacttcaag	1500
tggttacgta aagccaatga acttgattgg cgtaaaaagt acccatcact taacattatc	1560
gcggcttcaa ctgtattcca acaagacat aatggttata cccaccaaga tccaggtgca	1620
ttaaactcatt tggccgaaaa gaaaccagaa tacattcgtg aatatttacc agccgatgcc	1680
aacacgttat tagctgtcgg tgacgtcatt ttccggagcc aagaaaagat caactacgtg	1740
gttacgtcaa aacaccacg tcaacaatgg ttcagcattg aagaagctaa gcaattagtt	1800
gacaatggtc ttggtatcat tgattgggca agtacggacc aaggtagcga accagacatt	1860
gtctttgcag ctgctgggac ggaaccaacg cttgaaacgt tggtgccat ccaattacta	1920
cacgacagtt tcccagagat gaagattcgt ttcgtgaacg tggtcgacat cttgaagtta	1980
cgtagtctg aaaaggatcc gcggggcttg tcagatgctg agtttgacca ttactttact	2040
aaggacaaac cagtggcttt tgctttccac ggttacgaag acttagttcg tgacatcttc	2100

tttgatcgtc acaaccataa cttatacgtc cacggttacc gtgaaaatgg tgatattacc 2160

acaccattcg acgtacgggt catgaaccag atggaccgct tcgacttagc taagtcggca 2220

attgcggcgc aaccagcaat ggaaaacact ggtgcggcct tcgttcaatc catggataat 2280

atgcttgcta aacacaatgc ctatatccgg gatgccggaa ctgacttgcc agaagttaat 2340

gatttgcaat ggaagggttt aaaataa 2367

<210> 964

<211> 9613

<212> DNA

<213> Artificial Sequence

<220><223> pHR81-ILV5p-K9SB

<400> 964

aaacagtatg gaagaatgta agatggctaa gatttactac caagaagact gtaacttgct 60

cttgttggat ggtaagacta tcgccgttat cggttacggt tctcaaggct acgctcatgc 120

cttgaatgct aaggaatccg gttgtaacgt tatcattggt ttattcgaag gtgcggagga 180

gtggaaaaga gctgaagaac aaggtttcga agtctacacc gctgctgaag ctgctaagaa 240

ggctgacatc attatgatct tgatcccaga tgaaaagcag gctaccatgt acaaaaacga 300

catcgaacca aacttggaag ccggttaacat gttgatgttc gctcacggtt tcaacatcca 360

tttcggttgt attgttccac caaaggacgt tgatgtcact atgatcgctc caaagggtcc 420

aggtcacacc gttagatccg aatacgaaga aggtaaaggt gtcccatgct tggttgctgt 480

cgaacaagac gctactggca aggctttgga tatggctttg gcctacgctt tagccatcgg 540

tggtgctaga gccggtgtct tggaaactac cttcagaacc gaaactgaaa ccgacttggt 600

cggtgaacaa gctgttttat gtgggtggtgt ctgcgctttg atgcaggccg gttttgaaac 660

cttggttgaa gccggttacg acccaagaaa cgcttacttc gaatgtatcc acgaaatgaa 720

gttgatcggt gacttgatct accaatctgg ttctccggt atgcgttact ctatctccaa 780

cactgctgaa tacggtgact acattaccgg tccaaagatc attactgaag ataccaagaa 840

ggctatgaag aagattttgt ctgacattca agatgggtacc ttgccaagg acttcttggt 900

tgacatgtct gatgctggtt cccaggtcca cttcaaggct atgagaaaagt tggcctccga 960

acaccagct gaagtgtgct gtgaagaaat tagatccttg tactcctggt ccgacgaaga 1020

caagttgatt aacaactgag gccctgcagg ccagaggaaa ataatatcaa gtgctggaaa 1080

ctttttctct tggaattttt gcaacatcaa gtcatagtca attgaattga cccaatttca 1140

catttaagat tttttttttt tcatccgaca tacatctgta cactaggaag ccctgttttt 1200

ctgaagcagc ttcaaatata tataatTTTT acatatTTat tatgattcaa tgaacaatct	1260
aattaaatcg aaaacaagaa ccgaaacgcg aataaataat ttatttagat ggtgacaagt	1320
gtataagtcc tcatcgggac agctacgatt tctctttcgg ttttggtga gctactggtt	1380
gctgtgacgc agcggcatta gcgcggcggt atgagctacc ctcgtggcct gaaagatggc	1440
gggaataaag cggaactaaa aattactgac tgagccatat tgaggatcaat ttgtcaactc	1500
gtcaagtca cgttggtgga cgccccctt ccaacgaatc gtatatacta acatgcgcgc	1560
gcttcctata tacacatata catatatata tatatatata tgtgtgcgtg tatgtgtaca	1620
cctgtattta atttcttac tcgcgggttt ttttttttc tcaattcttg gcttctcttt	1680
tctcgagcgg accggatcct ccgcggtgcc ggagatcta tttaaatggc gcgcgcgacgt	1740
caggtggcac ttttcgggga aatgtgcgcg gaaccctat ttgtttattt ttctaatac	1800
attcaaatat gtatccgtc atgagacaat aacctgata aatgcttcaa taatattgaa	1860
aaaggaagag tatgagtatt caacatttcc gtgtcgccct tattcccttt ttgcgcat	1920
tttgccttcc tgttttctg caccagaaa cgttggtgaa agtaaaagat gctgaagatc	1980
agttgggtgc acgagtggt tacatcgaac tggatctcaa cagcggaag atccttgaga	2040
gttttcgccc cgaagaacgt tttcaatga tgagacttt taaagtctg ctatgtggcg	2100
cgttattatc ccgtattgac gccgggcaag agcaactcg tcgccgata cactattctc	2160
agaatgactt ggttagtac tcaccagtca cagaaaagca tcttacggat ggcatgacag	2220
taagagaatt atgcagtct gccataacca tgagtataa cactgcggcc aacttacttc	2280
tgacaacgat cggaggaccg aaggagctaa ccgctttttt gcacaacatg ggggatcatg	2340
taactcgct tgatcgttg gaaccggagc tgaatgaagc cataccaaac gacgagcgtg	2400
acaccacgat gcctgtagca atggcaacaa cgttcgcaa actattaact ggcaactac	2460
ttactctagc tttccggcaa caattaatag actggatgga ggccgataaa gtgcaggac	2520
cacttctgcg ctcgccctt ccggctggct ggtttattgc tgataaatct ggagccggtg	2580
agcgtgggtc tcgcgtatc attgcagcac tggggccaga tggttaagccc tcccgtatcg	2640
tagttatcta cagcagggg agtcaggcaa ctatggatga acgaaataga cagatcgctg	2700
agataggtgc ctactgatt aagcattggt aactgtcaga ccaagtttac tcatatatac	2760
tttagattga tttaaaactt catttttaat ttaaaaggat ctaggatgaag atcctttttg	2820
ataatctcat gacaaaatc ctttaacgtg agttttcgtt cactgagcg tcagaccccg	2880
tagaaaagat caaagatct tcttgagatc cttttttct gcgcgtaac tgctgcttgc	2940
aaacaaaaa accaccgcta ccagcgggtg tttgtttgcc ggatcaagag ctaccaactc	3000

ttttccgaa ggtaactggc ttcagcagag cgcagatacc aaatactgtt cttctagtgt	3060
agccgtagt tt agccaccac ttcaagaact ctgtagcacc gcctacatac ctgcctctgc	3120
taatcctgtt accagtggct gctgccagtg gcgataagtc gtgtcttacc gggttggact	3180
caagacgata gttaccggat aaggcgcagc ggtcgggctg aacgggggggt tcgtgcacac	3240
agcccagctt ggagcgaacg acctacaccg aactgagata cctacagcgt gagctatgag	3300
aaagcggcac gcttcccga gggagaaaagg cggacaggta tccggttaagc ggcagggtcg	3360
gaacaggaga gcgcacgagg gagcttcag ggggaaacgc ctggtatctt tatagtcctg	3420
tcgggtttcg ccacctctga cttgagcgtc gatTTTTgtg atgctcgtca ggggggcgga	3480
gcctatggaa aaacgccagc aacgcggcct ttttacggtt cctggccttt tgctggcctt	3540
ttgctcacat gttctttcct gcgttatccc ctgattctgt ggataaccgt attaccgcct	3600
ttgagtgagc tgataccgt cgccgcagcc gaacgaccga gcgcagcagc tcagtgcagc	3660
aggaagcggg agagcgccca atacgcaaac cgcctctccc cgcgcgttgg ccgattcatt	3720
aatgcagctg gcacgacagg tttcccact ggaaagcggg cagtgcgcgc aacgcaatta	3780
atgtgagtta gctcactcat taggcacccc aggcctttaca ctttatgctt ccggtcgtta	3840
tgttgtgtgg aattgtgagc ggataacaat ttcacacagg aaacagctat gaccatgatt	3900
acgccaagct tttctttcc aattttttt ttttcgtcat tataaaaac attacgaccg	3960
agattcccgg gtaataactg atataattaa attgaagctc taatttgtga gtttagtata	4020
catgcattta cttataatc agtttttttag ttttgctggc cgcattctct caaatatgct	4080
tcccagcctg cttttctgta acgttcaccc tctaccttag catcccttcc ctttgcaaat	4140
agtcctcttc caacaataat aatgtcagat cctgtagaga ccacatcatc cacggttcta	4200
tactgttgac ccaatgcgtc tcccttgta tctaaacca caccgggtgt cataatcaac	4260
caatcgtaac cttcatctct tccacccatg tctctttgag caataaagcc gataacaaaa	4320
tctttgtcgc tcttcgcaat gtcaacagta cccttagtat attctccagt agataggag	4380
cccttgcatg acaattctgc taacatcaaa aggcctctag gttcctttgt tacttcttct	4440
gccgcctgct tcaaaccgct aacaatacct gggcccaacca caccgtgtgc attcgtaatg	4500
tctgcccatt ctgctattct gtatacacc gcagagtact gcaatttgac tgtattacca	4560
atgtcagcaa attttctgtc ttggaagagt aaaaaattgt acttggcgga taatgccttt	4620
agcggcttaa ctgtgccctc catggaaaaa tcagtcaaga tatccacatg tgtttttagt	4680
aaacaaattt tgggacctaa tgcttcaact aactccagta attccttggt ggtacgaaca	4740

tccaatgaag cacacaagtt tgtttgcttt tcgtgcatga tattaaatag ctggcagca	4800
acaggactag gatgagtagc agcacgttcc ttatatgtag ctttcgacat gatttatctt	4860
cgtttcctgc aggtttttgt tctgtgcagt tgggttaaga atactgggca atttcatgtt	4920
tcttcaacac tacatatgcg tataatatacc aatctaagtc tgtgctcctt ctttcgttct	4980
tccttctgtt cggagattac cgaatcaaaa aaatttcaag gaaaccgaaa tcaaaaaaaaa	5040
gaataaaaaa aaaatgatga attgaaaagc ttgcatgcct gcaggtcgac tctagtatac	5100
tccgtctact gtacgataca cttccgctca ggtccttgtc ctttaacgag gccttaccac	5160
tcttttgta ctctattgat ccagctcagc aaaggcagtg tgatctaaga ttctatcttc	5220
gcgatgtagt aaaactagct agaccgagaa agagactaga aatgcaaaag gcacttctac	5280
aatggctgcc atcattatta tccgatgtga cgctgcattt tttttttttt tttttttttt	5340
tttttttttt tttttttttt tttttttttg taaaaatc ataaaaaag agaacttttt	5400
taagcaagga ttttctaac ttcttcggcg acagcatcac cgacttcggt ggtactgttg	5460
gaaccaccta aatcaccagt tctgatacct gcatccaaaa cttttttaac tgcactttca	5520
atggctttac cttcttcagg caagttcaat gacaatttca acatcattgc agcagacaag	5580
atagtggcga tagggttgac cttattcttt ggcaaatctg gagcggaacc atggcatggt	5640
tcgtacaaac caaatgcggt gttcttgtct ggcaaagagg ccaaggacgc agatggcaac	5700
aaaccaagg agcctgggat aacggaggct tcatcggaaga tgatatacc aaacatgttg	5760
ctggtgatta taataccatt taggtgggtt gggttcttaa ctaggatcat ggcggcagaa	5820
tcaatcaatt gatgttgaa tttcaatgta gggaattcgt tcttgatggt ttctccaca	5880
gtttttctcc ataactttga agaggccaaa acattagctt tatccaagga ccaaataaggc	5940
aatggtggct catgtttag ggccatgaaa gcggccattc ttgtgattct ttgcattct	6000
ggaacggtgt attgttcaat atccaagcg acaccatcac catcgtcttc ctttctctta	6060
ccaaagtaaa tacctccac taattctcta acaacaacga agtcagtacc tttagcaaat	6120
tgtggcttga ttggagataa gtctaaaaga gagtccgatg caaagttaca tggctttaag	6180
ttggcgtaca attgaagttc ttacggatt tttagtaaac cttgttcagg tctaacta	6240
ccggtacccc atttaggacc acccacagca cctaacaaaa cggcatcagc cttcttggag	6300
gcttccagcg cctcatctgg aagtggaaca cctgtagcat cgatagcagc accaccaatt	6360
aaatgatttt cgaatcgaa cttgacattg gaacgaacat cagaaatagc tttagaacc	6420
ttaatggctt cggtgtgat ttcttgacca acgtggtcac ctggcaaac gacgatcttc	6480
ttaggggcag acattacaat ggtatatcct tgaatatata aaaaaaaaaa aaaaaaaaaa	6540
aaaaaaaaa aatgcagctt ctcaatgata ttggaatagc ctttaggag atacgccta	6600

atatccgaca aactgtttta cagatttacg atcgtacttg ttacccatca ttgaattttg	6660
aacatccgaa cctgggagtt ttccctgaaa cagatagtat atttgaacct gtataataat	6720
atatagtcta gcgctttacg gaagacaatg tatgtatttc ggttcctgga gaaactattg	6780
catctattgc ataggtaatc ttgcacgtcg catccccggt tcattttctg cgtttccatc	6840
ttgcacttca atagcatatc ttgtttaacg aagcatctgt gcttcatttt gtagaacaaa	6900
aatgcaacgc gagagcgcta atttttcaaa caaagaatct gagctgcatt ttacagAAC	6960
agaaatgcaa cgcgaaagcg ctattttacc aacgaagaat ctgtgcttca tttttgtaaa	7020
acaaaaatgc aacgcgagag cgctaatttt tcaaacaaag aatctgagct gcatttttac	7080
agaacagaaa tgcaacgcga gagcgctatt ttaccaacaa agaactataa cttctttttt	7140
gttctacaaa aatgcatccc gagagcgcta tttttctaac aaagcatctt agattacttt	7200
ttttctcctt tgtgcgctct ataatgcagt ctcttgataa ctttttgcaC tgtaggtccg	7260
ttaaggttag aagaaggcta ctttgggtgc tattttctct tccataaaaa aagcctgact	7320
ccacttcccg cgtttactga ttactagcga agctgcgggt gcattttttc aagataaagg	7380
catcccgat tatattctat accgatgtgg attgcgcata ctttgtgaac agaaagtgat	7440
agcgttgatg attcttcatt ggtcagaaaa ttatgaacgg tttcttctat tttgtctcta	7500
tatactacgt ataggaaatg ttacattttt cgtattgttt tcgattcact ctatgaatag	7560
ttcttactac aatttttttg tctaaagagt aatactagag ataaacataa aaaatgtaga	7620
ggtcgagttt agatgcaagt tcaaggagcg aaaggtggat gggtaggtta tatagggata	7680
tagcacagag atatatagca aagagatact tttagcaat gtttgtggaa gcggtattcg	7740
caatatttta gtagctcgtt acagtcgggt gcgttttttg tttttgaaa gtgcgtcttc	7800
agagcgcttt tggttttcaa aagcgctctg aagtccctat actttctaga gaataggaac	7860
ttcggaatag gaacttcaaa gcgtttccga aaacgagcgc ttccgaaaat gcaacgcgag	7920
ctgcgcacat acagctcact gttcacgtcg cacctatata tgcgtgttgc ctgtatatat	7980
atatacatga gaagaacggc atagtgcgtg tttatgctta aatgcgtact tatatgcgtc	8040
tatttatgta ggatgaaagg tagtctagta cctcctgtga tattatccca ttccatgcgg	8100
ggtatcgtat gcttcttca gcactacct ttagctgttc tatagtctgc cactcctcaa	8160
ttggattagt ctcatcttc aatgctatca tttccttga tattggatca tatgcatagt	8220
accgagaaac tagaggatct cccattaccg acatttgggc gctatacgtg catatgttca	8280
tgtatgtatc tgtatttaaa acacttttgi attatttttc ctcatatatg tgtatagggt	8340
tatacggatg atttaattat tacttcacca ccttttatit caggetgata tcttagcctt	8400
gttactagtc accggtggcg gccgcacctg gtaaaacctc tagtggagta gtagatgtaa	8460

tcaatgaagc ggaagccaaa agaccagagt agaggcctat agaagaaact gcgatacctt 8520

ttgtgatggc taaacaaaca gacatctttt tatatgtttt tacttctgta tatcgtgaag 8580

tagtaagtga taagcgaatt tggctaagaa cgttgtaagt gaacaaggga cctcttttgc 8640

ctttcaaaaa aggattaaat ggagttaatc attgagattt agttttcgtt agattctgta 8700

tccttaata actcccttac ccgacgggaa ggcacaaaag acttgaataa tagcaaacgg 8760

ccagtagcca agaccaaata atactagagt taactgatgg tcttaaacag gcattacgtg 8820

gtgaactcca agaccaatat acaaaatc gataagtat tcttgccac caatttaagg 8880

agcctacatc aggacagtag taccattcct cagagaagag gtatacataa caagaaaatc 8940

gcgtgaacac cttatataac ttagcccggt attgagctaa aaaaccttgc aaaatttcct 9000

atgaataaga atacttcaga cgtgataaaa atttactttc taactcttct cacgtgccc 9060

ctatctgttc ttccgctcta ccgtgagaaa taaagcatcg agtacggcag ttcgctgtca 9120

ctgaactaaa acaataaggc tagttcgaat gatgaacttg cttgctgtca aacttctgag 9180

ttgccgtga tgtgacactg tgacaataaa ttcaaaccgg ttatagcggc ctccctccgt 9240

accggttctg ccacctcaa tagagctcag taggagtcag aacctctgcg gtggctgtca 9300

gtgactcctc cgcgtttcgt aagtgtgctg cgtgcacatt tcgcccgttc ccgtcatct 9360

tgcagcaggc ggaatttttc atcacgctgt aggacgcaa aaaaaataa ttaatcgtac 9420

aagaatcttg gaaaaaaat tgaaaaattt tgtataaaag ggaatgaccta acttgactca 9480

atggctttta caccagtat ttccctttc cttgtttgtt acaattatag aagcaagaca 9540

aaaacatata gacaacctat tcctaggagt tatattttt taccctacca gcaatataag 9600

taaaaaactg ttt 9613

<210> 965

<211> 12319

<212> DNA

<213> Artificial Sequence

<220><223> pBP2090

<400> 965

tccattacc gacatttggg cgctatacgt gcatatgttc atgtatgtat ctgtatttaa 60

aacacttttg tattattttt cctcatatat gtgtataggt ttatcggat gatttaatta 120

ttacttcacc accctttatt tcagctgat atcttagcct tgttactaga ttaatcatgt 180

aattagttat gtcacgtta cattcacgcc cccccccac atccgctcta accgaaaagg 240

aaggagttag acaacctgaa gtctagggtc ctatttattt ttttatagtt atgttagtat 300

taagaacgtt atttatattt caaatTTTTt ttttttttct gtacagacgc gtgtacgcat	360
gtaacattat actgaaaacc ttgcttgaga aggttttggg acgctcgaag gctttaattt	420
gcgggCGGCC gcacctggta aaacctctag tggagtagta gatgtaatca atgaagcgga	480
agccaaaaga ccagagtaga ggcctataga agaaactgcg ataccttttg tgatggctaa	540
acaaacagac atctttttat atgtttttac ttctgtatat cgtgaagtag taagtataa	600
gcgaatttgg ctaagaacgt tgtaagtga caaggacct cttttgcctt tcaaaaaagg	660
attaaatgga gttaatcatt gagatttagt tttcgtaga ttctgtatcc ctaaataact	720
cccttaccgc acgggaaggc acaaaagact tgaataatag caaacggcca gtagccaaga	780
ccaaataata ctagagttaa ctgatggctt taaacaggca ttacgtggtg aactccaaga	840
ccaatataca aaatatcgat aagttattct tgcccaccaa ttttaaggagc ctacatcagg	900
acagtagtac cattcctcag agaagaggta tacataacaa gaaaatcgcg tgaacacctt	960
atataactta gcccgttatt gagctaaaaa accttgcaaa atttcctatg aataagaata	1020
cttcagacgt gataaaaatt tactttctaa ctcttctcac gctgccccta tctgttcttc	1080
cgctctaccg tgagaaataa agcatcgagt acggcagttc gctgtcactg aactaaaaca	1140
ataaggctag ttgcaatgat gaacttgctt gctgtcaaac ttctgagttg ccgctgatgt	1200
gacactgtga caataaatc aaaccggtta tagcggcttc ctccggtacc gtttctgcca	1260
cctccaatag agctcagtag gagtcagaac ctctgcggtg gctgtcagtg actcatccgc	1320
gtttcgtaag ttgtgcgct gcacatttcg cccgttcccg ctcatcttgc agcaggcgga	1380
aattttcatc acgctgtagg acgcaaaaaa aaaataatta atcgtacaag aatcttgga	1440
aaaaaattga aaaattttgt ataaaaggga tgacctaaact tgactcaatg gcttttacac	1500
ccagtatttt ccttttcctt gtttgttaca attatagaag caagacaaaa acatatagac	1560
aacctattcc taggagttaa atttttttac cctaccagca atataagtaa aaaactgttt	1620
aaacagtatg gaagaatgta agatggctaa gatttactac caagaagact gtaacttgtc	1680
cttgttggat ggtaagacta tcgccgttat cggttacggt tctcaaggtc acgctcatgc	1740
cctgaatgct aaggaaatccg gttgtaacgt tatcattggg ttatacgaag gtgctaagga	1800
ttggaaaaga gctgaagaac aaggtttcga agtctacacc gctgctgaag ctgctaagaa	1860
ggctgacatc attatgatct tgatcaacga tgaagcag gctaccatgt acaaaaacga	1920
catcgaacca aacttggaag ccggtaacat gttgatgttc gctcacggtt tcaacatcca	1980
tttcggttgt attgttcac caaaggacgt tgatgtcact atgatcgctc caaagggtcc	2040
aggtcacacc gttagatccg aatacgaaga aggtaaagggt gtcccatgct tggttgctgt	2100

cgaacaagac gctactggca aggctttgga tatggctttg gcctacgctt tagccatcgg 2160

tggtgctaga gccggtgtct tggaaactac cttcagaacc gaaactgaaa ccgacttggt 2220

cggtgaacaa gctgttttat gtgggtggtgt ctgcgctttg atgcaggccg gttttgaaac 2280

cttggttgaa gccggttacg acccaagaaa cgcttacttc gaatgtatcc acgaaatgaa 2340

gttgatcggt gacttgatct accaatctgg tttctccggt atgcgttact ctatctccaa 2400

cactgctgaa tacggtgact acattaccgg tccaaagatc attactgaag ataccaagaa 2460

ggctatgaag aagattttgt ctgacattca agatggtacc tttccaagg acttcttggt 2520

tgacatgtct gatgctggtt cccaggcca cttcaaggct atgagaaaagt tggcctccga 2580

acaccagct gaagtgtcg gtgaagaaat tagatccttg tactcctggt ccgacgaaga 2640

caagttgatt aacaactgag gccctgcagg ccagaggaaa ataatatcaa gtgctggaaa 2700

ctttttctct tggaattttt gcaacatcaa gtcatagtca attgaattga cccaatttca 2760

catttaagat ttttttttt tcatccgaca tacatctgta cactaggaag cctgttttt 2820

ctgaagcagc ttcaaatata tataattttt acatatattat tatgattcaa tgaacaatct 2880

aattaaatcg aaaacaagaa ccgaaacgcg aataaataat ttatttagat ggtgacaagt 2940

glataagtcc tcatcgggac agctacgatt tctctttcgg ttttggtga gctactggtt 3000

gctgtgacgc agcggcatta gcgcggcggt atgagctacc ctcgtggcct gaaagatggc 3060

gggaataaag cggaactaaa aattactgac tgagccatat tgaggtaaat ttgtcaactc 3120

gtcaagtcac gtttggtgga cgcccccttt ccaacgaatc gtatatacta acatgcgcgc 3180

gcttcctata tacacatata catatatata tatatatata tgtgtgcgtg tatgtgtaca 3240

cctgtattta atttccttac tcgcggggtt tcttttttcc tcaattcttg gcttcctctt 3300

tctcgagcgg accggatcct cgcgaactcc aaaatgagct atcaaaaacg atagatcgat 3360

taggatgact ttgaaatgac tccgcagtgg actggcgggt aatttcaagc gtgagtaaaa 3420

tagtgcatga caaaagatga gctaggcttt tgtaaaaata tcttacgttg taaaatttta 3480

gaaatcatta tttccttcat atcattttgt catlgacctt cagaagaaaa gagccgacca 3540

ataatataaa taaataaata aaaataatat tccattatit ctaaacagat tcaatactca 3600

ttaaaaaact atatcaatta atttgaatta acgcggccgc ttaaccacag caaccaggac 3660

aacatititit gccagtittt tcaggcttcc aaaagtctgt tacggctccc ctagaagcag 3720

acgaaacgat gtgagcatat ttaccaagga taccgctga atagagcggg ggcaattcaa 3780

tggtctcttg acgatgtttt aactcttcat cggagatatc aaagtgtaat tccttagtgt 3840

cttggccaat agtgactatg tctctgttt gcaggtaggc gattggaccg ccatcttgtg	3900
cttcaggagc gatatgacc acgacaagac cataagtacc acctgagaag cggccatctg	3960
tcagaagggc aactttttca ccttgccctt taccaacaat cattgatgaa agggaaagca	4020
tttcaggcat accaggaccg ccttttggc ctacaaaacg tacgacaaca acatcaccat	4080
caacaatatc atcattcaag acagcttcaa tggcttcttc ttcagaatta aagaccttag	4140
caggaccgac atgacgacgc acttttacac cagaaacttt ggcaacggca cgtctggag	4200
ccaagttacc atggagaata atgagcggac catcttcacg tttaggattt tcaagcggca	4260
taataacctt ttgaccaggt gttaaatcat caaaagcctt caaattttca gcgactgttt	4320
tgccagtaca agtgatacgg tcacatgaa ggaagccatt ttttaaggaga tatttcataa	4380
ctgctggtag ccttcgacc ttgtaaaggt cttggaatac atattgacca gaaggtttca	4440
aatcagccaa atgaggaact ttttcttga aagtattgaa atcatcaagt gtcaattcca	4500
cattagcagc atgggcaata gctaagaggt gaagggttga gttggttgaa cctcccagag	4560
ccatagttag agtaatagca tcttcaaaag cttcacgcgt taaaatgtca gaaggtttta	4620
agcccatttc gagcattttg acaacagcgc gaccagcttc ttcaatatct gctttctttt	4680
ctgcggattc agccgggtga gaagatgaac ccggaaggct aagtcccaa acttcaatag	4740
ctgtcgccat tgtgttagca gtatacatc caccgcagcc tccaggaccg ggacaagcat	4800
tacattccaa agctttaact tcttctttgg tcatatcgcc gtggttccaa tggccgacac	4860
cttcaaagac agagactaaa tcgatatctt tgccgtctaa attaccaggt gcaattgttc	4920
cgccgtaaag aaaaatggct gggatatcca tgttagccat agcgataaca gaaccgggca	4980
tgtttttatc acaaccgcca atggctacaa aagcatccgc attatgacct cccatggctg	5040
cttcaataga atctgcaata atatcacgag atgtcaagga gaaacgcatt ccttgggttc	5100
ccatggcgat tccatcagaa accgtgattg ttccgaactg aactggccaa gcaccagctt	5160
ccttaacacc gactttggct agtttacc aaagcatgaa gtggatatta caaggtgtgt	5220
tttcagccca agttgaaatg acaccgacga taggtttttc aaagtcttca tcttgcatc	5280
cagttgcacg caacatagca cgattagggt atttaaccat tgaatcgtaa acagaactac	5340
gatttcttaa gtctttaaga gttttttgt cagtcatact cacgtggaac ttagattaga	5400
ttgctatgct tctctctaa cgagcaagaa gtaaaaaaag ttgtaataga acaagaaaaa	5460
tgaaactgaa gcttgagaaa ttgaagaccg tttattagct taaatatcaa tgggaggtca	5520
tcgaaagaga aaaaaatcaa gaaagaaact ctcaagaaaa agaaacgtga taaaaatttt	5580
tattgcctct ctcgacgaag agaaagaaac gaggcggtcc cttttttctt ttccaaacct	5640
ttagtacggg taattagcga caccctagag gaagaaagag gggaaattta gtatgctgtg	5700

cttgggtgtc ttgaagtggc acggcgatgc gcggagtcgc agaaaatctg gaagagtaaa 5760

aaggggtag aagcgttttg aagctatgcc ggcagatcta tttaaatggc gcgccgacgt 5820

caggtggcac ttttcgggga aatgtgcgcg gaaccctat ttgtttatit ttctaaatac 5880

attcaaatat gtatccgctc atgagacaat aaccctgata aatgcttcaa taatattgaa 5940

aaaggaagag tatgagtatt caacatttcc gtgtcgccct tattcccttt ttgcgggcat 6000

tttgccttcc tgtttttgct caccagaaa cgttggtgaa agtaaaagat gctgaagatc 6060

agttgggtgc acgagtgggt tacatcgaac tggatctcaa cagcggtaag atccttgaga 6120

gttttcgccc cgaagaacgt tttccaatga tgagcacttt taaagtctct ctatgtggcg 6180

cgggtattatc ccgtattgac gccgggcaag agcaactcgg tcgccgcata cactattctc 6240

agaatgactt gggtgaglac tcaccagtca cagaaaagca tcttacggat ggcatgacag 6300

taagagaatt atgcagtgtc gccataacca tgagtataa cactgcggcc aacttacttc 6360

tgacaacgat cggaggaccg aaggagctaa ccgctttttt gcacaacatg ggggatcatg 6420

taactcgctt tgatcgttgg gaaccggagc tgaatgaagc cataccaaac gacgagcgtg 6480

acaccacgat gctgtagca atggcaacaa cgttgcgcaa actattaact ggccaactac 6540

ttactctagc ttcccggcaa caattaatag actggatgga ggcggataaa gttgcaggac 6600

cacttctgcg ctcgccctt ccggctggct ggtttattgc tgataaatct ggagccggtg 6660

agcgtgggtc tcgcggtatc attgcagcac tggggccaga tggtaagccc tcccgtatcg 6720

tagttatcta cagcaggggg agtcaggcaa ctatggatga acgaaataga cagatcgctg 6780

agatagggtc ctactgatt aagcattggt aactgtcaga ccaagtttac tcatatatac 6840

tttagattga tttaaaactt catTTTTaat ttaaaaggat ctaggtgaag atcctttttg 6900

ataatctcat gacaaaaatc ccttaacgtg agttttcggt cactgagcg tcagaccccg 6960

tagaaaagat caaaggatct tcttgagatc ctttttttct gcgcgtaatc tgctgcttgc 7020

aaacaaaaaa accaccgcta ccagcgggtg tttgtttgcc ggatcaagag ctaccaactc 7080

tttttccgaa ggtaactggc ttcagcagag cgcagatacc aaatactgtt cttctagtgt 7140

agccgtagtt aggccaccac ttcaagaact ctgtagcacc gcctacatac ctgctctgc 7200

taatcctgtt accagtggct gctgccagtg gcgataagtc gtgtcttacc gggttggact 7260

caagacgata gttaccggtt aaggcgagc ggtcgggctg aacggggggt tcgtgcacac 7320

agcccagctt ggagcgaacg acctacaccg aactgagata cctacagcgt gagctatgag 7380

aaagcggcac gcttcccgaa gggagaaaagg cggacaggta tccggttaagc ggcagggtcg 7440

gaacaggaga gcgcacgagg gagcttccag ggggaaacgc ctggtatctt tatagtcctg 7500

tcgggtttcg ccacctctga cttgagcgtc gatTTTTgtg atgctcgtca ggggggcgga 7560

gcctatggaa aaacgccagc aacgcggcct ttttacggtt cctggccttt tgctggcctt	7620
ttgctcacat gttcttttct gcgttatccc ctgattctgt ggataaccgt attaccgcct	7680
ttgagtgage tgataaccgt cgccgcagcc gaacgaccga gcgcagcgag tcagtgageg	7740
aggaagcgga agagcgccca atacgcaaac cgctctccc cgcgcttgg ccgattcatt	7800
aatgcagctg gcacgacagg tttcccgaact ggaaagcggg cagtgagegc aacgcaatta	7860
atgtgagtta gtcactcat taggcacccc aggctttaca ctttatgctt ccggctcgta	7920
tgttgtgtgg aattgtgagc ggataacaat ttcacacagg aaacagctat gaccatgatt	7980
acgccaagct ttttctttcc aatttttttt ttttcgtcat tataaaaatc attacgaccg	8040
agattcccgg gtaataactg atataattaa attgaagctc taatttgtga gtttagtata	8100
catgcattta cttataatac agtttttttag ttttgctggc cgcatcttct caaatatgct	8160
tccagcctg tttttctgta acgttcaccc tctaccttag catcccttcc ctttgcaaat	8220
agtcctcttc caacaataat aatgtcagat cctgtagaga ccacatcatc cacggttcta	8280
tactgttgac ccaatgcgtc tcccttgta tctaaacca caccgggtgt cataatcaac	8340
caatcgtaac cttcatctct tccacccatg tctctttgag caataaagcc gataacaaaa	8400
tctttgtcgc tcttcgcaat gtcaacagta cccttagtat attctccagt agataggag	8460
cccttgcatg acaattctgc taacatcaaa aggctcttag gttcctttgt tacttcttct	8520
gccgcctgct tcaaaccgct aacaatacct ggcccacca caccgtgtgc attcgtaatg	8580
tctgcccatt ctgctattct gtatacccc gcagagtact gcaatttgac tgtattacca	8640
atgtcagcaa attttctgtc ttccaagagt aaaaaattgt acttggcgga taatgccttt	8700
agcggcttaa ctgtgccctc catggaaaaa tcagtcaaga tatccacatg tgtttttagt	8760
aaacaaattt tgggacctaa tgcttcaact aactccagta attccttggt ggtacgaaca	8820
tccaatgaag cacacaagtt tgtttgcttt tcgtgcatga tattaaatag cttggcagca	8880
acaggactag gatgagtagc agcacgttcc ttatatgtag ctttcgacat gatttatctt	8940
cgtttctcgc aggtttttgt tctgtgcagt tgggttaaga atactgggca atttcatgtt	9000
tcttcaacac tacatatgcg tataataacc aatctaagtc tgtgtcctt ccttcgttct	9060
tccttctgtt cggagattac cgaatcaaaa aaatttcaag gaaaccgaaa tcaaaaaaaaa	9120
gaataaaaaa aaaaatgatg attgaaaagc ttgcatgcct gcaggtcgac tctagtatac	9180
tccgtctact gtacgataca ctcccgctca ggctcttgct ctttaacgag gccttaccac	9240
tcttttgta ctctattgat ccagctcagc aaaggcagtg tgatctaaga ttctatcttc	9300

gcgatgtagt aaaactagct agaccgagaa agagactaga aatgcaaaag gcacttctac	9360
aatggctgcc atcattatta tccgatgtga cgctgcattt tttttttttt tttttttttt	9420
tttttttttt tttttttttt tttttttttg tacaatatc ataaaaaag agaacttttt	9480
taagcaagga ttttctaac ttcttcggcg acagcatcac cgacttcggt ggtactgttg	9540
gaaccaccta aatcaccagt tctgatacct gcatccaaaa cttttttaac tgcactttca	9600
atggctttac cttcttcagg caagtccaat gacaatttca acatcattgc agcagacaag	9660
atagtggcga tagggttgac cttattcttt ggcaaatctg gagcggaacc atggcatggt	9720
tcgtacaaac caaatgcggt gttcttgtct ggcaaagagg ccaaggacgc agatggcaac	9780
aaaccaagg agcctgggat aacggaggct tcatcggaaga tgatatacc aaacatgttg	9840
ctggtgatta taataccatt taggtgggtt gggttcttaa ctaggatcat ggcggcagaa	9900
tcaatcaatt gatgttgaaac tttcaatgta gggaattcgt tcttgatggt ttctccaca	9960
gtttttctcc ataactttga agaggccaaa acattagctt tatccaagga ccaaataaggc	10020
aatggtaggt catgtttag ggccatgaaa gcggccattc ttgtgattct ttgcatttct	10080
ggaacggtgt attgttcaat atccaagcg acaccatcac catcgtcttc ctttctctta	10140
ccaaagtaaa tacctccac taattctcta acaacaacga agtcagtacc tttagcaaat	10200
tgtggcttga ttggagataa gtctaaaaga gagtcggatg caaagttaca tggctttaag	10260
ttggcgtaca attgaagttc ttacggatt tttagtaaac cttgttcagg tctaactata	10320
ccggtacccc atttaggacc acccacagca cctaacaaaa cggcatcagc cttcttgag	10380
gcttcacg cctcatctgg aagtggaaca cctgtagcat cgatagcagc accaccaatt	10440
aaatgatatt cgaaatcgaa cttgacattg gaacgaacat cagaaatagc ttaagaacc	10500
ttaatggctt cggctgtgat ttcttgacca acgtgggtcac ctggcaaac gacgatcttc	10560
ttaggggcag acattacaat ggtatatcct tgaatatat ataaaaaaaa aaaaaaaaaa	10620
aaaaaaaaa aatgcagctt ctcaatgata ttgaatacg ctttaggag atacagccta	10680
atatccgaca aactgtttta cagatttacg atcgtacttg ttacccatca ttgaattttg	10740
aacatccgaa cctgggagtt ttccctgaaa cagatagtat atttgaacct gtataataat	10800
atatagtcta gcgctttacg gaagacaatg tatgtatttc ggttcctgga gaaactattg	10860
catctattgc ataggtaatc ttgcacgtcg catccccggt tcattttctg cgtttccatc	10920
ttgcacttca atagcatatc ttgtttaacg aagcatctgt gcttcatttt gtagaacaaa	10980
aatgcaacgc gagagcgcta atttttcaaa caaagaatct gagctgcatt ttacagaac	11040
agaaatgcaa cgcgaaagcg ctattttacc aacgaagaat ctgtgcttca tttttgtaaa	11100
acaaaaatgc aacgcgagag cgctaatttt tcaacaaaag aatctgagct gcatttttac	11160

agaacagaaa tgcaacgcga gagcgctatt ttaccaacaa agaattctata cttctttttt	11220
gttctacaaa aatgcatccc gagagcgcta tttttctaac aaagcatctt agattacttt	11280
ttttctcctt tgtgcgctct ataatgcagt ctcttgataa ctttttgcac tgtaggtcgg	11340
ttaaggtttag aagaaggcta ctttgggtgc tttttctct tccataaaaa aagcctgact	11400
ccacttcccg cgtttactga ttactagcga agctgcgggt gcattttttc aagataaagg	11460
catccccgat tatattctat accgatgtgg attgcgcata ctttgtgaac agaaagtgat	11520
agcgttgatg attcttcatt ggtcagaaaa ttatgaacgg tttcttctat tttgtctcta	11580
tatactacgt ataggaaatg tttacatttt cgtattgttt tgcattcact ctatgaatag	11640
ttcttactac aatttttttg tctaaagagt aatactagag ataaacataa aaaatgtaga	11700
ggtcgagttt agatgcaagt tcaaggagcg aaaggtggat gggtaggtta tatagggata	11760
tagcacagag atatatagca aagagatact tttgagcaat gtttgtggaa gcggtattcg	11820
caatatatta gtagctcgtt acagtccggt gcgttttttg ttttttgaaa gtgcgtcttc	11880
agagcgcttt tggttttcaa aagcgctctg aagttcctat actttctaga gaataggaac	11940
ttcggaatag gaacttcaaa gcgtttccga aaacgagcgc ttccgaaaaat gcaacgcgag	12000
ctgcgcacat acagctcact gtccacgtcg cacctatata tgcgtgttgc ctgtatatat	12060
atatacatga gaagaacggc atagtgcgtg tttatgctta aatgcgtact tatatgcgtc	12120
tatttatgta ggatgaaagg tagtctagta cctcctgtga tattatecca ttccatgcgg	12180
ggtatcgtat gcttccttca gcactaccct ttagctgttc tatatgctgc cactcctcaa	12240
ttggattagt ctcatccttc aatgctatca tttcctttga tattggatca tatgcatagt	12300
accgagaaac tagaggatc	12319