

(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2013-0138760
(43) 공개일자 2013년12월19일

(51) 국제특허분류(Int. Cl.)
C12N 1/21 (2006.01) C12N 15/54 (2006.01)
C12P 7/64 (2006.01)
(21) 출원번호 10-2013-7007481
(22) 출원일자(국제) 2011년08월26일
심사청구일자 없음
(85) 번역문제출일자 2013년03월25일
(86) 국제출원번호 PCT/US2011/049384
(87) 국제공개번호 WO 2012/027689
국제공개일자 2012년03월01일
(30) 우선권주장
61/377,248 2010년08월26일 미국(US)
(뒷면에 계속)

(71) 출원인
이 아이 듀폰 디 네모아 앤드 캄파니
미합중국 델라웨어주 (우편번호 19898) 월밍톤시
마아캣트 스트리트 1007
(72) 발명자
홍, 승-표
미국 19707 델라웨어주 호케신 피어슨스 리지 9
샤프, 파멜라, 엘.
미국 19808 델라웨어주 월밍톤 케이만 코트 232
(뒷면에 계속)
(74) 대리인
김영, 양영준, 양영환

전체 청구항 수 : 총 13 항

(54) 발명의 명칭 **고농도의 에이코사펜타엔산 생성을 위한 재조합 미생물 숙주 세포**

(57) 요약

건조 세포 중량의 중량 백분율로 측정하여, 25 중량 백분율 초과인 에이코사펜타엔산["EPA"], 즉 오메가-3 다중 불포화 지방산을 포함하는 미생물 오일을 생성시킬 수 있는 유지성 효모 야로위아 리폴리티카의 엔지니어링된 균주가 본 명세서에 개시된다.

(72) 발명자

수에, 지시웅

미국 19317-9728 펜실베이니아주 차즈 포드 하베이
레인 111

야다브, 나렌드라, 에스.

미국 19803 델라웨어주 월링톤 터커웨이 드라이브
126

장, 홍시양

미국 19317 펜실베이니아주 차즈 포드 로더필드 레인
3808

주, 쿤

미국 19382 펜실베이니아주 웨스트 체스터 리비어 로
드 544

(30) 우선권주장

61/428,277 2010년12월30일 미국(US)

61/479,921 2011년04월28일 미국(US)

특허청구의 범위

청구항 1

건조 세포 중량의 중량 백분율로 측정하여 적어도 25 중량 백분율의 에이코사펜타엔산을 포함하는 오일을 생성하는 재조합 미생물 숙주 세포.

청구항 2

제1항에 있어서, 오일이 전체 지방산의 중량 백분율로 측정하여 적어도 45 중량 백분율의 에이코사펜타엔산을 포함하는 재조합 미생물 숙주 세포.

청구항 3

제1항 또는 제2항에 있어서, 오일은 전체 지방산의 중량 백분율로 측정한 에이코사펜타엔산 대 전체 지방산의 중량 백분율로 측정한 리놀레산의 비가 적어도 2.4인 재조합 미생물 숙주 세포.

청구항 4

- (a) 적어도 하나의 델타-8 불포화화효소(desaturase)에 연결된 적어도 하나의 델타-9 연장효소를 갖는 폴리펩티드를 포함하는 적어도 하나의 멀티자임(multizyme);
- (b) 발현이 하향-조절되는 적어도 하나의 퍼옥시좀 생합성 인자 단백질; 및
- (c) 적어도 라이소포스포티드산 아실트랜스퍼라제["LPAAT"] 활성을 갖는 적어도 2개의 폴리펩티드;
- (d) 적어도 인지질:다이아실글리세롤 아실트랜스퍼라제["PDAT"] 활성을 갖는 적어도 하나의 폴리펩티드를 포함하는 재조합 미생물 숙주 세포.

청구항 5

제4항에 있어서, 적어도 하나의 돌연변이 델타-9 연장효소 폴리펩티드를 추가로 포함하는 재조합 미생물 숙주 세포로서, 상기 돌연변이 델타-9 연장효소 폴리펩티드는 서열 번호 1에 기재된 아미노산 서열을 포함하며, 서열 번호 1은 적어도 하나의 아미노산 돌연변이에 의해 서열 번호 3과 상이하며, 상기 돌연변이(들)가

- i) L35F 돌연변이;
- ii) L35M 돌연변이;
- iii) L35G 돌연변이;
- iv) L35G 돌연변이, 및 S9A, S9D, S9G, S9I, S9K, S9Q, Q12K, A21D, A21T, A21V, V32F, Y84C, Q107E, L108G, G127L, W132T, M143N, M143W, L161T, L161Y, W168G, I179M, I179R, C236N, Q244N, A254W 및 A254Y로 이루어진 군으로부터 선택되는 적어도 하나의 다른 돌연변이;
- v) L35G, A21V, L108G 및 I179R 돌연변이;
- vi) L35G, W132T 및 I179 돌연변이;
- vii) L35G, S9D, Y84C 및 I179R 돌연변이;
- viii) L35G, Y84C, I179R 및 Q244N 돌연변이;
- ix) L35G, A21V, W132T, I179R 및 Q244N 돌연변이;
- x) K58R 및 I257T 돌연변이;
- xi) D98G 돌연변이;
- xii) L130M 및 V243A 돌연변이; 및
- xiii) K58R, L35F, L35G, L35M, S9A, S9D, S9G, S9I, S9K, S9Q, Q12K, A21D, A21T, A21V, V32F, Y84C, D98G, Q107E, L108G, G127L, L130M, W132T, M143N, M143W, L161T, L161Y, W168G, I179M, I179R, C236N, V243A, Q244N, A254W, A254Y 및 I257T로 이루어진 군으로부터 선택되는 적어도 2개의 돌연변이를 포함하는 임의의 조합

으로 이루어진 군으로부터 선택되는 재조합 미생물 숙주 세포.

청구항 6

제5항에 있어서, 적어도 하나의 돌연변이 델타-9 연장효소 폴리펩티드가 L35G 치환을 포함하며, 돌연변이 델타-9 연장효소 폴리펩티드가 서열 번호 3의 델타-9 연장효소 활성과 비교하여 델타-9 연장효소 활성이 향상된 재조합 미생물 숙주 세포.

청구항 7

제4항에 있어서, 멀티자임이

(a) 서열 번호 4, 서열 번호 5, 서열 번호 6, 서열 번호 7, 서열 번호 8, 서열 번호 9 및 서열 번호 10으로 이루어진 군으로부터 선택되는 링커(linker); 및

(b) 서열 번호 12, 서열 번호 14 및 서열 번호 16으로 이루어진 군으로부터 선택되는 서열로 본질적으로 이루어진 서열로 이루어진 군으로부터 선택되는 특성을 갖는 재조합 미생물 숙주 세포.

청구항 8

제4항에 있어서, 적어도 2개의 라이소포스파티드산 아실트랜스퍼라제가

(a) 서열 번호 18, 서열 번호 20, 서열 번호 22, 서열 번호 23 및 서열 번호 25로 이루어진 군으로부터 선택되는 서열로 본질적으로 이루어진 서열; 및

(b) 클러스털 더블유(Clustal W) 정렬 방법에 기초하여, 서열 번호 18, 서열 번호 22 및 서열 번호 23으로 이루어진 군으로부터 선택되는 아미노산 서열과 비교하여, 적어도 43.9%의 아미노산 동일성을 가지며, 서열 번호 26 및 서열 번호 27로 이루어진 군으로부터 선택되는 적어도 하나의 1-아실-sn-글리세롤-3-포스페이트 아실트랜스퍼라제 패밀리 모티프를 추가로 포함하는 폴리펩티드로 이루어진 군으로부터 선택되는 재조합 미생물 숙주 세포.

청구항 9

제4항에 있어서, 적어도 하나의 인지질:다이아실글리세롤 아실트랜스퍼라제가

(a) 서열 번호 29 및 서열 번호 30으로 이루어진 군으로부터 선택되는 서열로 본질적으로 이루어진 서열; 및

(b) 클러스털 더블유 정렬 방법에 기초하여, 서열 번호 29 및 서열 번호 30으로 이루어진 군으로부터 선택되는 아미노산 서열과 비교하여, 적어도 90%의 아미노산 동일성을 갖는 폴리펩티드로 이루어진 군으로부터 선택되는 재조합 미생물 숙주 세포.

청구항 10

제4항에 있어서, 숙주 세포가 야로위아(*Yarrowia*) 속의 세포인 재조합 미생물 숙주 세포.

청구항 11

제4항 내지 제10항 중 어느 한 항에 있어서, 숙주 세포가 적어도 하나의 돌연변이 델타-5 불포화화효소 폴리펩티드를 추가로 포함하며, 상기 돌연변이 델타-5 불포화화효소 폴리펩티드가 하기로 이루어진 군으로부터 선택되는 재조합 미생물 숙주 세포:

a) 서열 번호 180[HxGx]에 기재된 아미노산 모티프 - 서열 번호 180[HxGx]은 서열 번호 181[HPGG]과 동일하지 않음 - ; 및 서열 번호 182[HxxxH]에 기재된 아미노산 모티프 - 서열 번호 182[HxxxH]는 서열 번호 183[HDAS H]과 동일하지 않음 - 를 포함하는 돌연변이 폴리펩티드; 및

b) 서열 번호 106[EgD5M 또는 코돈-최적화된 EgD5R*-34g158g], 서열 번호 108[EgD5R*-34g158g347s], 서열 번호 110[EgD5S-36s157g], 서열 번호 112[EaD5S-35a158g], 서열 번호 299[EgD5R*-34g157g], 서열 번호 301[EgD5R*-34g158a], 서열 번호 303[EgD5R*-34g158g], 서열 번호 329[EgD5S-36s156e], 서열 번호 331[EgD5S-36s158a], 서열 번호 333[EgD5S-36s158g], 서열 번호 363[EaD5S-35a158s] 및 서열 번호 365[EaD5S-35a159g]로 이루어진 군으로부터 선택되는 아미노산 서열을 갖는 돌연변이 폴리펩티드.

청구항 12

- a) 제1항 내지 제10항 중 어느 한 항의 숙주 세포를 배양하는 단계로서, 에이코사펜타엔산을 포함하는 미생물 오일이 생성되는 단계; 및
- b) 임의로, 단계 (a)의 미생물 오일을 회수하는 단계를 포함하는 에이코사펜타엔산을 포함하는 미생물 오일의 제조 방법.

청구항 13

제12항에 있어서, 단계 (b)의 회수된 오일이 추가로 가공되는 방법.

명세서

기술 분야

[0001] 본 출원은 미국 가출원 제61/377,248호(출원일: 2010년 8월 26일), 미국 가출원 제61/428,277호(출원일: 2010년 12월 30일) 및 미국 가출원 제61/479,921호(출원일: 2011년 4월 28일)의 이익을 주장하며, 이들 각각은 본 명세서에 그 전문이 참고로 포함된다.

[0002] 본 발명은 생명공학의 기술 분야이다. 더욱 구체적으로, 본 발명은 고농도의 에이코사펜타엔산, 즉 오메가-3 다중불포화 지방산["PUFA"]을 효율적으로 생성할 수 있는 엔지니어링된(engineered) 재조합 미생물 숙주 세포에 관한 것이다.

배경 기술

[0003] 에이코사펜타엔산["EPA"; *cis*-5, 8, 11, 14, 17-에이코사펜타엔산; 오메가-3]의 임상적 가치와 약제학적인 가치는 널리 공지되어 있다(미국 특허 출원 공개 제2009-0093543-A1호 참조). 유사하게, 천연의 미생물 공급원으로부터 또는 어유 및 해양 플랑크톤으로부터의 분리를 통하여 EPA를 생성하는 것과 대조적으로, 재조합 수단을 사용하여 미생물에서 EPA를 생성하는 것의 이점도 또한 널리 인지되어 있다.

[0004] 비록 EPA 생성을 담당하는 오메가-3/오메가-6 다중불포화 지방산["PUFA"] 생합성 경로의 다양한 부분을 식물 및 비-유지성(non-oleaginous) 효모 내로 도입하는 많은 예가 문헌에 최근에 보고되었지만, 출원인의 양수인에 의한 상당한 노력이 유지성 효모, 즉 야로위아 리폴리티카(*Yarrowia lipolytica*)의 사용에 집중되어 있다(미국 특허 제7,238,482호; 미국 특허 제7,932,077호; 미국 특허 출원 공개 제2009-0093543-A1호; 미국 특허 출원 공개 제2010-0317072-A1호). 유지성 효모는 천연적으로 오일을 합성하고 축적할 수 있는 효모(여기서, 오일 축적은 세포 건조 중량의 적어도 25%이다), 또는 그들이 오일을 합성하고 축적할 수 있도록 유전적으로 엔지니어링된 효모(여기서, 오일 축적은 세포 건조 중량의 적어도 25%이다)로 정의된다.

[0005] 더욱 구체적으로, 미국 특허 제7,932,077호에는 하기의 유전자의 발현에 의하여 감마-리놀렌산["GLA"; 오메가-6]의 공동-합성 없이 재조합 야로위아 리폴리티카 균주에서 전체 지방산["TFA"]의 9%의 EPA가 생성되는 것이 입증되어 있다: 델타-9 연장효소(elongase), 델타-8 불포화화효소(desaturase), 델타-5 불포화화효소, 델타-17 불포화화효소, 델타-12 불포화화효소 및 C_{16/18} 연장효소.

[0006] 미국 특허 출원 공개 제2009-0093543-A1호에는 하기의 유전자의 발현에 의하여 재조합 야로위아 리폴리티카 균주에서 TFA의 최대 55.6%의 EPA를 생성하는 최적화된 재조합 야로위아 리폴리티카 균주가 기재되어 있다: 델타-9 연장효소, 델타-8 불포화화효소, 델타-5 불포화화효소, 델타-17 불포화화효소, 델타-12 불포화화효소, C_{16/18} 연장효소 및 다이아실글리세롤 콜린포스포트랜스퍼라제.

[0007] 미국 특허 출원 공개 제2010-0317072-A1호에는 TFA의 최대 50%의 EPA를 포함하고, TFA의 중량 백분율로 측정된 EPA 대 TFA의 중량 백분율로 측정된 리놀렌산의 비가 적어도 3.1인 미생물 오일을 생성하는 최적화된 재조합 야로위아 리폴리티카 균주가 추가로 기재되어 있다. 미국 특허 출원 공개 제2009-0093543-A1호에 상세히 기재된 오메가-3/오메가-6 지방산 생합성 경로의 유전자의 발현에 더하여, 이들 향상된 균주는 다음에 의해 구별된다:

[0008] 1) 적어도 하나의 지방산 델타-8 불포화화효소에 연결된 적어도 하나의 지방산 델타-9 연장효소를 갖는 폴리펩티드["DGLA 신타제(synthase)"]를 포함하는 적어도 하나의 멀티자임(multizyme)의 포함;

[0009] 2) 말로닐 CoA 신테타제(synthetase) 또는 아실-CoA 라이소인지질 아실트랜스퍼라제["LPLAT"]로 이루어진 군으

로부터 선택되는 효소를 암호화하는 적어도 하나의 폴리뉴클레오티드의 임의의 포함; 및

[0010] 3) 발현이 하향 조절되는 적어도 하나의 퍼옥시좀 생합성 인자 단백질의 포함.

[0011] 상술된 개시내용에도 불구하고, 균주 개선은 높은 전체 지질 함량(즉, 높은 EPA 생산성)에 더하여, 전체 지방산의 중량 백분율로서 고농도의 EPA의 생성을 가능하게 하면서, 최종 오일 생성물 중의 중간체 지방산, 예컨대 리놀레산["LA"; 오메가-6] 및 부산물 지방산의 생성을 최소화시킬 EPA의 상업적 생성에 필요하다. 출원인은 향상된 최적화 야로위아 리폴리티카 균주를 엔지니어링시킴으로써 언급된 문제를 해결하였으며, 여기서, 향상은 건조 세포 중량의 중량 백분율로서 측정하여 적어도 25 중량%의 EPA를 포함하는 미생물 오일의 생성을 가능하게 하는 것이다.

발명의 내용

[0012] 제1 실시형태에서, 본 발명은 건조 세포 중량의 중량 백분율로 측정하여 적어도 25 중량%의 에이코사펜타엔산을 포함하는 오일을 생성하는 재조합 미생물 숙주 세포에 관한 것이다.

[0013] 제2 실시형태에서, 전체 지방산의 중량 백분율로 측정하여 적어도 45 중량%의 에이코사펜타엔산을 포함하는 오일이 본 명세서에 개시된다.

[0014] 바람직하게는 상기 오일 중 어느 하나는 전체 지방산의 중량 백분율로 측정한 에이코사펜타엔산 대 전체 지방산의 중량 백분율로 측정한 리놀레산의 비가 적어도 2.4이다.

[0015] 제3 실시형태에서,

[0016] (a) 적어도 하나의 델타-8 불포화효소에 연결된 적어도 하나의 델타-9 연장효소를 갖는 폴리펩티드를 포함하는 적어도 하나의 멀티자임;

[0017] (b) 발현이 하향 조절되는 적어도 하나의 퍼옥시좀 생합성 인자 단백질; 및

[0018] (c) 적어도 라이소포스파티드산 아실트랜스퍼라제["LPAAT"] 활성을 갖는 적어도 2개의 폴리펩티드;

[0019] (d) 적어도 인지질:다이아실글리세롤 아실트랜스퍼라제["PDAT"] 활성을 갖는 적어도 하나의 폴리펩티드를 포함하는 재조합 미생물 숙주 세포가 본 명세서에 개시된다.

[0020] 제4 실시형태에서, 재조합 미생물 숙주 세포는 적어도 하나의 돌연변이 델타-9 연장효소 폴리펩티드를 추가로 포함할 수 있으며, 여기서, 상기 돌연변이 델타-9 연장효소 폴리펩티드는 서열 번호 1에 기재된 아미노산 서열을 포함하며, 서열 번호 1은 적어도 하나의 아미노산 돌연변이에 의해 서열 번호 3과 상이하며, 상기 돌연변이(들)은

[0021] i) L35F 돌연변이;

[0022] ii) L35M 돌연변이;

[0023] iii) L35G 돌연변이;

[0024] iv) L35G 돌연변이, 및 S9A, S9D, S9G, S9I, S9K, S9Q, Q12K, A21D, A21T, A21V, V32F, Y84C, Q107E, L108G, G127L, W132T, M143N, M143W, L161T, L161Y, W168G, I179M, I179R, C236N, Q244N, A254W 및 A254Y로 이루어진 군으로부터 선택되는 적어도 하나의 다른 돌연변이;

[0025] v) L35G, A21V, L108G 및 I179R 돌연변이;

[0026] vi) L35G, W132T 및 I179 돌연변이;

[0027] vii) L35G, S9D, Y84C 및 I179R 돌연변이;

[0028] viii) L35G, Y84C, I179R 및 Q244N 돌연변이;

[0029] ix) L35G, A21V, W132T, I179R 및 Q244N 돌연변이;

[0030] x) K58R 및 I257T 돌연변이;

[0031] xi) D98G 돌연변이;

[0032] xii) L130M 및 V243A 돌연변이; 및

- [0033] xiii) K58R, L35F, L35G, L35M, S9A, S9D, S9G, S9I, S9K, S9Q, Q12K, A21D, A21T, A21V, V32F, Y84C, D98G, Q107E, L108G, G127L, L130M, W132T, M143N, M143W, L161T, L161Y, W168G, I179M, I179R, C236N, V243A, Q244N, A254W, A254Y 및 I257T로 이루어진 군으로부터 선택되는 적어도 2개의 돌연변이를 포함하는 임의의 조합으로 이루어진 군으로부터 선택된다.
- [0034] 바람직하게는, 적어도 하나의 돌연변이 델타-9 연장효소 폴리펩티드는 L35G 치환을 포함하며, 돌연변이 델타-9 연장효소 폴리펩티드는 서열 번호 3의 델타-9 연장효소 활성과 비교하여 델타-9 연장효소 활성이 향상된다.
- [0035] 바람직하게는, 적어도 하나의 멀티자임은
- [0036] (a) 서열 번호 4, 서열 번호 5, 서열 번호 6, 서열 번호 7, 서열 번호 8, 서열 번호 9 및 서열 번호 10으로 이루어진 군으로부터 선택되는 링커; 및
- [0037] (b) 서열 번호 12, 서열 번호 14 및 서열 번호 16으로 이루어진 군으로부터 선택되는 서열로 본질적으로 이루어진 아미노산 서열로 이루어진 군으로부터 선택되는 특성을 갖는다.
- [0038] 바람직하게는, 적어도 2개의 라이소포스파티드산 아실트랜스퍼라제 중 적어도 하나는
- [0039] (a) 서열 번호 18, 서열 번호 20, 서열 번호 22, 서열 번호 23 및 서열 번호 25로 이루어진 군으로부터 선택되는 서열로 본질적으로 이루어진 아미노산 서열, 및
- [0040] (b) 클러스털 더블유(Clustal W) 정렬 방법에 기초하여, 서열 번호 18, 서열 번호 22, 서열 번호 23으로 이루어진 군으로부터 선택되는 아미노산 서열과 비교하여, 적어도 43.9%의 아미노산 동일성을 가지며, 서열 번호 26 및 서열 번호 27로 이루어진 군으로부터 선택되는 적어도 하나의 1-아실-*sn*-글리세롤-3-포스페이트 아실트랜스퍼라제 패밀리 모티프를 추가로 포함하는 폴리펩티드로 이루어진 군으로부터 선택된다.
- [0041] 바람직하게는, 적어도 하나의 인지질:다이아실글리세롤 아실트랜스퍼라제는
- [0042] (a) 서열 번호 29 및 서열 번호 30으로 이루어진 군으로부터 선택되는 서열로 본질적으로 이루어진 아미노산 서열; 및
- [0043] (b) 클러스털 더블유 정렬 방법에 기초하여, 서열 번호 29 및 서열 번호 30으로 이루어진 군으로부터 선택되는 아미노산 서열과 비교하여, 적어도 90%의 아미노산 동일성을 갖는 폴리펩티드로 이루어진 군으로부터 선택된다.
- [0044] 바람직하게는, 숙주 세포는 야로위아 속의 세포이다.
- [0045] 제5 실시형태에서, 본 발명은
- [0046] a) 본 발명의 임의의 숙주 세포를 배양하는 단계로서, 에이코사펜타엔산을 포함하는 미생물 오일이 생성되는 단계; 및
- [0047] b) 임의로, 단계 (a)의 미생물 오일을 회수하는 단계를 포함하는 에이코사펜타엔산을 포함하는 미생물 오일의 제조 방법에 관한 것이다.
- [0048] 제6 실시형태에서, 본 발명은 본 발명의 방법에 의해 생성되는 오일의 추가의 가공에 관한 것이다.
- [0049] 생물학적 기탁
- [0050] 하기의 생물학적 물질은 미국 버지니아주 20110-2209, 머내서스 볼리바드 대학 10801 아메리칸 타입 컬처 콜렉션(ATCC)에 기탁되어 있으며, 하기에 명칭, 수탁 번호 및 기탁일이 있다.
- | 생물학적 물질 | 수탁 번호 | 기탁일 |
|------------------|----------------|--------------|
| 야로위아 리폴리티카 Y8412 | ATCC PTA-10026 | 2009년 5월 14일 |
- [0051]
- [0052] 상술된 생물학적 물질은 특허 절차 상의 미생물 기탁의 국제적 승인에 관한 부다페스트 조약 하에 기탁되었다. 열거된 기탁물은 적어도 30년간 표시된 국제기탁기관에 유지될 것이며, 그것을 개시한 특허의 허여시 공중에게 이용가능하게 될 것이다. 기탁물의 입수가능성은 정부 조치에 의해 허여된 특허권을 침해하는 방식으로 대상 발명이 실시되는 허가를 이루는 것은 아니다.
- [0053] 야로위아 리폴리티카 Y9502는 미국 특허 출원 공개 제2010-0317072-A1호에 기재된 방법에 따라, 야로위아 리폴리티카 Y8412로부터 유래된 것이다.
- [0054] 도면의 간단한 설명 및 서열 설명

- [0055] <도 1a 및 도 1b>
- [0056] 도 1a 및 도 1b는 오메가-3/오메가-6 지방산 생합성 경로를 예시한 것이며, 하기에서 이러한 경로의 설명을 고려하는 경우 함께 감안해야 한다.
- [0057] <도 2>
- [0058] 도 2는 야로위아 리폴리티카 균주 Y4305의 발효 과정 동안의 EPA % TFA와 LA % TFA 간의 관계를 그래프로 보여 준다(미국 특허 출원 공개 제2009-0093543-A1호).
- [0059] <도 3a 및 도 3b>
- [0060] 도 3a 및 도 3b는 야로위아 리폴리티카 ATCC #20362 유래의 다양한 야로위아 리폴리티카 균주의 개발을 도시한다.
- [0061] <도 4>
- [0062] 도 4는 (a) pZKUM; 및 (b) pZKL3-9DP9N의 플라스미드 맵을 제공한다.
- [0063] <도 5>
- [0064] 도 5는 (a) pY187; 및 (b) pZK16-ML8N의 플라스미드 맵을 제공한다.
- [0065] <도 6>
- [0066] 도 6은 (a) pZK16-MyL8N; 및 (b) pZK16-ML3의 플라스미드 맵을 제공한다.
- [0067] <도 7>
- [0068] 도 7은 (a) pZKMP-mL9DP; 및 (b) pZKMP-mL9DCB의 플라스미드 맵을 제공한다.
- [0069] <도 8>
- [0070] 도 8은 (a) pZKSL-5S5A5; 및 (b) pZP2-85m98F의 플라스미드 맵을 제공한다.
- [0071] <도 9>
- [0072] 도 9는 균주 Z5567 유래의 다양한 야로위아 리폴리티카 균주의 개발을 예시한 것이다.
- [0073] <도 10a>
- [0074] 도 10a는 pYPS234와의 상동성 재조합 반응을 개략적으로 나타내는 한편, 도 10b는 pYPS234의 플라스미드 맵을 제공한다.
- [0075] <도 11a>
- [0076] 도 11a는 pYPS233과의 상동성 재조합 반응을 개략적으로 나타내는 한편, 도 11b는 pYPS233의 플라스미드 맵을 제공한다.
- [0077] <도 12a>
- [0078] 도 12a는 pYPS241과의 상동성 재조합 반응을 개략적으로 나타내는 한편, 도 12b는 pYPS241의 플라스미드 맵을 제공한다.
- [0079] <도 13>
- [0080] 도 13은 (a) pZR5AU-555; 및 (b) pZR5AU-555M의 플라스미드 맵을 제공한다.
- [0081] <도 14>
- [0082] 도 14는 pZUFmEgD9ES의 플라스미드 맵이다.
- [0083] <도 15a, 15b, 15c, 15d, 15e, 15f, 15g 및 15h>
- [0084] 도 15a, 15b, 15c, 15d, 15e, 15f, 15g 및 15h는 클러스털 더블유 정렬 방법을 사용한 키오나 인테스티날리스(*Ciona intestinalis*)[서열 번호 133], 온코린치스 마이키스(*Oncorhynchus mykiss*)[서열 번호 134], 마르찬티아 폴리모르파(*Marchantia polymorpha*)[서열 번호 135], 피스코미트렐라 파텐스(*Physcomitrella patens*)[서열

번호 136], 마르찬티아 폴리모르파[서열 번호 137], 오스트레오코커스 타우리(*Ostreococcus tauri*)[서열 번호 138], 파블로바(*Pavlova*) 중 CCMP459[서열 번호 139], 파블로바 살리나(*Pavlova salina*)[서열 번호 140], 오스트레오코커스 타우리[서열 번호 141], 유글레나 아나바에나(*Euglena anabaena*)[서열 번호 34], 유글레나 그라실리스(*Euglena gracilis*)[서열 번호 32], 유트렙티엘라(*Eutreptiella*) 중 CCMP389[서열 번호 38], 아이소크라이시스 갈바나(*Isochrysis galbana*)[서열 번호 42], 탈라시오시라 슈도나나(*Thalassiosira pseudonana*)[서열 번호 142], 탈라시오시라 슈도나나[서열 번호 143], 모르티에렐라 알피나(*Mortierella alpina*)[서열 번호 144] 및 트라우스토키트리움(*Thraustochytrium*) 중 FJN-10[서열 번호 145]으로부터의 17개 지방산 연장효소의 정렬이다.

[0085] <도 16a>

[0086] 도 16a는 EgD9eS의 막 토폴로지(topology) 모델을 보여주며; 각각의 수직의 원통은 막-스패닝 세그먼트를 나타내는 한편 각각의 수평의 원통은 내막 리플릿(leaflet) 내에 또는 그 근처에 놓여 있는 소수성 스트레치(stretch)를 나타낸다.

[0087] <도 16b>

[0088] 도 16b는 L35F 돌연변이; L35M 돌연변이; L35G 돌연변이; L35G 돌연변이, 및 S9A, S9D, S9G, S9I, S9K, S9Q, Q12K, A21D, A21T, A21V, V32F, Y84C, Q107E, L108G, G127L, W132T, M143N, M143W, L161T, L161Y, W168G, I179M, I179R, C236N, Q244N, A254W 및 A254Y로 이루어진 군으로부터 선택되는 적어도 하나의 다른 돌연변이; L35G, A21V, L108G 및 I179R 돌연변이; L35G, W132T 및 I179R 돌연변이; L35G, S9D, Y84C 및 I179R 돌연변이; L35G, Y84C, I179R 및 Q244N 돌연변이; L35G, A21V, W132T, I179R 및 Q244N 돌연변이; K58R 및 I257T 돌연변이; D98G 돌연변이; L130M 및 V243A 돌연변이; 및 K58R, L35F, L35G, L35M, S9A, S9D, S9G, S9I, S9K, S9Q, Q12K, A21D, A21T, A21V, V32F, Y84C, D98G, Q107E, L108G, G127L, L130M, W132T, M143N, M143W, L161T, L161Y, W168G, I179M, I179R, C236N, V243A, Q244N, A254W, A254Y 및 I257T로 이루어진 군으로부터 선택되는 적어도 2개의 돌연변이를 포함하는 임의의 조합을 임의로 포함하는 유글레나 그라실리스 유래의 합성 돌연변이 델타-9 연장효소의 도식을 보여준다(즉, "EgD9eS-돌연변이체 공통서열(consensus)"; 서열 번호 1).

[0089] <도 17>

[0090] 도 17은 벡터 엔티아이(Vector NTI)(등록 상표)의 얼라인엑스(AlignX) 프로그램(미국 캘리포니아주 칼스배드 소재의 인비트로젠 코포레이션(Invitrogen Corporation))을 사용한 아이소크라이시스 갈바나의 델타-9 연장효소["IgD9e"] (서열 번호 42), 유트렙티엘라 중 CCMP389의 델타-9 연장효소["E389D9e"] (서열 번호 38), 유글레나 그라실리스의 델타-9 연장효소["EgD9e"] (서열 번호 32) 및 유글레나 아나바에나의 델타-9 연장효소["EaD9e"] (서열 번호 34)의 정렬이다.

[0091] <도 18>

[0092] 도 18은 유글레나 그라실리스 델타-5 불포화화효소의 예측된 토폴로지 모델이다.

[0093] <도 19a 및 19b>

[0094] 도 19a 및 19b는 유글레나 그라실리스로부터의 야생형 델타-5 불포화화효소 유전자(즉, EgD5; 서열 번호 184)와, S347R 돌연변이를 함유하는 변이체 야생형 유글레나 그라실리스 델타-5 불포화화효소 유전자(즉, EgD5R; 서열 번호 192)의 DNA 서열의 정렬을 보여준다.

[0095] <도 20a, 20b 및 20c>

[0096] 도 20a, 20b 및 20c는 플라스미드 pDMW367-M4의 구축을 예시한 것이다.

[0097] <도 21>

[0098] 도 21은 유글레나 그라실리스로부터의 야생형 델타-5 불포화화효소 유전자(즉, EgD5R; 서열 번호 192)의 5' 부분과, 야로위아 리폴리티카 코돈-최적화된 델타-5 불포화화효소 돌연변이 유전자(즉, EgD5M; 서열 번호 105)의 처음 204 bp의 서열 정렬을 보여준다.

[0099] <도 22>

[0100] 도 22는 (a) pEgD5M; 및 (b) pDMW367-5M에 대한 플라스미드 맵을 제공한다.

[0101] 본 발명은 하기의 상세한 설명 및 첨부된 서열 설명으로부터 더욱 완전하게 이해될 수 있으며, 이들은 본 출원

의 일부를 형성한다.

[0102] 하기의 서열들은 37 C.F.R. § 1.821-1.825("뉴클레오타이드 서열 및/또는 아미노산 서열 개시를 포함하는 특허출원에 관한 요건 - 서열 규정")를 따르며 세계지적재산권기구(World Intellectual Property Organization, WIPO) 표준 ST.25(1998), 및 EPO 및 PCT의 서열 목록 요건(규정 5.2 및 49.5(a-bis), 및 시행세칙의 섹션 208 및 부칙 C)에 부합한다. 뉴클레오타이드 및 아미노산 서열 데이터에 사용된 기호 및 체계는 37 C.F.R. § 1.822에 제시된 규정을 준수한다.

[0103] 서열 번호 1 내지 437은 표 1에 나타난 바와 같이, 프로모터, 유전자 또는 단백질(또는 그의 단편)을 암호화하는 ORF, 프라이머 또는 플라스미드이다.

표 1

유전자의 요약 및 단백질 서열 번호

설명	핵산 서열 번호	단백질 서열 번호
L35F 돌연변이; L35M 돌연변이; L35G 돌연변이; L35G 돌연변이, 및 S9A, S9D, S9G, S9I, S9K, S9Q, Q12K, A21D, A21T, A21V, V32F, Y84C, Q107E, L108G, G127L, W132T, M143N, M143W, L161T, L161Y, W168G, I179M, I179R, C236N, Q244N, A254W 및 A254Y 로 이루어진 군으로부터 선택되는 적어도 하나의 다른 돌연변이; L35G, A21V, L108G 및 I179R 돌연변이; L35G, W132T 및 I179R 돌연변이; L35G, S9D, Y84C 및 I179R 돌연변이; L35G, Y84C, I179R 및 Q244N 돌연변이; L35G, A21V, W132T, I179R 및 Q244N 돌연변이; K58R 및 I257T 돌연변이; D98G 돌연변이; L130M 및 V243A 돌연변이; 및 K58R, L35F, L35G, L35M, S9A, S9D, S9G, S9I, S9K, S9Q, Q12K, A21D, A21T, A21V, V32F, Y84C, D98G, Q107E, L108G, G127L, L130M, W132T, M143N, M143W, L161T, L161Y, W168G, I179M, I179R, C236N, V243A, Q244N, A254W, A254Y 및 I257T 로 이루어진 군으로부터 선택되는 적어도 2 개의 돌연변이를 포함하는 임의의 조합을 임의로 포함하는 유글레나 그라실리스 유래의 합성 돌연변이 델타-9 연장효소("EgD9eS-돌연변이체 공동서열")	--	1 (258 AA)
야로위아 리폴리티카에서의 발현을 위해 코돈-최적화된 유글레나 그라실리스 유래의 합성 델타-9 연장효소("EgD9eS")	2 (777 bp)	3 (258 AA)
멀티자임 링커 GAGPARPAGLPPATYYDSLAVMGS	--	4
멀티자임 링커 GPARPAGLPPATYYDSLAV	--	5
멀티자임 링커 PARPAGLPPATYYDSLAV	--	6
멀티자임 링커 PTRPAGPPPATYYDSLAV	--	7
멀티자임 링커 PGGPGKPSEIASLPPPIRPVGNPPAAYYDALAT	--	8
멀티자임 링커 PARPAGLPPATYYDSLAVSGRT	--	9
멀티자임 링커 PGGPGKPSEIASLPPPIRPVGNPPAAYYDALATGRT	--	10
EgD9eS/EgD8M 유전자 융합을 포함하는 DGLA 신타제	11 (2112 bp)	12 (703 AA)
EaD9eS/EaD8S 유전자 융합을 포함하는 DGLA 신타제	13 (2109 bp)	14 (702 AA)
E389D9eS/EgD8M 유전자 융합을 포함하는 DGLA 신타제	15 (2127 bp)	16 (708 AA)
모르티에렐라 알피나 라이소포스포티드산 아실트랜스퍼라제("MaLPAAT1")	17 (945 bp)	18 (314 AA)
야로위아 리폴리티카에서의 발현을 위해 코돈-최적화된 모르티에렐라 알피나 유래의 합성 LPAAT1("MaLPAAT1S")	19 (955 bp)	20 (314 AA)
야로위아 리폴리티카 라이소포스포티드산 아실트랜스퍼라제("YLPAAT1")	21 (849 bp)	22 (282 AA)
사카로마이세스 세레비지애(<i>Saccharomyces cerevisiae</i>) 라이소포스포티드산 아실트랜스퍼라제("ScLPAAT"; 또한 ORF "YDL052C"; 진뱅크(GenBank) 수탁 번호 NP_010231)	--	23 (303 AA)

[0104]

야로위아 리폴리티카에서의 발현을 위해 코돈-최적화된 사카로마이세스 세레비지에 유래의 합성 LPAAT("ScLPAATS")	24 (926 bp)	25 (303 AA)
1-아실- <i>sn</i> -글리세롤-3-포스페이트 아실트랜스퍼라제 모티프 NHxxxxD	--	26
1-아실- <i>sn</i> -글리세롤-3-포스페이트 아실트랜스퍼라제 모티프 EGTR	--	27
야로위아 리폴리티카 인지질:다이아실글리세롤 아실트랜스퍼라제("YIPDAT")	28 (1947 bp)	29 (648 AA)
사카로마이세스 세레비지에 PDAT(진뱅크 수탁 번호 P40345)	--	30 (661 AA)
유글레나 그라실리스 델타-9 연장효소("EgD9e")	31 (777 bp)	32 (258 AA)
유글레나 아나바에나 델타-9 연장효소("EaD9e")	33 (774 bp)	34 (258 AA)
야로위아 리폴리티카에서의 발현을 위해 코돈-최적화된 유글레나 아나바에나 유래의 합성 델타-9 연장효소("EaD9eS")	35 (774 bp)	36 (258 AA)
유타랩티엘라 종 CCMP389 델타-9 연장효소("E389D9e")	37 (792 bp)	38 (263 AA)
야로위아 리폴리티카에서의 발현을 위해 코돈-최적화된 유타랩티엘라 종 CCMP389 델타-9 연장효소 유래의 합성 델타-9 연장효소("E389D9eS")	39 (792 bp)	40 (263 AA)
아이소크라이시스 갈바나 델타-9 연장효소("IgD9e")	41 (1064 bp)	42 (263 AA)
유글레나 그라실리스 유래의 합성 돌연변이 델타-9 연장효소("EgD9eS-L35G")	43 (777 bp)	44 (258 AA)
야로위아 리폴리티카 콜린포스페이트 시티딜릴트랜스퍼라제 유전자("YIPCT")	45 (1101 bp)	46 (366 AA)
야로위아 리폴리티카 다이아실글리세롤 콜린포스포트랜스퍼라제 유전자("YICPT1")	47 (1185 bp)	48 (394 AA)
야로위아 리폴리티카에서의 발현을 위해 코돈-최적화된 리조비움 레구미노사룸(<i>Rhizobium leguminosarum</i>) 생물변이형 비시아(<i>bv. viciae</i>) 3841(진뱅크 수탁 번호 YP_766603) 유래의 합성 말로닐-CoA 신테타제("MCS")	49 (1518 bp)	50 (505 AA)
야로위아 리폴리티카 Pex1p(진뱅크 수탁 번호 CAG82178)	--	51 (1024 AA)
야로위아 리폴리티카 Pex2p(진뱅크 수탁 번호 CAG77647)	--	52 (381 AA)
야로위아 리폴리티카 Pex3p(진뱅크 수탁 번호 CAG78565)	--	53 (431 AA)
야로위아 리폴리티카 Pex3Bp(진뱅크 수탁 번호 CAG83356)	--	54 (395 AA)
야로위아 리폴리티카 Pex4p(진뱅크 수탁 번호 CAG79130)	--	55 (153 AA)
야로위아 리폴리티카 Pex5p(진뱅크 수탁 번호 CAG78803)	--	56 (598 AA)
야로위아 리폴리티카 Pex6p(진뱅크 수탁 번호 CAG82306)	--	57 (1024 AA)
야로위아 리폴리티카 Pex7p(진뱅크 수탁 번호 CAG78389)	--	58 (356 AA)
야로위아 리폴리티카 Pex8p(진뱅크 수탁 번호 CAG80447)	--	59

[0105]

		(671 AA)
야로위아 리폴리티카 Pex10p(진뱅크 수탁 번호 CAG81606)	--	60 (377 AA)
야로위아 리폴리티카 Pex12p(진뱅크 수탁 번호 CAG81532)	--	61 (408 AA)
야로위아 리폴리티카 Pex13p(진뱅크 수탁 번호 CAG81789)	--	62 (412 AA)
야로위아 리폴리티카 Pex14p(진뱅크 수탁 번호 CAG79323)	--	63 (380 AA)
야로위아 리폴리티카 Pex16p(진뱅크 수탁 번호 CAG79622)	--	64 (391 AA)
야로위아 리폴리티카 Pex17p(진뱅크 수탁 번호 CAG84025)	--	65 (225 AA)
야로위아 리폴리티카 Pex19p(진뱅크 수탁 번호 AAK84827)	--	66 (324 AA)
야로위아 리폴리티카 Pex20p(진뱅크 수탁 번호 CAG79226)	--	67 (417 AA)
야로위아 리폴리티카 Pex22p(진뱅크 수탁 번호 CAG77876)	--	68 (195 AA)
야로위아 리폴리티카 Pex26p(진뱅크 수탁 번호 NC_006072, 뉴클레오타이드 117230-118387의 안티센스 번역)	--	69 (386 AA)
야로위아 종에서 최적으로 발현되는 유전자에 대한 코돈-최적화된 번역 개시 부위	70 (10 bp)	--
His-풍부 모티프: Q(X) ₂ HH	--	71
His-풍부 모티프: H(X) ₂ HH	--	72
델타-9 연장효소 모티프 #1: Y-N-X-(L 또는 F)-X ₁ -S-X ₂ -S-F	--	73
델타-9 연장효소 모티프 #2: F-Y-X-S-K-X ₂ -(E 또는 D)-Y-X-D-(T 또는 S)-X ₂ -L	--	74
델타-9 연장효소 모티프 #3: L-(Q 또는 H)-X-E-H-H-X-G-A	--	75
델타-9 연장효소 모티프 #4: M-Y-X-Y-Y-X ₂ -(K 또는 R 또는 N)-F	--	76
델타-9 연장효소 모티프 #5: K-X-L-(I 또는 L 또는 M)-T-X ₂ -Q	--	77
델타-9 연장효소 모티프 #6: W-X-E-N-Y-X-Y	--	78
델타-9 연장효소 모티프 #7: Y-X-G-X-Y-X ₂ -L-F	--	79
야로위아 리폴리티카 델타-9 불포화효소 유전자("YID9")	80 (1449 bp)	81 (482 AA)
플라스미드 pZKUM	82 (4313 bp)	--
플라스미드 pZKL3-9DPN9N	83 (13565 bp)	--
플라스미드 pY187	84 (9632 bp)	--
플라스미드 pZK16-ML8N	85 (15262 bp)	--
플라스미드 pZK16-MyL8N	86 (15181 bp)	--
유글레나 그라실리스 유래의("EgD8S") 합성 돌연변이 델타-8 불포화효소("EgD8M")	87 (1272 bp)	88 (422 AA)
플라스미드 pZK16-ML3	89 (15105 bp)	--
야로위아 리폴리티카에서의 발현을 위해 코돈-최적화된 모르티에렐라 알피나 ELO3 유래의 합성 C _{16/18} 연장효소("ME3S")	90 (828 bp)	91 (275 AA)

[0106]

플라스미드 pZKMP-ML9DP	92 (13237 bp)	--
야로위아 리폴리티카 ALK2LM1 프로모터 영역 + N-말단 66 bp CDS	93 (899 bp)	--
야로위아 리폴리티카 DGAT2M 프로모터 영역	94 (722 bp)	--
플라스미드 pZKMP-ML9DCB	95 (13320 bp)	--
플라스미드 pZKSL-5S5A5	96 (13975 bp)	--
플라스미드 pZP2 -85m98F	97 (14619 bp)	--
993 bp 스테퍼(stuffer) 단편	98 (993 bp)	--
플라스미드 pYPS234	99 (7338 bp)	--
1019 bp 스테퍼 단편	100 (1019 bp)	--
플라스미드 pYPS233	101 (7364 bp)	--
플라스미드 pYSP241	102 (9211 bp)	--
야로위아 리폴리티카에서의 발현을 위해 코돈-최적화된 유글레나 아나바에나 UTEX 373 유래의 합성 델타-8 불포화화효소("EaD8S")	103 (1260 bp)	104 (420 AA)
야로위아 리폴리티카에서의 발현을 위해 코돈-최적화된 유글레나 그라실리스 유래의 합성 불연변이 델타-5 불포화화효소("EgD5M" 또는 "EgD5R*-34g158g") (즉, HgGG 및 HDAGH 모티프 포함)	105 (1350 bp)	106 (449 AA)
야로위아 리폴리티카에서의 발현을 위해 코돈-최적화된 유글레나 그라실리스 유래의 합성 불연변이 델타-5 불포화화효소("EgD5M1" 또는 "EgD5R*-34g158g347s")(즉, HgGG 및 HDAGH 모티프, 및 아미노산 위치 347에서 Ser 잔기 포함)	107 (1350 bp)	108 (449 AA)
유글레나 그라실리스 유래의 합성 불연변이 델타-5 불포화화효소("EgD5S-36s157g")(즉, HPGs 및 HDgSH 모티프 포함)	109 (1350 bp)	110 (449 AA)
유글레나 아나바에나 UTEX 373 유래의 합성 불연변이 델타-5 불포화화효소("EaD5S-35a158g")(즉, HaGG 및 HDgSH 모티프 포함)	111 (1365 bp)	112 (454 AA)
플라스미드 pZR5AU -555	113 (13926 bp)	--
플라스미드 pZR5AU -555M	114 (13926 bp)	--
플라스미드 pZUFmEgD9ES	115 (7769 bp)	--
프라이머 pZUFm_6980_012208f	116	--
프라이머 pZUFm_40_012208r	117	--
유글레나 그라실리스 유래의 합성 불연변이 델타-9 연장효소("EgD9eS-L35F")	118 (777 bp)	119 (258 AA)
플라스미드 pZuFmEgD9ES-L35F	120 (7769 bp)	--

[0107]

유글레나 그라실리스 유래의 합성 돌연변이 델타-9 연장효소("EgD9eS-K58R/1257T")	121 (777 bp)	122 (258 AA)
플라스미드 pZuFmEgD9ES-K58R/1257T	123 (7769 bp)	--
유글레나 그라실리스 유래의 합성 돌연변이 델타-9 연장효소("EgD9eS-L130M/V243A ₁ ")	124 (777 bp)	125 (258 AA)
플라스미드 pZuFmEgD9ES-L130M/V243A ₁	126	--
유글레나 그라실리스 유래의 합성 돌연변이 델타-9 연장효소("EgD9eS-D98G")	127 (777 bp)	128 (258 AA)
플라스미드 pZuFmEgD9ES-D98G	129 (7769 bp)	--
유글레나 그라실리스 유래의 합성 돌연변이 델타-9 연장효소("EgD9eS-L130M/V243A ₂ ")	130 (777 bp)	131 (258 AA)
플라스미드 pZuFmEgD9ES-L130M/V243A ₂	132 (7769 bp)	--
키오나 인테스티날리스 연장효소(진뱅크 수탁 번호 AAV67802)	--	133 (289 AA)
온코린처스 마이키스 연장효소(진뱅크 수탁 번호 AAV67803)	--	134 (295 AA)
마르찬티아 폴리모르파 연장효소(진뱅크 수탁 번호 AAT85662)	--	135 (290 AA)
피스코미트렐라 파텍스 연장효소(진뱅크 수탁 번호 AAL84174)	--	136 (290 AA)
마르찬티아 폴리모르파 연장효소(진뱅크 수탁 번호 BAE71130)	--	137 (348 AA)
오스트레오코쿠스 타우리 연장효소(진뱅크 수탁 번호 AAV67797)	--	138 (292 AA)
파블로바 종 CCMP459 연장효소(진뱅크 수탁 번호 AAV33630)	--	139 (277 AA)
파블로바 살리나 연장효소(진뱅크 수탁 번호 AAY15135)(유의어: 레베카 살리나(<i>Rebecca salina</i>))	--	140 (302 AA)
오스트레오코쿠스 타ური 연장효소(진뱅크 수탁 번호 AAV67798)	--	141 (300 AA)
탈라시오시라 슈도나나 연장효소(진뱅크 수탁 번호 AAV67800)	--	142 (358 AA)
탈라시오시라 슈도나나 연장효소(진뱅크 수탁 번호 AAV67799)	--	143 (272 AA)
모르티에렐라 알피나 연장효소(진뱅크 수탁 번호 AAF70417)	--	144 (318 AA)
트라우스토키티리움 종 FJN-10 연장효소(진뱅크 수탁 번호 ABC18314)	--	145 (271 AA)
프라이머 EgD9E 102_053008f	146	--
프라이머 EgD9E 760_053008r	147	--
플라스미드 pZuFmEgD9ES-L35G	148 (7769 bp)	--
유글레나 그라실리스 유래의 합성 돌연변이 델타-9 연장효소("EgD9eS-L35M/Q107E")	149 (777 bp)	150 (258 AA)
플라스미드 pZuFmEgD9ES-L35M/Q107E	151 (7769 bp)	--
유글레나 그라실리스 유래의 합성 돌연변이 델타-9 연장효소("EgD9eS-L35G")	152 (777 bp)	--
유글레나 그라실리스 유래의 합성 돌연변이 델타-9 연장효소("EgD9eS-L35G")	153 (777 bp)	--

[0108]

유글레나 그라실리스 유래의 합성 돌연변이 델타-9 연장효소("EgD9eS-L35G")	154 (777 bp)	--
부위 지정 돌연변이유발에 의해 EgD9eS-L35G를 돌연변이시키는데 사용되는 올리고뉴클레오타이드 프라이머 쌍	155-176	--
유글레나 그라실리스 유래의 합성 돌연변이 델타-9 연장효소("EgD9eS-A21V/L35G/L108G/I179R")	177 (777 bp)	178 (258 AA)
플라스미드 pZuFmEgD9ES-A21V/L35G/L108G/I179R	179 (7769 bp)	--
HxGx 모티브	--	180
HPGG 모티브	--	181
HxxxH 모티브	--	182
HDASH 모티브	--	183
유글레나 그라실리스 델타-5 불포화효소("EgD5")	184 (1350 bp)	185 (449 AA)
His-풍부 모티브: H(X) ₂ H	--	186
His-풍부 모티브: H(X) ₂ H	--	187
His-풍부 모티브: H(X) ₂ HH	--	188
His-풍부 모티브: H(X) ₂ HH	--	189
His-풍부 모티브: (H/Q)(X) ₂ HH	--	190
His-풍부 모티브: (H/Q)(X) ₂ HH	--	191
아미노산 위치 347에서 Arg를 포함하는 변이체 유글레나 그라실리스 델타-5 불포화효소("EgD5R")	192 (1350 bp)	193 (449 AA)
플라스미드 pDMW367	194 (8438 bp)	--
야로위아 리폴리티카에서의 발현을 위해 코돈-최적화된 유글레나 그라실리스 유래의 합성 델타-5 불포화효소("EgD5S")	195 (1350 bp)	196 (449 AA)
4개의 암호화 영역 제한 효소 부위가 제거된, 아미노산 위치 347에서 Arg를 포함하는 변형된 변이체 유글레나 그라실리스 델타-5 불포화효소("EgD5R*")	197 (1350 bp)	198 (449 AA)
플라스미드 pDMW367-M4	199 (8438 bp)	--
EgD5R 암호화 영역의 천연의 <i>EcoRI</i> 및 <i>BglII</i> 제한 효소 부위를 돌연변이시키고, pDMW367-M4를 생성하기 위해 사용되는 올리고뉴클레오타이드 프라이머 쌍	200-203	--
플라스미드 pDMW367-M2	204 (8438 bp)	--
EgD5R 암호화 영역의 천연의 <i>HindIII</i> 및 <i>NcoI</i> 제한 효소 부위를 돌연변이시키고, pDMW367-M4를 생성하기 위해 사용되는 올리고뉴클레오타이드 프라이머 쌍	205-208	--
부위-지정 돌연변이유발에 의해 EgD5R*의 HDASH 모티브의 Ala 잔기를 개별적으로 돌연변이시키기 위하여 사용되는 올리고뉴클레오타이드 프라이머	209-246	--
유글레나 그라실리스 유래의 합성 돌연변이 델타-5 불포화효소("EgD5R*-157g")(즉, HDgSH 모티브 포함)	--	247 (449 AA)
유글레나 그라실리스 유래의 합성 돌연변이 델타-5 불포화효소("EgD5R*-157s")(즉, HDsSH 모티브 포함)	--	248 (449 AA)
부위-지정 돌연변이유발에 의해 EgD5R*의 HDASH 모티브의 Ser 잔기를 개별적으로 돌연변이시키기 위해	249-286	--

[0109]

사용되는 올리고뉴클레오타이드 프라이머		
유글레나 그라실리스 유래의 합성 돌연변이 델타-5 불포화화효소("EgD5R*-158a")(즉, HDAaH 모티브 포함)	--	287 (449 AA)
유글레나 그라실리스 유래의 합성 돌연변이 델타-5 불포화화효소("EgD5R*-158g")(즉, HDAgH 모티브 포함)	--	288 (449 AA)
플라스미드 pDMW367M4 -157g	289 (8438 bp)	--
플라스미드 pDMW367M4 -158a	290 (8438 bp)	--
플라스미드 pDMW367M4 -158g	291 (8438 bp)	--
부위-지정 돌연변이유발에 의해 EgD5R*-157g, EgD5R*-158a 및 EgD5R*-158g의 HPGG 모티브의 Pro 잔기 또는 두번째 Gly 잔기 중 어느 하나를 개별적으로 돌연변이시키기 위해 사용되는 올리고뉴클레오타이드 프라이머	292-297	--
유글레나 그라실리스 유래의 합성 돌연변이 델타-5 불포화화효소("EgD5R*-34g157g")(즉, HgGG 및 HDgSH 모티브 포함)	298 (1350 bp)	299 (449 AA)
유글레나 그라실리스 유래의 합성 돌연변이 델타-5 불포화화효소("EgD5R*-34g158a")(즉, HgGG 및 HDAaH 모티브 포함)	300 (1350 bp)	301 (449 AA)
유글레나 그라실리스 유래의 합성 돌연변이 델타-5 불포화화효소("EgD5R*-34g158g") (즉, HgGG 및 HDAgH 모티브 포함)	302 (1350 bp)	303 (449 AA)
플라스미드 pEgD5M	304 (4070 bp)	--
EgD5M 을 포함하는 플라스미드 pDMW367-5M	305 (8438 bp)	--
플라스미드 pEgD5M1	306 (4070 bp)	--
EgD5M1 을 포함하는 플라스미드 pDMW367-5M1	307 (8438 bp)	--
야로위아 리폴리티카에서의 발현을 위해 코돈-최적화된 유글레나 그라실리스 유래의 합성 돌연변이 델타-5 불포화화효소("EgD5S-36s")(즉, HPGs 모티브 포함)	--	308 (449 AA)
플라스미드 pDMW369S	309 (8438 bp)	--
부위-지정 돌연변이유발에 의해 EgD5S-36s의 HDASH 모티브의 Asp, Ala 또는 Ser 잔기를 개별적으로 돌연변이시키기 위해 사용되는 올리고뉴클레오타이드 프라이머	310-327	--
야로위아 리폴리티카에서의 발현을 위해 코돈-최적화된 유글레나 그라실리스 유래의 합성 돌연변이 델타-5 불포화화효소("EgD5S-36s156e")(즉, HPGs 및 HeASH 모티브 포함)	328 (1350 bp)	329 (449 AA)
야로위아 리폴리티카에서의 발현을 위해 코돈-최적화된 유글레나 그라실리스 유래의 합성 돌연변이 델타-5 불포화화효소("EgD5S-36s158a")(즉, HPGs 및 HDAaH 모티브 포함)	330 (1350 bp)	331 (449 AA)
야로위아 리폴리티카에서의 발현을 위해 코돈-최적화된 유글레나 그라실리스 유래의 합성 돌연변이 델타-5 불포화화효소("EgD5S-36s158a")(즉, HPGs 및 HDAaH 모티브 포함)	332 (1350 bp)	333 (449 AA)

[0110]

불포화화효소("EgD5S-36s158g")(즉, HPGs 및 HDAgH 모티프 포함)		
야로위아 리폴리티카에서의 발현을 위해 코돈-최적화된 유글레나 아나바에나 유래의 합성 돌연변이 델타-5 불포화화효소("EaD5S-35a")(즉, HaGG 모티프 포함)	--	334 (454 AA)
유글레나 아나바에나 UTEX 373 델타-5 불포화화효소("EaD5")	335 (1362 bp)	336 (454 AA)
야로위아 리폴리티카에서의 발현을 위해 코돈-최적화된 유글레나 아나바에나 UTEX 373 유래의 합성 델타-5 불포화화효소("EaD5S")	337 (1362 bp)	338 (454 AA)
플라스미드 pZuFmEaD5S-A(S)	339 (8357 bp)	--
부위-지정 돌연변이유발에 의해 EgD5S-35a의 HDASH 모티프의 Asp, Ala 또는 Ser 잔기를 개별적으로 돌연변이시키기 위해 사용되는 올리고뉴클레오타이드 프라이머	340-361	--
야로위아 리폴리티카에서의 발현을 위해 코돈-최적화된 유글레나 아나바에나 유래의 합성 돌연변이 델타-5 불포화화효소("EaD5S-35a158s")(즉, HaGG 및 HDsSH 모티프 포함)	362 (1365 bp)	363 (454 AA)
야로위아 리폴리티카에서의 발현을 위해 코돈-최적화된 유글레나 아나바에나 유래의 합성 돌연변이 델타-5 불포화화효소("EaD5S-35a159g")(즉, HaGG 및 HDAgH 모티프 포함)	364 (1365 bp)	365 (454 AA)
프라이머 FBAIN-F	366	--
프라이머 Y1026	367	--
유글레나 그라실리스 유래의 합성 돌연변이 델타-9 연장효소("EgD9eS-L35G/W132T/I179R")	368 (777 bp)	369 (258 AA)
플라스미드 pZuFmEgD9ES-L35G/W132T/I179R	370 (7769 bp)	--
유글레나 그라실리스 유래의 합성 돌연변이 델타-9 연장효소("EgD9eS-S9D/L35G/Y84C/I179R")	371 (777 bp)	372 (258 AA)
플라스미드 pZuFmEgD9ES-S9D/L35G/Y84C/I179R	373 (7769 bp)	--
유글레나 그라실리스 유래의 합성 돌연변이 델타-9 연장효소("EgD9eS-L35G/Y84C/I179R/Q244N")	374 (777 bp)	375 (258 AA)
플라스미드 pZuFmEgD9ES-L35G/Y84C/I179R/Q244N	376 (7769 bp)	--
유글레나 그라실리스 유래의 합성 돌연변이 델타-9 연장효소("EgD9eS-A21V/L35G/W132T/I179R/Q244N")	377 (777 bp)	378 (258 AA)
플라스미드 pZuFmEgD9ES-A21V/L35G/W132T/I179R/Q244N	379 (7769 bp)	--
HDgnH 모티프	--	380
HDAnH 모티프	--	381
HefaH 모티프	--	382
HefH 모티프	--	383
HemgH 모티프	--	384

[0111]

HeAgH 모티브	--	385
HDfgH 모티브	--	386
HDygH 모티브	--	387
HDscH 모티브	--	388
HDAcH 모티브	--	389
HDcSH 모티브	--	390
HDdSH 모티브	--	391
HDcSH 모티브	--	392
HDfSH 모티브	--	393
HDhSH 모티브	--	394
HDiSH 모티브	--	395
HDkSH 모티브	--	396
HDISH 모티브	--	397
HDmSH 모티브	--	398
HDnSH 모티브	--	399
HDpSH 모티브	--	400
HDqSH 모티브	--	401
HDrSH 모티브	--	402
HDtSH 모티브	--	403
HDvSH 모티브	--	404
HDwSH 모티브	--	405
HDySH 모티브	--	406
HDAcH 모티브	--	407
HDAdH 모티브	--	408
HDAcH 모티브	--	409
HDAfH 모티브	--	410
HDAhH 모티브	--	411
HDAiH 모티브	--	412
HDAkH 모티브	--	413
HDAIH 모티브	--	414
HDAmH 모티브	--	415
HDAhH 모티브	--	416
HDApH 모티브	--	417
HDAqH 모티브	--	418
HDArH 모티브	--	419
HDAtH 모티브	--	420
HDAvH 모티브	--	421
HDAwH 모티브	--	422
HDAYH 모티브	--	423
HDxxH 모티브	--	424
HgGG 모티브	--	425
HhGG 모티브	--	426
HPGs 모티브	--	427
HaGG 모티브	--	428
HDgSH 모티브	--	429
HDsSH 모티브	--	430
HDAaH 모티브	--	431
HDAgH 모티브	--	432
HeASH 모티브	--	433
HDxSH 모티브	--	434
HDAxH 모티브	--	435
HPGx 모티브	--	436
HxGG 모티브	--	437

[0112]

발명을 실시하기 위한 구체적인 내용

[0113]

본 명세서에 개시된 모든 특허, 특허 출원서 및 공개문헌은 그들 전문이 참고로 포함된다.

[0114]

본 개시내용에서, 다수의 용어 및 약어가 사용된다. 표준 3-문자 코드 또는 1-문자 코드가 아미노산을 지칭하기 위해 사용된다. 하기의 정의가 제공된다.

[0115]

"오픈 리딩 프레임(open reading frame)"은 "ORF"로 약칭된다.

[0116]

"중합효소 연쇄 반응"은 "PCR"로 약칭된다.

[0117]

"아메리칸 타입 컬처 콜렉션(American Type Culture Collection)"은 "ATCC"로 약칭된다.

[0118]

"다중불포화 지방산(들)"은 "PUFA(s)"로 약칭된다.

[0119]

"트라이아실글리세롤"은 "TAG"로 약칭된다.

[0120]

"코-엔자임(Co-enzyme) A"는 "CoA"로 약칭된다.

- [0121] "전체 지방산"은 "TFA"로 약칭된다.
- [0122] "지방산 메틸 에스테르"는 "FAME"로 약칭된다.
- [0123] "건조 세포 중량"은 "DCW"로 약칭된다.
- [0124] "중량 백분율"은 "wt%"로 약칭된다.
- [0125] 본 명세서에서 사용되는 용어 "발명" 또는 "본 발명"은 본 명세서의 특허청구범위 및 상세한 설명에 기재된 본 발명의 모든 태양 및 실시형태를 지칭하고자 하며, 임의의 특정 실시형태 또는 태양에 제한되는 것으로 해석되어서는 안된다.
- [0126] 용어 "식품", "약제", "유아용 조제식", "식이 보조제", "동물 사료" 및 "수산양식(aquaculture) 사료"는 미국 특허 출원 공개 제2010-0317072-A1호에 정의된 바와 같다.
- [0127] 본 명세서에서 사용되는 용어 "바이오매스"는 구체적으로 상업적으로 유의미한 양으로 EPA를 생성하는 재조합 생성 숙주의 발효로부터 소비되거나 사용되는 세포 물질을 지칭한다. 바람직한 생성 숙주는 유지성 효모, 바람직하게는 야로위아 속, 더욱 바람직하게는 야로위아 리폴리티카의 재조합 균주이다. 바이오매스는 전체 세포, 전체 세포 용해물, 균질화된 세포, 부분 가수분해된 세포 물질 및/또는 부분 정제된 세포 물질(예컨대, 미생물에 의해 생성되는 오일)의 형태일 수 있다.
- [0128] 용어 "지질"은 임의의 지방-용해성(즉, 지방 친화성), 자연-발생 분자를 지칭한다. 지질의 일반적인 개관은 미국 특허 출원 공개 제2009-0093543-A1호(상기 문헌의 표 2 참조)에 제공되어 있다.
- [0129] 용어 "오일"은 25℃에서 액체인 지질 물질을 지칭하며; 오일은 소수성이나 유기 용매 중에서는 용해성이다. 유지성 유기체에서, 오일은 전체 지질의 주요 부분을 구성한다. "오일"은 주로 트리아실글리세롤["TAG"]로 이루어지나, 다른 중성 지질, 인지질 및 유리 지방산도 또한 함유할 수 있다. 오일 중의 지방산 조성 및 전체 지질의 지방산 조성은 일반적으로 유사하며; 이에 따라, 전체 지질 중의 지방산의 농도의 증가 또는 감소는 오일 중의 지방산의 농도의 증가 또는 감소와 부합할 것이며, 그 역도 또한 그러하다.
- [0130] "중성 지질"은 저장 지방으로서 지질체 중의 세포에서 통상적으로 관찰되는 지질을 지칭하고, 이렇게 지칭되는 것은 세포 pH에서 지질이 전하를 띠지 않은 기를 보유하기 때문이다. 일반적으로, 이들은 물에 대한 친화성이 없이 완전히 비-극성이다. 중성 지질은 일반적으로, 글리세롤과 지방산의 모노-, 다이-, 및/또는 트리아에스테르를 지칭하고, 또한 모노아실글리세롤, 다이아실글리세롤 또는 트리아실글리세롤로 각각 지칭되거나, 집합적으로 아실글리세롤로 지칭된다. 아실글리세롤로부터 유리 지방산을 유리시키기 위해서는 가수분해 반응이 일어나야만 한다.
- [0131] 용어 "트리아실글리세롤"["TAG"]은 글리세롤 분자에 에스테르화된 3개의 지방 아실 잔기로 구성된 중성 지질을 지칭한다. TAG는 장쇄 PUFA 및 포화 지방산뿐 아니라 보다 짧은 쇠의 포화 및 불포화 지방산을 포함할 수 있다.
- [0132] 본 명세서에서, 용어 "전체 지방산"["TFA"]은 예를 들어, 바이오매스 또는 오일일 수 있는 주어진 시료에서 염기 에스테르 교환 방법(당업계에 공지되어 있는)에 의해 지방산 메틸 에스테르["FAME"]로 유도체화될 수 있는 모든 세포 지방산의 합을 말한다. 따라서, 전체 지방산은 중성 지질 분획(다이아실글리세롤, 모노아실글리세롤 및 TAG 포함) 및 극성 지질 분획(예컨대, 포스파티딜콜린 및 포스파티딜에탄올아민 분획 포함)으로부터의 지방산을 포함하나, 유리 지방산은 포함하지 않는다.
- [0133] 용어 세포의 "전체 지질 함량"은 건조 세포 중량["DCW"]의 백분율로서의 TFA의 측정치이지만, 전체 지질 함량은 DCW의 백분율로서의 FAME["FAME % DCW"]의 측정치로 근사치를 계산할 수 있다. 따라서, 전체 지질 함량["TFA % DCW"]은 예를 들어, 100 밀리그램의 DCW당 전체 지방산의 밀리그램과 등가이다.
- [0134] 전체 지질 중의 지방산의 농도는 본 명세서에서 TFA의 중량 백분율["% TFA"], 예컨대, 100 밀리그램의 TFA당 주어진 지방산의 밀리그램으로 표현된다. 본 명세서의 개시내용에서 구체적으로 달리 기재되지 않는 한, 전체 지질에 대한 주어진 지방산의 백분율에 대한 참조는 % TFA로서의 지방산의 농도와 등가이다(예컨대, 전체 지질의 EPA%는 EPA % TFA와 등가이다).
- [0135] 일부 경우에, 건조 세포 중량의 그의 중량 백분율["% DCW"]로서 세포 중의 주어진 지방산(들)의 함량을 표현하는 것이 유용하다. 따라서, 예를 들어, EPA 생산성의 측정치[EPA % DCW]는 하기의 식에 따라 결정할 것이다: $(EPA \% TFA) * (TFA \% DCW) / 100$. 그러나, 세포 중의 주어진 지방산(들)의 함량은 건조 세포 중량의 그의 중

량 백분율["% DCW"]로서 하기의 식으로 근사치를 계산할 수 있다: $(\text{EPA \% TFA}) * (\text{FAME \% DCW}) / 100$.

- [0136] 용어 "지질 프로파일" 및 "지질 조성"은 상호교환가능하며, 특정 지질 분획, 예컨대 전체 지질 또는 오일에 함유되는 개별 지방산의 양을 지칭하며, 여기서 이러한 양은 TFA의 wt%로 표현된다. 혼합물에 존재하는 각 개별 지방산의 합은 100이어야 한다.
- [0137] 용어 "추출된 오일"은 다른 세포 물질, 예컨대 오일을 합성하는 미생물로부터 분리된 오일을 지칭한다. 추출된 오일은 매우 다양한 방법을 통해 수득되며, 그 중 가장 간단한 것에는 단독의 물리적 수단이 포함된다. 예를 들어, 다양한 압착 구성(press configuration)(예컨대, 스크루(screw), 익스펠러(expeller), 피스톤, 비드 비터(bead beaters) 등)를 사용한 기계적 분쇄(mechanical crushing)에 의해, 세포 물질로부터 오일을 분리할 수 있다. 대안적으로, 오일 추출은 다양한 유기 용매(예컨대, 헥산)로의 처리를 통하여, 효소적 추출을 통하여, 삼투압 충격(osmotic shock)을 통하여, 초음파 추출을 통하여, 초임계 유체 추출(예컨대, CO₂ 추출)을 통하여, 비누화를 통하여, 그리고 이들 방법의 조합을 통하여 일어날 수 있다. 추출된 오일은 정제되거나 추가로 농축될 수 있다. 본 명세서에 기재된 추출된 오일은 적어도 45 EPA % TFA를 포함할 것이다.
- [0138] 용어 "배합된 오일"은 목적하는 조성을 수득하기 위하여 본 명세서에 기재된 추출된 오일을 오일의 임의의 조합 또는 개별 오일과 혼합하거나 배합함으로써 수득되는 오일을 말한다. 따라서, 예를 들어, 상이한 미생물로부터의 오일의 유형들을 함께 혼합하여 목적하는 PUFA 조성물을 수득할 수 있다. 대안적으로, 또는 부가적으로, 본 명세서에 개시된 PUFA-함유 오일을 어유, 식물유, 또는 둘 모두의 혼합물과 배합하여 목적하는 조성물을 수득할 수 있다.
- [0139] 용어 "지방산"은 약 C₁₂ 내지 C₂₂의 다양한 쇠 길이의 장쇄 지방족 산(알칸산)을 지칭하지만, 더 긴 쇠 길이의 산과 더 짧은 쇠 길이의 산이 둘 모두 공지되어 있다. 주된 쇠 길이는 C₁₆ 내지 C₂₂이다. 지방산의 구조는 "X:Y"의 단순한 표기 체계로 표현되며, 여기에서, X는 특정 지방산에서의 전체 탄소 원자["C"]의 개수이고, Y는 이중결합의 개수이다. "포화 지방산" 대 "불포화 지방산", "단일불포화 지방산" 대 "다중불포화 지방산"["PUFA"] 및 "오메가-6 지방산"[" ω -6" 또는 " n -6"] 대 "오메가-3 지방산"[" ω -3" 또는 " n -3"] 간의 구별에 관한 추가의 상세사항은 본 명세서에 참고로 포함되는 미국 특허 제7,238,482호에 제공되어 있다.
- [0140] 본 명세서에서 PUFA를 기재하는데 사용되는 명명법은 표 2에 제공되어 있다. "약칭 표기"라는 표제의 열에서, 오메가-참조 체계는 탄소의 개수, 이중결합의 개수 및 오메가 탄소와 가장 가까운 이중결합의 위치를 오메가 탄소(이러한 목적을 위해 1로 번호를 매김)로부터 카운팅하여, 표시하기 위해 사용된다. 표의 나머지는 오메가-3 및 오메가-6 지방산의 일반명 및 그들의 전구체, 본 명세서에서 사용될 약어 및 각 화합물의 화학명을 요약한 것이다.

표 2

다중불포화 지방산 및 전구체의 명명법

일반명	약어	화학명	약칭 표기
미리스트산	--	테트라데칸산	14:0
팔미트산	팔미테이트	헥사데칸산	16:0
팔미톨레산	--	9-헥사데센산	16:1
스테아르산	--	옥타데칸산	18:0
올레산	--	시스-9-옥타데센산	18:1
리놀레산	LA	시스-9,12-옥타데카다이엔산	18:2 ω-6
γ-리놀렌산	GLA	시스-6,9,12-옥타데카트라이엔산	18:3 ω-6
에이코사다이엔산	EDA	시스-11,14-에이코사다이엔산	20:2 ω-6
다이호모-γ-리놀렌산	DGLA	시스-8,11,14-에이코사트라이엔산	20:3 ω-6
아라키돈산	ARA	시스-5,8,11,14-에이코사테트라엔산	20:4 ω-6
α-리놀렌산	ALA	시스-9,12,15-옥타데카트라이엔산	18:3 ω-3
스테아리돈산	STA	시스-6,9,12,15-옥타데카테트라엔산	18:4 ω-3
에이코사트라이엔산	ETrA	시스-11,14,17-에이코사트라이엔산	20:3 ω-3
에이코사-테트라엔산	ETA	시스-8,11,14,17-에이코사테트라엔산	20:4 ω-3
에이코사-펜타엔산	EPA	시스-5,8,11,14,17-에이코사펜타엔산	20:5 ω-3
도코사-테트라엔산	DTA	시스-7,10,13,16-도코사테트라엔산	22:4 ω-6
도코사-펜타엔산	DPAAn-6	시스-4,7,10,13,16-도코사펜타엔산	22:5 ω-6
도코사-펜타엔산	DPA	시스-7,10,13,16,19-도코사펜타엔산	22:5 ω-3
도코사-헥사엔산	DHA	시스-4,7,10,13,16,19-도코사헥사엔산	22:6 ω-3

[0141]

[0142]

용어 "PUFA 생합성 경로"는 올레산을 오메가-6 지방산, 예컨대 LA, EDA, GLA, DGLA, ARA, DTA 및 DPAn-6, 및 오메가-3 지방산, 예컨대 ALA, STA, ETrA, ETA, EPA, DPA 및 DHA로 전환시키는 대사 과정을 지칭한다. 이러한 과정은 문헌에 널리 기재되어 있다(예컨대, 미국 특허 제7,932,077호 및 미국 특허 출원 공개 제2009-0093543-A1호 참조). 간단하게 설명하자면, 이 과정은 소포체 막에 존재하는 "PUFA 생합성 경로 효소"로 지칭되는 일련의 특별한 연장 및 불포화화 효소를 통하여, 탄소 원자의 부가를 통한 탄소 사슬의 연장 및 이중 결합의 부가를 통한 분자의 불포화화를 수반한다. 더욱 구체적으로, "PUFA 생합성 경로 효소"는 PUFA의 생합성과 관련이 있는 다음을 포함하는 효소(및 상기 효소를 암호화하는 유전자) 중 임의의 것을 지칭한다: 델타-4 불포화화효소, 델타-5 불포화화효소, 델타-6 불포화화효소, 델타-12 불포화화효소, 델타-15 불포화화효소, 델타-17 불포화화효소, 델타-9 불포화화효소, 델타-8 불포화화효소, 델타-9 연장효소, C_{14/16} 연장효소, C_{16/18} 연장효소, C_{18/20} 연장효소 및/또는 C_{20/22} 연장효소.

[0143]

용어 "델타-9 연장효소/ 델타-8 불포화화효소 경로"는 적어도 하나의 델타-9 연장효소 및 적어도 하나의 델타-8 불포화화효소를 포함하여, 중간체 지방산으로서 EDA 및/또는 ETrA와 함께, 각각 LA 및 ALA로부터 DGLA 및/또는 ETA의 생합성을 가능하게 하는 PUFA 생합성 경로를 지칭할 것이다. 다른 불포화화효소 및 연장효소의 발현과 함께, ARA, DTA, DPAn-6, EPA, DPA 및 DHA도 합성될 수 있다.

[0144]

용어 "전환 효율" 및 "기질 전환 백분율"은 특정 효소, 예컨대 불포화화효소, 연장효소 또는 멀티자임이 기질을 생성물로 전환시킬 수 있는 효율을 말한다. 전환 효율은 하기의 식에 따라 측정된다: ([생성물]/[기질+생성물])*100, 여기서, '생성물'은 중간체 생성물 및 그로부터 유래되는 경로에서의 모든 생성물을 포함한다.

[0145]

용어 "불포화화효소"는 하나 이상의 지방산을 불포화화시켜, 즉, 하나 이상의 지방산에 이중 결합을 도입시켜, 대상 지방산 또는 전구체를 생성할 수 있는 폴리펩티드를 지칭한다. 본 명세서에서 특정 지방산을 언급하기 위해 오메가-참조 체계가 사용되고는 있지만, 델타-체계를 사용하여 기질의 카복실 말단으로부터 카운팅하여 불포화화효소의 활성을 표시하는 것이 보다 편리하다. 델타-8 불포화화효소, 델타-5 불포화화효소, 델타-17 불포화화효소 및 델타-12 불포화화효소가 본 명세서에서 특히 흥미로운 것이다. 다른 유용한 불포화화효소는 델타-4 불포화화효소, 델타-6 불포화화효소, 델타-15 불포화화효소 및 델타-9 불포화화효소를 포함할 수 있다.

- [0146] 용어 "연장효소"는 지방산 탄소쇄를 연장시켜, 연장효소가 작용하는 지방산 기질보다 2개의 탄소가 더 긴 산을 생성할 수 있는 폴리펩티드를 지칭한다. 이러한 연장 과정은 미국 특허 제7,659,120호에 기재된 바와 같이, 지방산 신타제와 연계된 다단계 메카니즘에서 일어난다. 연장효소 계에 의해 촉매작용되는 반응의 예는 GLA로부터 DGLA로의 전환, STA로부터 ETA로의 전환, ARA로부터 DTA로의 전환 및 EPA로부터 DPA로의 전환이다. 일반적으로, 연장효소의 기질 선택성은 다소 광범위하나,쇄 길이, 및 불포화 정도 및 유형 둘 모두에 의해 구분된다. 예를 들어, C_{14/16} 연장효소는 C₁₄ 기질(예컨대, 미리스트산)을 사용할 것이며, C_{16/18} 연장효소는 C₁₆ 기질(예컨대, 팔미테이트)을 사용할 것이며, C_{18/20} 연장효소는 C₁₈ 기질(예컨대, GLA, STA)을 사용할 것이며, C_{20/22} 연장효소[델타-5 연장효소 또는 C20 연장효소로도 지칭됨]는 C₂₀ 기질(예컨대, ARA, EPA)을 사용할 것이다. 본 명세서의 목적을 위하여, 2가지의 별개의 C_{18/20} 연장효소의 유형을 정의할 수 있다: 델타-6 연장효소는 각각 GLA 및 STA로부터 DGLA 및 ETA로의 전환을 촉매작용시킬 것이나, 델타-9 연장효소는 각각 LA 및 ALA로부터 EDA 및 ETrA로의 전환을 촉매작용시킬 수 있다.
- [0147] "C₁₈로부터 C₂₀으로의 연장 전환 효율"은 C_{18/20} 연장효소가 C₁₈ 기질(즉, LA, ALA, GLA, STA)을 C₂₀ 생성물(즉, EDA, ETrA, DGLA, ETA)로 전환시킬 수 있는 효율을 말한다. 이들 C_{18/20} 연장효소는 델타-9 연장효소 또는 델타-6 연장효소 중 어느 하나일 수 있다.
- [0148] 용어 "델타-9 연장 전환 효율"은 델타-9 연장효소가 C₁₈ 기질(즉, LA, ALA)을 C₂₀ 생성물(즉, EDA, ETrA)로 전환시킬 수 있는 효율을 말한다.
- [0149] 용어 "EgD9e"는 본 명세서에서 서열 번호 31에 의해 암호화되는 유글레나 그라실리스로부터 분리된 델타-9 연장효소(서열 번호 32)를 지칭한다. 유사하게, 용어 "EgD9eS"는 야로위아 리폴리티카에서의 발현을 위해 코돈-최적화된 유글레나 그라실리스 유래의 합성 델타-9 연장효소(즉, 서열 번호 2 및 3)를 지칭한다. EgD9e 및 EgD9eS뿐 아니라 델타-9 연장효소 모티프에 관한 추가의 상세사항은 미국 특허 제7,645,604호에 기재되어 있다.
- [0150] 용어 "EaD9e"는 본 명세서에서 서열 번호 33에 의해 암호화된 유글레나 아나바에나로부터 분리된 델타-9 연장효소(서열 번호 34)를 지칭한다. 유사하게, 용어 "EaD9eS"는 야로위아 리폴리티카에서의 발현을 위해 코돈-최적화된 유글레나 아나바에나 유래의 합성 델타-9 연장효소(즉, 서열 번호 35 및 36)를 지칭한다. EaD9e 및 EaD9eS와 관련된 추가의 상세사항은 미국 특허 제7,794,701호에 기재되어 있다.
- [0151] 용어 "E389D9e"는 본 명세서에서 서열 번호 37에 의해 암호화된 유트렙티엘라 중 CCMP389로부터 분리된 델타-9 연장효소(서열 번호 38)를 지칭한다. 유사하게, 용어 "E389S9eS"는 야로위아 리폴리티카에서의 발현을 위해 코돈-최적화된 유트렙티엘라 중 CCMP389 유래의 합성 델타-9 연장효소(즉, 서열 번호 39 및 40)를 지칭한다. E389D9e 및 E389D9eS와 관련된 추가의 상세사항은 미국 특허 제7,645,604호에 기재되어 있다.
- [0152] 용어 "IgD9e"는 본 명세서에서 서열 번호 41에 의해 암호화된, 아이소크라이시스 갈바나로부터 분리된 델타-9 연장효소(서열 번호 42; NCBI 수탁 번호 AAL37626(GI 17226123))를 지칭한다.
- [0153] 용어 "돌연변이 델타-9 연장효소" 또는 "돌연변이 EgD9eS"는 야로위아 리폴리티카에서의 발현을 위해 코돈-최적화된 유글레나 그라실리스 유래의 합성 델타-9 연장효소(즉, EgD9eS[서열 번호 2 및 3])에 대하여 적어도 하나의 돌연변이를 갖는 델타-9 연장효소를 지칭한다. "돌연변이"가 임의의 결실, 삽입 및 점 돌연변이(또는 그들의 조합)를 포함할 수 있지만, 바람직한 실시형태에서, 돌연변이 EgD9eS는 서열 번호 1에 기재되어 있으며(도 16b), 여기서 서열 번호 1은 적어도 하나의 아미노산 돌연변이에 의해 서열 번호 3과 상이하며, 상기 돌연변이(들)은 a) L35F 돌연변이; b) L35M 돌연변이; c) L35G 돌연변이; d) L35G 돌연변이, 및 S9A, S9D, S9G, S9I, S9K, S9Q, Q12K, A21D, A21T, A21V, V32F, Y84C, Q107E, L108G, G127L, W132T, M143N, M143W, L161T, L161Y, W168G, I179M, I179R, C236N, Q244N, A254W 및 A254Y로 이루어진 군으로부터 선택되는 적어도 하나의 다른 돌연변이; e) L35G, A21V, L108G 및 I179R 돌연변이; f) L35G, W132T 및 I179R 돌연변이; g) L35G, S9D, Y84C 및 I179R 돌연변이; h) L35G, Y84C, I179R 및 Q244N 돌연변이; i) L35G, A21V, W132T, I179R 및 Q244N 돌연변이; j) K58R 및 I257T 돌연변이; k) D98G 돌연변이; l) L130M 및 V243A 돌연변이; 및 m) K58R, L35F, L35G, L35M, S9A, S9D, S9G, S9I, S9K, S9Q, Q12K, A21D, A21T, A21V, V32F, Y84C, D98G, Q107E, L108G, G127L, L130M, W132T, M143N, M143W, L161T, L161Y, W168G, I179M, I179R, C236N, V243A, Q244N, A254W, A254Y 및 I257T로 이루어진 군으로부터 선택되는 적어도 2개의 돌연변이를 포함하는 임의의 조합으로 이루어진 군으로부터 선택된다. 열거된 각각의 치환에 대하여, 처음의 문자는 EgD9eS의 아미노산(서열 번호 3)에 해당하며, 두번째 문자는 돌연변이체(서열 번호 1)의 동일한 위치에서 관찰되는 아미노산에 해당하며, 즉, L35F는 위치 35에서 EgD9eS의

Leu[L]로부터 EgD9eS 돌연변이체의 Phe[F]로의 변화를 나타낸다. 이러한 명명법은 본 명세서에 기재된 델타-9 연장효소 단백질 내의 돌연변이를 지칭하기 위해 본 명세서에 사용되며; 유사한 명명법을 사용하여 뉴클레오티드 서열 내의 치환을 기재한다(즉, C62T는 위치 62에서 EgD9eS(서열 번호 2)의 시토신[C]으로부터 EgD9eS 돌연변이체의 티민[T]으로의 변화를 나타낸다).

[0154] 돌연변이 EgD9eS는 폴리펩티드 서열의 차이에도 불구하고, 효소 활성의 비교시, EgD9eS에 대하여 "델타-9 연장효소 활성이 향상"될 것이다. 따라서, 돌연변이 EgD9eS 서열은 EgD9eS의 효소 활성과 비교하여 증가된 효소 활성(즉, EgD9eS의 효소 활성의 적어도 약 101-110%, 바람직하게는 적어도 약 110-125%, 더욱 바람직하게는 적어도 약 125-150%, 가장 바람직하게는 약 150% 초과)을 가질 것이다. 바람직한 범위가 상술되어 있지만, 유용한 전환 효율의 예에는 50% 내지 적어도 150%, 예컨대 51%, 52%, 53%, 54%, 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%, 99%, 100%, 101%, 102%, 103%, 104%, 105%, 106%, 107%, 108%, 109%, 110%, 111%, 112%, 113%, 114%, 115%, 116%, 117%, 118%, 119%, 120%, 121%, 122%, 123%, 124%, 125%, 126%, 127%, 128%, 129%, 130%, 131%, 132%, 133%, 134%, 135%, 136%, 137%, 138%, 139%, 140%, 141%, 142%, 143%, 144%, 145%, 146%, 147%, 148%, 149% 및 150%의 임의의 정수 백분율이 포함된다.

[0155] 용어 "EgD9eS-L35G"는 본 명세서에서 서열 번호 43에 의해 암호화되는 EgD9eS(서열 번호 3)에 대하여 단일의 L35G 돌연변이를 갖는 합성 돌연변이 델타-9 연장효소(서열 번호 44)를 지칭한다.

[0156] 용어 "멀티자임" 또는 "융합 단백질"은 적어도 2개의 독립적이며 분리될 수 있는 효소 활성을 갖는 단일의 폴리펩티드를 지칭하며, 여기서, 제1 효소 활성은 바람직하게는 제2 효소 활성과 연결된다(미국 특허 출원 공개 제 2008-0254191-A1호). 적어도 2개의 독립적이며 분리될 수 있는 효소 활성 사이의 "링커"는 단일의 폴리펩티드로 이루어질 수 있지만, 링커는 또한 하나의 아미노산 잔기, 예컨대 Pro 또는 적어도 하나의 Pro를 포함하는 폴리펩티드로 이루어질 수 있다. 바람직한 링커는 서열 번호 4, 서열 번호 5, 서열 번호 6, 서열 번호 7, 서열 번호 8, 서열 번호 9 및 서열 번호 10으로 이루어진 군으로부터 선택된다.

[0157] 용어 "DGLA 신타제"는 멀티자임을 지칭하며, 여기서, 델타-9 연장효소는 델타-8 불포화화효소에 연결된다. 용어 "EgD9eS/EgD8M"은 링커 서열(즉, 서열 번호 4[GAGPARPAGLPATYYDSLAVMGS]; 미국 특허 출원 공개 제 2008-0254191-A1호)을 사용하여 델타-9 연장효소 "EgD9eS"(미국 특허 제 7,645,604호)를 델타-8 불포화화효소 "EgD8M"(미국 특허 제 7,709,239호)에 연결시킴으로써 생성되는 DGLA 신타제(서열 번호 11 및 12)를 지칭한다. 유사하게, 용어 "EaD9eS/EaD8S"는 서열 번호 4에 기재된 링커 서열을 사용하여 델타-9 연장효소 "EaD9eS"(미국 특허 제 7,794,701호)를 델타-8 불포화화효소 "EaD8S"(미국 특허 제 7,790,156호)에 연결시킴으로써 생성되는 DGLA 신타제(서열 번호 13 및 14)를 지칭한다. 그리고, 용어 "E389D9eS/EgD8M"은 서열 번호 4에 기재된 링커 서열을 사용하여 델타-9 연장효소 "E389D9eS"(미국 특허 제 7,645,604호)를 델타-8 불포화화효소 "EgD8M"(상기)에 연결시킴으로써 생성되는 DGLA 신타제(서열 번호 15 및 16)를 지칭한다.

[0158] 용어 "아실트랜스퍼라제"는 공여체 지질로부터 수용체 지질 분자로의 아실기의 전달을 담당하는 효소를 지칭한다.

[0159] 용어 "아실-CoA:라이소인지질 아실트랜스퍼라제" 또는 "라이소인지질 아실트랜스퍼라제"["LPLAT"]는 sn-2 위치에서 다양한 라이소인지질 기질을 아실화시키는 능력을 갖는 넓은 부류의 아실트랜스퍼라제를 지칭한다. 더욱 구체적으로, LPLAT는 LPA로부터 포스파티드산["PA"]으로의 전환을 촉매작용시키는 능력을 갖는 라이소포스파티드산["LPA"] 아실트랜스퍼라제["LPAAT"], LPC로부터 포스파티딜콜린["PC"]으로의 전환을 촉매작용시키는 능력을 갖는 라이소포스파티딜콜린["LPC"] 아실트랜스퍼라제["LPCAT"], LPE로부터 포스파티딜에탄올아민["PE"]으로의 전환을 촉매작용시키는 능력을 갖는 라이소포스파티딜에탄올아민["LPE"] 아실트랜스퍼라제["LPEAT"], LPS로부터 포스파티딜세린["PS"]으로의 전환을 촉매작용시키는 능력을 갖는 라이소포스파티딜세린["LPS"] 아실트랜스퍼라제["LPSAT"], LPG로부터 포스파티딜글리세롤["PG"]로의 전환을 촉매작용시키는 능력을 갖는 라이소포스파티딜글리세롤["LPG"] 아실트랜스퍼라제["LPGAT"] 및 LPI로부터 포스파티딜이노시톨["PI"]로의 전환을 촉매작용시키는 능력을 갖는 라이소포스파티딜이노시톨["LPI"] 아실트랜스퍼라제["LPIAT"]를 포함한다.

[0160] 용어 "적어도 라이소포스파티드산 아실트랜스퍼라제["LPAAT"] 활성을 갖는 폴리펩티드"는 하기의 반응을 촉매작용시킬 수 있는 효소를 지칭할 것이다: 아실-CoA + 1-아실-sn-글리세롤 3-포스페이트 → CoA + 1,2-다이아실-sn-글리세롤 3-포스페이트(EC 2.3.1.51). 따라서, "LPAAT"는 미국 특허 출원 공개 제 2010-0317072-A1호 및 미국 특허 출원 공개 제 2010-0317882-A1호에 기재된 단백질을 지칭하며, 이는 1) LPAAT 활성을 가지며, 클러스털

더블유 정렬 방법을 기준으로, 서열 번호 18(MaLPAAT1), 서열 번호 22(YILPAAT1) 및 서열 번호 23(ScLPAAT1)으로 이루어진 군으로부터 선택되는 아미노산 서열과 비교시 적어도 약 43.9%의 아미노산 동일성을 공유하고/거나; 2) LPAAT 활성을 가지며, NH_{xxxxD}(서열 번호 26) 및 EGTR(서열 번호 27)로 이루어진 군으로부터 선택되는 적어도 하나의 1-아실-*sn*-글리세롤-3-포스페이트 아실트랜스퍼라제 패밀리를 갖는다.

- [0161] LPAAT 폴리펩티드의 예에는 하기 ScLPAAT, ScLPAATS, MaLPAAT1, MaLPAAT1S 및 YILPAAT1이 포함된다.
- [0162] 용어 "ScLPAAT"는 사카로마이세스 세레비지에(ORF "YDL052C")로부터 분리된 LPAAT(서열 번호 23)를 지칭한다. 그에 반해, 용어 "ScLPAATS"는 야로위아 리폴리티카에서의 발현을 위해 코돈-최적화된 사카로마이세스 세레비지에 유래의 합성 LPAAT를 지칭한다(즉, 서열 번호 24 및 25)(미국 특허 출원 공개 제2010-0317882-A1호).
- [0163] 용어 "MaLPAAT1"은 서열 번호 17에 기재된 뉴클레오티드 서열에 의해 암호화된, 모르티에렐라 알피나로부터 분리된 LPAAT(서열 번호 18)를 지칭한다. 그에 반해, 용어 "MaLPAAT1S"는 야로위아 리폴리티카에서의 발현을 위해 코돈-최적화된 모르티에렐라 알피나 유래의 합성 LPAAT를 지칭한다(즉, 서열 번호 19 및 20)(미국 특허 제 7,879,591호).
- [0164] 용어 "YILPAAT1"은 서열 번호 21에 기재된 뉴클레오티드 서열에 의해 암호화되는 야로위아 리폴리티카로부터 분리된 LPAAT(서열 번호 22)를 지칭한다.
- [0165] 용어 "적어도 인지질:다이아실글리세롤 아실트랜스퍼라제[\"PDAT\"] 활성을 갖는 폴리펩티드"는 인지질(예컨대, 포스파티딜콜린)의 *sn*-2 위치로부터 1,2-다이아실글리세롤[E.C.2.3.1.158]의 *sn*-3 위치로 지방 아실기를 전달시켜, 라이소인지질 및 TAG를 야기할 수 있는 효소를 지칭할 것이다. PDAT 및 다이아실글리세롤 아실트랜스퍼라제(DAG AT)[E.C. 2.3.1.20] 둘 모두가 TAG 생합성의 최종 단계에 수반되지만, 오직 PDAT만이 아실-CoA-독립적 메커니즘을 통하여 TAG를 합성할 수 있다. 서열 번호 30에 기재된 대표적인 PDAT 효소는 사카로마이세스 세레비지에에서 LRO1 유전자에 의해 암호화된다(문헌[Dahlqvist et al., *Proc. Natl. Acad. Sci. USA* 97:6487 (2000)]).
- [0166] 용어 "YIPDAT"는 서열 번호 28에 기재된 뉴클레오티드 서열에 의해 암호화되는 야로위아 리폴리티카로부터 분리된 PDAT(서열 번호 29)를 지칭한다(미국 특허 제7,901,928호).
- [0167] 용어 "콜린 포스페이트 시티딜트랜스퍼라제"는 하기의 화학 반응을 촉매작용시키는 포스파티딜콜린[\"PC\"] 생합성 경로의 효소(EC 2.7.7.15)를 지칭한다: 시티딘 트라이포스페이트[\"CTP\"] + 콜린 포스페이트 ⇌ 다이포스페이트 + 시티딘 다이포스페이트-콜린[\"CDP-콜린\"]. 따라서, 이러한 효소는 인-함유 뉴클레오티드 기를 전달시켜, 글리세로인지질 메커니즘에서 역할을 수행할 수 있는 트랜스퍼라제(즉, 뉴클레오티드트랜스퍼라제)이다.
- [0168] 용어 "YIPCT"는 서열 번호 45에 의해 암호화되는, 야로위아 리폴리티카로부터 분리된 콜린포스페이트 시티딜트랜스퍼라제(서열 번호 46)를 지칭한다.
- [0169] 용어 "다이아실글리세롤 콜린포스포트랜스퍼라제"는 CDP-콜린 및 1,2-다이아실글리세롤로부터 포스파티딜콜린의 합성을 촉매작용시키는 포스파티딜콜린[\"PC\"] 생합성 경로의 효소(EC 2.7.8.2)를 지칭한다.
- [0170] 용어 "YICPT1"은 서열 번호 47에 의해 암호화되는, 야로위아 리폴리티카로부터 분리된 다이아실글리세롤 콜린포스포-트랜스퍼라제(서열 번호 48)를 지칭한다. YICPT1은 국제 출원 공개 WO 2006/052870호(또한, 진뱅크 수탁 번호 XM_501703(YALIO10989g) 참조)에 기재되어 있다.
- [0171] "말로닐-CoA 신테타제"[EC 6.2.1.-]는 하기의 효소 반응을 촉매작용시킨다: 말로네이트 + ATP + CoA → 말로닐-CoA + AMP + 피로포스페이트(PPi). 효소를 먼저 말로네이트-성장 슈도모나스 플루오레스스(*Pseudomonas fluorescens*)로부터 정제하였지만(문헌[Kim, Y.S. and S.K. Bang, *J. Biol. Chem.*, 260:5098-5104 (1985)]), 다양한 리조비아(*Rhizobia*) 상동체가 콩과식물 뿌리혹(legume nodule) 내의 박테로이드로부터 분리되었다(예컨대, 문헌[Kim, Y.S. and H.Z. Chae, *Biochem. J.*, 273:511-516 (1991)] 및 문헌[Kim, Y.S. and S.W. Kang, *Biochem. J.*, 297:327-333 (1994)] 참조).
- [0172] 용어 "MCS"는 야로위아 리폴리티카에서의 발현을 위해 코돈-최적화된 리조비움 레구미노사룸 생물변이형 비시에 3841(진뱅크 수탁 번호 YP_766603) 유래의 말로닐-CoA 신테타제를 암호화하는 합성 유전자를 지칭한다(즉, 서열 번호 49 및 50)(미국 특허 출원 공개 제2010-0159558-A1호).
- [0173] 용어 "퍼옥시좀"은 모든 진핵 세포에서 관찰되는 혼한 세포소기관을 지칭한다. 그들은 그들의 내용물을 세포질로부터 분리하고, 후술되는 기능에 필수적인 다양한 막 단백질을 함유하는 단일의 지질 이중층 막을 갖는다.

퍼옥시좀은 "확장된 셔틀 메카니즘(extended shuttle mechanism)"을 통하여 단백질을 선택적으로 유입시킨다. 더욱 구체적으로, 퍼옥신(peroxin)으로 지칭되는 적어도 32개의 공지되어 있는 퍼옥시좀 단백질이 존재하며, 이는 퍼옥시좀 막을 통하여 ATP 가수분해의 수단에 의해 단백질을 유입시키는 과정에 참여한다. 일단 세포 단백질이 퍼옥시좀 내로 유입되면, 그들은 전형적으로 몇몇 분해 수단으로 처리된다. 예를 들어, 퍼옥시좀은 산화 효소, 에센대 카탈라제, D-아미노산 옥시다제 및 요산 옥시다제를 포함하며, 이는 세포에 독성이 있는 기질의 분해를 가능하게 한다. 대안적으로, 퍼옥시좀은 지방산 분자를 분해시켜, 유리 아세틸-CoA 분자를 생성하며, 이는 β -산화로 지칭되는 과정에서 세포질로 다시 유출된다.

[0174] 용어 "퍼옥시좀 생합성 인자 단백질", "퍼옥신" 및 "Pex 단백질"은 상호교환가능하며, 퍼옥시좀 생합성에 수반되고/거나 퍼옥시좀 막을 통한 ATP 가수분해의 수단에 의한 세포 단백질의 유입 과정에 참여하는 단백질을 지칭한다. 임의의 이들 단백질을 암호화하는 유전자의 두문자어는 "Pex 유전자"이다. 명명법 체계는 문헌[Distel et al., *J. Cell Biol.*, 135:1-3 (1996)]에 기재되어 있다. 적어도 32개의 상이한 Pex 유전자가 지금까지 다양한 진핵 유기체에서 동정되었다. 17개의 상이한 진균류 종의 게놈 서열의 가상 환경에서의 분석을 수행한 문헌[Kiel, J. A. K. W., et al. *Traffic*, 7:1291-1303 (2006)]에 의한 검토에 기초하여, 하기의 Pex 단백질을 동정하였다: Pex1p, Pex2p, Pex3p, Pex3Bp, Pex4p, Pex5p, Pex5Bp, Pex5Cp, Pex5/20p, Pex6p, Pex7p, Pex8p, Pex10p, Pex12p, Pex13p, Pex14p, Pex15p, Pex16p, Pex17p, Pex14/17p, Pex18p, Pex19p, Pex20p, Pex21p, Pex21Bp, Pex22p, Pex22p-유사 및 Pex26p. 집합적으로, 이들 단백질은 본 명세서에서 "Pex 유전자"에 의해 암호화되는 "Pex 단백질"로 지칭될 것이다.

[0175] 적어도 하나의 퍼옥시좀 생합성 인자 단백질에서의 또는 그와 관련된 용어 "하향 조절된"은 야생형 단백질의 활성과 비교하여, 천연의 Pex 단백질의 활성의 감소 또는 제거를 지칭한다. 하향 조절은 전형적으로, 천연의 Pex 유전자가 그 유전자의 일부 내의 삽입, 결실 또는 표적화된 돌연변이를 지칭하는 "파괴(disruption)"를 가져, 유전자가 게놈으로부터 결실되고 단백질이 번역되지 않게 하는 완전한 유전자 낙아웃(knockout), 또는 삽입, 결실, 아미노산 치환 또는 다른 표적화된 돌연변이를 갖는 번역된 Pex 단백질 중 어느 하나가 야기되는 경우 발생한다. 하향 조절되는 Pex 단백질은 하향 조절되지 않은 Pex 단백질에 대하여 손상된 활성을 가질 것이며, 비-기능성일 수 있다. Pex 단백질 발현을 낮추거나 결여되게 하는 하향 조절은 또한, 조절 서열, 전사 및 번역 인자 및/또는 신호 전달 경로의 조작을 통해, 또는 센스, 안티센스 또는 RNAi 기술 등의 사용에 의해 유발할 수 있다.

[0176] 용어 "보존된 도메인" 또는 "모티프"는 진화적으로 관련된 단백질의 정렬된 서열을 따라 특정 위치에서 보존된 일련의 아미노산을 의미한다. 다른 위치의 아미노산은 상동성 단백질들 간에 달라질 수 있는 반면에, 특정 위치에서 고도로 보존된 아미노산은 이들 아미노산이 단백질의 구조, 안정성, 또는 활성에 있어서 중요할 수 있음을 나타낸다. 이들은 단백질 상동체의 패밀리 of 정렬된 서열에서의 이들의 높은 보존도에 의해 확인되기 때문에, 이들은 새로 결정된 서열을 갖는 단백질이 이전에 동정된 단백질 패밀리에 속하는지 여부를 결정하기 위한 식별자(identifier) 또는 "표시자(signature)"로 사용될 수 있다.

[0177] 용어 "미생물 숙주 세포" 및 "미생물 숙주 유기체"는 본 명세서에서 상호교환가능하게 사용되며, 외래 또는 이종 유전자를 받아들일 수 있으며 그 유전자를 발현할 수 있는 미생물을 지칭한다. "재조합 미생물 숙주 세포"는 재조합으로 엔지니어링되는 미생물 숙주 세포를 지칭한다.

[0178] 일반적으로, 용어 "유지성"은 유기체의 에너지를 오일의 형태로 저장하는 경향이 있는 유기체를 지칭한다(문헌[Weete, In: *Fungal Lipid Biochemistry*, 2nd Ed., Plenum, 1980]). 이러한 과정 동안, 유지성 미생물의 세포 오일 함량은 일반적으로 S자 모양의 곡선을 따르며, 여기서, 지질의 농도는 그것이 후기 대수 성장기 또는 초기 정지 성장기에서 최대에 도달할 때까지 증가하였다가 후기 정지기 및 사멸기 동안에 점차적으로 감소한다(문헌[Yongmanitchai and Ward, *Appl. Environ. Microbiol.*, 57:419-25 (1991)]). 용어 "유지성"은 본 출원의 목적을 위해, 미생물의 건조 세포 중량["DCW"]의 적어도 약 25%를 오일로 축적할 수 있는 미생물을 지칭한다.

[0179] 용어 "유지성 효모"는 오일을 만들 수 있는 효모로 분류되는 이러한 유지성 미생물을 지칭하며, 다시 말하면, 여기서, 효모의 DCW의 약 25%의 과량으로 오일이 축적될 수 있다. 유지성 효모의 예에는 하기의 속이 포함되나, 이에 제한되는 것을 의미하는 것은 아니다: 야로위아, 칸디다(*Candida*), 로도토룰라(*Rhodotorula*), 로도스포리툼(*Rhodospiridium*), 크립토크커스(*Cryptococcus*), 트리코스포론(*Trichosporon*) 및 리포마이세스(*Lipomyces*). 효모의 DCW의 약 25% 과량의 오일을 축적하는 능력은 재조합 엔지니어링(engineering)의 결과를 통하여 또는 유기체의 천연의 능력을 통하여 이루어질 수 있다.

[0180] 용어 "발효가능한 탄소원"은 미생물이 대사작용시켜, 에너지를 도출할 탄소원을 의미한다. 전형적인 탄소원은

단당류, 이당류, 올리고당류, 다당류, 알칸, 지방산, 지방산의 에스테르, 글리세롤, 모노글리세리드, 다이글리세리드, 트라이글리세리드, 이산화탄소, 메탄올, 포름알데히드, 포르메이트 및 탄소-함유 아민을 포함하나 이에 한정되지 않는다.

- [0181] 용어 "폴리뉴클레오티드", "폴리뉴클레오티드 서열", "핵산 서열", "핵산 단편" 및 "분리된 핵산 단편"은 본 명세서에서 상호교환가능하게 사용된다. 이들 용어는 뉴클레오티드 서열 등을 포함한다. 폴리뉴클레오티드는 합성, 비-천연 뉴클레오티드 염기 또는 변경된 뉴클레오티드 염기를 임의로 함유하는 단일- 또는 이중-가닥인 RNA 또는 DNA의 폴리머일 수 있다. DNA의 폴리머 형태인 폴리뉴클레오티드는 cDNA, 게놈 DNA, 합성 DNA, 또는 그들의 혼합물의 하나 이상의 세그먼트로 이루어질 수 있다. 뉴클레오티드(통상적으로 그들의 5'-모노포스페이트 형태로 발견됨)는 하기와 같이 1-문자 표기에 의해 지칭된다: 아데닐레이트 또는 데옥시아데닐레이트(각각 RNA 또는 DNA에 대하여)에 대하여 "A", 시티딜레이트 또는 데옥시시티딜레이트에 대하여 "C", 구아닐레이트 또는 데옥시구아닐레이트에 대하여 "G", 유리딜레이트에 대하여 "U", 데옥시티미딜레이트에 대하여 "T", 퓨린(A 또는 G)에 대하여 "R", 피리미딘(C 또는 T)에 대하여 "Y", G 또는 T에 대하여 "K", A 또는 C 또는 T에 대하여 "H", 이노신에 대하여 "I", 그리고 임의의 뉴클레오티드에 대하여 "N".
- [0182] 아미노산 또는 뉴클레오티드 서열의 "상당한 부분"은, 당업자에 의한 서열의 수동 평가에 의해, 또는 BLAST(기본 국소 정렬 검색 도구(Basic Local Alignment Search Tool); 문헌[Altschul, S. F., et al., *J. Mol. Biol.*, 215:403-410(1993)])와 같은 알고리즘을 사용하는 컴퓨터-자동화 서열 비교 및 동정에 의해, 미국 특허 출원 공개 제2010-0317072-A1호에 기재된 바와 같이 폴리펩티드 또는 유전자를 추정상으로 동정하기에 충분한 폴리펩티드의 아미노산 서열 또는 유전자의 뉴클레오티드 서열을 포함하는 부분이다.
- [0183] 용어 "상보성"은 서로 혼성화될 수 있는 뉴클레오티드 염기 간의 관계를 기술하기 위해 사용된다. 예를 들어, DNA에 있어서 아데노신은 티민에 상보성이고 시토신은 구아닌에 상보성이다.
- [0184] "코돈 축퇴성(Codon degeneracy)"은 암호화된 폴리펩티드의 아미노산 서열에 영향을 주지 않으면서 뉴클레오티드 서열의 변이를 허용하는 유전 암호의 성질을 지칭한다. 주어진 아미노산을 특정하기 위한 뉴클레오티드 코돈의 사용에 있어서 특정 숙주 세포에서 나타나는 "코돈-편향(codon-bias)"이 당업자에게 주지되어 있다. 그러므로, 숙주 세포 내의 발현을 향상시키기 위해 유전자를 합성할 경우에는, 그의 코돈 사용 빈도가 숙주 세포가 선호하는 코돈의 사용 빈도에 근접하도록 유전자를 고안하는 것이 바람직하다.
- [0185] "합성 유전자"는 숙주 세포의 코돈 편향을 반영하기 위하여 뉴클레오티드 서열의 최적화에 기초하여 최적의 유전자 발현을 위해 맞춤화된다. 당업자는 코돈 사용이 숙주가 선호하는 코돈에 편향된다면, 성공적인 유전자 발현이 가능함을 인식한다. 바람직한 코돈의 결정은 서열 정보를 이용할 수 있는 숙주 세포로부터 유래된 유전자의 조사에 기초할 수 있다. 예를 들어, 야로위아 리폴리티카에 대한 코돈 사용 프로파일은 미국 특허 제 7,125,672호에 제공되어 있다.
- [0186] "유전자"는 특정 단백질을 발현하며, 단독의 암호화 영역을 지칭하거나 암호화 영역 업스트림 및/또는 다운스트림의 조절 서열(예컨대, 암호화 영역의 전사 시작 부위의 업스트림 5' 미번역 영역, 3' 비-암호화 영역)을 포함할 수 있는 핵산 단편을 지칭한다. "천연 유전자"는 그 자신의 조절 서열과 함께 자연에서 발견되는 유전자를 지칭한다. "키메라 유전자"는 자연에서 함께 발견되지 않는 조절 및 암호화 서열을 포함하는, 천연 유전자가 아닌 임의의 유전자를 지칭한다. 따라서, 키메라 유전자는 상이한 공급원으로부터 유래된 조절 서열과 암호화 서열, 또는 동일한 공급원으로부터 유래되었지만, 자연에서 발견되는 것과 상이한 방식으로 배열된 조절 서열과 암호화 서열을 포함할 수 있다. "내인성 유전자"는 유기체의 게놈에서 그의 자연의 위치에 있는 천연 유전자를 지칭한다. "외래" 유전자는 유전자 전달에 의해 숙주 유기체 내로 도입된 유전자를 지칭한다. 외래 유전자는 비-천연 유기체에 삽입된 천연 유전자, 천연 숙주 내의 새로운 위치에 도입된 천연 유전자, 또는 키메라 유전자를 포함할 수 있다. "트랜스유전자(transgene)"는 형질전환 절차에 의해 게놈 내로 도입된 유전자이다. "코돈-최적화된 유전자"는 그의 코돈 사용 빈도가 숙주 세포의 선호하는 코돈 사용의 빈도를 모방하도록 설계된 유전자이다.
- [0187] "암호화 서열"은 특정 아미노산 서열을 암호화하는 DNA 서열을 말한다. "조절 서열"은 암호화 서열의 전사 시작 부위의 업스트림, 5' 미번역 영역 및 3' 비암호화 영역에 위치하고, 전사, RNA 프로세싱 또는 안정성, 또는 관련 암호화 서열의 번역에 영향을 미칠 수 있는 뉴클레오티드 서열을 지칭한다. 조절 서열은 프로모터, 인핸서(enhancer), 사일런서(silencer), 5' 미번역 리더 서열(leader sequence), 인트론, 폴리아데닐화 인식 서열, RNA 프로세싱 부위, 이펙터(effector) 결합 부위 및 스템-루프(stem-loop) 구조를 포함할 수 있다.

- [0188] "프로모터"는 암호화 서열 또는 기능적 RNA의 발현을 조절할 수 있는 DNA 서열을 지칭한다. 일반적으로, 프로모터 서열은 암호화 서열의 5' 업스트림이다. 프로모터는 천연 유전자로부터 그 전체가 유래될 수도 있거나, 자연에서 발견되는 상이한 프로모터로부터 유래된 상이한 요소로 구성될 수도 있거나, 심지어 합성 DNA 세그먼트를 포함할 수도 있다. 상이한 프로모터가 상이한 조직 또는 세포 유형에서, 또는 상이한 세포 성장 및/또는 발생 단계에서, 또는 상이한 환경 또는 생리 조건에 반응하여 유전자의 발현을 유도할 수 있음을 담당자는 이해할 것이다. 대부분의 세포 유형에서 대부분의 시간에 유전자가 발현되도록 하는 프로모터는 통상적으로 "구성적 프로모터"라고 지칭된다. 대부분의 경우에 조절 서열의 정확한 경계가 완벽하게 규정되지 않았으므로, 상이한 길이의 DNA 단편이 동일한 프로모터 활성을 가질 수 있음이 추가로 인정된다.
- [0189] 용어 "3' 비-암호화 서열" "전사 종결자(terminator)" 및 "종결자"는 암호화 서열의 3' 다운스트림에 위치한 DNA 서열을 지칭한다. 이는 폴리아데닐화 인식 서열 및 mRNA 프로세싱 또는 유전자 발현에 영향을 미칠 수 있는 조절 신호를 암호화하는 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 암호화 서열의 번역에 영향을 미칠 수 있다.
- [0190] 용어 "작동가능하게 연결된"은 하나의 기능이 나머지에 의해 영향을 받도록 하는, 단일 핵산 단편 상에서의 핵산 서열의 연계(association)를 지칭한다. 예를 들어, 프로모터는 그것이 암호화 서열의 발현에 영향을 미칠 수 있는 경우에 암호화 서열에 작동가능하게 연결된다. 다시 말하면, 암호화 서열은 프로모터의 전사적 조절 하에 있다. 암호화 서열은 센스(sense) 또는 안티센스(antisense) 배향으로 조절 서열에 작동가능하게 연결될 수 있다.
- [0191] 용어 "재조합"은, 예를 들어 유전 공학 기술에 의한 분리된 핵산 세그먼트의 조작에 의하거나 화학적 합성에 의한, 다르게 분리되어 있는 2개의 서열 세그먼트의 인공의 조합을 지칭한다.
- [0192] 본 명세서에 사용되는 용어 "발현"은 센스(mRNA) 또는 안티센스 RNA의 전사 및 안정적인 축적을 지칭한다. 또한, 발현은 mRNA의 폴리펩티드로의 번역도 포함한다.
- [0193] "형질전환"은 핵산 분자의 숙주 유기체로의 전달을 지칭한다. 핵산 분자는 자가 복제되는 플라스미드일 수 있거나; 그것이 숙주 유기체의 게놈 내로 통합될 수 있다. 형질전환된 핵산 단편을 함유하는 숙주 유기체는 "트랜스제닉" 또는 "재조합" 또는 "형질전환된" 유기체 또는 "형질전환체"로 지칭된다.
- [0194] "안정적인 형질전환"은 핵산 단편을 핵 및 세포소기관 게놈 둘 모두를 포함하는 숙주 유기체의 게놈 내로 전달하여, 유전적으로 안정적인 유전을 야기하는 것을 말한다(즉, 핵산 단편은 "안정적으로 통합된다"). 반대로, "일시적 형질전환"은, 숙주 유기체의 핵, 또는 DNA-함유 소기관 내로 핵산 단편을 전달하여, 통합 또는 안정적인 유전성이 없는 유전자 발현을 유발하는 것을 지칭한다.
- [0195] 용어 "플라스미드" 및 "벡터"는 종종 세포의 중추적 대사의 일부가 아닌 유전자를 운반하는 보통 원형의 이중 가닥 DNA 단편 형태의 염색체 외 요소를 지칭한다. 이러한 요소는 자가 복제되는 서열, 게놈 통합 서열, 파지 또는 뉴클레오티드 서열을 가질 수 있으며, 임의의 공급원으로부터 유래되는 선형 또는 원형의 단일- 또는 이중-가닥 DNA 또는 RNA일 수 있고, 여기서, 다수의 뉴클레오티드 서열은 세포 내로 발현 카세트(들)를 도입할 수 있는 독특한 구조물 내로 결합되거나 재조합된다.
- [0196] 용어 "발현 카세트"는, 선택된 유전자의 암호화 서열 및 선택된 유전자 생성물의 발현을 위해 필요한 암호화 서열에 선행하는 조절 서열(5' 비-암호화 서열) 및 후행하는 조절 서열(3' 비-암호화 서열)을 포함하는 DNA의 단편을 지칭한다. 따라서, 발현 카세트는 전형적으로, 1) 프로모터; 2) 암호화 서열(즉, ORF); 및 3) 통상 진핵 생물에서의 폴리아데닐화 부위를 함유하는 종결자로 이루어진다. 발현 카세트(들)는 통상적으로 벡터 내에 포함되어, 클로닝 및 형질전환을 용이하게 한다. 정확한 조절 서열이 각각의 숙주에 대하여 사용되는 한, 상이한 발현 카세트가 박테리아, 효모, 식물 및 포유류 세포를 비롯한 상이한 유기체 내로 형질전환될 수 있다.
- [0197] 용어 "서열 분석 소프트웨어"는 뉴클레오티드 또는 아미노산 서열을 분석하는데 유용한 임의의 컴퓨터 알고리즘 또는 소프트웨어 프로그램을 지칭한다. "서열 분석 소프트웨어"는 상업적으로 입수할 수 있거나 독립적으로 개발할 수 있다. 통상적인 서열 분석 소프트웨어는 하기의 것들을 포함할 것이나 이에 한정되지 않는다: 1.) GCG 프로그램 모음(위스콘신 패키지 버전 9.0, 위스콘신주 메디슨 소재의 제네틱스 컴퓨터 그룹(GCG: Genetics Computer Group)); 2) BLASTP, BLASTN, BLASTX(문헌[Altschul et al., *J. Mol Biol.*, 215:403-410 (1990)]); 3) DNASTAR(미국 위스콘신주 메디슨 소재의 디엔에이스타, 인코포레이티드(DNASTAR, Inc.)); 4) 시퀀처(Sequencher)(미국 미시간주 앤 아버 소재의 진 코즈 코포레이션(Gene Codes Corporation)); 및 5) 스미스-워

터만(Smith-Waterma) 알고리즘을 통합하는 파스타(FASTA) 프로그램(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]). 본 출원과 관련하여, 서열 분석 소프트웨어가 분석을 위해 사용된 경우에는, 그 분석 결과가 달리 명시되지 않는 한, 참조한 프로그램의 "디폴트 값"에 기초할 것임을 이해할 것이다. 본 명세서에 사용되는 바와 같이, "디폴트 값"은 최초 초기화할 때에 소프트웨어에 원래 로딩된 값 또는 파라미터의 임의의 세트를 의미할 것이다.

[0198] 핵산 또는 폴리펩티드 서열과 관련하여 "서열 동일성" 또는 "동일성"은, 특정 비교창에 걸쳐 최대 상응도로 정렬되는 경우에 2개 서열 중의 동일한 핵산 염기 또는 아미노산 잔기를 지칭한다. 따라서, "서열 동일성의 백분율" 또는 "동일성 백분율"은 비교창에 걸쳐 최적으로 정렬된 2개의 서열을 비교함으로써 결정된 값을 지칭하며, 여기서 비교창 내의 폴리뉴클레오타이드 또는 폴리펩티드 서열의 부분은 2개 서열의 최적 정렬을 위한 기준 서열(삽입 또는 결실을 포함하지 않음)과 비교하여 삽입 또는 결실(즉, 갭)을 포함할 수 있다. 양자 모두의 서열 내에서 동일한 핵산 염기 또는 아미노산 잔기가 나타나는 위치의 개수를 결정하여 일치하는 위치의 개수를 산출하고, 일치하는 위치의 개수를 비교창 내의 위치의 총 개수로 나누고, 그 결과에 100을 곱하여 서열 동일성의 백분율을 산출함으로써 백분율을 계산한다.

[0199] "동일성 백분율" 및 "유사성 백분율"의 결정 방법은 공중 이용가능한 컴퓨터 프로그램에 코드화되어 있다. 동일성 백분율 및 유사성 백분율은 하기의 문헌에 기재된 것을 포함하나 이에 한정되지 않는 공지의 방법에 의해 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)].

[0200] 서열 정렬 및 동일성 또는 유사성 백분율 계산은, 레이저진 바이오인포매틱스 컴퓨팅 스위트(LASERGENE bioinformatics computing suite)(미국 위스콘신주 메디슨 소재의 디엔에이스타 인코포레이티드)의 맥앨라인(MegAlign)(상표명) 프로그램을 포함하나 이에 한정되지 않는, 상동성 서열을 검출하도록 설계된 다양한 비교 방법을 사용하여 결정할 수 있다. 대안적으로, "BLASTN 정렬 방법"은 디폴트 파라미터를 사용하여 뉴클레오타이드 서열을 비교하기 위해 국립생물정보센터[NCBI]에 의해 제공되는 알고리즘이며, "BLASTP 정렬 방법"은 디폴트 파라미터를 사용하여 단백질 서열을 비교하기 위해 NCBI에 의해 제공되는 알고리즘이다.

[0201] 본 명세서에 사용되는 표준 재조합 DNA 및 분자 클로닝 기술은 당업계에 주지되어 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)(이후에는 "Maniatis"); 문헌[Silhavy, T. J., Bennis, M. L. and Enquist, L. W., *Experiments with Gene Fusions*, Cold Spring Harbor Laboratory: 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, Hoboken, NJ (1987)]에 기재되어 있다.

[0202] 도 1a 및 도 1b에는 후술되는 바와 같이, EPA 생성을 위한 다수의 경로가 함께 기재되어 있다. 모든 경로는 델타-12 불포화화효소에 의해 올레산이 처음의 오메가-6 지방산인 리놀렌산[LA]으로 초기 전환되는 것을 필요로 한다. 이어서, "델타-9 연장효소/델타-8 불포화화효소 경로" 및 기질로서 LA를 사용하여, 장쇄 오메가-6 지방산이 하기와 같이 형성된다: 1) LA는 델타-9 연장효소에 의해 에이코사다이엔산[EDA]으로 전환되고; 2) EDA는 델타-8 불포화화효소에 의해 다이호모- γ -리놀렌산[DGLA]으로 전환되고; 3) DGLA는 델타-5 불포화화효소에 의해 아라키돈산[ARA]으로 전환되고; 4) ARA는 C_{20/22} 연장효소에 의해 도코사테트라엔산[DTA]으로 전환되고; 5) DTA는 델타-4 불포화화효소에 의해 도코사펜타엔산[DPA-6]으로 전환된다.

[0203] "델타-9 연장효소/델타-8 불포화화효소 경로"는 또한 하기와 같이 기질로서 α -리놀렌산[ALA]을 사용하여 장쇄 오메가-3 지방산을 생성할 수 있다: 1) LA는 델타-15 불포화화효소에 의해 처음의 오메가-3 지방산인 ALA로 전환되고; 2) ALA는 델타-9 연장효소에 의해 에이코사트라이엔산[ETra]으로 전환되고; 3) ETra는 델타-8 불포화화효소에 의해 에이코사테트라엔산[ETA]으로 전환되고; 4) ETA는 델타-5 불포화화효소에 의해 에이코사펜타엔산[EPA]으로 전환되고; 5) EPA는 C_{20/22} 연장효소에 의해 도코사펜타엔산[DPA]으로 전환되고; 6) DPA는 델타-4 불포화화효소에 의해 도코사헥사엔산[DHA]으로 전환된다. 임의로, 오메가-6 지방산은 오메가-3 지방산

으로 전환될 수 있다. 예를 들어, ETA 및 EPA는 델타-17 불포화화효소 활성에 의해 각각 DGLA 및 ARA로부터 생성된다. 유리하게는, 본 명세서의 목적을 위하여, 델타-9 연장효소/델타-8 불포화화효소 경로는 상당한 양의 γ -리놀렌산["GLA"]이 결여된 EPA 오일의 생성을 가능하게 한다.

[0204] 오메가-3/오메가-6 지방산의 생합성을 위한 대안적인 경로는 델타-6 불포화화효소 및 $C_{18/20}$ 연장효소, 즉, "델타-6 불포화화효소/델타-6 연장효소 경로"를 사용한다. 더욱 구체적으로, LA 및 ALA는 델타-6 불포화화효소에 의해 각각 GLA 및 스테아리돈산["STA"]으로 전환된 다음; $C_{18/20}$ 연장효소는 GLA로부터 DGLA로 전환시키고/거나 STA로부터 ETA로 전환시킬 수 있다.

[0205] 제조합 미생물 숙주 세포에서의 EPA의 경제적인 상업적 생성은 EPA 농도["EPA % TFA"], 전체 지질 함량["TFA % DCW"] 및 EPA 생산성["EPA % DCW"]을 포함하는 다양한 변수의 고려를 필요로 한다. 더욱이, 최종 오일 생성물 중의 중간체 지방산 및 부산물 지방산의 생성을 줄여, 목적하는 지방산, 즉, EPA의 생성을 최대화시키는 것이 바람직하다.

[0206] 중간체 지방산은 다른 대사 경로 효소의 작용에 의하여 EPA로 추가로 전환될 수 있는 지방산(예컨대, 올레산, LA, ALA, EDA, DGLA, ETA)이다. 대조적으로, 부산물 지방산(예컨대, 스시아돈산(sciaidonic acid), 주니페론산(juniperonic acid))은 EPA로 추가로 전환될 수 있는 중간체 지방산 또는 EPA 중 어느 것도 아닌 임의의 생성되는 지방산을 말한다.

[0207] 미국 특허 출원 공개 제2009-0093543-A1호에는 적어도 약 43.3 EPA % TFA와 함께, 약 23.6 미만의 LA % TFA(1.83의 EPA:LA 비)를 포함하는 미생물 오일을 생성하는 능력을 갖는 제조합 야로위아 리폴리티카의 최적화된 균주가 기재되어 있다. 바람직한 균주는 Y4305였으며, 그의 최대 생성은 55.6 EPA % TFA와 함께, EPA:LA 비는 3.03이었다. 일반적으로, 미국 특허 출원 공개 제2009-0093543-A1호의 EPA 균주는 하기의 오메가-3/오메가-6 지방산 생합성 경로의 유전자를 포함한다:

[0208] a) 델타-9 연장효소를 암호화하는 적어도 하나의 유전자; 및

[0209] b) 델타-8 불포화화효소를 암호화하는 적어도 하나의 유전자; 및

[0210] c) 델타-5 불포화화효소를 암호화하는 적어도 하나의 유전자; 및

[0211] d) 델타-17 불포화화효소를 암호화하는 적어도 하나의 유전자; 및

[0212] e) 델타-12 불포화화효소를 암호화하는 적어도 하나의 유전자; 및

[0213] f) $C_{16/18}$ 연장효소를 암호화하는 적어도 하나의 유전자; 및

[0214] g) 임의로, 다이아실글리세롤 콜린포스포트랜스퍼라제(CPT1)를 암호화하는 적어도 하나의 유전자.

[0215] 상술된 효소 기능을 갖는 바람직한 유전자의 예는 미국 특허 출원 공개 제2009-0093543-A1호의 표 3에 기재되어 있다. 이들 유전자를 제한하고자 하지 않으며; 대신에, 미국 특허 출원 공개 제2009-0093543-A1호의 유전자는 유용한 참고물질로 사용되어 델타-9 연장효소, 델타-8 불포화화효소, 델타-5 불포화화효소, 델타-17 불포화화효소, 델타-12 불포화화효소, $C_{16/18}$ 연장효소 및/또는 CPT1 기능을 갖는 오메가-3/오메가-6 지방산 생합성 경로의 적절한 유전자의 선택을 안내할 것이다.

[0216] 당업자는 각 특정 숙주 세포가 주어진 아미노산을 특정하기 위한 뉴클레오티드 코돈의 사용에서 "코돈-편향"을 나타낼 것임을 인식할 것이다. 따라서, 각각의 특정 델타-9 연장효소, 델타-8 불포화화효소, 델타-5 불포화화효소, 델타-17 불포화화효소, 델타-12 불포화화효소, $C_{16/18}$ 연장효소 및/또는 CPT1 유전자를 그의 코돈 사용 빈도가 EPA의 생성을 위해 엔지니어링될 제조합 미생물 숙주 세포가 선호하는 코돈 사용 빈도에 접근하도록 설계하는 것이 바람직할 것이다.

[0217] 동일하거나 유사한 기능 또는 활성을 가진 폴리펩티드를 다른 종으로부터 동정함에 있어서 여러 수준의 서열 동일성이 유용하다는 것이 당업자에게 주지되어 있다. 바람직한 핵산 단편, 즉, 델타-9 연장효소, 델타-8 불포화화효소, 델타-5 불포화화효소, 델타-17 불포화화효소, 델타-12 불포화화효소, $C_{16/18}$ 연장효소 및/또는 CPT1 폴리펩티드를 암호화하는 분리된 폴리뉴클레오티드는 미국 특허 출원 공개 제2009-0093543-A1호의 표 3에 기재된 것과 적어도 약 70-80% 동일한 폴리펩티드를 암호화하는 한편, 더욱 바람직한 핵산 단편은 적어도 약 80-85% 또는 적어도 약 85-90%, 또는 심지어는 적어도 약 90-95% 동일한 아미노산 서열을 암호화한다.

- [0218] 미국 특허 출원 공개 제2010-0317072-A1호에는 EPA % TFA 및 EPA:LA의 비에 기초하여, 미국 특허 출원 공개 제2009-0093543-A1호에 기재된 균주와 비교하여 미생물 오일을 생성하는 능력이 향상된 재조합 야로위아 리폴리티카의 최적화된 균주가 기재되어 있다. 바람직한 균주는 Y9502였으며, 그의 최대 생성은 57 EPA % TFA와 함께, 4.49의 EPA:LA 비 및 21.3 EPA % DCW의 EPA 생산성이었다. 미국 특허 출원 공개 제2009-0093543-A1호에 상세히 기재된 오메가-3/오메가-6 지방산 생합성 경로의 유전자의 발현에 더하여, 이들 향상된 균주는 다음에 의해 구별된다:
- [0219] 1) 적어도 하나의 지방산 델타-8 불포화화효소에 연결된 적어도 하나의 지방산 델타-9 연장효소를 갖는 폴리펩티드["DGLA 신타제"]를 포함하는 적어도 하나의 멀티자임의 포함;
- [0220] 2) 임의로, 말로닐 CoA 신테타제["MCS"] 또는 아실-CoA 라이소인지질 아실트랜스퍼라제["LPLAT"]로 이루어진 군으로부터 선택되는 효소를 암호화하는 적어도 하나의 폴리뉴클레오티드의 포함;
- [0221] 3) 발현이 하향-조절되는 적어도 하나의 피옥시좀 생합성 인자 단백질의 포함;
- [0222] 4) 적어도 약 50 EPA % TFA의 생성; 및
- [0223] 5) 적어도 약 3.1의 EPA:LA의 비를 가짐.
- [0224] 멀티자임 링커는 바람직하게는 서열 번호 4, 서열 번호 5, 서열 번호 6, 서열 번호 7, 서열 번호 8, 서열 번호 9 및 서열 번호 10으로 이루어진 군으로부터 선택되며; 멀티자임은 바람직하게는 본질적으로 EgD9eS/EgD8M(서열 번호 12), EaD9eS/EaD8S(서열 번호 14) 및 E389D9eS/EgD8M(서열 번호 16)으로 이루어진 군으로부터 선택되는 서열로 이루어진 서열이다. 발현이 하향 조절되는 적어도 하나의 피옥시좀 생합성 인자 단백질은 바람직하게는 Pex1p(서열 번호 51), Pex2p(서열 번호 52), Pex3p(서열 번호 53), Pex3Bp(서열 번호 54), Pex4p(서열 번호 55), Pex5p(서열 번호 56), Pex6p(서열 번호 57), Pex7p(서열 번호 58), Pex8p(서열 번호 59), Pex10p(서열 번호 60), Pex12p(서열 번호 61), Pex13p(서열 번호 62), Pex14p(서열 번호 63), Pex16p(서열 번호 64), Pex17p(서열 번호 65), Pex19p(서열 번호 66), Pex20p(서열 번호 67), Pex22p(서열 번호 68) 및 Pex26p(서열 번호 69)로 이루어진 군으로부터 선택되며, Pex3p 낙-아웃이 특히 바람직하다.
- [0225] 상기 서열이 재조합 야로위아 숙주 세포에서 사용하기에 바람직하지만, 이들 유전자는 제한하고자 하는 것이 아니다. 재조합 미생물 숙주 세포 내에서 발현되는 특정 불포화화효소 및 연장효소에 관하여 이전에 논의된 바와 같이, 주어진 아미노산을 특정하기 위한 뉴클레오티드 코돈의 사용에서 "코돈-편향"이 고려되어야 한다. 따라서, EPA의 생성을 위해 엔지니어링되어야 하는 임의의 재조합 미생물 숙주 세포에 대한 바람직한 Pex 유전자 낙아웃 및 DGLA 신타제 멀티자임의 설계에 관해서는 미국 특허 출원 공개 제2010-0317072-A1호에 기재된 교시를 적용할 수 있다. 바람직한 핵산 단편, 즉, DGLA 신타제 및/또는 Pex 폴리펩티드를 암호화하는 분리된 폴리뉴클레오티드는 상술된 것과 적어도 약 70 내지 80% 동일한 폴리펩티드를 암호화하는 한편, 더욱 바람직한 핵산 단편은 적어도 약 80 내지 85% 또는 적어도 약 85 내지 90% 또는 심지어는 적어도 약 90 내지 95% 동일한 아미노산 서열을 암호화한다.
- [0226] 증가된 EPA 생산성(즉, 증가된 EPA % DCW로 측정)에 기초하여, 미국 특허 출원 공개 제2009-0093543-A1호 및 미국 특허 출원 공개 제2010-0317072-A1호에 기재된 균주와 비교하여 미생물 오일을 생성하는 능력이 향상된 추가의 향상된 최적화 재조합 미생물 숙주 세포가 본 명세서에 제공된다. 적어도 하나의 멀티자임(여기서, 상기 멀티자임은 미국 특허 출원 공개 제2010-0317072-A1호에 기재된 바와 같이 적어도 하나의 델타-8 불포화화효소에 연결된 적어도 하나의 델타-9 연장효소를 갖는 폴리펩티드를 포함한다)을 포함하는 오메가-3/오메가-6 지방산 생합성 경로의 유전자를 발현하고, 발현이 하향-조절되는 적어도 하나의 피옥시좀 생합성 인자 단백질(미국 특허 출원 공개 제2010-0317072-A1호에 기재된 바와 같은)을 포함하는 것에 더하여, 본 명세서에 개시된 향상된 재조합 미생물 숙주 세포는 하기에 의해 구별된다:
- [0227] 1) 적어도 라이소포스파티드산 아실트랜스퍼라제["LPAAT"] 활성을 갖는 적어도 2개의 폴리펩티드의 포함;
- [0228] 2) 적어도 인지질:다이아실글리세롤 아실트랜스퍼라제["PDAT"] 활성을 갖는 적어도 하나의 폴리펩티드의 포함;
- [0229] 3) 임의로, 적어도 하나의 돌연변이 델타-9 연장효소 폴리펩티드의 포함(여기서, 상기 돌연변이 델타-9 연장효소 폴리펩티드는 서열 번호 1에 기재된 아미노산 서열을 포함하며, 서열 번호 1은 적어도 하나의 아미노산 돌연변이에 의해 서열 번호 3과 상이하며, 상기 돌연변이(들)은
- [0230] i) L35F 돌연변이;

- [0231] ii) L35M 돌연변이;
- [0232] iii) L35G 돌연변이;
- [0233] iv) L35G 돌연변이, 및 S9A, S9D, S9G, S9I, S9K, S9Q, Q12K, A21D, A21T, A21V, V32F, Y84C, Q107E, L108G, G127L, W132T, M143N, M143W, L161T, L161Y, W168G, I179M, I179R, C236N, Q244N, A254W 및 A254Y로 이루어진 군으로부터 선택되는 적어도 하나의 다른 돌연변이;
- [0234] v) L35G, A21V, L108G 및 I179R 돌연변이;
- [0235] vi) L35G, W132T 및 I179 돌연변이;
- [0236] vii) L35G, S9D, Y84C 및 I179R 돌연변이;
- [0237] viii) L35G, Y84C, I179R 및 Q244N 돌연변이;
- [0238] ix) L35G, A21V, W132T, I179R 및 Q244N 돌연변이;
- [0239] x) K58R 및 I257T 돌연변이;
- [0240] xi) D98G 돌연변이;
- [0241] xii) L130M 및 V243A 돌연변이; 및
- [0242] xiii) K58R, L35F, L35G, L35M, S9A, S9D, S9G, S9I, S9K, S9Q, Q12K, A21D, A21T, A21V, V32F, Y84C, D98G, Q107E, L108G, G127L, L130M, W132T, M143N, M143W, L161T, L161Y, W168G, I179M, I179R, C236N, V243A, Q244N, A254W, A254Y 및 I257T로 이루어진 군으로부터 선택되는 적어도 2개의 돌연변이를 포함하는 임의의 조합으로 이루어진 군으로부터 선택됨); 및
- [0243] 4) DCW의 wt%로 측정하여 적어도 25wt%의 EPA를 포함하는 미생물 오일의 생성.
- [0244] 상술된 바와 같이, 본 명세서에 개시된 향상된 재조합 미생물 숙주 세포는 군주가 적어도 LPAAT 활성을 갖는 적어도 2개의 폴리펩티드 및 적어도 PDAT 활성을 갖는 적어도 하나의 폴리펩티드를 갖는다는 점에서 독특하며; 이에 따라, 데 노보(de novo) 글리세로인지질 생합성 경로의 성분의 상향-조절이 요망된다.
- [0245] 생물학적 막의 주요 성분인 글리세로인지질은 *sn*-1 위치 및 *sn*-2 위치에서 R 기로 부착된 지방산을 갖는 글리세롤 코어(core), 및 포스포다이에스테르 결합을 통하여 *sn*-3 위치에서 결합된 극성 헤드기(head group)를 포함한다(미국 특허 출원 공개 제2010-0317882-A1호 참조). 글리세로인지질은 다양한 포스포릴 헤드기로부터 야기될 뿐 아니라 상이한 쇠 길이 및 그들의 지방산의 포화도의 결과로서의 엄청난 다양성을 갖는다. 일반적으로, 포화 지방산 및 단일불포화 지방산은 *sn*-1 위치에서 에스테르화되는 한편, PUFA는 *sn*-2 위치에서 에스테르화된다.
- [0246] 글리세로인지질 생합성은 복잡하다. 하기의 표 3은 원래 케네디(Kennedy) 및 웨이스(Weiss)에 의해 기재된 데 노보 경로 내의 단계를 요약한 것이다(문헌[*J. Biol. Chem.*, 222:193-214 (1956)]):

표 3

테 노보 글리세로인지질 생합성의 일반 반응

반응	촉매작용 효소
<i>sn</i> -글리세롤-3-포스페이트 → 라이소포스파티드산 (1-아실- <i>sn</i> -글리세롤 3-포스페이트 또는 "LPA")	글리세롤-3-포스페이트 아실트랜스퍼라제(GPAT)[E.C. 2.3.1.15]는 처음의 아실-CoA를 <i>sn</i> -글리세롤 3-포스페이트의 <i>sn</i> -1 위치로 에스테르화시킨다.
LPA → 포스파티드산 (1,2-다이아실글리세롤 포스페이트 또는 "PA")	라이소포스파티드산 아실트랜스퍼라제(LPAAT)[E.C. 2.3.1.51]는 2 번째 아실-CoA를 LPA의 <i>sn</i> -2 위치로 에스테르화시킨다.
PA → 1,2-다이아실글리세롤("DAG") 또는 PA → 시티딘 다이포스페이트 다이아실글리세롤("CDP-DG")	포스파티드산 포스파타제[E.C. 3.1.3.4]는 PA로부터 포스페이트를 제거하며; DAG는 이후에 포스파티딜콜린["PC"], 포스파티딜에탄올아민["PE"], 또는 트라이아실글리세롤["TAG"]로 전환될 수 있고, 여기서, TAG 합성은 다이아실글리세롤 아실트랜스퍼라제(DGAT)[E.C. 2.3.1.20] 또는 인지질:다이아실글리세롤 아실트랜스퍼라제(PDAT)[E.C.2.3.1.158] 중 어느 하나를 필요로 한다. CDP-다이아실글리세롤 신타제[EC 2.7.7.41]는 피로포스페이트의 제거와 함께 PA 및 시티딘 트라이포스페이트의 축합을 야기하며; CDP-DG는 이후에 포스파티딜-이노시톨["PI"], 포스파티딜세린["PS"], 포스파티딜글리세롤["PG"] 또는 카르디올리핀["CL"]으로 전환될 수 있다.

[0247]

[0248]

글리세로인지질은 그들의 테 노보 합성 후에, *sn*-2 위치에서 지방 아실 조성의 신속한 턴오버(turnover)를 겪을 수 있다. 이러한 "리모델링", 또는 "아실 교정(editing)"은 막 구조 및 기능, 스트레스 조건에 대한 생물학적 반응 및 생명과학 응용에서 지방산 조성 및 양의 조작에 중요하다. 구체적으로, 리모델링은 글리세로인지질의 탈아실화 및 생성되는 라이소인지질의 이후의 재아실화의 원인이 된다(즉, 아실-CoA 지방산이 세포 아실-CoA 풀(pool)로부터 제거되고, 다양한 라이소인지질 기질이 인지질 풀에서 *sn*-2 위치에서 아실화되도록).

[0249]

다양한 연구에 의해, 트랜스제닉 유기체의 오일에서 원하는 지방산의 양을 증가시키기 위한 아실-CoA:라이소인지질 아실트랜스퍼라제["LPLAT"]와 PUFA 생합성 유전자의 공동-발현의 유익한 효과는 전체 오일 함량을 증가시키거나 원하는 지방산의 함량을 선택적으로 증가시키는 것인데, 이는 지방산 생합성이 아실-CoA 풀과 인지질 풀 사이의 아실기의 신속한 교환을 필요로 하기 때문인 것으로 결론지어졌다(즉, 불포화화는 주로 인지질의 *sn*-2 위치에서 발생하는 한편, 연장은 아실-CoA 풀에서 발생한다)(국제 출원 공개 WO 2004/076617호, WO 2004/087902호, WO 2006/069936호, WO 2006/052870호, WO 2009/001315호, WO 2009/014140호 참조). 그러나, 미국 특허 출원 공개 2010-0317882-A1호에 개시된 연구는 높은 수준의 EPA의 생성을 위해 엔지니어링된 유지성 유기체에서의 LPAAT 및 LPCAT의 효과를 시험하기 위해 수행한 처음의 연구였다.

[0250]

더욱 구체적으로, 미국 특허 출원 공개 제2010-0317882-A1호의 실시예 3, 4, 7 및 8은 델타-9 연장효소/델타-8 불포화화효소 경로의 발현을 통하여 전체 지질과 비교하여 약 51% EPA의 생성을 위해 엔지니어링된 야로위아 리폴리티카의 재조합 균주 Y8406에서의 ScLPAAT1S, MalLPAAT1S 및 YILPAAT1(하기 표 4에 기재된 바와 같음)의 과발현의 효과를 비교한 것이다.

표 4

LPAAT 및 야로위아 리폴리티카 Y8406의 형질전환 균주에서의 LPAAT 과발현의 영향의 요약

유기체	참조	야생형 약어 및 서열 번호	코돈-최적화 약어* 및 서열 번호	대조군에 대한 과발현의 영향		
				LA % TFA	EPA % TFA	델타-9 연장효소 전환 효율**
사카로마이세스- 세레비지에	ORF "YDL052C"; 진뱅크 수탁 번호 NP_010231	"ScLPAAT1" (서열 번호 23)	"ScLPAAT1S" (서열 번호 24 및 25)	대조군의 72- 74% (플라스미드 발현)	대조군의 104- 106%(플라스미드 발현)	대조군의 108- 109%(플라스미드 발현)
모르티에넬라 알피나	미국 특허 제 7,879,591 호	"MaLPAAT1" (서열 번호 17 및 18)	"MaLPAAT1S" (서열 번호 19 및 20)	대조군의 55- 56% (염색체 통합)	대조군의 112% (염색체 통합)	대조군의 118-119% (염색체 통합)
야로위아 리폴리티카	ORF "YAL10E18964g"; 진뱅크 수탁 번호 XP_504127; 미국 특허 제 7,189,559 호	"YILPAAT1" (서열 번호 21 및 22)	--	대조군의 63% (염색체 통합); 대조군의 76% (플라스미드 발현)	대조군의 115%(염색체 통합); 대조군의 101% (플라스미드 발현)	대조군의 115% (염색체 통합); 대조군의 107% (플라스미드 발현)

* 코돈-최적화는 미국 특허 제 7,125,672 호에 기재된 바와 같은 야로위아 리폴리티카에서의 발현을 위한 것이다. 효율적인 효모
변역 개시 서열을 포함시키고, 최적화 유전자 발현을 수득하기 위하여, 이러한 유기체 내의 코돈-최적화된 합성 유전자를 종종
변역 개시 코돈 'ATG' 주위에 하기의 공통(consensus) 서열을 포함하도록 변경시켰다: MAMMATIGNIIS(서열 번호 70), 여기서,
사용되는 핵산 축퇴 코드는 하기와 같다: M=A/C; S=C/G; H=A/C/T; 및 N=A/C/G/T.

** 전환 효율을 하기 식에 따라 계산하였다: 생성물(들)/(생성물(들)+기질)*100, 여기서, 생성물은 생성물 및 생성물 유도체 둘
모두를 포함한다. 이는 기질을 생성물로 전환시키는 효소의 능력의 척도이다.

[0251]

[0252]

MaLPAAT1 및 YILPAAT1은 34.0%의 서열 동일성을 공유하는 한편, ScLPAAT 및 YILPAAT1은 43.9%의 서열 동일성을 공유하며, 모든 3가지 야생형 단백질은 문헌[Lewin, T.W. et al. (*Biochemistry*, 38:5764-5771 (1999))] 및 문헌[Yamashita et al. (*Biochim. Biophys. Acta*, 1771:1202-1215 (2007))]에 기재된 바와 같이, NH₄xxxD(서열 번호 26) 및 EGTR(서열 번호 27)로 기재된 1-아실-sn-글리세롤-3-포스페이트 아실트랜스퍼라제 패밀리에 속하는 효소이다.

[0253]

(a) MaLPAAT1S 또는 YILPAAT1 중 어느 하나를 지니는 선형화된 DNA를 염색체 통합에 의해 형질전환시켜, 안정적인 통합을 야기하고, 형질전환체를 상대적으로 풍부한 비-선택 배지에서 성장시킨 후에, 또는 (b) 자가 복제 서열을 지니는 원형 플라스미드 DNA 및 YILPAAT 또는 ScLPAAT1S 중 어느 하나를 형질전환시키고, 형질전환체를 선택 배지에서 성장시킨 후에[즉, 표 4에서 "염색체 통합" 또는 "플라스미드 발현" 중 어느 하나로 표시], 야로위아 리폴리티카에서의 ScLPAAT1S, MaLPAAT1S 및 YILPAAT1의 과발현을 분석하였다. LPAAT 과발현이 염색체 통합을 통한 것과 반대로 플라스미드 발현을 통해 발생하는 경우 결과는 최소화되었으며, 이 현상은 염색체 통합의 "위치 효과" 및/또는 상이한 성장 조건 때문일 것 같으며; 대안적으로, 플라스미드의 소실도 또한 관찰되는 결과에 기여할 수 있다. 그러나, 임의의 경우에, 표 4의 결과는 LPAAT 과발현이 이들 각각의 파라미터를 대조군과 비교하는 경우, TFA의 중량%["LA % TFA"]로서의 LA(18:2)의 농도의 유의미한 감소, TFA의 중량%로서의 EPA의 농도["EPA % TFA"]의 증가 및 델타-9 연장효소(들)의 전환 효율의 증가를 야기하는 것을 나타냈다.

[0254]

이에 따라, 상기 요약된 결과에 기초하여, EPA % DCW를 생성하는 능력이 향상된 재조합 미생물 숙주 세포의 향상된 최적화 균주는 하기로 이루어진 군으로부터 선택되는 적어도 2개의 LPAAT를 포함한다:

[0255]

(a) 본질적으로, 서열 번호 18(MaLPAAT1), 서열 번호 20(MaLPAAT1S), 서열 번호 22(YILPAAT1), 서열 번호 23(ScLPAAT1) 및 서열 번호 25(ScLPAAT1S)로 이루어진 군으로부터 선택된 서열로 이루어진 서열; 및

[0256]

(b) 클러스털 더블유 정렬 방법에 기초하여, 서열 번호 18(MaLPAAT1), 서열 번호 22(YILPAAT1) 및 서열 번호 23(ScLPAAT1)으로 이루어진 군으로부터 선택되는 아미노산 서열과 비교하여, 적어도 43.9%의 아미노산 동일성을 가지며, 서열 번호 26 및 서열 번호 27로 이루어진 군으로부터 선택되는 적어도 하나의 1-아실-sn-글리세롤-3-포스페이트 아실트랜스퍼라제 패밀리에 속하는 효소를 추가로 포함하는 폴리펩티드.

[0257]

명확하게 하기 위하여, 적어도 2개의 LPAAT는 1) 단일의 종으로부터 분리된 특정 LPAAT에 대한 동일한 암호화 서열의 2 카피(copy); 또는 2) 중 "A"로부터 분리된 LPAAT에 대한 하나의 암호화 서열 및 중 "B"로부터 분리된 LPAAT에 대한 하나의 암호화 서열 중 어느 하나여서, 집합적으로 2개의 LPAAT를 야기할 수 있다.

[0258]

본 발명의 최적화된 재조합 미생물 숙주 세포는 또한 PDAT 활성을 갖는 적어도 하나의 폴리펩티드를 포함할 것이다. 달퀴스트(Dahlqvist) 등(문헌[*Proc. Nat. Acad. Sci. (USA)*, 97:6487-6492 (2000)]) 및 오엘커스(Oelkers) 등(문헌[*J. Biol. Chem.*, 275:15609-15612 (2000)])은 TAG 합성이 아실-CoA-독립적 PDAT 효소(단백질의 레시틴:콜레스테롤 아실트랜스퍼라제 패밀리와 구조적으로 관련된)를 통하여 아실-CoA의 부재 하에 발생할 수 있음을 처음으로 인식하였다. 더욱 구체적으로, 달퀴스트 등 및 오엘커스 등은 PDAT를 암호화하는 사카로마이세스 세레비지에 LRO1 유전자(서열 번호 30; "ScPDAT")의 과발현이 TAG 함량의 증가를 야기하는 한편, ScPDAT

의 결실이 TAG 합성의 유의미한 감소를 야기함을 입증하였다. 이러한 연구 후에, 미국 특허 제7,267,976호는 PDAT를 암호화하는 야로위아 리폴리티카 ATCC #90812 유전자(즉, 본 명세서에서 서열 번호 28 및 29)의 클로닝, 과발현 및 낙아웃을 기재하고 있으며, 이는 ScPDAT와 47.1%의 아미노산 서열 동일성을 공유하는 것으로 결정되었다. PDAT가 파괴된 야로위아 리폴리티카 균주는 야생형 균주와 비교하여 더 낮은 오일 함량["TFA % DCW"]을 갖는 것으로 관찰된 한편(약 29-38%), PDAT2 및 DGAT2 둘 모두에서 파괴를 갖는 균주는 대조군과 비교하여 오직 17-27%의 오일 함량을 갖는 것으로 결정되었다. 이어서, 야로위아 리폴리티카 PDAT를 그의 천연의 PDAT 및 DGAT2 유전자에서 파괴를 갖는 사카로마이세스 세레비지에 균주에서 발현시켰고; TFA % DCW는 대조군과 비교하여 형질전환 균주에서 배가되었다.

[0259] 상기 논의에 기초하여, 당업자는 PDAT가 전체 지질 함량을 변경시키는 데에서 수행하는 역할을 인식할 것이다. 이에 따라, 본 명세서에 기재된 재조합 미생물 숙주 세포는 하기로 이루어진 군으로부터 선택되는 적어도 하나의 PDAT를 포함할 것이다:

[0260] (a) 본질적으로, 서열 번호 29 및 서열 번호 30으로 이루어진 군으로부터 선택되는 서열로 이루어진 서열; 및

[0261] (b) 클러스털 더블유 정렬 방법에 기초하여, 서열 번호 29 및 서열 번호 30으로 이루어진 군으로부터 선택되는 아미노산 서열과 비교하여, 적어도 약 90%의 아미노산 동일성, 또는 더욱 바람직하게는 적어도 약 95%의 아미노산 동일성을 갖는 폴리펩티드.

[0262] 최근에, 상업적으로 유용한 재조합 미생물 숙주 세포에서 PUFA 생합성 경로로의 통합에 매우 적절한 활성이 높은 델타-9 연장효소 돌연변이체를 동정하기 위해 이.아이. 듀폰 드 네무어 앤드 컴퍼니(E.I. duPont de Nemours & Company)는 상당한 노력을 쏟았는데, 이는 이전의 연구에 의해 델타-9 연장(또한 델타-6 연장)이 인지질과 아실-CoA 풀 사이의 아실기의 불량한 전달 때문에 장쇄 PUFA 생합성의 장애물인 것으로 나타났기 때문이다. 미국 가출원 제61/377248호[대리인 문서 번호 CL4783USPRV, 출원일: 2010년 8월 26일, 본 명세서에 그 전문이 참고로 포함]에 기재되고, 본 명세서에서 실시예 10A 내지 10J로 기재된 바와 같이, LA로부터 EDA로의 전환에 기초하여, 서열 번호 3의 효소 활성과 비교하여, 효소 활성의 최대 45%의 향상을 야기하는 특정 돌연변이가 돌연변이 EgD9eS 폴리펩티드(즉, EgD9eS[서열 번호 3]) 유래의)에서 동정되었다.

[0263] 더욱 구체적으로, 델타-9 연장효소 내의 적절한 돌연변이를 동정하기 위한 방법을 표적으로 하는 이유는 델타-9 연장효소로부터의 임의의 결정 구조가 없기 때문에, 그리고 아이소크라이시스 갈바나 델타-9 연장효소["IgD9e"; 서열 번호 41 및 42]의 변이형 히스티딘-박스["His-box"] 모티프 내의 Gln 잔기의 중요성에 관한 오직 단일의 연구(문헌[Qi, B., et al., FEBS Lett., 547:137-139 (2003)]) 때문에 이상적이지 않다. 델타-9 연장효소 활성으로 동정된 처음의 PUFA-특이적 연장효소인 IgD9e는 델타-6 연장효소에 존재하는 고도로 보존된 His-Xaa-Xaa-His-His["HxxHH"; 서열 번호 72] 모티프 대신에, Gln-Xaa-Xaa-His-His["QxxHH"; 서열 번호 71] 모티프를 갖는 것으로 관찰되었다. 쿼, 비.(Qi, B.) 등에 의해, Gln의 His, Ala 또는 Phe 잔기로의 치환이 거기서 분석된 각각의 돌연변이 IgD9e 폴리펩티드에서 더 낮은 델타-9 연장효소 활성을 야기하는 것이 입증되었으며, 이에 따라, "히스티딘 박스 내의 글루타민 잔기...가 최적의 효소 촉매작용에 필수적인 것으로 보인다"고 결론지었다.

[0264] 쿼 등의 연구에 더하여, 7개의 모티프가 IgD9e(서열 번호 42), 유글레나 그라실리스 델타-9 연장효소["EgD9e"; 서열 번호 32; 미국 특허 제7,645,604호] 및 유티렙티엘라 종 CCMP389 델타-9 연장효소["E389D9e"; 서열 번호 38; 미국 특허 제7,645,604호] 간에 그리고 그들 중에서 보존되는 것으로 알려져 있다. 이들 모티프는 미국 특허 제7,645,604호에 기재되었으며, Y-N-X-(L 또는 F)-X₄-S-X₂-S-F(서열 번호 73); F-Y-X-S-K-X₂-(E 또는 D)-Y-X-D-(T 또는 S)-X₂-L(서열 번호 74); L-(Q 또는 H)-X-F-H-H-X-G-A(서열 번호 75); M-Y-X-Y-Y-X₇-(K 또는 R 또는 N)-F(서열 번호 76); K-X-L-(I 또는 L 또는 M)-T-X₂-Q(서열 번호 77); W-X-F-N-Y-X-Y(서열 번호 78); 및 Y-X-G-X-V-X₂-L-F(서열 번호 79)를 포함하며; 여기서, X는 임의의 아미노산일 수 있으며, 밑줄이 그어진 아미노산은 델타-9 연장효소에 독특한 것일 수 있다.

[0265] 이에 따라, 델타-9 연장효소를 암호화하는 돌연변이 서열의 라이브러리를 주형으로서 EgD9eS(서열 번호 2)를 사용하여 오류-유발 PCR["ePCR"]에 의해 합성으로 엔지니어링하였으며, 여기서, EgD9eS는 키메라 FBAINm::EgD9eS::Pex20 유전자를 포함하는 플라스미드 구축물 내에 함유되어 있다. 이어서, ePCR 라이브러리를 야로위아 리폴리티카 내로 형질전환시키고, GC 분석 및 EDA의 생성에 기초하여 향상된 델타-9 연장효소 활성에 대하여 스크리닝하였다.

- [0266] 완전히 비-기능성인 돌연변이 델타-9 연장효소(즉, 검출가능한 델타-9 연장효소 활성을 갖지 않는) 또는 야생형 효소, EgD9eS에 대하여 델타-9 연장효소 활성이 실질적으로 감소된 돌연변이 델타-9 연장효소를 야기하는 많은 클론을 동정하였다. 그러나, 놀랍게도, LA로부터 EDA로의 전환 효율($\frac{[EDA]}{[LA+EDA]} \times 100$ 으로 계산)을 향상시키는 다양한 돌연변이가 동정되었다. 구체적으로, 4개의 상이한 돌연변이 델타-9 연장효소 유전자를 포함하는(즉, EgD9eS의 단백질 서열[서열 번호 3]과 비교시, 각각 K58R/I257T 돌연변이, L35F 돌연변이, D98G 돌연변이 및 L130M/V243A 돌연변이를 포함하는) 5개 개별 형질전환체가 동정되었으며, 여기서, 델타-9 연장효소 전환 활성은 야생형 EgD9eS의 105% 내지 117% 범위이고(하기 표 5), 이는 5 내지 17% 향상에 상응한다. 따라서, 이러한 연구에 의해, EgD9eS의 델타-9 연장효소 활성이 실제로 단백질 엔지니어링에 의해 향상될 수 있었던 것이 입증되었다.
- [0267] 이어서, 상기 EgD9eS ePCR 라이브러리로부터 수득되는 초기의 데이터를 사용하여, 부위-포화 라이브러리의 생성을 위한 적절한 표적이었던 EgD9eS 내의 2개의 상이한 아미노산 잔기(즉, 잔기 35 및 107)를 합리적으로 동정하였다. 다시, 생성된 돌연변이 EgD9eS의 델타-9 연장효소 활성에서의 각각의 돌연변이의 영향을 스크리닝하였으며, 이에 따라, LA로부터 EDA로의 전환 효율을 향상시키는 2개의 추가의 돌연변이의 동정이 가능하게 되었다. 구체적으로, 돌연변이 델타-9 연장효소 내의 L35G 돌연변이 또는 L35M/Q107E 돌연변이 중 어느 하나를 포함하는 형질전환 균주가 동정되었으며, 여기서, 델타-9 연장효소 전환 활성은 야생형 EgD9eS의 142% 내지 145% 또는 132% 중 어느 하나였으며(하기의 표 5), 이는 32% 내지 45% 향상에 상응한다.
- [0268] L35G 돌연변이의 동정 후에, 50개의 상이한 아미노산 잔기를 표적화하는 후속의 라이브러리를 슬로노맥스(SlonoMax)(등록 상표) 기술 및 표적으로서 EgD9eS-L35G 유전자를 사용하여 생성하였다. 모 연장효소, 즉, EgD9eS-L35G(하기의 표 5)와 비교하여 96% 내지 141%의 델타-9 연장효소 전환 활성을 야기하는(-4% 내지 41%의 향상에 상응) 각각 L35G 돌연변이와 조합된 25개의 상이한 돌연변이를 동정하였다.
- [0269] 마지막으로, 최근의 연구에서, 슬로노맥스(등록 상표) 라이브러리 내에서 동정된 다수의 유익한 돌연변이를 조합(또는 스택킹(stack))하려고 시도하였으며, 이에 의해, 합성의 코돈-최적화된 EgD9eS 서열 내의 적절한 개별 아미노산 돌연변이를 "스택킹"하였다. 이에 따라, 예를 들어, 서열 번호 3에 대하여 A21V, L35G, W132T, I179R 및 Q244N 돌연변이를 포함하는 돌연변이 델타-9 연장효소가 EgD9eS와 비교하여 123%(23% 향상에 상응)의 델타-9 연장효소 전환 활성의 향상(하기의 표 5)을 야기하는 것으로 입증되었다.

표 5

델타-9 연장효소 활성이 증가된 돌연변이체의 요약

라이브러리 생성 방법	생성된 아미노산 치환	돌연변이 유전자에 대한 명칭	상대 활성
cPCR	L35F	EgD9eS-L35F	115% ^a
	K58R 및 I257T	EgD9eS-K58R/I257T	105% ^a
	L130M 및 V243A	EgD9eS-L130M/V243A	106%-111% ^a
	D98G	EgD9eS-D98G	117% ^a
부위-포화	L35G	EgD9eS-L35G	142%-145% ^a
	L35M 및 Q107E	EgD9eS-L35M/Q107E	132% ^a
슬로노맥스 (등록 상표)	L35G 및 S9A	EgD9eS-L35G/S9A	126% ^b
	L35G 및 S9D	EgD9eS-L35G/S9D	141% ^b
	L35G 및 S9G	EgD9eS-L35G/S9G	129% ^b
	L35G 및 S9I	EgD9eS-L35G/S9I	113% ^b
	L35G 및 S9K	EgD9eS-L35G/S9K	122% ^b
	L35G 및 S9Q	EgD9eS-L35G/S9Q	111% ^b
	L35G 및 Q12K	EgD9eS-L35G/Q12K	123% ^b
	L35G 및 A21D	EgD9eS-L35G/A21D	118% ^b
	L35G 및 A21T	EgD9eS-L35G/A21T	110% ^b
	L35G 및 A21V	EgD9eS-L35G/A21V	118% ^b
	L35G 및 V32F	EgD9eS-L35G/V32F	104% ^b
	L35G 및 Y84C	EgD9eS-L35G/Y84C	144% ^b
	L35G 및 L108G	EgD9eS-L35G/L108G	104% ^b
	L35G 및 G127L	EgD9eS-L35G/G127L	104% ^b
	L35G 및 W132T	EgD9eS-L35G/W132T	100% ^b
	L35G 및 M143N	EgD9eS-L35G/M143N	96% ^b
	L35G 및 M143W	EgD9eS-L35G/M143W	106% ^b
	L35G 및 L161T	EgD9eS-L35G/L161T	131% ^b
	L35G 및 L161Y	EgD9eS-L35G/L161Y	119% ^b
	L35G 및 W168G	EgD9eS-L35G/W168G	115% ^b
	L35G 및 I179M	EgD9eS-L35G/I179M	104% ^b
	L35G 및 I179R	EgD9eS-L35G/I179R	141% ^b
	L35G 및 C236N	EgD9eS-L35G/C236N	102% ^b
	L35G 및 Q244N	EgD9eS-L35G/Q244N	134% ^b
	L35G 및 A254W	EgD9eS-L35G/A254W	112% ^b
	L35G 및 A254Y	EgD9eS-L35G/A254Y	116% ^b
조합	L35G 및 W132T 및 I179R	EgD9eS-L35G/W132T/I179R	110% ^a
조합	S9D 및 L35G 및 Y84C 및 I179R	EgD9eS-S9D/L35G/Y84C/I179R	108% ^a
조합	A21V 및 L35G 및 L108G 및 I179R	EgD9eS-A21V/L35G/L108G/I179R	104% ^a
조합	L35G 및 Y84C 및 I179R 및 Q244N	EgD9eS-L35G/Y84C/I179R/Q244N	111% ^a
조합	A21V 및 L35G 및 W132T 및 I179R 및 Q244N	EgD9eS-A21V/L35G/W132T/I179R/Q244N	123% ^a

^a "상대 활성"은 서열 번호 3에 기재된 EgD9eS의 델타-9 연장효소 활성에 대한 각각의 돌연변이 EgD9eS의 델타-9 연장효소 활성을 지칭한다.

^b "상대 활성"은 서열 번호 44에 기재된 EgD9eS-L35G의 델타-9 연장효소 활성에 대한 각각의 돌연변이 EgD9eS의 델타-9 연장효소 활성을 지칭한다.

[0270]

[0271]

이에 따라, 상술된 연구의 결과로서, 본 명세서에 기재된 재조합 미생물 숙주 세포는 적어도 하나의 돌연변이 델타-9 연장효소 폴리펩티드를 포함할 수 있으며, 여기서, 상기 돌연변이 델타-9 연장효소 폴리펩티드는 서열 번호 1에 기재된 아미노산 서열을 포함하며, 서열 번호 1은 도 16b에 나타난 바와 같이, 적어도 하나의 아미노산 돌연변이에 의해 서열 번호 3과 상이하며, 상기 돌연변이(들)은

[0272]

i) L35F 돌연변이;

[0273]

ii) L35M 돌연변이;

[0274]

iii) L35G 돌연변이;

[0275]

iv) L35G 돌연변이, 및 S9A, S9D, S9G, S9I, S9K, S9Q, Q12K, A21D, A21T, A21V, V32F, Y84C, Q107E, L108G,

G127L, W132T, M143N, M143W, L161T, L161Y, W168G, I179M, I179R, C236N, Q244N, A254W 및 A254Y로 이루어진 군으로부터 선택되는 적어도 하나의 다른 돌연변이;

[0276] v) L35G, A21V, L108G 및 I179R 돌연변이;

[0277] vi) L35G, W132T 및 I179 돌연변이;

[0278] vii) L35G, S9D, Y84C 및 I179R 돌연변이;

[0279] viii) L35G, Y84C, I179R 및 Q244N 돌연변이;

[0280] ix) L35G, A21V, W132T, I179R 및 Q244N 돌연변이;

[0281] x) K58R 및 I257T 돌연변이;

[0282] xi) D98G 돌연변이;

[0283] xii) L130M 및 V243A 돌연변이; 및

[0284] xiii) K58R, L35F, L35G, L35M, S9A, S9D, S9G, S9I, S9K, S9Q, Q12K, A21D, A21T, A21V, V32F, Y84C, D98G, Q107E, L108G, G127L, L130M, W132T, M143N, M143W, L161T, L161Y, W168G, I179M, I179R, C236N, V243A, Q244N, A254W, A254Y 및 I257T로 이루어진 군으로부터 선택되는 적어도 2개의 돌연변이를 포함하는 임의의 조합으로 이루어진 군으로부터 선택된다.

[0285] 바람직한 실시형태에서, 돌연변이 EgD9eS는 서열 번호 3에 대하여 적어도 L35G 돌연변이를 포함한다. 예를 들어, 서열 번호 3에 대하여 단일의 L35G 돌연변이를 갖는 서열 번호 44에 기재된, 본 명세서에 "EgD9eS-L35G"로 기재되는 돌연변이 델타-9 연장효소 폴리펩티드의 델타-9 연장효소 활성은 EgD9eS의 델타-9 연장효소 활성과 비교하여 142 내지 145%였으며, 이는 42 내지 45% 향상에 상응한다.

[0286] 적어도 2개의 LPAAT, 적어도 하나의 PDAT 및 (임의로) 상술된 바와 같은 적어도 하나의 돌연변이 델타-9 연장효소뿐 아니라 EPA 생합성을 위한 델타-9 연장효소/델타-8 불포화화효소 오메가-3/오메가-6 지방산 생합성 경로 (즉, 적어도 하나의 델타-9 연장효소, 적어도 하나의 델타-8 불포화화효소, 적어도 하나의 델타-5 불포화화효소, 적어도 하나의 델타-17 불포화화효소, 적어도 하나의 델타-12 불포화화효소 및 적어도 하나의 C_{16/18} 연장효소를 포함하며, 적어도 하나의 델타-8 불포화화효소에 연결된 적어도 하나의 델타-9 연장효소를 포함하는 적어도 하나의 멀티자임이 존재하는)를 암호화하는 유전자를 발현하는 것에 더하여, 적어도 25 EPA % DCW를 생성하는 본 발명의 재조합 숙주 세포는 임의로 다양한 다른 이중 유전자를 발현할 수 있다. 이들은 예를 들어, 후술되는 콜린포스페이트 시티딜트랜스퍼라제["PCT"], 다이아실글리세롤 콜린포스포트랜스퍼라제["CPT1"], 말로닐 CoA 신테타제["MCS"] 및/또는 델타-9 불포화화효소를 암호화하는 유전자를 포함할 수 있다.

[0287] 포스파티딜콜린["PC"] 생합성 경로는 3 단계를 포함한다:

[0288] (i) 콜린 키나제[EC 2.7.1.32]에 의해 촉매작용되는, ATP + 콜린 → ADP + O-포스포콜린;

[0289] (ii) 콜린 포스페이트 시티딜트랜스퍼라제["PCT"; EC 2.7.7.15]에 의해 촉매작용되는 시티딘 트라이포스페이트["CTP"] + 콜린 포스페이트 ⇌ 다이포스페이트 + 시티딘 다이포스페이트-콜린["CDP-콜린"]; 및

[0290] (iii) 다이아실글리세롤 콜린포스포트랜스퍼라제["CPT1"; EC 2.7.8.2]에 의해 촉매작용되는 CDP-콜린 + 1,2-다이아실글리세롤 ⇌ 시티딘-5'-모노포스페이트["CMP"] + 포스파티딜콜린.

[0291] 미국 특허 출원 공개 제2009-0093543-A1호에는 CPT1을 암호화하는 적어도 하나의 유전자의 임의의 공동-발현이 기재되어 있다. 본 명세서에서, PCT를 암호화하는 적어도 하나의 유전자의 임의의 공동-발현이 제시된다. 이들 효소 중 어느 하나 또는 둘 모두의 발현은 PC 생합성 경로를 상향조절하여, PC의 생합성의 증가를 야기할 것이다. PUFA는 PC(및 다른 글리세로인지질)의 sn-2 위치에서 에스테르화된다. 따라서, 이들 효소의 발현 증가는 PUFA가 재조합 미생물 숙주 세포에서 저장될 수 있는(다시 말하면, PUFA의 주요 저장은 TAG의 형태일 것이지만) 추가의 메커니즘을 제공할 수 있다. PC의 생성 증가는 또한 하기의 협동 작용을 통하여 세포 내에서의 이후의 "리모델링" 또는 "아실 교정"을 용이하게 할 수 있다: 1) 포스포리파제, 예컨대 PC의 sn-2 위치로부터 지방산을 유리시키는 포스포리파제 A₂; 및, 2) LPLAT, 예컨대, sn-2 위치에서 LPC를 재아실화시키는 LPCAT. 이는 세포 아실-CoA 풀로부터 아실-CoA 지방산의 제거와, 인지질 풀에서의 sn-2 위치에서 다양한 라이소인지질 기질의 아실화를 용이하게 한다.

- [0292] 상술된 특정 이론에 제한되지 않고, 본 개시내용에 따라 엔지니어링된 재조합 숙주 세포 내에서 PCT 및/또는 CPT1 중 어느 하나가 과발현되는 것이 바람직할 수 있다. PCT 및/또는 CPT1을 암호화하는 유전자(들)는 숙주 세포에 천연이거나 이종일 수 있다. 예를 들어, 야로위아의 최적화된 균주에서, 서열 번호 45에 기재된 야로위아 PCT 유전자(서열 번호 46의 단백질을 암호화하는) 및/또는 서열 번호 47에 기재된 야로위아 CPT1 유전자(서열 번호 48의 단백질을 암호화하는) 또는 서열 번호 46 및/또는 서열 번호 48에 대하여 서열 및 기능의 실질적인 유사성을 공유하는 관련 효소를 발현하는 것이 바람직하다.
- [0293] 미국 특허 출원 공개 제2010-0317072-A1호는 EPA 생합성을 위해 엔지니어링된 재조합 숙주 세포에서 말로닐 CoA 신테타제["MCS"]를 암호화하는 적어도 하나의 폴리뉴클레오타이드의 임의의 공동-발현을 기재하고 있다. 이전의 연구에 의해, 야로위아 리폴리티카에서의 PUFA의 생성의 엔지니어링과 관련된 많은 유전적 돌연변이가 유기산 "부산물"의 축적을 증가시키는 것으로 결정되었으며, 이는 미생물 발효 동안에 추가로 사용될 수 없는 것이다 (말로네이트가 축적되는 전체 유기산의 약 45%를 차지한다). 특히, 미국 특허 출원 공개 제2010-0159558-A1호에는 하기의 효소 반응을 촉매작용시키는 EPA를 생성하는 야로위아 리폴리티카의 재조합 균주에서의 이중 MCS[EC 6.2.1.-]의 발현이 기재되어 있다: 말로네이트 + ATP + CoA → 말로닐-CoA + AMP + 피로포스페이트 (PPi). 부산물(즉, 말로네이트)로부터 말로닐-CoA로의 전환에 의하여, 이러한 기질은 유기체 내의 지방산의 합성 동안에 이용가능하게 되었다. 따라서, 약 94%의 말로네이트(g/g(DCW))의 부산물 생성의 감소에 더하여, 이중 MCS의 발현은 또한 유기체 내의 탄소 및 에너지 소모를 방지하고, 발효 공정 동안 최적의 pH 범위를 유지하는데 필요한 염기의 양을 감소시키고, 발효 폐증기 내의 중화를 필요로 하는 부산물 유기산의 양을 감소시키는데 도움이 될 수 있다. 바람직한 MCS는 야로위아 리폴리티카에서의 발현을 위해 코돈-최적화된 리조비움 레구미노사룸 생물변이형 비시에 3841(진뱅크 수탁 번호 YP_766603)(즉, 서열 번호 49 및 50) 유래의 것이었다.
- [0294] 본 개시내용에 따라 엔지니어링된 야로위아 또는 다른 재조합 숙주 세포의 최적화된 균주에서 상기 서열 번호 49 및 50에 기재된 MCS 또는 서열 번호 50에 대하여 서열 및 기능의 실질적인 유사성을 공유하는 관련 효소를 발현하는 것이 바람직할 수 있다.
- [0295] 본 명세서에 기재된 재조합 미생물 숙주 세포에서 임의로 발현될 수 있는 다른 유전자는 델타-9 불포화화효소이다. 당업자에게 명백할 것처럼, 이러한 특정 효소의 과발현은 스테아르산[18:0]으로부터 올레산[18:1]으로의 전환을 증가시켜, PUFA 생합성 경로로의 탄소의 더 큰 "푸싱(pushing)"을 야기할 것이다.
- [0296] 본 명세서에 기재된 재조합 미생물 숙주 세포는 적어도 하나의 돌연변이 델타-5 불포화화효소를 추가로 포함할 수 있다. 델타-6, 델타-8 및 델타-4 불포화화효소와 함께, 델타-5 불포화화효소는 장쇄 PUFA "프론트-엔드(front-end)" 불포화화효소로 알려져 있다(여기서, 불포화화는 메틸-지정 불포화화와 반대로, 지방산의 아실기의 카복실 말단과 기존의 이중 결합 사이에서 발생한다). 이들 불포화화효소는 3개의 His 박스[His(X)₃₋₄H(서열 번호 186 및 187), H(X)₂₋₃HH(서열 번호 188 및 189) 및 H/Q(X)₂₋₃HH(서열 번호 190 및 191)]을 특징으로 하며, 사이토크롬 b₅ 융합 슈퍼패밀리의 구성원인데, 이는 이들이 이들의 N-말단에 전자 공여체로 사용되는 융합된 사이토크롬 b₅ 도메인을 갖기 때문이다. 사이토크롬 b₅ 도메인은 또한 나머지 사이토크롬 b₅ 도메인 서열의 차이에도 불구하고, 보존된 헴(heme)-결합 모티프(즉, HPGG 서열[서열 번호 181])를 포함한다. 이전에 델타-5 불포화화효소의 특징으로 동정된 추가의 보존된 시그니처(signature) 모티프도 또한 히스티딘이 풍부한 것으로 보이지만(즉, HDASH 서열[서열 번호 183]), 효소 활성화에 대한 HDASH 모티프의 중요성은 아직 밝혀지지 않았다.
- [0297] 본 발명의 일부 실시형태에서, 적어도 하나의 돌연변이 델타-5 불포화화효소는 하기로 이루어진 군으로부터 선택될 수 있다:
- [0298] a) 서열 번호 180[HxGx]에 기재된 아미노산 모티프 - 여기서, 서열 번호 180[HxGx]은 서열 번호 181[HPGG]와 동일하지 않다 - 및 서열 번호 182[HxxxH]에 기재된 아미노산 모티프 - 여기서, 서열 번호 182[HxxxH]는 서열 번호 183[HDASH]와 동일하지 않다 - 를 포함하는 돌연변이 폴리펩티드;
- [0299] b) 서열 번호 106[EgD5M 또는 코돈-최적화된 EgD5R*-34g158g], 서열 번호 108[EgD5R*-34g158g347s], 서열 번호 110[EgD5S-36s157g], 서열 번호 112[EaD5S-35a158g], 서열 번호 299[EgD5R*-34g157g], 서열 번호 301[EgD5R*-34g158a], 서열 번호 303[EgD5R*-34g158g], 서열 번호 329[EgD5S-36s156e], 서열 번호 331[EgD5S-36s158a], 서열 번호 333[EgD5S-36s158g], 서열 번호 363[EaD5S-35a158s], 서열 번호 365[EaD5S-35a159g]로 이루어진 군으로부터 선택되는 아미노산 서열을 갖는 돌연변이 폴리펩티드.
- [0300] 본 명세서에 기재된 재조합 미생물 숙주 세포는 적어도 약 25 EPA % DCW, 바람직하게는 적어도 약 25-30 EPA %

DCW, 더욱 바람직하게는 적어도 약 30-32.5 EPA % DCW, 더욱 바람직하게는 적어도 약 32.5-35 EPA % DCW, 가장 바람직하게는 적어도 약 35-40 EPA % DCW를 포함하는 미생물 오일을 생성할 수 있을 것이다. 발효 분야의 숙련자는 발효 시행 그 자체, 배지 조건, 공정 파라미터, 대량화 등뿐 아니라 배양물을 시료 추출하는 특정 시점에 따라, 특정 재조합 미생물 숙주 세포의 오일 프로파일에서 다양성이 발생할 것임을 인식할 것이다. 따라서, 특정 유전형을 갖는 특정 재조합 미생물 숙주는 최적의 조건 하에서 배양되는 경우 적어도 약 25 EPA % DCW를 포함하는 미생물 오일을 생성할 수 있을 수 있으나, 항상 적어도 약 25 EPA % DCW를 포함하는 미생물 오일을 생성하지는 않을 것이다(예컨대, 발효 기간이 충분하지 않은 경우). 따라서, 본 발명의 논의는 적절한 조건 하에서 배양되는 경우 적어도 약 25 EPA % DCW를 생성하는 유기체의 "능력"을 언급한다.

[0301] 당업자에게 명백할 것처럼, 적어도 약 25 EPA % DCW를 생성할 수 있는 다수의 상이한 최적화된 재조합 균주는 본 명세서에 기재된 방법을 사용하여 엔지니어링될 수 있다. 상업적인 목적을 위한 바람직한 균주의 선택은 전체 지방산의 백분율로서의 EPA의 농도["EPA % TFA"] 및 전체 지질 함량["TFA % DCW"] 둘 모두를 고려할 것인데, 이는 둘 모두의 인자가 건조 세포 중량의 백분율로서의 EPA의 세포 함량["EPA % DCW"]에 영향을 미치기 때문이다. 다시 말하면, EPA % DCW는 하기의 식으로 계산한다: $(\text{EPA \% TFA}) \times (\text{TFA \% DCW}) / 100$. 예를 들어, 40 EPA % TFA를 생성하며 62.5 TFA % DCW를 갖는 균주, 45 EPA % TFA를 생성하며 55.55 TFA % DCW를 갖는 균주, 50 EPA % TFA를 생성하며 50 TFA % DCW를 갖는 균주, 55 EPA % TFA를 생성하며 45.45 TFA % DCW를 갖는 균주, 60 EPA % TFA를 생성하며 41.67 TFA % DCW를 갖는 균주 및 65 EPA % TFA를 생성하며 38.46 TFA % DCW를 갖는 균주는 모두 25 EPA % DCW를 생성한다.

[0302] 바람직한 실시형태에서, 향상된 재조합 미생물 숙주 세포는 적어도 25 EPA % DCW를 포함하는 오일을 생성할 수 있을 것이며, 적어도 45 EPA % TFA를 생성할 것이다. 더욱 바람직하게는, 오일은 적어도 약 47-50 EPA % TFA, 바람직하게는 적어도 약 50-55 EPA % TFA, 더욱 바람직하게는 적어도 약 55-60 EPA % TFA, 더욱 바람직하게는 적어도 약 60-70 EPA % TFA, 가장 바람직하게는 적어도 약 70-80 EPA % TFA를 포함할 것이다.

[0303] 다른 실시형태에서, 향상된 재조합 미생물 숙주 세포는 적어도 25 EPA % DCW를 포함하는 오일을 생성할 수 있을 것이며, 향상된 재조합 미생물 숙주 세포 내의 또는 그로부터 추출되거나 비농축된 오일 내의 지질 프로파일은 EPA % TFA 대 LA % TFA의 비가 적어도 약 2.4일 것이다. 미국 특허 출원 공개 제2009-0093543-A1호에서 이전에 논의된 바와 같이, 중간체 지방산, 즉, LA의 농도 감소(EPA:LA 비의 증가 야기)는 PUFA 생합성 경로를 통한 탄소의 "푸싱"이 더 커지게 할 것이며, EPA 합성의 증가를 가능하게 할 것이다. 바람직한 실시형태에서, EPA:LA의 비는 적어도 약 2.4-2.75, 더욱 바람직하게는 적어도 약 2.75-3.25, 더욱 바람직하게는 적어도 약 3.25-4, 가장 바람직하게는 적어도 약 4-5.5일 것이다.

[0304] 다양한 박테리아, 효모, 조류, 유글레나류(euglenoids), 부등편모조류(stramenopiles), 진균류 및 그들의 혼합물을 비롯한 다양한 미생물 숙주 세포는 천연적으로 미생물 오일을 생성한다. 그리고, EPA는 사용되는 특정 미생물 유기체[예컨대, 종속영양 규조류 사이클로텔라(*Cyclotella*) 종 및 니치시아(*Nitzschia*) 종](미국 특허 제 5,244,921호); 슈도모나스(*Pseudomonas*), 알테로모나스(*Alteromonas*) 또는 쉬아넬라(*Shewanella*) 종(미국 특허 제 5,246,841호); 피티움(*Pythium*) 속의 사상 진균(미국 특허 제 5,246,842호); 또는 모르티에렐라 엘롱가타(*Mortierella elongata*), 엠. 엑시구아(*M. exigua*), 또는 엠. 하이그로필라(*M. hygrophila*)(미국 특허 제 5,401,646호)]의 천연의 능력에 기초하여, 많은 상이한 과정을 통해 미생물에 의해 생성될 수 있다. EPA를 천연적으로 생성하는 미생물을 기술하는 유용한 보고서는 문헌[Z. Wen and F. Chen, In *Single Cell Oils*; C. Ratledge and Z. Cohen, Eds.; AOCs Publishing, 2005; Chapter 10, entitled "Prospects for EPA production using microorganisms"]의 보고서이다.

[0305] 본 명세서의 목적을 위하여, 재조합 미생물 숙주 세포는 유전 공학의 도구를 통해 유전적으로 조작될 수 있고, 적어도 25 EPA % DCW를 포함하는 미생물 오일을 생성할 수 있는 세포여야 한다.

[0306] EPA를 생성하는 천연의 능력이 없는 미생물은 적절한 PUFA 생합성 경로 유전자, 예컨대 델타-5 불포화화효소, 델타-6 불포화화효소, 델타-12 불포화화효소, 델타-15 불포화화효소, 델타-17 불포화화효소, 델타-9 불포화화효소, 델타-8 불포화화효소, 델타-9 연장효소, C_{14/16} 연장효소, C_{16/18} 연장효소 및 C_{18/20} 연장효소의 특정 조합의 도입에 의하여 PUFA 생합성 경로를 발현하도록 엔지니어링될 수 있지만, 도입된 특정 효소(및 그들 효소를 암호화하는 유전자)가 본 명세서에서 본 발명을 결코 제한하는 것이 아님을 인식할 것이다. 예를 들어, 공공의 문헌 및 특허 문헌은 EPA 생합성을 위하여 하기의 것을 엔지니어링하는 수단을 교시한다(비록 낮은 생성 수준일지라도): 에스케리키아 콜라이(*Escherichia coli*)(문헌[Orikasa, A. et al., *Cell Mol. Biol.* 50:625-630 (2004)]), 사카로마이세스 세레비시아(*Saccharomyces cereviasiae*)(문헌[Tavares, S., et al., *AEM*,

77(5)1854-1861 (2011)]).

- [0307] 바람직한 실시형태에서, 미생물 숙주 세포는 유지성이어서, 그들은 그들의 DCW의 약 25% 과량을 오일로 축적한다. 유지성 미생물 숙주 세포는 예컨대, 모르티에렐라, 트라우스토키트리움, 스킴조키트리움(*Schizochytrium*) 및 유지성 효모로 이루어진 군으로부터 선택되는 속의 구성원일 수 있다. 유지성 효모는 오일을 합성하고 축적할 수 있으며, 여기서, 전체 오일 함량은 DCW의 약 25% 초과, 더욱 바람직하게는 DCW의 약 30% 초과, 가장 바람직하게는 DCW의 약 40% 초과로 포함될 수 있다. 대안적인 실시형태에서, 비-유지성 효모, 예컨대 사카로마이세스 세레비지애와 같은 효모(국제 출원 공개 WO 2006/102342호)는 그것이 DCW의 25% 초과와 오일을 생성할 수 있도록 유전적으로 변형되어 유지성이 될 수 있다.
- [0308] 통상적으로 유지성 효모로 동정되는 속에는 야로위아, 칸디다, 로도토룰라, 로도스포리둠, 크립토크커스, 트리코스포론 및 리포마이세스가 포함되나 이에 한정되지 않는다. 더욱 구체적으로, 예시적인 오일-합성 효모에는 로도스포리둠 토룰로이데스(*Rhodospiridium toruloides*), 리포마이세스 스타케이이(*Lipomyces starkeyii*), 엘. 리포페러스(*L. lipoferus*), 칸디다 레브카우피(*Candida revkaufi*), 씨. 풀케리마(*C. pulcherrima*), 씨. 트로피칼리스(*C. tropicalis*), 씨. 유틸리스(*C. utilis*), 트리코스포론 풀란스(*Trichosporon pullans*), 티. 쿠타네움(*T. cutaneum*), 로도토룰라 글루티누스(*Rhodotorula glutinus*), 알. 그라미니스(*R. graminis*) 및 야로위아 리폴리티카(이전에 칸디다 리폴리티카(*Candida lipolytica*)로 분류)가 포함된다.
- [0309] 유지성 효모 야로위아 리폴리티카가 가장 바람직하며; 추가의 실시형태에서, ATCC #20362, ATCC #8862, ATCC #18944, ATCC #76982 및/또는 LGAM S(7)1로 명명되는 야로위아 리폴리티카 균주가 가장 바람직하다(문헌 [Papanikolaou S., and Aggelis G., *Bioresour. Technol.* 82(1):43-9 (2002)]).
- [0310] 상기 언급된 임의의 속은 적어도 25 EPA % DCW를 포함하는 미생물 오일을 생성할 수 있는 숙주 세포를 생성하기 위한 본 명세서의 개시내용에 따른 재조합 엔지니어링에 적절할 수 있다. 따라서, 유전자 조작이 적절한 PUFA 생합성 경로의 도입 또는 상향 조절에 제한되지 않는 것이 명백할 것이며; 대신에, 숙주 유기체가 추가로 유전적으로 조작되어 축적되는 전체 지질을 변경시키고, 글리세로인지질 생합성을 변경시키고, 세포를 통과하는 탄소 유동을 변경시키고, PUFA 분해를 (직접적으로 또는 간접적으로) 야기하는 경로를 변경시킬 수 있는 등등이다.
- [0311] EPA의 구조적 형태가 제한되지 않는 것이 알려져 있으며; 이에 따라, 예를 들어, EPA는 에스테르화 형태로 또는 FFA로서 전체 지질에 존재할 수 있다. 바람직하게는, 적어도 하나의 PUFA는 TAG의 형태이다.
- [0312] 비록 수많은 재조합 미생물 숙주 세포가 본 명세서의 개시내용에 따라 EPA를 생성하도록 엔지니어링될 수 있지만, 본 발명은 야로위아 리폴리티카에서 입증되었다. 그러나, 당업자는 본 발명의 방법이 본 발명이 입증하는 종 또는 속의 사용에 제한되지 않는 것을 인식할 것이다. 대신에, 적어도 25 EPA % DCW를 포함하는 미생물 오일을 생성할 수 있는 임의의 유지성 효모 또는 임의의 다른 적절한 미생물은 본 발명의 방법에 사용하기에 동등하게 적절할 것이다.
- [0313] 외래 유전자의 높은 수준의 발현을 지시하는 조절 서열을 함유하는 미생물 발현 체계 및 발현 벡터가 당업자에게 잘 알려져 있다. 이들 중 임의의 것을 사용하여 바람직한 불포화화효소, 연장효소, DGLA 신타제, LPAAT, PDAT, PCT, CPT1 및 MCS 단백질을 암호화하는 키메라 유전자를 구축할 수 있다. 그 다음, 이들 키메라 유전자를 표준 형질전환 방법을 사용하여 미생물 숙주 세포 내로 도입시켜, 암호화된 효소의 높은 수준의 발현을 제공할 수 있다.
- [0314] 미생물 숙주 세포의 형질전환에 유용한 벡터(예컨대, 구축물, 플라스미드) 및 DNA 발현 카세트는 당업계에 잘 알려져 있다. 구축물에 존재하는 서열의 특이적 선택은 목적하는 발현 생성물, 숙주 세포의 성질 및 형질전환된 세포를 형질전환되지 않은 세포에 대하여 분리하기 위한 제한된 수단에 좌우된다. 그러나, 전형적으로, 벡터는 적어도 하나의 발현 카세트, 선택가능한 마커 및 자가 복제 또는 염색체 통합을 가능하게 하는 서열을 함유한다. 적절한 발현 카세트는 전형적으로 프로모터, 선택된 유전자의 암호화 서열 및 종결자를 포함한다. 둘 모두의 조절 영역이 형질전환된 숙주 세포로부터의 유전자로부터 유래되는 경우가 가장 바람직하다.
- [0315] 둘 이상의 유전자를 별개의 복제 벡터로부터 발현시키는 경우에는, 각 벡터가 상이한 선별 수단을 갖는 것이 바람직하며, 안정적인 발현을 유지하고 구축물 사이에서의 요소의 재배열을 방지하기 위해 다른 구축물과의 상동성이 없어야 한다. 조절 영역, 선별 수단 및 도입된 구축물을 증가시키는 방법의 현명한 선택은, 도입된 모든 유전자가 목적하는 생성물의 합성을 제공하는데 필요한 수준으로 발현되도록 실험적으로 결정할 수 있다.
- [0316] 대상 유전자(들)를 포함하는 구축물 또는 벡터는 임의의 표준 기술에 의해 숙주 세포 내로 도입될 수 있다. 이

들 기술로는 형질전환(예컨대, 아세트산리튬 형질전환(문헌[*Methods in Enzymology*, 194:186-187 (1991)]), 바이올리스틱 충격, 전기천공, 미세주입, 또는 대상 유전자(들)를 숙주 세포 내로 도입하는 임의의 다른 방법)이 포함된다. 일 예로서, 미국 특허 제4,880,741호 및 제5,071,764호 및 문헌[Chen, D. C. et al. (*Appl. Microbiol. Biotechnol.*, 48(2):232-235 (1997))]에는 선형화된 DNA 단편에 기초한 야로위아 리폴리티카를 위한 통합 기술이 기재되어 있다.

[0317] 편의상, DNA 서열을 취하는 임의의 방법(예컨대, 발현 카세트)에 의해 조작되는 숙주 세포는 본 명세서에서 "형질전환된", "형질전환체" 또는 "재조합체"라고 지칭된다. 형질전환된 숙주는 적어도 하나의 카피의 발현 카세트를 가질 것이며, 발현 카세트가 계놈 내로 통합되는지 아니면 다중 카피수를 갖는 염색체의 요소 상에 존재하는지의 여부에 따라 2개 이상을 가질 수 있다. 형질전환된 숙주 세포는 다양한 선택 기술에 의해, 예컨대 미국 특허 제7,238,482호 및 미국 특허 제7,259,255호에 기재된 바와 같이 확인될 수 있다.

[0318] 본 명세서에서 사용하기에 바람직한 선택 방법에는 카나마이신, 하이그로마이신 및 아미노 글리코시드 G418에 대한 내성뿐 아니라, 우라실, 류신, 라이신, 트립토판 또는 히스티딘이 결여된 배지에서의 성장 능력이 있다. 대안적인 실시형태에서, 5-플루오로오로트산(5-플루오로우라실-6-카복실산 1수화물; "5-FOA")이 효모 Ura⁻ 돌연변이체의 선택을 위해 사용되거나(미국 특허 출원 공개 제2009-0093543-A1호), 또는 설폰닐 우레아 제조제 내성을 부여하는 천연의 아세트하이드록시산 신타제(또는 아세트락테이트 신타제; E.C. 4.1.3.18)(국제 출원 공개 WO 2006/052870호)가 형질전환체의 선택을 위해 사용된다. 부위-특이적 재조합효소 체계의 사용에 의하여, 다수의 순차적인 형질전환에서 그들을 사용하기 위한 바람직한 선택 마커의 쌍의 독특한 "재활용" 방법도 또한 미국 특허 출원 공개 제2009-0093543-A1호에 교시되어 있다.

[0319] 당업자에게 잘 알려져 있는 바와 같이, 단순히 유전자(예컨대, 불포화화효소, 연장효소, DGLA 신타제, LPAAT, PDAT, PCT, CPT1, MCS)를 클로닝 벡터에 삽입하는 것에 의해서는 원하는 비율, 농도, 양 등에서의 그의 발현이 보장되지는 않는다. 전사, RNA 안정성, 번역, 단백질 안정성 및 단백질 위치, 산소 제한 및 숙주 세포로부터의 분비의 면을 조절하는 다수의 상이한 유전 요소를 조작하는 것이 바람직할 수 있다. 더욱 구체적으로, 유전자 발현은 관련있는 프로모터 및 종결자 서열의 성질; 클로닝된 유전자의 카피수; 유전자가 플라스미드-운반형(plasmid-borne)인지 숙주 세포의 계놈 내로 통합되는지 여부; 합성된 외래 단백질의 최종 세포내 위치; 숙주 유기체에서의 번역의 효율; 숙주 세포 내의 클로닝된 유전자 단백질의 고유 안정성; 및 클로닝된 유전자 내의 코돈 사용의 빈도가 숙주 세포의 바람직한 코돈 사용 빈도에 접근되도록 클로닝된 유전자 내의 코돈 사용을 변경시킴으로써 조절될 수 있다. 이들 몇몇 과발현 방법이 후술될 것이며, 재조합 미생물 숙주 세포의 유전적 조작 동안 예컨대, 불포화화효소, 연장효소, DGLA 신타제, LPAAT, PDAT, PCT, CPT1 및 MCS를 암호화하는 유전자를 과발현시키기 위한 수단으로서 유용하다.

[0320] 원하는 유전자(들)의 발현은 mRNA 또는 암호화된 단백질 중 어느 하나로부터의 불안정화 서열의 제거/결실에 의하여 또는 안정화 서열의 mRNA로의 첨가에 의하여 발현을 증가시키기 위하여 더 강력한 프로모터(조절형 또는 구성적 프로모터 중 어느 하나)의 사용을 통하여 전사 수준에서 증가될 수 있다(미국 특허 제4,910,141호).

[0321] 미생물 숙주 세포에서 이중 유전자가 발현되게 추진하는데 유용한 프로모터는 다양하며, 당업자에게 공지되어 있다. 발현은 유도된 방식 또는 구성적 방식으로 달성될 수 있다. 유도된 발현은 대상 유전자에 작동가능하게 연결된 조절가능한 프로모터의 활성을 유도함으로써 달성될 수 있으며, 구성적 발현은 대상 유전자에 작동가능하게 연결된 구성적 프로모터의 사용에 의해 달성될 수 있다. 사실상, 불포화화효소, 연장효소, DGLA 신타제, LPAAT, PDAT, PCT, CPT1 및 MCS 유전자의 발현을 지시할 수 있는 임의의 프로모터(즉, 천연, 합성 또는 키메라)가 적절할 것이지만, 숙주 종으로부터의 전사 및 번역 영역이 특히 유용하다.

[0322] 일반적으로, 종결자는 프로모터가 수득되는 유전자의 3' 영역 또는 상이한 유전자로부터 유래될 수 있다. 매우 다수의 종결자가 공지되어 있으며, 그들이 유래된 것과 동일한 속과 종, 및 그와 상이한 속과 종에서 이용하는 경우 둘 모두에 다양한 숙주에서 만족스럽게 기능한다. 종결자는 통상 임의의 특정 특성 때문이 아니라, 편의성의 문제로 선택된다. 바람직하게는, 종결자는 효모 유전자 유래의 것이다. 또한, 종결자는 당업자가 종결자를 설계하고 합성하기 위하여 입수할 수 있는 정보를 이용할 수 있음에 따라, 합성될 수 있다. 종결자가 불필요할 수 있지만, 이는 매우 바람직하다.

[0323] 비록 제한하고자 하는 것은 아니지만, 야로위아 속의 재조합 미생물 숙주 세포에서 사용하기에 바람직한 프로모터 및 종결자는 미국 특허 공개 제2009-0093543-A1호, 미국 특허 공개 제2010-0068789-A1호, 미국 특허 공개 제2011-0059496-A1호, 미국 가특허 출원 제61/469,933호(대리인 문서 번호 CL4736USPRV(출원일: 2011년 3월 31일)), 미국 가특허 출원 제61/470,539호(대리인 문서 번호 CL5380USPRV(출원일: 2011년 4월 1일)), 미국 가특

허 출원 제61/471,736호(대리인 문서 번호 CL5381USPRV(출원일: 2011년 4월 5일)) 및 미국 가특허 출원 제 61/472,742호(대리인 문서 번호 CL5382USPRV(출원일: 2011년 4월 7일))에 교시된 것이며, 각각의 개시내용은 본 명세서에 참고로 포함된다.

- [0324] 추가의 카피(즉, 1 초과 카피)의 PUFA 생합성 경로 불포화효소, 연장효소, DGLA 신타제, LPAAT, PDAT, PCT, CPT1 및 MCS 유전자를 재조합 미생물 숙주 세포 내로 도입하여, 이에 의해 EPA 생성 및 축적을 증가시킬 수 있다. 구체적으로, 추가의 카피의 유전자는 단일의 발현 구축물 내에 클로닝될 수 있고/거나; 추가의 카피의 클로닝된 유전자(들)는 플라스미드 카피수의 증가에 의해 또는 클로닝된 유전자의 게놈 내로의 다수의 통합(하기)에 의해 숙주 세포 내로 도입될 수 있다.
- [0325] 본 명세서의 방법에 따라 최적화된 재조합 미생물 숙주 세포를 생성하는 경우, 다양한 불포화효소, 연장효소, DGLA 신타제, LPAAT, PDAT, PCT, CPT1 및 MCS의 카피가 종종 언급되는 것을 주의하는 것이 중요하다. 예를 들어, 2 카피의 델타-9 연장효소가 요구된다면, 이는 1) 단일의 종으로부터 분리된 특정 델타-9 연장효소에 대한 동일한 2 카피의 암호화 서열; 또는 2) 종 "A"로부터 분리된 델타-9 연장효소에 대한 하나의 암호화 서열 및 종 "B"로부터 분리된 델타-9 연장효소에 대한 하나의 암호화 서열(이에 따라, 집합하여 2개의 델타-9 연장효소가 야기됨)을 말할 수 있다.
- [0326] 일반적으로, 일단 재조합 미생물 숙주 세포에서의 발현에 적절한 DNA 카세트(예컨대, 프로모터, ORF 및 종결자를 포함하는 키메라 유전자 포함)가 수득되면, 그를 숙주 세포에서 자가 복제될 수 있는 플라스미드 벡터 내에 배치하거나, 또는 숙주 세포의 게놈 내로 직접 통합시킨다. 발현 카세트의 통합은 숙주 게놈 내에서 무작위로 발생할 수 있거나, 숙주 유전자좌와의 재조합을 표적화하기에 충분한 숙주 게놈과의 상동성 영역을 함유하는 구축물의 사용을 통하여 표적화시킬 수 있다. 비록 본 명세서에서 필요로 하지 않지만, 전사 및 번역 조절 영역의 전부 또는 일부가 구축물이 표적화되는 내인성 유전자좌에 의해 제공될 수 있다.
- [0327] 엔지니어링된 재조합 야로위아 리폴리티카 숙주 세포에 대하여, 이러한 미생물 숙주에서 유전자를 발현하는 바람직한 방법은 선형 DNA 단편을 숙주의 게놈 내로 통합하는 것에 의한 것이다. 높은 유전자 발현 수준이 바람직한 경우, 게놈 내의 다수의 위치로의 통합이 특히 유용할 수 있다. 바람직한 유전자좌는 미국 특허 공개 제 2009-0093543-A1호에 교시된 것을 포함한다.
- [0328] 추가로, 유레트젝(Juretzek) 등(문헌[Yeast, 18:97-113 (2001)])은 야로위아 리폴리티카에서의 통합된 DNA 단편의 안정성이 사용되는 개별 형질전환체, 수용 균주 및 표적화 플랫폼에 좌우됨을 언급하고 있다. 따라서, 당업자는 특정 재조합 미생물 숙주의 다수의 형질전환체를 스크리닝하여, 원하는 발현 수준 및 패턴을 나타내는 균주를 수득해야 하는 것을 인식할 것이다. 이러한 스크리닝은 DNA 블롯의 서던 분석(문헌[Southern, J. Mol. Biol., 98:503 (1975)]), mRNA 발현의 노던 분석(문헌[Kroczeck, J. Chromatogr. Biomed. Appl., 618 (1-2):133-145 (1993)]), 단백질 발현의 웨스턴 분석, PUFA 생성물의 표현형 분석 또는 GC 분석에 의해 달성될 수 있다.
- [0329] 본 개시내용의 형질전환된 재조합 미생물 숙주 세포는 키메라 유전자(예컨대, 불포화효소, 연장효소, DGLA 신타제, LPAAT, PDAT, PCT, CPT1 및 MCS를 암호화하는)의 발현을 최적화시키고, 가장 경제적이며 가장 큰 수율의 EPA를 생성하는 조건 하에서 성장시킨다. 일반적으로, 탄소원의 유형 및 양, 질소원의 유형 및 양, 탄소-대-질소 비, 상이한 미네랄 이온의 양, 산소 수준, 성장 온도, pH, 바이오매스 생성 단계의 기간, 오일 축적 단계의 기간 및 세포 수집 시간 및 방법을 변경함으로써 배지 조건을 최적화시킬 수 있다. 예를 들어, 야로위아 리폴리티카는 일반적으로 복합 배지, 예컨대 효모 추출물-펩톤-텍스트로스 브로쓰["YPD"] 또는 성장에 필수적인 성분이 결여되어 목적하는 발현 카세트가 선택되게 하는 정의된 최소 배지(예컨대, 이스트 니트로젠 베이스(미국 미시간주 디트로이트 소재의 디프코 래보러터리즈(DIFCO Laboratories)))에서 성장시킨다.
- [0330] 본 명세서에 기재된 방법 및 숙주 세포를 위한 발효 배지는 적절한 탄소원, 예컨대 미국 특허 제7,238,482호 및 미국 특허 공개 제2011-0059204-A1호에 교시된 것을 함유해야 한다. 비록 본 발명에 사용되는 탄소원이 매우 다양한 탄소-함유 공급원을 포함할 수 있는 것으로 여겨지지만, 바람직한 탄소원은 당(예컨대, 글루코스, 전화당(invert sucrose), 프룩토스 및 그들의 조합), 글리세롤 및/또는 지방산(예컨대, 10 내지 22개 탄소를 함유하는 것)이다.
- [0331] 질소는 무기 공급원(예컨대, (NH₄)₂SO₄) 또는 유기 공급원(예컨대, 우레아 또는 글루타메이트) 공급원으로부터 공급될 수 있다. 적절한 탄소원과 질소원에 더하여, 발효 배지는 적절한 미네랄, 염, 보조인자, 완충제, 비타민 및 재조합 미생물 숙주 세포의 성장과 EPA 생성을 위한 효소 경로의 촉진에 적절한 것으로 당업자에게 공지된

다른 성분도 함유해야 한다. 지질과 PUFA의 합성을 촉진시키는 몇가지 금속 이온, 예를 들어 Fe^{+2} , Cu^{+2} , Mn^{+2} , Co^{+2} , Zn^{+2} 및 Mg^{+2} 에 특히 주목한다(문헌[Nakahara, T. et al., *Ind. Appl. Single Cell Oils*, D. J. Kyle and R. Colin, eds. pp 61-97 (1992)]).

[0332] 본 명세서에 기재된 방법 및 숙주 세포에 바람직한 성장 배지는 통상적인 상업적으로 제조된 배지, 예를 들어 이스트 니트로겐 베이스(미국 미시간주 디트로이트 소재의 디프코 래보러터리즈)이다. 다른 정의된 성장 배지 또는 합성 성장 배지를 사용할 수도 있고, 예컨대 야로위아 리폴리티카의 성장에 적절한 배지는 미생물학 또는 발효 과학 분야의 전문가에게 공지되어 있을 것이다. 전형적으로, 발효에 적절한 pH 범위는 약 pH 4.0 내지 pH 8.0이며, pH 5.5 내지 pH 7.5가 초기 성장 조건의 범위로서 바람직하다. 발효는 호기성 또는 혐기성 조건 하에 행해질 수 있으며, 미세호기성(microaerobic) 조건이 바람직하다.

[0333] 전형적으로, 높은 수준의 PUFA가 유지성 효모 세포에 축적되기 위해서는 2-단계 공정이 필요한데, 이는 대사 상태가 성장과 지방의 합성/저장 간에 "균형을 이루어야" 하기 때문이다. 따라서, 가장 바람직하게는, 야로위아 리폴리티카에서 EPA를 생성시키기 위해 2-단계 발효 공정이 필요하다. 이러한 접근법은 다양한 적절한 발효 공정 설계(즉, 회분식, 유가식(fed-batch) 및 연속식) 및 성장 동안의 고려사항이 미국 특허 제7,238,482호에 기재되어 있다.

[0334] 미국 특허 출원 공개 제2009-0093543-A1호의 실시예 10은 또한 재조합 야로위아 리폴리티카 균주 Y4305의 2-L 발효에 필요한 파라미터의 상세한 설명을 제공한다(그의 최대 생성은 162시간의 기간에 걸친 12.1 EPA % DCW[즉, 55.6 EPA % TFA, 3.03의 EPA % TFA 대 LA % TFA 비]였다). 본 개시내용은 씨드 배양물을 생성하기 위하여 동결된 배양물로부터 접종원을 준비하고, 처음에 신속한 성장을 촉진시키는 조건 하에서 효모를 높은 세포 밀도로 배양한 다음, 효모를 배양하여 지질 및 PUFA 축적을 촉진시키는(연속 공급 글루코스 및 질소에 대한 결핍을 통하여) 수단의 기재를 포함한다. 온도(30 내지 32°C로 조절), pH(5 내지 7로 조절), 용존 산소 농도 및 글루코스 농도를 비롯한 공정 변수를 표준 운영 조건마다 감시하고 조절하여, 일관된 공정 성능 및 최종 PUFA 오일 품질을 보장하였다.

[0335] 특히, 미국 특허 출원 공개 제2009-0093543-A1호의 실시예 10의 데이터를 사용하여, 본 명세서의 도 2에 나타내고 하기의 표 6에 요약된 바와 같이, EPA % TFA 및 LA % TFA가 발효 과정에 걸쳐 어떻게 달라지는 지를 나타내는 그래프를 생성할 수 있다.

표 6

발효 과정에 걸친 EPA % TFA와 LA % TFA의 관계

발효 시간	EPA % TFA	LA % TFA	EPA:LA 비	DCW의 EPA%
68.3	33.2	26.6	1.25	6.18
99.7	46.9	22.6	2.08	10.18
123.7	53.6	19.7	2.72	11.68

[0336]

[0337] 특히, EPA % TFA가 약 40 내지 125시간의 발효 동안 증가되고, LA % TFA가 약 40 내지 125시간의 발효 동안 감소되고, EPA:LA 비가 증가됨을 주목해야 한다. 야로위아 리폴리티카 균주 Y4305가 12.1 EPA % DCW를 생성할 수 있지만, 재조합 미생물 숙주 세포의 오일 프로파일은 발효 시행 그 자체, 배지 조건, 공정 파라미터, 대량화 등 뿐만 아니라 배양물이 시료 추출되는 특정 시점에 좌우될 것임이 명백하다. 따라서, 엔지니어링된 균주는 다양한 상이한 지질 함량 및 조성을 갖는 미생물 오일을 생성할 수 있었다(즉, EPA % TFA, LA % TFA 및 EPA:LA 비에 기초하여).

[0338] 또한, 이들 인자는 본 명세서에 기재된 재조합 미생물 숙주 세포를 배양하는 경우, 엔지니어링된 숙주 세포의 완전한 잠재력을 실현하고, 임의의 특정 발효 시행에서 적어도 25 EPA % DCW를 달성하는 것으로 고려되어야 한다.

[0339] 본 명세서의 일부 태양에서, 주요 생성물은 재조합 미생물 바이오매스이다. 이와 같이, 미생물 바이오매스로부터의 EPA-함유 오일의 분리 및 정제는 필요하지 않을 수 있다(다시 말하면, 여기서, 전체 세포 바이오매스가 생성물이다). 그러나, 특정 최종 용도 및/또는 생성물 형태는 미생물 바이오매스로부터 EPA-함유 오일의 부분적인 및/또는 완전한 분리/정제를 필요로 하여, 부분 정제된 미생물 바이오매스, 정제된 오일 및/또는 정제된 EPA를 야기할 수 있다. 이들 태양에 관한 추가의 상세사항에 대해서는 미국 특허 출원 공개 제2010-0317072-A1호

를 참조한다.

- [0340] 개선되고/거나 정제된 EPA를 함유하는 오일을 수소화시켜, 이에 의해 다양한 용융 특성 및 질감을 갖는 지방을 야기할 수 있다. 스프레드(spread), 제과 지방(confectionary fat), 하드 버터(hard butter), 마가린, 베이킹 쇼트닝 등을 비롯한 많은 가공된 지방은 실온에서 다양한 고형도를 필요로 하며, 공급원 오일의 물리적 특성의 변경을 통해서만 생성될 수 있다. 이는 촉매적 수소화를 통해 가장 통상적으로 달성된다(추가의 상세사항 및 참고문헌에 대해서는 국제 출원 공개 WO 2006/052870호 참조).
- [0341] EPA를 포함하는 유지성 효모 바이오매스를 포함하는 식품, 유아용 조제식, 기능성 식품, 의료식, 의료용 영양제, 식이 보조제, 약제학적 조성물, 동물 사료 및 개인 관리 제품이 미국 특허 출원 공개 제2010-0317072-A1호에 교시되어 있으며, 이들 이용은 본 명세서에서 EPA 그 자체를 포함하는 재조합 미생물 바이오매스 또는 그로부터 분리된 EPA를 포함하는 미생물 오일 중 어느 하나에 동일하게 적용가능하다.
- [0342] 가공 및 포물레이션(formulation) 분야의 숙련자는 소정의 양 및 조성의 재조합 미생물 바이오매스, 부분 정제된 바이오매스, 정제된 오일, 및/또는 정제된 EPA를 표적 중 및/또는 최종 용도에 따라 특정 생성물에 첨가할 수 있는 방법을 이해할 것이다. 더욱 구체적으로, "유효한" 양이 제품 포물레이션에 혼입되어야 하지만, 이러한 양은 식품 또는 사료 제품, 제품이 보충하고자 하는 식이 또는 의료식 또는 의료용 영양제가 바로잡거나 치료하고자 하는 의학적 병상에 좌우될 것이다.
- [0343] 가장 바람직하게는, 유효량의 EPA는 오메가-3/오메가-6 PUFA 섭취와 관련된 바람직한 건강상의 특징을 제공하기에 충분할 것이다. 전형적으로, 제품으로 혼입되는 EPA의 양은 몇가지 예를 들면, 가공 조건, 전형적인 취급 및 보관 조건, 제품 중의 EPA의 안정성 및 표적 중의 생체이용가능성/생체흡수 효율과 관련된 소실을 고려한다.
- [0344] 가공 및 포물레이션 분야의 숙련자는 본 명세서에 기재된 재조합 미생물 숙주 세포로부터 생성된 미생물 오일을 농축시켜, 이에 의해 전체 지질 분획 중의 EPA의 농도를 증가시켜, 이것이 적어도 약 55-60%, 적어도 약 60-65%, 적어도 약 65-70%, 적어도 약 70-85%, 적어도 약 85-90%, 적어도 약 90-95% EPA 또는 심지어 95-99% EPA를 포함하도록 하는 과정이 친숙할 것이다. 본 명세서에 기재된 정제된 오일을 다른 정제된 지방산(예컨대, LA, GLA, EDA, DGLA, ARA, DTA, DPA_n-6, ALA, STA, ETrA, ETA, DPA 및 DHA), 또는 대안의 지방산을 바람직한 농도로 포함하는 오일과 배합하는 수단도 또한 당업자에게 잘 알려져 있다. 이들 교시는 독특하게 맞춤화된 지방산 프로파일을 포함하는 오일을 용이하게 생성하는 것을 가능하게 한다.
- [0345] 바람직한 실시형태의 설명
- [0346] 유지성 효모 야로위아 리폴리티카의 다양한 재조합 균주는 하기 실시예에 기술된 바와 같이 25 초과 EPA % DCW를 생성하는 것으로 본 명세서에서 입증되었다. 표 7은 플라스크 검정법에 의해 결정시 유전형, 전체 지질 함량 및 지질 조성에 기초하여, 재조합 야로위아 리폴리티카 균주 Z1978(이전에 22.5 EPA % DCW를 생성하도록 엔지니어링된)에 대한 이들 일부 재조합 균주의 요약을 제공한다.
- [0347] 표 내의 각각의 구획은 단일의 형질전환 내에서 생성되는 하나의 균주 또는 균주들을 나타내며(다시 말하면, 균주 Z1977, Z1978, Z1979, Z1980 및 Z1981은 단일의 형질전환으로부터의 개별 콜로니였다), 이에 따라 동일한 유전형을 공유하는 것으로 예상된다. 균주 Z1978 유래의 균주의 유전형은 형질전환체에서 발현되는 추가의 유전자뿐 아니라 이들 유전자의 카피수를 강조표시하는 약칭 표기를 사용하여 균주 Z1978의 유전형에 대하여 요약한 것이다. 따라서, 예를 들어, 균주 L250은 균주 Z1978과 비교하여, Y1LPAAT1을 포함하는 하나의 추가의 발현 카세트 및 Y1PDAT를 포함하는 하나의 추가의 발현 카세트로 형질전환되었으며; 균주 L258은 균주 Z1978과 비교하여, Y1LPAAT1을 포함하는 2개의 추가의 발현 카세트 및 Y1PDAT를 포함하는 2개의 추가의 발현 카세트로 형질전환되었다.
- [0348] 표는 세포의 전체 건조 세포 중량["DCW"], 세포의 전체 지질 함량["TFA % DCW"], TFA의 중량 백분율["% TFA"]로서의 각 지방산의 농도, 건조 세포 중량의 백분율로서의 EPA 함량["EPA % DCW"] 및 EPA % TFA 대 LA % TFA의 비["EPA:LA 비"]를 요약한 것이다. 지방산은 16:0(팔미테이트), 16:1(팔미톨레산), 18:0(스테아르산), 18:1(올레산), 18:2(리놀레산), ALA(알파-리놀렌산), EDA(에이코사다이엔산), DGLA(다이호모-감마-리놀렌산), ARA(아라키돈산), ETrA(에이코사트라이엔산), ETA(에이코사테트라엔산), EPA(에이코사펜타엔산) 등이 있다.

표 7

플라스크 검정법에 의한 다양한 제조합 야로위아 리폴리타 균주에서의 전체 지질 함량 및 조성

균주	DCW(g/L)	% TFA											기타	EPA % DCW	EPA:LA 비		
		TFA % DCW	16:0	16:1	18:0	18:1	18:2	AL A	EDA	DGLA	ARA	EtrA				ETA	EPA
Z1977	3.8	34.3	2	0.5	1.9	4.6	11.2	0.7	3.1	3.3	0.9	0.7	2.2	59.1	9.9	20.3	5.28
Z1978	3.9	38.3	2.4	0.4	2.4	4.8	11.1	0.7	3.2	3.3	0.8	0.6	2.1	58.7	9.5	22.5	5.29
Z1979	3.7	33.7	2.3	0.4	2.4	4.1	10.5	0.6	3.2	3.6	0.9	0.6	2.2	59.4	9.8	20	5.66
Z1980	3.6	32.7	2.1	0.4	2.2	4	10.8	0.6	3.1	3.5	0.9	0.7	2.2	59.5	10	19.5	5.51
Z1981	3.5	34.3	2.2	0.4	2.1	4.2	10.6	0.6	3.3	3.4	1	0.8	2.2	58.5	10.7	20.1	5.52
균주 Z1978 에 대한 유전형 부가: YILPAAT1, YIPDAT																	
L250	4.4	51.5	2	0.7	2.8	6.1	16.7	0.9	3.3	4.9	0.7	0.6	3.2	50.4	7.4	26	3.02
균주 Z1978 에 대한 유전형 부가: 2 YILPAAT1, 2 YIPDAT																	
L258	5	57.1	2.3	0.9	3.4	7.8	18.7	0.9	4	5.3	0.8	0.6	3.2	45.2	6.6	25.8	2.42
균주 Z1978 에 대한 유전형 부가: EgD8M, MCS, MaLPAAT1S, 2 YILPAAT1, 2 YIPDAT																	
Z5565	4.8	56.1	2.1	0.8	2.8	6.8	17.3	0.8	3.8	5.2	1.1	0.8	3.4	47.4	7.1	26.6	2.74
Z5567	4.9	56.2	1.9	0.7	2.6	6.2	16.7	0.7	3.8	5.6	1.1	1	3.6	48.3	7.2	27.1	2.89
Z5575	4.7	53.8	1.8	0.7	2.4	5.7	15.3	0.6	3.6	5.9	1.2	1	3.6	50.4	7.4	27.1	3.29
Z5576	4.9	55.6	2.3	0.9	2.8	6.9	16.9	0.7	3.6	5.5	1.2	0.9	3.3	47.2	7.4	26.2	2.79
균주 Z1978 에 대한 유전형 부가: EgD8M, MCS, 3 YILPAAT1, 2 YIPDAT																	
Z5620	4.5	52.8	2.1	0.7	2.8	6.6	16.1	0.7	3.6	5.7	1.1	0.7	3.3	49	6.9	25.9	3.04
Z5623	4.3	51.7	2.3	0.8	2.4	6	15.9	0.7	3.8	5.2	1.1	0.7	3.1	50	7.3	25.8	3.14
Z5625	4.6	52.7	2.1	0.7	2.7	6.2	16.6	0.7	3.9	5.4	1.1	0.8	3.2	49.1	6.8	25.9	2.96
균주 Z1978 에 대한 유전형 부가: ME3S, MCS, MaLPAAT1S, 2 YILPAAT1, 2 YIPDAT																	
Z5581	4.7	56.3	1.9	0.7	2.6	6.1	16.5	0.7	3.7	5.6	1.2	1	3.5	48.7	7.2	27.4	2.95
Z5582	4.8	55.6	1.9	0.7	2.5	6.1	16.4	0.7	3.7	5.7	1.1	0.9	3.6	48.9	7.2	27.2	2.98

[0349]

Z5583	4.9	56.8	2	0.7	2.6	6.2	16.7	0.8	3.7	5.4	1	1	3.7	48.4	7.2	27.5
Z5584	4.9	55.3	2	0.7	2.7	6.5	16.1	0.7	3.7	5.7	1.1	1	3.6	48.6	7.1	26.8
균주 Z1978 에 대한 유전형 부가: YIPCT, YID9, MaLPAAT1S, 2 YILPAAT1, 2 YIPDAT																
Z5570	4.8	55	2	0.8	2.5	6.1	16.4	0.7	3.7	5.5	1.2	1	3.4	48.6	7.4	26.8
Z5571	4.8	54.1	2.2	0.8	2.4	6.5	16.7	0.7	3.8	5.5	1.1	0.9	3.2	48.3	7.2	26.2
Z5572	4.9	54	2.1	0.8	2.5	6.5	16.7	0.7	3.7	5.5	1.1	0.9	3.3	48.4	7.2	26.1
Z5574	5	53.8	1.8	0.7	2.4	5.7	15.3	0.6	3.6	5.9	1.2	1	3.6	50.4	7.4	27.1
균주 Z1978 에 대한 유전형 부가: YICPT1, YID9, MaLPAAT1S, 2 YILPAAT1, 2 YIPDAT																
Z5585	4.6	56.6	1.9	0.7	2.6	5.6	16.4	0.7	3.5	5.5	1.1	1	3.5	49.4	7.3	28
Z5627	4.8	52	1.9	0.7	2.6	6.2	16.1	0.6	4	5.6	1.2	0.9	3.2	49.3	6.9	25.6

[0350]

[0351]

적어도 LPAAT 활성을 갖는 적어도 2개의 폴리펩티드 및 적어도 PDAT 활성을 갖는 적어도 하나의 폴리펩티드를 포함하는 균주 내에서, EPA % DCW는 25.6 내지 28 범위이며, EPA % TFA는 45.2 내지 50.4 범위이며, EPA % TFA 대 LA % TFA의 비("EPA:LA 비")는 2.42 내지 3.29 범위이다.

[0352]

이들 모든 균주가 25 초과의 EPA % DCW를 생성하였지만, 전체 발효 시간을 단축시킴으로써 동일한 균주를 사용하여 25 미만의 EPA % DCW를 생성할 수 있음을 주목할 가치가 있다. 도 2에 관하여 상술된 제조합 야로위아 리폴리타 균주 Y4305의 성능과 유사하게, 이들 표 7의 균주는 발효 시간이 단축된다면, 더 적은 EPA % TFA, 더 많은 LA % TFA를 생성하며 EPA:LA 비가 감소되는 것으로 예상될 것이다. 따라서, 당업자는 이들 엔지니어링된 야로위아 리폴리타 균주가 발효에서의 특정 시료 추출 시점에 따라, 다양한 EPA:LA 비와 함께, 다양한 농도의 EPA를 갖는 미생물 오일을 생성할 수 있음을 인식할 것이다.

[0353]

EPA, LA 및 올레산은 TFA의 대략 70-75%로 포함된다. 또한, 본 명세서에 기재된 향상된 최적화 제조합 야로위아 리폴리타 균주는 약 0.5% 미만의 GLA 또는 DHA를 가지며(약 0.1%까지의 검출가능한 수준을 갖는 장비를 사용하여 GC 분석으로 측정하여), 약 8% 미만의 포화 지방산 함량을 갖는 것으로 구별된다. 이러한 낮은 백분율의 포화 지방산(다시 말하면, 16:0 및 18:0)은 인간 및 동물에게 상당한 건강상의 이익을 야기한다.

[0354]

실시예

[0355]

본 발명은 하기 실시예에서 추가로 정의된다. 이들 실시예는 본 발명의 바람직한 실시형태를 나타내면서도 단지 예로써만 주어지는 것으로 이해해야 할 것이다. 상기 논의 및 이러한 실시예로부터, 당업자는 본 발명의 본질적 특징을 확인할 수 있으며, 본 발명의 사상 및 범주를 벗어나지 않는 한 본 발명을 다양하게 변경하고 수정하여 다양한 용도 및 조건에 적합하게 할 수 있다.

[0356]

일반 방법

[0357]

본 실시예에 사용되는 표준 제조합 DNA 및 분자 클로닝 기술은 당업계에 주지되어 있으며, 1) 문헌[Sambrook, J., Fritsch, E. F. and Maniatis, T. *Molecular Cloning: A Laboratory Manual*; Cold Spring Harbor

Laboratory: Cold Spring Harbor, NY (1989) (Maniatis)]; 2) 문헌[T. J. Silhavy, M. L. Bannan, and L. W. Enquist, Experiments with Gene Fusions; Cold Spring Harbor Laboratory: Cold Spring Harbor, NY (1984)]; 및 3) 문헌[Ausubel, F. M. et al., Current Protocols in Molecular Biology, published by Greene Publishing Assoc. and Wiley-Interscience, Hoboken, NJ (1987)]에 기재되어 있다.

- [0358] 미생물 배양의 유지 및 성장에 적절한 재료 및 방법은 당업계에 잘 알려져 있다. 하기의 실시예에 사용하기에 적절한 기술은 문헌[Manual of Methods for General Bacteriology (P. Gerhardt, R.G.E. Murray, R.N. Costilow, E.W. Nester, W.A. Wood, N.R. Krieg and G.B. Phillips, Eds), American Society for Microbiology: Washington, D.C. (1994)]; 또는 문헌[Thomas D. Brock in Biotechnology: A Textbook of Industrial Microbiology, 2nd ed., Sinauer Associates: Sunderland, MA (1989)]에 기재된 바와 같이 찾을 수 있다. 미생물 세포의 성장 및 유지에 사용되는 모든 시약, 제한 효소 및 재료는 달리 특정되지 않는 한, 디프코 래보러터리즈(미국 미시간주 디트로이트 소재), 뉴 잉글랜드 바이오랩스, 인코포레이티드(New England Biolabs, Inc.)(미국 매사추세츠주 베벌리 소재), 집코(GIBCO)/BRL(미국 메릴랜드주 게이더스버그 소재) 또는 시그마-알드리치 케미컬 컴퍼니(Sigma-Aldrich Chemical Company)(미국 미주리주 세인트 루이스 소재)로부터 취득하였다. 에스케리키아 콜라이 균주를 전형적으로 루리아-베르타니(Luria Bertani["LB"]) 플레이트 상에서 37 °C에서 성장시켰다.
- [0359] 일반적 분자 클로닝을 표준 방법에 따라 수행하였다(상기 문헌[Sambrook et al.]). PCR 또는 부위-지정 돌연변이 유발이 서브클로닝에 수반되는 경우, 구축물을 시퀀싱하여, 서열 내로 오류가 도입되지 않음을 확인하였다. PCR 생성물을 프로메가(Promega)의 pGEM-T-easy(미국 위스콘신주 매디슨 소재) 또는 pCR 4 TOPO(인비트로젠(Invitrogen), 미국 캘리포니아주 샌 디에고 소재) 벡터로 클로닝하였다.
- [0360] 약어의 의미는 하기와 같다: "sec"는 초를 의미하고, "min"은 분을 의미하며, "h"는 시간을 의미하고, "d"는 일 을 의미하며, " μ l"은 마이크로리터를 의미하고, "mL"은 밀리리터를 의미하며, "L"은 리터를 의미하며, " μ M"은 마이크로몰을 의미하고, "mM"은 밀리몰을 의미하며, "M"은 몰을 의미하고, "mmol"은 밀리몰을 의미하며, " μ mole"는 마이크로몰을 의미하고, "g"는 그램을 의미하며, " μ g"는 마이크로그램을 의미하고, "ng"는 나노그램을 의미하며, "U"는 유닛을 의미하고, "bp"는 염기쌍을 의미하며, "kB"는 킬로베이스를 의미하고, "DCW"는 건조 세포 중량을 의미하며, "TFA"는 전체 지방산을 의미하고, "FAME"는 지방산 메틸 에스테르를 의미한다.
- [0361] 발현 카세트에 대한 명명법
- [0362] 발현 카세트의 구조는 "X::Y::Z"의 간단한 표기 체계에 의해 나타낼 것이며, 여기서, X는 프로모터 단편을 말하며, Y는 유전자 단편을 말하고, Z는 종결자 단편을 말하며, 이들 모두는 서로 작동가능하게 연결된다.
- [0363] 야로위아 리폴리티카의 형질전환 및 배양
- [0364] 야로위아 리폴리티카 균주 ATCC #20362를 아메리칸 타입 컬처 콜렉션(미국 미들랜드주 록빌 소재)으로부터 구입 하였다. 야로위아 리폴리티카 균주를 전형적으로 하기에 나타낸 방법에 따라 몇몇 배지에서 28 내지 30°C에서 성장시켰다. 아가(agar) 플레이트를 표준 방법에 따라, 각 액체 배지에 20 g/L의 아가를 첨가함으로써 요구되는 바와 같이 준비하였다.
- [0365] YPD 아가 배지(리터당): 10 g의 효모 추출물[디프코], 20 g의 박토(Bacto)
- [0366] 웹톤[디프코] 및 20 g의 글루코스.
- [0367] 기본 최소 배지["MM"](리터당): 20 g의 글루코스, 1.7 g의 이스트 니트로젠 베이스(아미노산 없음), 1.0 g의 프롤린 및 pH 6.1(조정할 필요 없음).
- [0368] 최소 배지 + 5-플루오로오로트산["MM + 5-FOA"](리터당): 20 g의 글루코스, 6.7 g의 이스트 니트로젠 베이스, 75 mg의 우라실, 75 mg의 우리딘
- [0369] 및 100 mg/L 내지 1000 mg/L 범위의 농도에 대한 FOA 활성 시험에 기초하여(공급처로부터 받은 각각의 배지에서 변이가 발생하기 때문에), 적절한 양의 FOA(미국 캘리포니아주 오렌지 소재의 자이모 리서치 코퍼레이션(Zymo Research Corp.)).
- [0370] 고 글루코스 배지["HGM"](리터당): 80의 글루코스, 2.58 g의 KH_2PO_4 및
- [0371] 5.36 g의 K_2HPO_4 , pH 7.5(조정할 필요가 없음).

- [0372] 발효 배지["FM"] (리터당): 6.70 g/L의 이스트 니트로겐 베이스, 6.00 g의
- [0373] KH_2PO_4 , 2.00 g의 K_2HPO_4 , 1.50 g의 $\text{MgSO}_4 \cdot 7\text{H}_2\text{O}$, 20 g의 글루코스 및 5.00 g의 효모 추출물(BBL).
- [0374] 야로위아 리폴리티카의 형질전환을 본 명세서에 참고로 포함되는 미국 특허 출원 공개 제2009-0093543-A1호에 기재된 바와 같이 수행하였다.
- [0375] 야로위아 리폴리티카의 지방산 분석
- [0376] 지방산["FA"] 분석을 위하여, 세포를 원심분리에 의해 수집하고, 지질을 문헌[Bligh, E. G. & Dyer, W. J. (*Can. J. Biochem. Physiol.*, 37:911-917 (1959))]에 기재된 바와 같이 추출하였다. 지방산 메틸 에스테르["FAME"]를 소듐 메톡시드를 사용한 지질 추출물의 에스테르 교환에 의해 제조하고(문헌[Roughan, G., and Nishida I., *Arch Biochem Biophys.*, 276(1):38-46 (1990)]), 이어서 30-m X 0.25 mm(내경) HP-INNOWAX(휴렛-팩카드)(Hewlett-Packard) 컬럼이 장착된 휴렛-팩카드 6890 GC로 분석하였다. 오븐 온도는 3.5°C/분으로 170°C (25분 유지)로부터 185°C까지였다.
- [0377] 직접적인 염기 에스테르 교환을 위하여, 야로위아 세포(0.5 mL 배양물)를 수집하고, 증류수에 1회 세정하고, 스피드-백(Speed-Vac)에서 5 내지 10분 동안 진공 하에 건조시켰다. 소듐 메톡시드(1%의 100 μl) 및 공지된 양의 C15:0 트리아실글리세롤(C15:0 TAG; 카달로그 번호 T-145, 미국 미네소타주 엘리시안 소재의 누-체크 프랩)을 시료에 첨가한 다음, 시료를 볼텍싱(vortexd)하고, 50°C에서 30분 동안 흔들었다. 3방울의 1 M NaCl 및 400 μl 의 헥산을 첨가한 후에, 시료를 볼텍싱하고, 회전시켰다. 상층을 제거하고, GC로 분석하였다.
- [0378] 대안적으로, 문헌[Lipid Analysis, William W. Christie, 2003]에 기재된 염기-촉매작용된 에스테르 교환 방법의 변형법을 발효 또는 플라스크 시료 중 어느 하나로부터의 브로쓰(broth) 시료의 통상적인 분석을 위해 사용하였다. 구체적으로, 브로쓰 시료를 실온의 물에서 빠르게 해동시킨 다음, 0.22 μm 코닝(Corning)(등록 상표) 코스타(Costar)(등록 상표) 스핀(Spin)-X(등록 상표) 원심분리용 튜브 필터(카달로그 번호 8161)가 있는 타르를 칠한 2 mL 미량 원심분리용 튜브에 칭량해 넣었다(0.1 mg까지). 이전에 결정된 DCW에 따라 시료(75 - 800 μl)를 사용하였다. 에펜도르프(Eppendorf) 5430 원심분리기를 사용하여, 시료를 14,000 rpm에서 5 내지 7분 동안 또는 브로쓰를 제거하는데 필요한 만큼 원심분리하였다. 필터를 제거하고, 액체를 빼내고, 약 500 μl 의 탈이온수를 필터에 첨가하여 시료를 세정하였다. 물을 제거하기 위한 원심분리 후에, 필터를 다시 제거하고, 액체를 빼내고, 필터를 다시 삽입하였다. 그 다음, 튜브를 원심분리기에 다시 삽입하고, 이번에는 상층을 개방하여 약 3 내지 5분 동안 건조시켰다. 이어서, 필터를 튜브의 대략 절반에서 절단하고, 새로운 2 mL 둥근 바닥 에펜도르프 튜브(카달로그 번호 22 36 335-2)에 삽입하였다.
- [0379] 필터를 오직 절단된 필터 용기의 가장자리와 접촉하며 시료 또는 필터 물질과는 접촉하지 않는 적절한 도구를 사용하여 튜브의 바닥으로 압축시켰다. 톨루엔 중의 공지되어 있는 양의 C15:0 TAG(상기)를 첨가하고, 새로 제조한 메탄올 용액 중의 1% 소듐 메톡시드 500 μl 를 첨가하였다. 시료 펠렛을 적절한 도구를 사용하여 철저히 부수고, 튜브를 닫고, 30분 동안 50°C 히트 블록(heat block)(VWR 카달로그 번호 12621-088)에 배치하였다. 이어서, 튜브가 적어도 5분 동안 냉각되게 하였다. 이어서, 헥산 400 μl 및 수용액 중의 1 M NaCl 500 μl 를 첨가하고, 튜브를 6초 동안 2회 볼텍싱시키고, 1분 동안 원심분리하였다. 대략 150 μl 의 상(유기)층을 인서트(insert)가 있는 GC 바이얼에 두고, GC로 분석하였다.
- [0380] GC 분석을 통해 기록된 FAME 피크를 공지되어 있는 지방산의 피크와 비교시 그들의 정제 시간에 의해 확인하고, FAME 피크 영역을 공지되어 있는 양의 내부 표준물질(C15:0 TAG)의 피크 영역과 비교함으로써 정량화하였다. 따라서, 임의의 지방산 FAME의 근사량(μg)["FAME μg "]은 식: (특정 지방산에 대한 FAME 피크의 영역/표준물질 FAME 피크의 영역) * (표준물질 C15:0 TAG의 μg)에 따라 계산하고, 임의의 지방산의 양(μg)["FA μg "]은 식: (특정 지방산에 대한 FAME 피크의 영역/표준물질 FAME 피크의 영역) * (표준물질 C15:0 TAG의 μg) * 0.9503에 따라 계산하는데, 이는 1 μg 의 C15:0 TAG가 0.9503 μg 의 지방산과 같기 때문이다. 0.9503 전환 인자가 0.95 내지 0.96 범위인 대부분의 지방산에 대하여 결정된 값의 근사치임을 주목해야 한다.
- [0381] TFA의 wt %로서 각각의 개별 지방산의 양을 요약하는 지질 프로파일을 개별 FAME 피크 영역을 모든 FAME 피크 영역의 합으로 나누고 이에 100을 곱함으로써 결정하였다.
- [0382] 플라스크 검정법에 의한 야로위아 리폴리티카에서의 전체 지질 함량 및 조성의 분석
- [0383] 야로위아 리폴리티카의 특정 균주에서의 전체 지질 함량 및 조성의 상세한 분석을 위하여, 플라스크 검정법을 하기와 같이 행하였다. 구체적으로, 하나의 루프의 새로 스트리킹한 세포를 3 mL FM 배지에 접종하고, 250 rpm

및 30℃에서 하룻밤 성장시켰다. OD_{600nm}를 측정하고, 세포의 분취물을 125 mL 플라스크에서 25 mL FM 배지 중에 최종 0.3의 OD_{600nm}로 첨가하였다. 250 rpm 및 30℃의 웨이킹 인큐베이터에서 2일 후에, 6 mL의 배양물을 원심분리에 의해 수집하고, 125 mL 플라스크에서 25 mL HGM에 재현탁화시켰다. 250 rpm 및 30℃의 웨이킹 인큐베이터에서 5일 후에, 1 mL 분취물을 지방산 분석(상기)을 위해 사용하고, 건조 세포 중량["DCW"] 결정을 위하여 10 mL을 건조시켰다.

[0384] DCW 결정을 위하여, 10 mL 배양물을 베크만(Beckman) GS-6R 원심분리기 내의 베크만 GH-3.8 로터에서 4000 rpm에서 5분 동안 원심분리함으로써 수집하였다. 펠렛을 25 mL의 물에 재현탁화시키고, 상술된 바와 같이 다시 수집하였다. 세정된 펠렛을 20 mL의 물에 재현탁화시키고, 사전-칭량된 알루미늄 팬으로 옮겼다. 세포 현탁액을 80℃에서 진공 오븐에서 하룻밤 건조시켰다. 세포의 중량을 결정하였다.

[0385] 세포의 전체 지질 함량["TFA % DCW"]을 계산하고, TFA의 중량 백분율["% TFA"]로서의 각각의 지방산의 농도 및 건조 세포 중량의 백분율로서의 EPA 함량["EPA % DCW"]을 표로 만든 데이터와 함께 고려하였다.

[0386] 실시예 1

[0387] 적어도 약 38.3 TFA % DCW와 함께 적어도 약 58.7 EPA % TFA를 생성하는 야로위아 리폴리티카 균주 Z1978의 생성

[0388] 본 실시예는 델타-9 연장효소/델타-8 불포화화효소 경로의 발현을 통하여 38.3 TFA % DCW와 함께 약 58.7 EPA % TFA를 생성할 수 있는 야로위아 리폴리티카 ATCC #20362 유래의 균주 Z1978의 구축을 기재하고 있다.

[0389] 야로위아 리폴리티카 균주 Y9502의 유전형

[0390] 균주 Y9502의 생성은 미국 특허 출원 공개 제2010-0317072-A1호에 기재되어 있다. 야로위아 리폴리티카 ATCC #20362로부터 유래된 균주 Y9502는 델타-9 연장효소/델타-8 불포화화효소 경로의 발현을 통하여 전체 지질과 비교하여 약 57.0%의 EPA를 생성할 수 있었다(도 3a).

[0391] 야생형 야로위아 리폴리티카 ATCC #20362에 대한 균주 Y9502의 최종 유전형은 Ura⁺, Pex3⁻, unknown 1⁻, unknown 2⁻, unknown 3⁻, unknown 4⁻, unknown 5⁻, unknown 6⁻, unknown 7⁻, unknown 8⁻, unknown9⁻, unknown 10⁻, YAT1::ME3S::Pex16, GPD::ME3S::Pex20, YAT1::ME3S::Lip1, FBAINm::EgD9eS::Lip2, EXP1::EgD9eS::Lip1, GPAT::EgD9e::Lip2, YAT1::EgD9eS::Lip2, FBAINm::EgD8M::Pex20, EXP1::EgD8M::Pex16, FBAIN::EgD8M::Lip1, GPD::Ead8S::Pex16 (2 카피), YAT1::E389D9eS/EgD8M::Lip1, YAT1::EgD9eS/EgD8M::Aco, FBAINm::Ead9eS/Ead8S::Lip2, GPD::Fmd12::Pex20, YAT1::Fmd12::Oct, EXP1::Fmd12S::Aco, GPDIN::Fmd12::Pex16, EXP1::EgD5M::Pex16, FBAIN::EgD5SM::Pex20, GPDIN::EgD5SM::Aco, GPM::EgD5SM::Oct, EXP1::EgD5SM::Lip1, YAT1::Ead5SM::Oct, FBAINm::Pad17::Aco, EXP1::Pad17::Pex16, YAT1::Pad17S::Lip1, YAT1::YICPT1::Aco, YAT1::MCS::Lip1, FBA::MCS::Lip1, YAT1::MaLPAAT1S::Pex16이었다. 약어는 하기와 같다: Fmd12는 푸사리움 모닐리포르메(*Fusarium moniliforme*) 델타-12 불포화화효소 유전자이며[미국 특허 제 7,504,259호]; Fmd12S는 푸사리움 모닐리포르메 유래의 코돈-최적화된 델타-12 불포화화효소 유전자이고[미국 특허 제 7,504,259호]; ME3S는 모르티에텔라 알피나 유래의 코돈-최적화된 C_{16/18} 연장효소 유전자이며[미국 특허 제 7,470,532호]; EgD9e는 유글레나 그라실리스 델타-9 연장효소 유전자이고[미국 특허 제 7,645,604호]; EgD9eS는 유글레나 그라실리스 유래의 코돈-최적화된 델타-9 연장효소 유전자이며[미국 특허 제 7,645,604호]; EgD8M은 유글레나 그라실리스 유래의[미국 특허 제 7,256,033호] 합성 돌연변이 델타-8 불포화화효소 유전자[미국 특허 제 7,709,239호]이고; Ead8S는 유글레나 아나바에나 유래의 코돈-최적화된 델타-8 불포화화효소 유전자이며[미국 특허 제 7,790,156호]; E389D9eS/EgD8M은 유트렙티엘라 종 CCMP389 유래의 코돈-최적화된 델타-9 연장효소 유전자("E389D9eS")(미국 특허 제 7,645,604호)를 델타-8 불포화화효소 "EgD8M"(상기)[미국 특허 출원 공개 제2008-0254191-A1호]에 연결시킴으로써 생성되는 DGLA 신타제이고; EgD9eS/EgD8M은 델타-9 연장효소 "EgD9eS"(상기)를 델타-8 불포화화효소 "EgD8M"(상기)[미국 특허 출원 공개 제2008-0254191-A1호]에 연결시킴으로써 생성되는 DGLA 신타제이며; Ead9eS/EgD8M은 유글레나 아나바에나 유래의 코돈-최적화된 델타-9 연장효소 유전자("Ead9eS")[미국 특허 제 7,794,701호]를 델타-8 불포화화효소 "EgD8M"(상기)[미국 특허 출원 공개 제2008-0254191-A1호]에 연결시킴으로써 생성되는 DGLA 신타제이고; EgD5M 및 EgD5SM은 유글레나 그라실리스 유래의[미국 특허 제 7,678,560호] 돌연변이 HPG(서열 번호 427) 모티프를 포함하는 합성 돌연변이 델타-5 불포화화효소 유전자[미국 특허 출원 공개 제2010-0075386-A1호]이며; Ead5SM은 유글레나 아나바에나 유래의[미국 특허 제 7,943,365호] 돌연변이 HaGG(서열 번호 428) 모티프를 포함하는 합성 돌연변이 델타-5 불포화화효소 유전자[미국 특허 출원 공개 제2010-0075386-A1호]이고; Pad17은 피티움 아파니데르마툼(*Pythium aphanidermatum*) 델타-

17 불포화화효소 유전자이며[미국 특허 제7,556,949호]; PaD17S는 피티움 아파니테르마툼 유래의 코돈-최적화된 델타-17 불포화화효소 유전자이고[미국 특허 제7,556,949호]; Y1CPT1은 야로위아 리폴리티카 다이아실글리세롤 콜린포스포트랜스퍼라제 유전자이며[미국 특허 제7,932,077호]; MCS는 리조비움 레구미노사룸 생물변이형 비시에 3841 유래의 코돈-최적화된 말로닐-CoA 신테타제 유전자이며[미국 특허 출원 공개 제2010-0159558-A1호], MaLPAAT1S는 모르티에렐라 알피나 유래의 코돈-최적화된 라이소포스포티드산 아실트랜스퍼라제 유전자이다[미국 특허 제7,879,591호].

[0392] 균주 Y9502 내의 전체 지질 함량 및 조성의 상세한 분석을 위하여, 세포를 총 7일 동안 2단계로 성장시키는 플라스크 검정법을 행하였다. 분석에 기초하여, 균주 Y9502는 3.8 g/L DCW, 37.1 TFA % DCW, 21.3 EPA % DCW를 생성하였으며, 지질 프로파일은 하기와 같았고, 여기서, 각 지방산의 농도는 TFA의 중량 백분율["% TFA"]로서의 것이다: 16:0(팔미테이트)-2.5, 16:1(팔미톨레산)-0.5, 18:0(스테아르산)-2.9, 18:1(올레산)-5.0, 18:2(LA)-12.7, ALA-0.9, EDA-3.5, DGLA-3.3, ARA-0.8, ETrA-0.7, ETA-2.4, EPA-57.0, 기타-7.5.

[0393] 균주 Y9502로부터의 야로위아 리폴리티카 균주 Z1978의 생성

[0394] 균주 Y9502로부터의 균주 Z1978의 발생은 도 3b에 나타나 있으며, 본 명세서에 참고로 포함되는 미국 가출원 제 61/377248호 및 제61/428,277호에 기재되어 있다.

[0395] 구체적으로, 균주 Y9502에서 *Ura3* 유전자를 파괴시키기 위하여, *SalI/PacI*-분해된 구축물 pZKUM(도 4a; 서열 번호 82; 본 명세서에 참고로 포함되는 미국 특허 출원 공개 제2009-0093543-A1호의 표 15에 기재)을 사용하여 일반 방법에 따라 *Ura3* 돌연변이 유전자를 균주 Y9502의 *Ura3* 유전자 내로 통합시켰다. 총 27개의 형질전환체(8개의 형질전환체를 포함하는 제1 그룹, 8개의 형질전환체를 포함하는 제2 그룹 및 11개의 형질전환체를 포함하는 제3 그룹으로부터 선택되는)를 최소 배지 + 5-플루오로오로토산["MM + 5-FOA"] 선택 플레이트에서 성장시키고, 30°C에서 2 내지 5일 동안 유지시켰다. 추가의 실험에 의해, 오직 제3 형질전환체 그룹만이 실제 *Ura*- 표현형을 갖는 것으로 결정되었다.

[0396] *Ura*- 세포를 MM + 5-FOA 플레이트로부터 긁어내고, 일반 방법에 따라 지방산 분석으로 처리하였다. 이러한 방식으로, GC 분석에 의해, 각각 그룹 3의 MM + 5-FOA 플레이트 상에서 성장시킨 pZKUM-형질전환체 #1, #3, #6, #7, #8, #10 및 #11에서 TFA의 28.5%, 28.5%, 27.4%, 28.6%, 29.2%, 30.3% 및 29.6%의 EPA가 있는 것으로 나타났다. 이들 7개의 균주를 각각 균주 Y9502U12, Y9502U14, Y9502U17, Y9502U18, Y9502U19, Y9502U21 및 Y9502U22(집합하여, Y9502U)로 명명하였다.

[0397] 이어서, 하나의 델타-9 불포화화효소 유전자, 하나의 콜린-포스페이트 시티딜릴-트랜스퍼라제 유전자 및 하나의 델타-9 연장효소 돌연변이 유전자가 균주 Y9502U의 야로위아 YALIOF32131p 유전자좌(진뱅크 수탁 번호 XM_506121) 내로 통합되도록 구축물 pZKL3-9DP9N(도 4b; 서열 번호 83)을 생성하였다. 델타-9 불포화화효소 돌연변이 유전자는 EgD9eS[서열 번호 3]에 대하여 L35G 돌연변이를 포함하였다(본 명세서에 참고로 포함되는 미국 가출원 제61/377248호[대리인 문서 번호 CL4783USPRV, 출원일: 2010년 8월 26일]에 기재된 바와 같음; 또한 실시예 10A 내지 10F도 참고한다). 따라서, pZKL3-9DP9N 플라스미드에는 하기의 성분이 포함되어 있다:

표 8

플라스미드 pZKL3-9DP9N(서열 번호 83)의 설명

서열 번호 83 내의 RE 부위 및 뉴클레오타이드	단편 및 키메라 유전자 성분의 설명
<i>AscI/BsiWI</i> (887-4)	YALI0F32131p 유전자좌의 884 bp 5' 부분(진뱅크 수탁 번호 XM_506121, 도면에서 "Lip3-5"로 표시)
<i>PacI/SphI</i> (4396-3596)	YALI0F32131p 유전자좌의 801 bp 3' 부분(진뱅크 수탁 번호 XM_506121, 도면에서 "Lip3-3"으로 표시)
<i>SwaI/BsiWI</i> (11716 - 1)	하기를 포함하는 YAT1::EgD9eS-L35G::Pex20: • YAT1: 야로위아 리폴리티카 YAT1 프로모터(도면에서 "YAT"로 표시; 미국 특허 출원 공개 제 2010-0068789-A1 호); • EgD9eS-L35G: 유글레나 그라실리스 유래의 델타-9 연장효소 유전자("EgD9eS"; 미국 특허 제 7,645,604 호)의 합성 돌연변이체(서열 번호 43; 미국 가출원 제 61/377248 호)(도면에 "EgD9eS-24"로 표시); • Pex20: 야로위아 <i>Pex20</i> 유전자로부터의 <i>Pex20</i> 종결자 서열(진뱅크 수탁 번호 AF054613)
<i>PmeI/SwaI</i> (8759-11716)	하기를 포함하는 GPDIN::YID9::Lip1: • GPDIN: 야로위아 리폴리티카 GPDIN 프로모터(미국 특허 제 7,459,546 호; 도면에서 "GPDPro+Intron"으로 표시); • YID9: 야로위아 리폴리티카 델타-9 불포화화효소 유전자(진뱅크 수탁 번호 XM_501496; 서열 번호 80)(도면에서 "YID9D"로 표시); • Lip1: 야로위아 <i>Lip1</i> 유전자로부터의 <i>Lip1</i> 종결자 서열(진뱅크 수탁 번호 Z50020)
<i>ClalI/PmeI</i> (6501-8759)	하기를 포함하는 EXP1::YIPCT::Pex16: • EXP1: 야로위아 리폴리티카 수송 단백질(EXP1) 프로모터(도면에서 "EXP"로 표시; 국제 출원 공개 제 WO 2006/052870 호); • YIPCT: 야로위아 리폴리티카 콜린-포스페이트 시티딜릴-트랜스퍼라제["PCT"] 유전자(진뱅크 수탁 번호 XM_502978; 서열 번호 45); • Pex16: 야로위아 <i>Pex16</i> 유전자로부터의 <i>Pex16</i> 종결자 서열(진뱅크 수탁 번호 U75433)
<i>SalI/EcoRI</i> (6501-4432)	야로위아 <i>Ura3</i> 유전자(진뱅크 수탁 번호 AJ306421)

[0398]

[0399]

pZKL3-9DP9N 플라스미드를 *AscI*/*SphI*로 분해한 다음, 균주 Y9502U17의 형질전환을 위해 사용하였다. 형질전환된 세포를 최소 배지["MM"] 플레이트 상에 플레이팅하고, 30°C에서 3 내지 4일 동안 유지시켰다. 단일의 콜로니를 MM 플레이트 상으로 다시 스트리킹한 다음, 30°C에서 액체 MM으로 접종하고, 250 rpm/분으로 2일 동안 웨이킹하였다. 세포를 원심분리에 의해 수집하고, 고 글루코스 배지["HGM"]에 재현탁화시킨 다음, 250 rpm/분에서 5일 동안 웨이킹하였다. 세포를 상기 지방산 분석으로 처리하였다.

[0400]

GC 분석에 의해, pZKL3-9DP9N을 갖는 선택된 96개의 Y9502U17 균주의 대부분이 TFA의 50 내지 56%의 EPA를 생성하는 것으로 나타났다. TFA의 약 59.0%, 56.6%, 58.9%, 56.5% 및 57.6%의 EPA를 생성하였던 5개의 균주(즉, #31, #32, #35, #70 및 #80)를 각각 균주 Z1977, Z1978, Z1979, Z1980 및 Z1981로 명명하였다.

[0401]

야생형 야로위아 리폴리티카 ATCC #20362에 대한 이들 pZKL3-9DP9N 형질전환된 균주의 최종 유전형은 Ura+, Pex3-, unknown 1-, unknown 2-, unknown 3-, unknown 4-, unknown 5-, unknown 6-, unknown 7-, unknown 8-, unknown 9-, unknown 10-, unknown 11-, YAT1::ME3S::Pex16, GPD::ME3S::Pex20, YAT1::ME3S::Lip1, FBAINm::EgD9eS::Lip2, EXP1::EgD9eS::Lip1, GPAT::EgD9eS::Lip2, YAT1::EgD9eS::Lip2, YAT1::EgD9eS-L35G::Pex20, FBAINm::EgD8M::Pex20, EXP1::EgD8M::Pex16, FBAIN::EgD8M::Lip1, GPD::EaD8S::Pex16(2 카피), YAT1::E389D9eS/EgD8M::Lip1, YAT1::EgD9eS/EgD8M::Aco, FBAINm::EaD9eS/EaD8S::Lip2, GPDIN::YID9::Lip1, GPD::Fmd12::Pex20, YAT1::Fmd12::Oct, EXP1::Fmd12S::Aco, GPDIN::Fmd12::Pex16, EXP1::EgD5M::Pex16, FBAIN::EgD5SM::Pex20, GPDIN::EgD5SM::Aco, GPM::EgD5SM::Oct, EXP1::EgD5SM::Lip1, YAT1::EaD5SM::Oct, FBAINm::PaD17::Aco, EXP1::PaD17::Pex16, YAT1::PaD17S::Lip1, YAT1::Y1CPT1::Aco, YAT1::MCS::Lip1, FBA::MCS::Lip1, YAT1::MaLPAAT1S::Pex16, EXP1::YIPCT::Pex16이었다.

[0402]

균주 Z1977, Z1978, Z1979, Z1980 및 Z1981에서의 YALI0F32131p 유전자좌(진뱅크 수탁 번호 XM_50612)의 낙아웃은 pZKL3-9DP9N으로의 형질전환에 의해 생성되는 이들 임의의 EPA 균주에서 확인되지 않았다.

[0403]

균주 Z1977, Z1978, Z1979, Z1980 및 Z1981의 YPD 플레이트로부터의 세포를 성장시키고, 전체 지질 함량 및 조성에 대하여 분석하였다. 구체적으로, 플라스크 검정법을 일반 방법에 기재된 바와 같이 행하였다.

- [0404] 따라서, 바람직한 실시형태의 설명(상기)에서의 표 7에는 플라스크 검정법에 의해 결정되는 바와 같은, 균주 Z1977, Z1978, Z1979, Z1980 및 Z1981의 총 DCW, TFA % DCW, 각 지방산의 농도["% TFA"] 및 EPA % DCW가 요약되어 있다.
- [0405] 미국 가출원 제61/377248호(대리인 문서 번호 CL4783USPRV, 출원일: 2010년 8월 26일)의 출원 이후에, 균주 Z1978을 부분 계놈 시퀀싱으로 처리하였다. 본 연구에 의해, 미국 가출원 제61/428,277호(대리인 문서 번호 CL5267USPRV, 출원일: 2010년 12월 30일)에 기재된 바와 같이, 야로위아 계놈 내에 6개의 델타-5 불포화화효소 유전자가 통합되는 대신에, 엔지니어링된 균주는 실제로 오직 4개만을 갖는 것이 결정되었다.
- [0406] 더욱 구체적으로, 2개의 별개의 플라스미드 단편(또는 그의 부분)은 하기에 추가로 기재되는 바와 같이, 균주 Z1978에서 검출되지 않았다.
- [0407] (1) 구축물 pZKL2-5mB89C(미국 특허 출원 공개 제2010-0317072-A1호, 상기 문헌의 서열 번호 131 참조)는 하나의 델타-5 불포화화효소 유전자를 균주 Y8069U의 *Lip2* 유전자좌에 통합시켜, 이에 의해 더 높은 수준의 EPA 생성을 가능하게 하고자 하였다. 그러나, 계놈의 시퀀싱에 의해, GPDIN::EgD5SM::Aco 키메라 유전자 및 pZKL2-5mB89C 단편의 *Lip2.3N* 말단 부분을 검출할 수 없었다. DNA 재배열은 Y8154 균주의 생성 동안 GPDIN::EgD5SM::Aco 카세트의 소실을 야기할 수 있었다(도 3a).
- [0408] (2) 구축물 pZKL1-2SR9G85(미국 특허 출원 공개 제2010-0317072-A1호, 상기 문헌의 서열 번호 132 참조)는 하나의 델타-5 불포화화효소 유전자를 균주 Y8154U1의 *Lip1* 유전자좌로 통합시키려고 하였다. 그러나, 계놈 시퀀싱 또는 PCR 증폭 중 어느 것도 균주 Z1978에서 델타-5 불포화화효소 유전자를 검출할 수 없었다. DNA 재배열은 균주 Y8269의 생성 동안 GPM::EgD5SM::Oct 카세트의 소실을 야기할 수 있었다(도 3a).
- [0409] 추가로, 구축물 pZSCP-Ma83(미국 특허 출원 공개 제2010-0317072-A1호, 상기 문헌의 서열 번호 133 참조) 및 구축물 pZP2-85m98F(미국 특허 출원 공개 제2010-0317072-A1호, 상기 문헌의 서열 번호 135 참조)가 둘 모두 YAL10B21890g 유전자좌 내로 통합되는 것으로 결정되었다.
- [0410] 따라서, 야생형 야로위아 리폴리티카 ATCC #20362에 대한 균주 Z1978의 실제 유전형은 하기와 같았다: Ura⁺, Pex3⁻, unknown 1⁻, unknown 2⁻, unknown 3⁻, unknown 4⁻, YAL10E12947g⁻, unknown6⁻, YAL10B21890g⁻, unknown 8⁻, unknown 10⁻, unknown 11⁻, YAT1::ME3S::Pex16, GPD::ME3S::Pex20, YAT1::ME3S::Lip1, FBAINm::EgD9eS::Lip2, EXP1::EgD9eS::Lip1, GPAT::EgD9e::Lip2, YAT1::EgD9eS::Lip2, YAT1::EgD9eS-L35G::Pex20, FBAINm::EgD8M::Pex20, EXP1::EgD8M::Pex16, FBAIN::EgD8M::Lip1, GPD::EaD8S::Pex16 (2 카피), YAT1::E389D9eS/EgD8M::Lip1, YAT1::EgD9eS/EgD8M::Aco, FBAINm::EaD9eS/EaD8S::Lip2, GPDIN::Y1D9::Lip1, GPD::FmD12::Pex20, YAT1::FmD12::Oct, EXP1::FmD12S::Aco, GPDIN::FmD12::Pex16, EXP1::EgD5M::Pex16, FBAIN::EgD5SM::Pex20, EXP1::EgD5SM::Lip1, YAT1::EaD5SM::Oct, FBAINm::PaD17::Aco, EXP1::PaD17::Pex16, YAT1::PaD17S::Lip1, YAT1::Y1CPT1::Aco, YAT1::MCS::Lip1, FBA::MCS::Lip1, YAT1::MaLPAAT1S::Pex16, EXP1::Y1PCT::Pex16.
- [0411] 야로위아 리폴리티카 균주 Y9502 및 균주 Z1978의 비교
- [0412] 균주 Z1978에서 발현되는 이중 유전자는 하나의 델타-9 불포화화효소 유전자, 하나의 콜린-포스페이트 시티딜릴-트랜스퍼라제 유전자 및 하나의 델타-9 연장효소 돌연변이체(즉, 서열 번호 43 및 44에 기재된 EgD9eS-L35G)의 추가의 발현만이 균주 Y9502에서 발현되는 것과 상이하다. LA 및 ALA로부터 EPA로의 전체 델타-9 연장효소 전환 효율["전환%"]을 Y9502 및 Z1978 균주에 대하여 하기의 식에 따라 표 9에서 계산하였다: ([생성물]/[기질 + 생성물])*100, 여기서, 생성물은 EDA % TFA, ETrA % TFA, DGLA % TFA, ETA % TFA, ARA % TFA 및 EPA % TFA의 합이었으며, 기질은 LA % TFA, ALA % TFA, EDA % TFA, ETrA % TFA, DGLA % TFA, ETA % TFA, ARA % TFA 및 EPA % TFA의 합이었다.

표 9

형질전환체 야로위아 리폴리티카 균주 Y9502 및 Z1978에서의 전체 지질 함량 및 조성, 및 델타-9 연장효소 활성의 비교

		야로위아 리폴리티카 균주 Y9502	야로위아 리폴리티카 균주 Z1978
DCW(g/L)		3.8	3.9
TFA % DCW		37.1	38.3
% TFA	16:0	2.5	2.4
	16:1	0.5	0.4
	18:0	2.9	2.4
	18:1	5	4.8
	18:2	12.7	11.1
	ALA	0.9	0.7
	EDA	3.5	3.2
	DGLA	3.3	3.3
	ARA	0.8	0.8
	ETrA	0.7	0.6
	ETA	2.4	2.1
	EPA	57	58.7
기타		7.5	9.5
EPA % DCW		21.3	22.5
LA로부터 EPA로의 전체 전환%		83.3	85.3

[0413]

[0414] 상술된 바와 같이, 전체 델타-9 연장효소 전환 효율이 균주 Y9502에서 83.3%인 것으로 결정된 한편, 효율은 균주 Z1978에서 향상되었다(즉, 85.3%).

[0415] 실시예 2

[0416] 적어도 약 57.1 TFA % DCW와 함께, 적어도 약 45.2 EPA % TFA를 생성하는 야로위아 리폴리티카 균주 L258의 생성

[0417] 본 실시예는 델타-9 연장효소/델타-8 불포화화효소 경로의 발현을 통하여 57.1 TFA % DCW와 함께 약 45.2 EPA % TFA를 생성할 수 있는 야로위아 리폴리티카 균주 Z1978(실시예 1) 유래의 균주 L258의 구축을 기재하고 있다.

[0418] 균주 L258(도 3b)의 구축은 중간 균주 Z1978U, L250 및 L250U의 구축을 필요로 하였다.

[0419] 균주 Z1978U(*Ura3*-)의 생성

[0420] *Ura3* 유전자를 파괴하기 위하여, 구축물 pZKUM(도 4a; 서열 번호 82; 미국 특허 출원 공개 제2009-0093543-A1호의 표 15에 기재)을 사용하여 *Ura3* 돌연변이 유전자를 균주 Y9502의 pZKUM 형질전환(실시예 1)에 대해 기재된 것과 유사한 방식으로, 균주 Z1978의 *Ura3* 유전자에 통합시켰다. 총 16개 형질전환체(8개의 형질전환체를 포함하는 제1 "B" 그룹 및 8개의 형질전환체를 포함하는 제2 "C" 그룹으로부터 선택)를 성장시키고, *Ura*- 표현형을 갖는 것으로 확인되었다.

[0421] GC 분석에 의해, MM + 5-FOA 플레이트 상에서 성장시킨, B 그룹 pZKUM-형질전환 균주 #1, #2, #3 및 #4 각각에서 TFA의 30.8%, 31%, 30.9% 및 31.3%의 EPA의 존재가 나타났다. 이들 4개의 균주를 각각 균주 Z1978BU1, Z1978BU2, Z1978BU3 및 Z1978BU4로 명명하였다.

[0422] GC 분석에 의해, MM + 5-FOA 플레이트 상에서 성장시킨, C 그룹 pZKUM-형질전환 균주 #1, #2, #5 및 #6 각각에서 TFA의 34.4%, 31.9%, 31.2% 및 31%의 EPA의 존재가 나타났다. 이들 4개의 균주를 각각 균주 Z1978CU1, Z1978CU2, Z1978CU3 및 Z1978CU4로 명명하였다.

[0423] 균주 Z1978BU1, Z1978BU2, Z1978BU3, Z1978BU4, Z1978CU1, Z1978CU2, Z1978CU3 및 Z1978CU4 균주를 집합하여 균주 Z1978U로 명명하였다.

[0424] 야로위아 리폴리티카 균주 L250의 생성

[0425] 하나의 라이소포스파티드산 아실트랜스퍼라제 유전자["LPAAT"] 및 인지질:다이아실글리세롤 아실트랜스퍼라제 유전자["PDAT"]가 균주 Z1978U의 게놈 내로 통합되도록 구축물 pY187(도 5a; 서열 번호 84)을 생성하였다. pY187 플라스미드는 하기의 성분을 함유하였다:

표 10

플라스미드 pY187(서열 번호 84)의 성분

서열 번호 84 내의 RE 부위 및 뉴클레오타이드	단편 및 키메라 유전자 성분의 설명
<i>Cla</i> I/ <i>Swa</i> I 6929-333	하기를 포함하는 YAT1::YIPDAT::Lip1: ● YAT1: 야로위아 리폴리티카 YAT1 프로모터(미국 특허 출원 공개 제 2010-0068789-A1 호); ● YIPDAT: 야로위아 리폴리티카 인지질:다이아실글리세롤 아실트랜스퍼라제 유전자(서열 번호 28; 미국 특허 제 7,901,928 호; 진뱅크 수탁 번호 XM_504038); ● Lip1: 야로위아 <i>Lip1</i> 유전자로부터의 종결자 서열(진뱅크 수탁 번호 Z50020)
645-1525	● ColE1 플라스미드 복제 원점
1595-2455	● 앰피실린-내성 유전자
<i>Sph</i> I/ <i>Apa</i> I 5337-3247	하기를 포함하는 FBAINm::YILPAAT1::Lip1(상보성): ● FBAINm: 야로위아 리폴리티카 FBAIN 프로모터(미국 특허 제 7,202,356 호); ● YILPAAT1: 야로위아 리폴리티카 라이소포스파티드산 아실트랜스퍼라제 유전자(서열 번호 21; 미국 특허 제 7,189,559 호; 진뱅크 수탁 번호 XP_504127); ● Lip1: 야로위아 <i>Lip1</i> 유전자로부터의 종결자 서열(진뱅크 수탁 번호 Z50020)
<i>Bsi</i> WI/ <i>Pac</i> I 6905-5407	야로위아 리폴리티카 <i>Ura3</i> 유전자(진뱅크 수탁 번호 AJ306421)

[0426]

[0427]

pY187 플라스미드를 *Swa*I/*Apa*I로 분해하고, 6.7 kbp의 큰 단편을 아가로스 겔로부터 정제한 다음, 일반 방법에 따라 균주 Z1978CU4의 형질전환을 위해 사용하였다. 형질전환된 세포를 MM 플레이트 상에 플레이팅하고, 30℃에서 5일 동안 유지시켰다. 이어서, 단일의 콜로니(19개)를 MM 플레이트 상에 다시 스트리킹하였다. 이들 균주의 전체 지질 함량 및 지방산 조성을 일반 방법에 따라 플라스크 검정법에 의해 평가하였다.

[0428]

분석에 기초하여 균주 L250은 4.4 g/L DCW, 51.5 TFA % DCW 및 26 EPA % DCW를 생성하였다. 지질 프로파일은 하기와 같고, 여기서, 각 지방산의 농도는 TFA의 중량 백분율["% TFA"]로서의 것이다: 16:0(팔미테이트)-2.0, 16:1(팔미톨레산)-0.7, 18:0(스테아르산)-2.8, 18:1(올레산)-6.1, 18:2(LA)-16.7, ALA-0.9, EDA-3.3, DGLA-4.9, ARA-0.7, ETrA-0.6, ETA-3.2, EPA-50.4 및 기타-7.4(또한 상기 바람직한 실시형태의 설명에서 표 7 참조).

[0429]

균주 L250U(*Ura3*-)의 생성

[0430]

Ura3 유전자를 파괴하기 위하여, 구축물 pZKUM(도 4a; 서열 번호 82; 미국 특허 출원 공개 제2009-0093543-A1호의 표 15에 기재)을 사용하여 *Ura3* 돌연변이 유전자를 균주 Y9502의 pZKUM 형질전환(실시예 1)에 대해 기재된 것과 유사한 방식으로, 균주 L250의 *Ura3* 유전자에 통합시켰다. 총 12개의 5-FOA 내성 콜로니를 성장시키고, *Ura*- 표현형을 갖는 것으로 동정되었다. 균주 #2 및 균주 #3을 각각 L250U2 및 L250U3으로 명명하였다(집합하여, 균주 L250U).

[0431]

야로위아 리폴리티카 균주 L258의 생성

[0432]

플라스미드 pY187(표 10; 서열 번호 84)을 사용하여 추가의 카피의 YILPAAT 유전자(서열 번호 21) 및 YIPDAT 유전자(서열 번호 28)를 균주 L250U의 야로위아 계통 내로 통합시켰다. 플라스미드 pY187의 정제된 6.7 kbp의 큰 단편을 일반 방법에 따라 균주 L250U2의 형질전환을 위해 사용하였다. 형질전환된 세포를 MM 플레이트 상에 플레이팅하고, 30℃에서 5일 동안 유지시켰다. 이어서, 단일의 콜로니를 MM 플레이트 상에 다시 스트리킹하였다. 세포를 일반 방법에 따라 플라스크 검정법에 의해 전체 지질 함량 및 조성을 평가하였다.

[0433]

분석에 기초하여 균주 L258은 5.0 g/L DCW, 57.1 TFA % DCW 및 25.8 EPA % DCW를 생성하였다. 지질 프로파일은 하기와 같았고, 여기서, 각 지방산의 농도는 TFA의 중량 백분율["% TFA"]로서의 것이다: 16:0(팔미테이트)-2.3, 16:1(팔미톨레산)-0.9, 18:0(스테아르산)-3.4, 18:1(올레산)-7.8, 18:2(LA)-18.7, ALA-0.9, EDA-4.0, DGLA-5.3, ARA-0.8, ETrA-0.6, ETA-3.2, EPA-45.2 및 기타-6.6(또한 상기 바람직한 실시형태의 설명에서 표 7 참조).

[0434]

야생형 야로위아 리폴리티카 ATCC #20362에 대한 균주 L258의 최종 유전형은 하기와 같다: *Ura*+, *Pex3*-, unknown 1-, unknown 2-, unknown 3-, unknown 4-, YALI0E12947g-, unknown6-, YALI0B21890g-, unknown 8-,

unknown 10-, unknown 11-, unknown 12-, unknown 13-, YAT1::ME3S::Pex16, GPD::ME3S::Pex20, YAT1::ME3S::Lip1, FBAINm::EgD9eS::Lip2, EXP1::EgD9eS::Lip1, GPAT::EgD9eS::Lip2, YAT1::EgD9eS::Lip2, YAT1::EgD9eS-L35G::Pex20, FBAINm::EgD8M::Pex20, EXP1::EgD8M::Pex16, FBAIN::EgD8M::Lip1, GPD::Ead8S::Pex16 (2 카피), YAT1::E389D9eS/EgD8M::Lip1, YAT1::EgD9eS/EgD8M::Aco, FBAINm::Ead9eS/Ead8S::Lip2, GPDIN::Y1D9::Lip1, GPD::FmD12::Pex20, YAT1::FmD12::Oct, EXP1::FmD12S::Aco, GPDIN::FmD12::Pex16, EXP1::EgD5M::Pex16, FBAIN::EgD5SM::Pex20, EXP1::EgD5SM::Lip1, YAT1::Ead5SM::Oct, FBAINm::PaD17::Aco, EXP1::PaD17::Pex16, YAT1::PaD17S::Lip1, YAT1::Y1CPT1::Aco, YAT1::MCS::Lip1, FBA::MCS::Lip1, EXP1::Y1PCT::Pex16, YAT1::MaLPAAT1S::Pex16, FBAINm::Y1LPAAT1::Lip1 (2 카피), YAT1::Y1PDAT::Lip1 (2 카피).

[0435] 실시예 3

[0436] 적어도 약 53 TFA % DCW와 함께 적어도 약 47 EPA % TFA를 생성하는 야로위아 리폴리티카 균주 Z5565, Z5567, Z5575 및 Z5576의 생성

[0437] 본 실시예는 델타-9 연장효소/델타-8 불포화효소 경로의 발현을 통하여, 53 초과 TFA % DCW와 함께, 약 47 EPA % TFA를 생성할 수 있는 야로위아 리폴리티카 균주 L258(실시예 2) 유래의 균주 Z5565, Z5567, Z5575 및 Z5576의 구축을 기재하고 있다.

[0438] 균주 Z5565, Z5567, Z5575 및 Z5576(도 3b)의 발생은 중간 균주 L258U의 구축을 필요로 하였다.

[0439] 균주 L258U(*Ura3*-)의 생성

[0440] *Ura3* 유전자를 파괴하기 위하여, 구축물 pZKUM(도 4a; 서열 번호 82; 미국 특허 출원 공개 제2009-0093543-A1호의 표 15에 기재)을 사용하여 *Ura3* 돌연변이 유전자를 균주 Y9502의 pZKUM 형질전환(실시예 1)에 대해 기재된 것과 유사한 방식으로, 균주 L258의 *Ura3* 유전자에 통합시켰다. 총 20개의 형질전환체를 성장시키고, *Ura*-표현형을 갖는 것으로 동정되었다.

[0441] GC 분석에 의해, MM + 5-FOA 플레이트 상에서 성장시킨 pZKUM-형질전환 균주 #7 및 #9에서 각각 37.6% 및 37.2%의 EPA의 존재가 나타났다. 이들 2개의 균주를 각각 L258U5 및 L258U6으로 명명하고, 집합하여 균주 L258U로 명명하였다.

[0442] 야로위아 리폴리티카 균주 Z5565, Z5567, Z5575 및 Z5576의 생성

[0443] 구축물 pZK16-ML8N(도 5b; 서열 번호 85)은 미국 특허 출원 공개 제2010-0317072-A1호, 상기 문헌의 표 15에 기재되어 있다. 키메라 YAT1::EgD8M::Pex20 유전자 내의 하나의 델타-8 불포화효소, 키메라 FBA::MCS::Lip1 유전자 내의 하나의 말로날-CoA 신데타제 및 키메라 YAT1::MaLPAAT1S::Pex16 유전자 내의 하나의 라이소포스파티드산 아실트랜스퍼라제가 야로위아 YALIOB14795p 유전자좌(진뱅크 수탁 번호 XM_500900) 내로 통합되도록 이를 생성하였다.

[0444] pZK16-ML8N 플라스미드를 *AscI*/*SphI*으로 분해한 다음, 일반 방법에 따라 개별적으로 균주 L258U5 및 L258U6의 형질전환을 위해 사용하였다. 형질전환된 세포를 MM 플레이트 상에 플레이팅하고, 30℃에서 5 내지 6일 동안 유지시켰다. 단일의 콜로니를 MM 플레이트 상으로 다시 스트리킹한 다음, 30℃에서 액체 MM으로 접종하고, 250 rpm/분으로 2일 동안 셰이킹하였다. 세포를 원심분리에 의해 수집하고, HGM에 재현탁화시킨 다음, 250 rpm/분에서 5일 동안 셰이킹하였다. 세포를 일반 방법에 따라 지방산 분석으로 처리하였다.

[0445] GC 분석에 의해, pZK16-ML8N을 갖는 선택된 48개의 L258U5 균주 중 7개가 TFA의 48% 초과 EPA를 생성하는 것으로 나타났다. TFA의 약 49.7% 및 50.9%의 EPA를 생성하였던 2개의 균주(즉, #3 및 #36)를 각각 Z5565 및 Z5567로 명명하였다.

[0446] GC 분석에 의해, pZK16-ML8N을 갖는 선택된 48개의 L258U6 균주 대부분이 TFA의 49% 초과 EPA를 생성하는 것으로 나타났다. TFA의 약 53.7% 및 50.2%의 EPA를 생성하였던 2개의 균주(즉, #2 및 #5)를 각각 Z5575 및 Z5576으로 명명하였다.

[0447] 야생형 야로위아 리폴리티카 ATCC #20362에 대한 이들 pZK16-ML8N 형질전환 균주의 최종 유전형은 하기와 같았다: *Ura*+, *Pex3*-, unknown 1-, unknown 2-, unknown 3-, unknown 4-, YALIOE12947g-, unknown6-, YALIOB21890g-, unknown 8-, unknown 10-, unknown 11-, unknown 12-, unknown 13-, unknown 14-, YAT1::ME3S::Pex16, GPD::ME3S::Pex20, YAT1::ME3S::Lip1, FBAINm::EgD9eS::Lip2, EXP1::EgD9eS::Lip1,

GPAT::EgD9e::Lip2, YAT1::EgD9eS::Lip2, YAT1::EgD9eS-L35G::Pex20, FBAINm::EgD8M::Pex20, EXP1::EgD8M::Pex16, FBAIN::EgD8M::Lip1, YAT1::EgD8M::Pex20, GPD::EaD8S::Pex16 (2 카피), YAT1::E389D9eS/EgD8M::Lip1, YAT1::EgD9eS/EgD8M::Aco, FBAINm::EaD9eS/EaD8S::Lip2, GPDIN::Y1D9::Lip1, GPD::FmD12::Pex20, YAT1::FmD12::Oct, EXP1::FmD12S::Aco, GPDIN::FmD12::Pex16, EXP1::EgD5M::Pex16, FBAIN::EgD5SM::Pex20, EXP1::EgD5SM::Lip1, YAT1::EaD5SM::Oct, FBAINm::PaD17::Aco, EXP1::PaD17::Pex16, YAT1::PaD17S::Lip1, YAT1::Y1CPT1::Aco, YAT1::MCS::Lip1, FBA::MCS::Lip1 (2 카피), EXP1::Y1PCT::Pex16, YAT1::MaLPAAT1S::Pex16 (2 카피), FBAINm::Y1LPAAT1::Lip1 (2 카피), YAT1::Y1PDAT::Lip1 (2 카피).

- [0448] 균주 Z5565, Z5567, Z5575 및 Z5576에서의 YALI0B14795p 유전자좌(진뱅크 수탁 번호 XM_500900)의 낙아웃은 pZK16-ML8N으로의 형질전환에 의해 생성되는 이들 EPA 균주 중 임의의 것에서 확인되지 않았다.
- [0449] 플라스크 검정법에 의한 전체 지질 함량 및 조성의 분석
- [0450] 균주 Z5565, Z5567, Z5575 및 Z5576의 YPD 플레이트로부터의 세포를 성장시키고, 일반 방법에 따라, 전체 지질 함량 및 조성에 대하여 분석하였다.
- [0451] 바람직한 실시형태의 설명에서의 표 7(상기)에는 균주 Z5565, Z5567, Z5575 및 Z5576의 총 DCW, TFA % DCW, TFA의 중량 백분율 ["% TFA"]로서의 각 지방산의 농도 및 EPA % DCW가 요약되어 있다. 평균 DCW는 4.8 g/L이었으며, 평균 TFA % DCW는 55.4이었으며, 평균 EPA % TFA는 48.3이었으며, 평균 EPA % DCW는 26.75이었다.
- [0452] 실시예 4
- [0453] 적어도 약 51 TFA % DCW와 함께 적어도 약 49 EPA % TFA를 생성하는 야로위아 리폴리티카 균주 Z5620, Z5623 및 Z5625의 생성
- [0454] 본 실시예에는 델타-9 연장효소/델타-8 불포화화효소 경로의 발현을 통하여 51 초과 TFA % DCW와 함께, 약 49 EPA % TFA를 생성할 수 있는 야로위아 리폴리티카 균주 L258U(실시예 3) 유래의 균주 Z5620, Z5623 및 Z5625의 구축이 기재되어 있다.
- [0455] 하나의 델타-8 불포화화효소 유전자, 하나의 말로닐-CoA 신테타제 유전자 및 하나의 라이소포스파티드산 아실트랜스퍼라제 유전자가 균주 L258U의 야로위아 YALI0B14795p 유전자좌(진뱅크 수탁 번호 XM_500900) 내로 통합되도록 구축물 pZK16-MyL8N을 생성하였다. 더욱 구체적으로, 구축물 pZK16-MyL8N(도 6a; 서열 번호 86)은 야로위아 리폴리티카 LPAAT 유전자("Y1LPAAT"; 서열 번호 21) 및 Lip1 종결자가 야로위아 리폴리티카에서의 발현을 위해 코돈-최적화된 모르티에렐라 알피나 LPAAT 유전자("MaLPAAT1S"; 서열 번호 19) 및 Pex16 종결자를 대체하는 것을 제외하고는 pZK16-ML8N(도 5b; 서열 번호 85; 실시예 3)과 동일하였다. pZK16-MyL8N 플라스미드에는 하기의 성분이 포함되어 있다:

표 11

플라스미드 pZK16-MyL8N(서열 번호 86)의 설명

서열 번호 86 내의 RE 부위 및 뉴클레오타이드	단편 및 키메라 유전자 성분의 설명
<i>AscI/BstWI</i> (6634-4730)	YALI0B14795p 유전자좌의 1904 bp 5' 부분(진뱅크 수탁 번호 XM_500900, 도면에서 "Y8716-5"로 표시)
<i>PacI/SphI</i> (11143-9342)	YALI0B14795p 유전자좌의 1801 bp 3' 부분(진뱅크 수탁 번호 XM_500900, 도면에서 "Y8716-3"으로 표시)
<i>SwaI/BstWI</i> (2387 - 4730)	하기를 포함하는 YAT1::EgD8M::Pex20: • YAT1: 야로위아 리폴리티카 YAT1 프로모터(미국 특허 출원 공개 제 2010-0068789-A1 호); • EgD8M: 유글레나 그라실리스 유래의("EgD8S", 미국 특허 제 7,256,033 호) 합성 돌연변이 델타-8 불포화화효소(서열 번호 87; 미국 특허 제 7,709,239 호); • Pex20: 야로위아 <i>Pex20</i> 유전자로부터의 Pex20 종결자 서열(진뱅크 수탁 번호 AF054613)
1-2387	하기를 포함하는 FBA::MCS::Lip1: • FBA: 야로위아 리폴리티카 FBA 프로모터(미국 특허 제 7,202,356 호); • MCS: 리조비움 레구미노사룸 생물변이형 비시에 3841 유래의 코돈-최적화된 말로닐-CoA 신테타제 유전자(서열 번호 49)(미국 특허 출원 공개 제 2010-0159558-A1 호); • Lip1: 야로위아 <i>Lip1</i> 유전자로부터의 Lip1 종결자 서열(진뱅크 수탁 번호 Z50020)
13244-1	하기를 포함하는 YAT1::YILPAAT::Lip1: • YAT1: 야로위아 리폴리티카 YAT1 프로모터(미국 특허 출원 공개 제 2010-0068789-A1 호); • YILPAAT1: 야로위아 리폴리티카 라이소포스포파티드산 아실트랜스퍼라제 유전자(서열 번호 21; 미국 특허 제 7,189,559 호; 진뱅크 수탁 번호 XP_504127) • Lip1: 야로위아 <i>Lip1</i> 유전자로부터의 Lip1 종결자 서열(진뱅크 수탁 번호 Z50020)
<i>SalI/EcoRI</i> (12794 - 11175)	야로위아 Ura3 유전자(진뱅크 수탁 번호 AJ306421)

[0456]

[0457]

pZK16-MyL8N 플라스미드를 *AscI/SphI*으로 분해한 다음, 일반 방법에 따라 균주 L258U6의 형질전환을 위해 사용하였다. 형질전환 세포를 MM 플레이트 상에 플레이팅하고, 30℃에서 5 내지 6일 동안 유지시켰다. 단일의 콜로니를 MM 플레이트 상으로 다시 스트리킹한 다음, 30℃에서 액체 MM으로 접종하고, 250 rpm/분으로 2일 동안 셰이킹하였다. 세포를 원심분리에 의해 수집하고, HGM에 재현탁화시킨 다음, 250 rpm/분에서 5일 동안 셰이킹하였다. 세포를 상기 일반 방법에 기재된 바와 같이 지방산 분석으로 처리하였다.

[0458]

GC 분석에 의해, pZK16-MyL8N으로 형질전환된 L258U6의 선택된 거의 모든 48개 균주가 TFA의 49% 초과 EPA를 생성하는 것으로 나타났다. TFA의 약 52.8%, 53.0% 및 49.9%의 EPA를 생성하였던 3개의 균주(즉, #5, #21 및 #48)를 각각 Z5620, Z5623 및 Z5625로 명명하였다.

[0459]

야생형 야로위아 리폴리티카 ATCC #20362에 대한 3개의 pZK16-MyL8N 형질전환 균주의 최종 유전형은 하기와 같았다: Ura⁺, Pex3⁻, unknown 1⁻, unknown 2⁻, unknown 3⁻, unknown 4⁻, YAL10E12947g⁻, unknown6⁻, YAL10B21890g⁻, unknown 8⁻, unknown 10⁻, unknown 11⁻, unknown 12⁻, unknown 13⁻, unknown 14⁻, YAT1::ME3S::Pex16, GPD::ME3S::Pex20, YAT1::ME3S::Lip1, FBAINm::EgD9eS::Lip2, EXP1::EgD9eS::Lip1, GPAT::EgD9e::Lip2, YAT1::EgD9eS::Lip2, YAT1::EgD9eS-L35G::Pex20, FBAINm::EgD8M::Pex20, EXP1::EgD8M::Pex16, FBAIN::EgD8M::Lip1, YAT1::EgD8M::Pex20, GPD::EaD8S::Pex16 (2 카피), YAT1::E389D9eS/EgD8M::Lip1, YAT1::EgD9eS/EgD8M::Aco, FBAINm::EaD9eS/EaD8S::Lip2, GPDIN::Y1D9::Lip1, GPD::FmD12::Pex20, YAT1::FmD12::Oct, EXP1::FmD12S::Aco, GPDIN::FmD12::Pex16, EXP1::EgD5M::Pex16, FBAIN::EgD5SM::Pex20, EXP1::EgD5SM::Lip1, YAT1::EaD5SM::Oct, FBAINm::PaD17::Aco, EXP1::PaD17::Pex16, YAT1::PaD17S::Lip1, YAT1::Y1CPT1::Aco, YAT1::MCS::Lip1, FBA::MCS::Lip1 (2 카피), EXP1::Y1PCT::Pex16, YAT1::MaLPAAT1S::Pex16, YAT1::Y1LPAAT::Lip1, FBAINm::Y1LPAAT1::Lip1 (2 카피), YAT1::Y1PDAT::Lip1 (2 카피).

[0460]

균주 Z5620, Z5623 및 Z5625에서의 YALI0B14795p 유전자좌(진뱅크 수탁 번호 Z5620XM_500900)의 낙아웃은 pZK16-MyL8N으로의 형질전환에 의해 생성되는 이들 EPA 균주 중 임의의 것에서 확인되지 않았다.

[0461]

플라스크 검정법에 의한 전체 지질 함량 및 조성의 분석

- [0462] 균주 Z5620, Z5623 및 Z5625의 YPD 플레이트로부터의 세포를 성장시키고, 일반 방법에 따라 전체 지질 함량 및 조성에 대하여 분석하였다.
- [0463] 바람직한 실시형태의 설명 내의 표 7(상기)은 균주 Z5620, Z5623 및 Z5625의 총 DCW, TFA % DCW, TFA의 중량 백분율["% TFA"]로서의 각 지방산의 농도 및 EPA % DCW를 요약한 것이다. 평균 DCW는 4.5 g/L였고, 평균 TFA % DCW는 52.4였으며, 평균 EPA % TFA는 48.4였고, 평균 EPA % DCW는 25.9였다.
- [0464] 실시예 5
- [0465] 적어도 약 55 TFA % DCW와 함께 적어도 약 48 EPA % TFA를 생성하는 야로위아 리폴리티카 균주 Z5581, Z5582, Z5583 및 Z5584의 생성
- [0466] 본 실시예에는 델타-9 연장효소/델타-8 불포화화효소 경로의 발현을 통하여 55 초과 TFA % DCW와 함께, 약 48 EPA % TFA를 생성할 수 있는 야로위아 리폴리티카 균주 L258U(실시예 3) 유래의 균주 Z5581, Z5582, Z5583 및 Z5584의 구축이 기재되어 있다.
- [0467] 하나의 말로닐-CoA 신테타제 유전자 및 하나의 라이소포스파티드산 아실트랜스퍼라제 유전자 및 하나의 C_{16/18} 연장효소 유전자가 균주 L258U의 야로위아 YALI0B14795p 유전자좌(진뱅크 수탁 번호 XM_500900) 내로 통합되도록 구축물 pZK16-ML3(도 6b, 서열 번호 89)을 생성하였다. pZK16-ML3플라스미드에는 하기의 성분이 포함되어 있다:

표 12

플라스미드 pZK16-ML3(서열 번호 89)의 설명	
서열 번호 89 내의 RE 부위 및 뉴클레오티드	단편 및 키메라 유전자 성분의 설명
<i>AscI/BsrWI</i> (1905-1)	YALI0B14795p 유전자좌의 1904 bp 5' 부분(진뱅크 수탁 번호 XM_500900, 도면에서 "Y8716-5"로 표시)
<i>PacI/SphI</i> (6414-4613)	YALI0B14795p 유전자좌의 1801 bp 3' 부분(진뱅크 수탁 번호 XM_500900, 도면에서 "Y8716-3"으로 표시)
<i>SwaI/BsrWI</i> (12920 - 1)	하기를 포함하는 GPAT::ME3S::Pex20: ● GPAT: 야로위아 리폴리티카 GPAT 프로모터(PCT 공개 WO 2006/031937 호); ● ME3S: 모르티에렐라 알피나 유래의 코돈-최적화된 C_{16/18} 연장효소 유전자(서열 번호 90)(미국 특허 제 7,470,532 호); ● Pex20: 야로위아 <i>Pex20</i> 유전자로부터의 Pex20 종결자 서열(진뱅크 수탁 번호 AF054613)
<i>PmeI/SwaI</i> (10534-12920)	하기를 포함하는 FBA::MCS::Lip1: ● FBA: 야로위아 리폴리티카 FBA 프로모터(미국 특허 제 7,202,356 호); ● MCS: 리조비움 레구미노사움 생물변이형 비시에 3841 유래의 코돈-최적화된 말로닐-CoA 신테타제 유전자(서열 번호 49)(미국 특허 출원 공개 제 2010-0159558-A1 호); ● Lip1: 야로위아 <i>Lip1</i> 유전자로부터의 Lip1 종결자 서열(진뱅크 수탁 번호 Z50020)
<i>ClaI/PmeI</i> (8515-10534)	하기를 포함하는 YAT1::MaLPAATIS::Pex16: ● YAT1: 야로위아 리폴리티카 YAT1 프로모터(미국 특허 출원 공개 제 2010-0068789-A1 호); ● MaLPAATIS: 모르티에렐라 알피나 유래의 코돈-최적화된 라이소포스파티드산 아실트랜스퍼라제 유전자(서열 번호 19)(미국 특허 제 7,879,591 호); ● Pex16: 야로위아 <i>Pex16</i> 유전자로부터의 Pex16 종결자 서열(진뱅크 수탁 번호 U75433)
<i>SalI/EcoRI</i> (8065 - 6446)	야로위아 Ura3 유전자(진뱅크 수탁 번호 AJ306421)

- [0468]
- [0469] pZK16-ML3 플라스미드를 *AscI/SphI*으로 분해한 다음, 일반 방법에 따라 균주 L258U5의 형질전환을 위해 사용하였다. 형질전환 세포를 MM 플레이트 상에 플레이팅하고, 30°C에서 4 내지 5일 동안 유지시켰다. 단일의 콜로니를 MM 플레이트 상으로 다시 스트리킹한 다음, 30°C에서 액체 MM으로 접종하고, 250 rpm/분으로 2일 동안 셰이킹하였다. 세포를 원심분리에 의해 수집하고, HGM에 재현탁화시킨 다음, 250 rpm/분에서 5일 동안 셰이킹하였다. 세포를 일반 방법에 따라 지방산 분석으로 처리하였다.
- [0470] GC 분석에 의해, pZK16-ML3으로 형질전환된 L258U5의 선택된 48개 균주 중 19개가 TFA의 50% 초과 EPA를 생성하는 것으로 나타났다. TFA의 약 50.9%, 52.4%, 51.5% 및 51.7%의 EPA를 생성하였던 4개의 균주(즉, #16,

#42, #46 및 #47)를 각각 Z5581, Z5582, Z5583 및 Z5584로 명명하였다.

- [0471] 야생형 야로위아 리폴리티카 ATCC #20362에 대한 이들 pZK16-ML3 형질전환된 균주의 최종 유전형을 하기와 같았다: Ura⁺, Pex3⁻, unknown 1⁻, unknown 2⁻, unknown 3⁻, unknown 4⁻, YALIOE12947g⁻, unknown6⁻, YALIOB21890g⁻, unknown 8⁻, unknown 10⁻, unknown 11⁻, unknown 12⁻, unknown 13⁻, unknown 14⁻, YAT1::ME3S::Pex16, GPD::ME3S::Pex20, YAT1::ME3S::Lip1, GPAT::ME3S::Pex20, FBAINm::EgD9eS::Lip2, EXP1::EgD9eS::Lip1, GPAT::EgD9eS::Lip2, YAT1::EgD9eS::Lip2, YAT1::EgD9eS-L35G::Pex20, FBAINm::EgD8M::Pex20, EXP1::EgD8M::Pex16, FBAIN::EgD8M::Lip1, GPD::EaD8S::Pex16 (2 카피), YAT1::E389D9eS/EgD8M::Lip1, YAT1::EgD9eS/EgD8M::Aco, FBAINm::EaD9eS/EaD8S::Lip2, GPDIN::Y1D9::Lip1, GPD::FmD12::Pex20, YAT1::FmD12::Oct, EXP1::FmD12S::Aco, GPDIN::FmD12::Pex16, EXP1::EgD5M::Pex16, FBAIN::EgD5SM::Pex20, EXP1::EgD5SM::Lip1, YAT1::EaD5SM::Oct, FBAINm::PaD17::Aco, EXP1::PaD17::Pex16, YAT1::PaD17S::Lip1, YAT1::Y1CPT1::Aco, YAT1::MCS::Lip1, FBA::MCS::Lip1 (2 카피), EXP1::Y1PCT::Pex16, YAT1::MaLPAAT1S::Pex16 (2 카피), FBAINm::Y1LPAAT1::Lip1 (2 카피), YAT1::Y1PDAT::Lip1 (2 카피).
- [0472] Z5581, Z5582, Z5583 및 Z5584에서의 YALIOB14795 유전자좌(진뱅크 수탁 번호 XM_500900)의 낙아웃은 pZK16-ML3으로의 형질전환에 의해 생성되는 이들 EPA 균주 중 임의의 것에서 확인되지 않았다.
- [0473] 플라스크 검정법에 의한 전체 지질 함량 및 조성의 분석
- [0474] 균주 Z5581, Z5582, Z5583 및 Z5584의 YPD 플레이트로부터의 세포를 성장시키고, 일반 방법에 따라 전체 지질 함량 및 조성에 대하여 분석하였다.
- [0475] 바람직한 실시형태의 설명에서의 표 7(상기)에는 균주 Z5581, Z5582, Z5583 및 Z5584의 총 DCW, TFA % DCW, TFA의 중량 백분율 ["% TFA"]로서의 각 지방산의 농도 및 EPA % DCW가 요약되어 있다. 평균 DCW는 4.8 g/L이었으며, 평균 TFA % DCW는 56이었으며, 평균 EPA % TFA는 48.65이었으며, 평균 EPA % DCW는 27.2이었다.
- [0476] 실시예 6
- [0477] 적어도 약 54 TFA % DCW와 함께 적어도 약 48 EPA % TFA를 생성하는 야로위아 리폴리티카 균주 Z5570, Z5571, Z5572 및 Z5574의 생성
- [0478] 본 실시예에는 델타-9 연장효소/델타-8 불포화화효소 경로의 발현을 통하여 54 초과 TFA % DCW와 함께, 약 48 EPA % TFA를 생성할 수 있는 야로위아 리폴리티카 균주 L258U(실시예 3) 유래의 균주 Z5570, Z5571, Z5572 및 Z5574의 구축이 기재되어 있다.
- [0479] 하나의 라이소포스파티드산 아실트랜스퍼라제 유전자, 하나의 델타-9 불포화화효소 유전자 및 하나의 콜린포스페이트 시티딜틸트랜스퍼라제 유전자가 균주 L258U의 야로위아 YALIOF02211g 유전자좌(진뱅크 수탁 번호 XP_504895) 내로 통합되도록 구축물 pZKMP-ML9DP(도 7a, 서열 번호 92)를 생성하였다. pZKMP-ML9DP 플라스미드에는 하기의 성분이 포함되어 있다:

표 13

플라스미드 pZKMP-ML9DP(서열 번호 92)의 설명

서열 번호 92 내의 RE 부위 및 뉴클레오타이드	단편 및 키메라 유전자 성분의 설명
<i>AscI/BsiWI</i> (696-1)	YAL10F02211g 유전자좌의 695 bp 5' 부분(진뱅크 수탁 번호 XP_504895; 도면에서 "yM1DP-5"로 표시)
<i>PacI/SphI</i> (4201-3404)	YAL10F02211g 유전자좌의 797 bp 3' 부분(진뱅크 수탁 번호 XP_504895; 도면에서 "yM1DP-3"으로 표시)
<i>SwaI/BsiWI</i> (11068 - 1)	하기를 포함하는 ALK2LM1::MaLPAAT1S::Pex20: ● ALK2LM1: 야로위아 리폴리티카 ALK2LM1 프로모터 + 야로위아 리폴리티카 <i>AIK2</i> 유전자의 N-말단 66 bp 암호화 영역(서열 번호 93; 미국 가특허 출원 제 61/471,746 호(출원일: 2011 년 4 월 5 일)(대리인 문서 번호, CL5381USPRV)); ● MaLPAAT1S: 모르티에펠라 알피나 유래의 코돈-최적화된 라이소포스포타이드산 아실트랜스퍼라제 유전자(서열 번호 19)(미국 특허 제 7,879,591 호); ● Pex20: 야로위아 <i>Pex20</i> 유전자로부터의 Pex20 종결자 서열(진뱅크 수탁 번호 AF054613)
<i>PmeI/SwaI</i> (8560-11068)	하기를 포함하는 DGAT2M::YID9::Lip1: ● DGAT2M: 야로위아 리폴리티카 DGAT2M 프로모터(서열 번호 94; 미국 가특허 출원 제 61/469,933 호(출원일: 2011 년 3 월 31 일)(대리인 문서 번호 CL4736USPRV)); ● YID9: 야로위아 리폴리티카 델타-9 불포화효소 유전자(도면에서 "YID9DS"로 표시; 진뱅크 수탁 번호 XM_501496.1; 서열 번호 80); ● Lip1: 야로위아 <i>Lip1</i> 유전자로부터의 Lip1 종결자 서열(진뱅크 수탁 번호 Z50020)
<i>ClaI/PmeI</i> (6302-8560)	하기를 포함하는 EXP1::YIPCT::Pex16: ● EXP1: 야로위아 리폴리티카 수송 단백질(EXP1) 프로모터(도면에서 "EXP"로 표시; 국제 출원 공개 제 WO 2006/052870 호); ● YIPCT: 야로위아 리폴리티카 콜린포스페이트 시티딜릴트랜스퍼라제 유전자(진뱅크 수탁 번호 XP_502978; 서열 번호 45); ● Pex16: 야로위아 <i>Pex16</i> 유전자로부터의 Pex16 종결자 서열(진뱅크 수탁 번호 U75433)
<i>SalI/EcoRI</i> (5852 - 4233)	야로위아 <i>Ura3</i> 유전자(진뱅크 수탁 번호 AJ306421)

[0480]

[0481]

pZKMP-ML9DP 플라스미드를 *AscI/SphI*으로 분해한 다음, 일반 방법에 따라 균주 L258U5의 형질전환을 위해 사용하였다. 형질전환 세포를 MM 플레이트 상에 플레이팅하고, 30℃에서 5일 동안 유지시켰다. 단일의 콜로니를 MM 플레이트 상으로 다시 스트리킹한 다음, 30℃에서 액체 MM으로 접종하고, 250 rpm/분으로 2일 동안 셰이킹하였다. 세포를 원심분리에 의해 수집하고, HGM에 재현탁화시킨 다음, 250 rpm/분에서 5일 동안 셰이킹하였다. 세포를 일반 방법에 따라 지방산 분석으로 처리하였다.

[0482]

GC 분석에 의해, pZKMP-ML9DP로 형질전환된 L258U5의 선택된 48개 균주 중 22개가 TFA의 50% 초과 EPA를 생성하는 것으로 나타났다. TFA의 약 53.5%, 51.8%, 52.9% 및 51.8%의 EPA를 생성하는 4개의 균주(즉, #17, #25, #40 및 #46)는 각각 Z5570, Z5571, Z5572 및 Z5574로 명명하였다.

[0483]

야생형 야로위아 리폴리티카 ATCC #20362에 대한 이들 pZKMP-ML9DP 형질전환 균주의 최종 유전형은 하기와 같다: *Ura*⁺, *Pex3*⁻, unknown 1⁻, unknown 2⁻, unknown 3⁻, unknown 4⁻, YAL10E12947g⁻, unknown6⁻, YAL10B21890g⁻, unknown 8⁻, unknown 10⁻, unknown 11⁻, unknown 12⁻, unknown 13⁻, unknown 14⁻, YAT1::ME3S::Pex16, GPD::ME3S::Pex20, YAT1::ME3S::Lip1, FBAINm::EgD9eS::Lip2, EXP1::EgD9eS::Lip1, GPAT::EgD9e::Lip2, YAT1::EgD9eS::Lip2, YAT1::EgD9eS-L35G::Pex20, FBAINm::EgD8M::Pex20, EXP1::EgD8M::Pex16, FBAIN::EgD8M::Lip1, GPD::EaD8S::Pex16 (2 카피), YAT1::E389D9eS/EgD8M::Lip1, YAT1::EgD9eS/EgD8M::Aco, FBAINm::EaD9eS/EaD8S::Lip2, DGAT2M::YID9::Lip1, GPDIN::YID9::Lip1, GPD::FmD12::Pex20, YAT1::FmD12::Oct, EXP1::FmD12S::Aco, GPDIN::FmD12::Pex16, EXP1::EgD5M::Pex16, FBAIN::EgD5SM::Pex20, EXP1::EgD5SM::Lip1, YAT1::EaD5SM::Oct, FBAINm::PaD17::Aco, EXP1::PaD17::Pex16, YAT1::PaD17S::Lip1, YAT1::Y1CPT1::Aco, YAT1::MCS::Lip1, FBA::MCS::Lip1, EXP1::YIPCT::Pex16 (2 카피), YAT1::MaLPAAT1S::Pex16, ALK2LM1::MaLPAAT1S::Pex20, FBAINm::Y1LPAAT1::Lip1 (2 카피), YAT1::Y1PDAT::Lip1 (2 카피).

[0484]

균주 Z5570, Z5571, Z5572 및 Z5574에서의 YAL10F02211g 유전자좌(진뱅크 수탁 번호 XP_504895)의 낙아아웃은 pZKMP-ML9DP로의 형질전환에 의해 생성된 이들 EPA 균주 중 임의의 것에서 확인되지 않았다.

- [0485] 플라스크 검정법에 의한 전체 지질 함량 및 조성의 분석
- [0486] 균주 Z5570, Z5571, Z5572 및 Z5574의 YPD 플레이트로부터의 세포를 성장시키고, 일반 방법에 따라, 전체 지질 함량 및 조성에 대하여 분석하였다.
- [0487] 바람직한 실시형태의 설명에서의 표 7(상기)에는 균주 Z5570, Z5571, Z5572 및 Z5574의 총 DCW, TFA % DCW, TFA의 중량 백분율["% TFA"]로서의 각 지방산의 농도 및 EPA % DCW가 요약되어 있다. 평균 DCW는 4.9 g/L이었으며, 평균 TFA % DCW는 54.2이었고, 평균 EPA % TFA는 48.9이었으며, 평균 EPA % DCW는 26.55이었다.
- [0488] 실시예 7
- [0489] 적어도 약 52 TFA % DCW와 함께 적어도 약 49 EPA % TFA를 생성하는 야로위아 리폴리티카 균주 Z5585 및 Z5627의 생성
- [0490] 본 실시예에는 델타-9 연장효소/델타-8 불포화화효소 경로의 발현을 통하여 52 초과의 TFA % DCW와 함께, 약 49 EPA % TFA를 생성할 수 있는 야로위아 리폴리티카 균주 L258U(실시예 3) 유래의 균주 Z5585 및 Z5627의 구축이 기재되어 있다.
- [0491] 하나의 라이소포스파티드산 아실트랜스퍼라제 유전자, 하나의 델타-9 불포화화효소 유전자 및 하나의 다이아실글리세롤 콜린포스포트랜스퍼라제 유전자가 균주 L258U의 야로위아 YAL10F02211g 유전자좌(진뱅크 수탁 번호 XP_504895) 내로 통합되도록 구축물 pZKMP-ML9DCB(도 7b, 서열 번호 95)를 생성하였다. pZKMP- ML9DCB 플라스미드에는 하기의 성분이 포함되어 있다:

표 14

플라스미드 pZKMP-ML9DCB(서열 번호 95)의 설명	
서열 번호 95 내의 RE 부위 및 뉴클레오타이드	단편 및 키메라 유전자 성분의 설명
<i>AscI/BstWI</i> (7715-7020)	YAL10F02211g 유전자좌의 695 bp 5' 부분(진뱅크 수탁 번호 XP_504895), 도면에 "yM1DP--5"로 표시)
<i>PacI/SphI</i> (11220-10423)	YAL10F02211g 유전자좌의 797 bp 3' 부분(진뱅크 수탁 번호 XP_504895), 도면에 "yM1DP--3"으로 표시)
<i>SwaI/BstWI</i> (4850 - 7020)	하기를 포함하는 ALK2LM1::MaLPAATIS::Pex20: ● ALK2LM1: 야로위아 리폴리티카 ALK2LM1 프로모터 + 야로위아 리폴리티카 <i>AIK2</i> 유전자의 N-말단 66 bp 암호화 영역(서열 번호 93; 미국 가특허 출원 제 61/471,746 호(출원일: 2011 년 4 월 5 일)(대리인 문서 번호 CL5381USPRV)); ● MaLPAATIS: 모르티에렐라 알피나 유래의 코돈-최적화된 라이소포스파티드산 아실트랜스퍼라제 유전자(서열 번호 19)(미국 특허 제 7,879,591 호); ● Pex20: 야로위아 <i>Pex20</i> 유전자로부터의 Pex20 종결자 서열(진뱅크 수탁 번호 AF054613)
<i>PmeI/SwaI</i> (2342-4850)	하기를 포함하는 DGAT2M::YID9::Lip1: ● DGAT2M: 야로위아 리폴리티카 DGAT2M 프로모터(서열 번호 94; 미국 가특허 출원 제 61/469,933 호(출원일: 2011 년 3 월 31 일)(대리인 문서 번호 CL4736USPRV)); ● YID9: 야로위아 리폴리티카 델타-9 불포화화효소 유전자(도면에서 "YID9DS"로 표시; 진뱅크 수탁 번호 XM_501496.1; 서열 번호 80); ● Lip1: 야로위아 <i>Lip1</i> 유전자로부터의 Lip1 종결자 서열(진뱅크 수탁 번호 Z50020)
<i>Clal/PmeI</i> (1-2342)	하기를 포함하는 EXP1::YICPT1::OCT: ● EXP1: 야로위아 리폴리티카 수송 단백질(EXP1) 프로모터(도면에서 "EXP"로 표시; 국제 출원 공개 제 WO 2006/052870 호); ● YICPT1: 야로위아 리폴리티카 다이아실글리세롤 콜린포스포트랜스퍼라제 유전자(서열 번호 47; 국제. 출원 공개 WO 2006/052870 호; 진뱅크 수탁 번호 XM_501703); ● OCT: 야로위아 <i>OCT</i> 유전자의 OCT 종결자 서열(진뱅크 수탁 번호 X69988)
<i>SalI/EcoRI</i> (12871 - 11252)	야로위아 Ura3 유전자(진뱅크 수탁 번호 AJ306421)

- [0492]
- [0493] pZKMP-ML9DCB 플라스미드를 *AscI/SphI*으로 분해한 다음, 일반 방법에 따라 개별적으로 균주 L258U5 및 L258U6의 형질전환을 위해 사용하였다. 형질전환 세포를 MM 플레이트 상에 플레이팅하고, 30℃에서 5 내지 6일 동안 유지시켰다. 단일의 콜로니를 MM 플레이트 상으로 다시 스트리킹한 다음, 30℃에서 액체 MM으로 접종하고, 250 rpm/분으로 2일 동안 셰이킹하였다. 세포를 원심분리에 의해 수집하고, HGM에 재현탁화시킨 다음, 250 rpm/분

에서 5일 동안 셰이킹하였다. 세포를 일반 방법에 따라 지방산 분석으로 처리하였다.

- [0494] GC 분석에 의해, pZKMP-ML9DCB로 형질전환된 L258U5의 선택된 50개 균주 중 21개가 TFA의 50% 초과 EPA를 생성하는 것으로 나타났다. TFA의 약 52.3% 및 51.9%의 EPA를 생성하였던 2개의 균주(즉, #1 및 #1B)를 각각 Z5585 및 Z5627로 명명하였다.
- [0495] 야생형 야로위아 리폴리티카 ATCC #20362에 대한 이들 pZKMP-ML9DCB 형질전환 균주의 최종 유전형은 하기와 같았다: Ura⁺, Pex3⁻, unknown 1⁻, unknown 2⁻, unknown 3⁻, unknown 4⁻, YALIOE12947g⁻, unknown6⁻, YALIOB21890g⁻, unknown 8⁻, unknown 10⁻, unknown 11⁻, unknown 12⁻, unknown 13⁻, unknown 14⁻, YAT1::ME3S::Pex16, GPD::ME3S::Pex20, YAT1::ME3S::Lip1, FBAINm::EgD9eS::Lip2, EXP1::EgD9eS::Lip1, GPAT::EgD9e::Lip2, YAT1::EgD9eS::Lip2, YAT1::EgD9eS-L35G::Pex20, FBAINm::EgD8M::Pex20, EXP1::EgD8M::Pex16, FBAIN::EgD8M::Lip1, GPD::EaD8S::Pex16 (2 카피), YAT1::E389D9eS/EgD8M::Lip1, YAT1::EgD9eS/EgD8M::Aco, FBAINm::EaD9eS/EaD8S::Lip2, DGAT2M::Y1D9::Lip1, GPDIN::Y1D9::Lip1, GPD::FmD12::Pex20, YAT1::FmD12::Oct, EXP1::FmD12S::Aco, GPDIN::FmD12::Pex16, EXP1::EgD5M::Pex16, FBAIN::EgD5SM::Pex20, EXP1::EgD5SM::Lip1, YAT1::EaD5SM::Oct, FBAINm::PaD17::Aco, EXP1::PaD17::Pex16, YAT1::PaD17S::Lip1, YAT1::Y1CPT1::Aco, EXP1::Y1CPT1::OCT, YAT1::MCS::Lip1, FBA::MCS::Lip1, EXP1::Y1PCT::Pex16, YAT1::MaLPAAT1S::Pex16, ALK2LM1::MaLPAAT1S::Pex20, FBAINm::Y1LPAAT1::Lip1 (2 카피), YAT1::Y1PDAT::Lip1 (2 카피).
- [0496] 균주 Z5585 및 Z5627에서의 YALIOF02211g 유전자좌(진뱅크 수탁 번호 XP_504895)의 낙아웃은 pZKMP-ML9DCB로의 형질전환에 의해 생성된 이들 EPA 균주 중 임의의 것에서 확인되지 않았다.
- [0497] 플라스크 검정법에 의한 전체 지질 함량 및 조성의 분석
- [0498] 균주 Z5585 및 Z5627의 YPD 플레이트로부터의 세포를 성장시키고, 일반 방법에 따라 전체 지질 함량 및 조성에 대하여 분석하였다.
- [0499] 바람직한 실시형태의 설명 내의 표 7(상기)은 균주 Z5585 및 Z5627의 DCW, TFA % DCW, TFA의 중량 백분율["% TFA"]로서의 각 지방산의 농도 및 EPA % DCW를 요약한 것이다. 평균 DCW는 4.7 g/L이었으며, 평균 TFA % DCW는 54.3이었고, 평균 EPA % TFA는 49.4이었으며, 평균 EPA % DCW는 26.8이었다.
- [0500] 실시예 8
- [0501] 적어도 약 45 TFA % DCW와 함께 적어도 약 36 ETA % TFA를 생성하는 야로위아 리폴리티카 균주 YOS9607 및 YOS9608의 생성
- [0502] 본 실시예에는 플라스크 검정법에서 45 초과 TFA % DCW와 함께, 36 초과 ETA % TFA를 생성할 수 있는 야로위아 리폴리티카 균주 Z5567(실시예 3) 유래의 균주 YOS9607 및 YOS9608의 구축이 기재되어 있다. 균주 Z5567 내의 원래의 4개의 델타-5 불포화화효소 유전자를 결실시켜, 균주 YOS9607 및 YOS9608을 생성하여, 이에 따라 EPA의 생성 없이 ETA를 생성할 수 있게 되었다.
- [0503] 균주 YOS9607 및 YOS9608의 발생은 중간 균주 Z5567U, YOS9601 및 YOS9602(도 9)의 구축을 필요로 한다.
- [0504] 균주 Z5567U(Ura3-)의 생성
- [0505] *Ura3* 유전자를 파괴하기 위하여, 구축물 pZKUM(도 4a; 서열 번호 82; 미국 특허 출원 공개 제2009-0093543-A1호의 표 15에 기재)을 사용하여 *Ura3* 돌연변이 유전자를 균주 Y9502의 pZKUM 형질전환(실시예 1)에 대해 기재된 것과 유사한 방식으로, 균주 Z5567의 *Ura3* 유전자에 통합시켰다. 그룹 C의 총 19개의 형질전환체를 성장시키고, *Ura*⁻ 표현형을 갖는 것으로 확인하였다.
- [0506] GC 분석에 의해, MM + 5-FOA 플레이트 상에서 성장시킨, pZKUM-형질전환 균주 #6, #11, #13, #15 및 #16에서 36.9%, 37.0%, 35.6%, 36.8% 및 36.0%의 EPA의 존재가 나타났다. 이들 5개의 균주를 각각 Z5567U14, Z5567U19, Z5567U21, Z5567U23 및 Z5567U24로 명명하고, 집합하여 균주 Z5567U로 명명하였다.
- [0507] 균주 YOS9607 및 YOS9608의 생성
- [0508] 균주 Z5567 내의 4개의 델타-5 불포화화효소 유전자를 처음에 2가지 상이한 구축물로부터 염색체로 통합시켰다: 키메라 EXP1::EgD5M::Pex16, FBAIN::EgD5SM::Pex20 및 YAT1::EaD5SM::OCT 유전자를 포함하는 pZKSL-5S5A5 (도 8a; 서열 번호 96) 및 키메라 EXP1::EgD5SM::Lip1 유전자를 포함하는 pZP2-85m98F(도 8b; 서열 번호 97). 이

들 키메라 유전자를 제거하기 위하여 3가지 개별 상동성 재조합 사건이 필요하였다.

[0509]

먼저, 균주 Z5567U의 게놈 내의 키메라 FBAIN::EgD5SM 유전자 및 *Leu* 유전자의 대부분(즉, pZKSL-5S5A5로부터의)을 상동성 재조합(도 10a)에 의해, 플라스미드 pYPS234(도 10b; 서열 번호 99) 내의 993 bp 스테퍼 DNA 단편(서열 번호 98)으로 대체하였으며, 여기서, 993 bp 스테퍼는 야로위아 카르니틴/아실 카르니틴 운반 유전자의 5' 및 3' 부분을 포함한다. 더욱 구체적으로, pYSP234 플라스미드는 하기의 성분을 함유한다.

표 15

플라스미드 pYPS234(서열 번호 99)의 설명

서열 번호 99 내의 RE 부위 및 뉴클레오티드	단편 및 키메라 유전자 성분의 설명
<i>SwaI/PacI</i> (1-1498)	Ura3: 야로위아 Ura3 유전자(진뱅크 수탁 번호 AJ306421)
2494-3354	Amp: 에스케리키아 콜라이에서의 선택을 위한 앰피실린-내성 유전자
<i>BsiWI/PmeI</i> (4239-4964)	YAT1: 야로위아 리플리터카 YAT1 프로모터(도면에서 "YAT"로 표시; 미국 특허 출원 공개 제 2010-0068789-A1 호)
4968-5320	Leu 단편: Leu2 유전자의 353 bp 단편(진뱅크 수탁 번호 AF260230)
5327-6319	180 스테퍼: 야로위아 카르니틴/아실 카르니틴 운반 유전자(진뱅크 수탁 번호 XP 501358)의 5' 및 3' 부분을 포함하는 993 bp DNA 단편(서열 번호 98)
<i>BamHI/BsiWI</i>	Pex20: 야로위아 Pex20 유전자로부터의 Pex20 종결자 서열(진뱅크 수탁 번호 AF054613)
<i>HindIII/SwaI</i> (6638-7338)	Lys5-5': 야로위아 <i>Lys5</i> 유전자의 720 bp 5' 부분(진뱅크 수탁 번호 M34929; 도면에서 "lys5 5' 영역"으로 표시)

[0510]

Lys5-5' DNA 단편 내에서 처음의 교차 사건이 발생하였으며, 두번째 교차 사건은 YAT1 프로모터 영역 내에서 발생하였다. 그의 게놈 내에 *Leu- Ura-* 표현형 및 3개의 델타-5 불포화화효소 유전자를 갖는 균주 YOS9601을 이러한 상동성 재조합으로부터 생성하였다.

[0511]

[0512]

이어서, 균주 YOS9601의 게놈 내의 키메라 EXP1::EgD5M::Pex16 및 YAT1::EaD5SM::OCT 유전자를 상동성 재조합(도 11a)에 의해 플라스미드 pYPS233(도 11b; 서열 번호 101) 내의 1019 bp 스테퍼 DNA 단편(서열 번호 100)으로 대체하였으며, 여기서, 1019 bp 스테퍼는 야로위아 *ALK2* 유전자의 5' 및 3' 부분을 포함하였다. pYSP233 플라스미드는 하기의 성분을 포함하였다:

표 16

플라스미드 pYPS233(서열 번호 101)의 설명

서열 번호 101 내의 RE 부위 및 뉴클레오티드	단편 및 키메라 유전자 성분의 설명
<i>SwaI/PacI</i> (1-1498)	Ura3: 야로위아 Ura3 유전자(진뱅크 수탁 번호 AJ306421)
2494-3354	Amp: 에스케리키아 콜라이에서의 선택을 위한 앰피실린-내성 유전자
<i>SphI/PacI</i> (4214-4901)	Lys5-3: 야로위아 <i>Lys5</i> 유전자의 684 bp 3' 부분(진뱅크 수탁 번호 M34929; 도면에서 "Lys5-3' 영역"으로 표시)
<i>PacI/MluI</i> (4904-5227)	Pex16: 야로위아 <i>Pex16</i> 유전자로부터의 Pex16 종결자 서열(진뱅크 수탁 번호 U75433)
<i>MluI/BsiWI</i> (5229-6263)	171 스테퍼: 야로위아 <i>ALK2</i> 유전자(진뱅크 수탁 번호 BAA31434)의 5' 및 3' 부분을 포함하는 1019 bp DNA 단편(서열 번호 100)
<i>BsiWI/PmeI</i> (6263-6988)	YAT1: 야로위아 리플리터카 YAT1 프로모터(도면에서 "YAT"로 표시; 미국 특허 출원 공개 제 2010-0068789-A1 호)
<i>PmeI/SwaI</i> (6988/1)	Leu 단편: Leu2 유전자의 353 bp 단편(진뱅크 수탁 번호 AF260230)

[0513]

[0514]

Lys5-3' DNA 단편 내에 처음의 교차 사건이 발생하였으며, 두번째 교차 사건은 3' *Leu* 또는 YAT1 프로모터 영역 중 어느 하나에서 발생하였다. 게놈 내에 남아있는 *Leu- Ura-* 표현형 및 하나의 기능성 델타-5 불포화화효소 유전자를 갖는 균주 YOS9602를 이러한 상동성 재조합으로부터 생성하였다.

[0515]

마지막으로, 균주 YOS9602의 게놈 내의 키메라 EXP1::EgD5SM::Lip1 유전자를 상동성 재조합(도 12a)에 의해 플라스미드 pYSP241 내의 기능성 *Leu2* 유전자(도 12b; 서열 번호 102)로 대체하였다. pYSP241 플라스미드는 하기의

의 성분을 포함하였다:

표 17

플라스미드 pYPS241(서열 번호 102)의 설명

서열 번호 102 내의 RE 부위 및 뉴클레오티드	단편 및 키메라 유전자 성분의 설명
<i>Clal/SphI</i> (1-2134)	LeuL: 아이소프로필말레이트 탈수소효소를 암호화하는 야로위아 Leu2 유전자(진뱅크 수탁 번호 AF260230)
<i>SwaI/PmeI</i> (3893-2134)	하기를 포함하는 EaD8S::Pex20: - EaD8S: 야로위아 리폴리티카에서의 발현을 위해 코돈-최적화된("EaD8S"; 서열 번호 103) 유글레나 아나바에나 유래의(미국 특허 제 7,789,156 호) 합성 델타-8 불포화화효소; - Pex20: 야로위아 <i>Pex20</i> 유전자로부터의 Pex20 종결자 서열(진뱅크 수탁 번호 AF054613)
<i>SphI/Clal</i> (8106-9209)	B21890 아암(arm) A: 야로위아 ORF B21890의 업스트림 DNA 서열(진뱅크 수탁 번호 XP_501199).
<i>SwaI/PaeI</i> (3893-5390)	Ura3: 야로위아 Ura3 유전자(진뱅크 수탁 번호 AJ306421)
6386-7246	Amp: 에스케리키아 콜라이에서의 선택을 위한 암피실린-내성 유전자

[0516]

[0517]

염색체 B의 위치 2870148 내지 2871250의 영역에서 첫번째 교차 사건이 발생하였으며, 두번째 교차 사건이 플라스미드 pYSP241의 EaD8S 영역 내에서 발생하여, 균주 YOS9607(Ura-) 및 YOS9608(Ura-)이 생성되었다. 균주 YOS9607 및 YOS9608(동일한 유전형을 갖는 2개의 개별 콜로니에 해당)을 이러한 상동성 재조합으로부터 생성하였으며, 각각 Ura- 표현형을 가지며, 게놈 내에 델타-5 불포화화효소 유전자를 갖지 않았다.

[0518]

야생형 야로위아 리폴리티카 ATCC #20362에 대한 YOS9607 및 YOS9608 균주의 최종 유전형은 하기와 같았다: Ura-, Pex3-, unknown 1-, unknown 2-, unknown 3-, unknown 4-, YAL10E12947g-, unknown6-, YAL10B21890g-, unknown 8-, unknown 10-, unknown 11-, unknown 12-, unknown 13-, unknown 14-, YAT1::ME3S::Pex16, GPD::ME3S::Pex20, YAT1::ME3S::Lip1, FBAINm::EgD9eS::Lip2, EXP1::EgD9eS::Lip1, GPAT::EgD9e::Lip2, YAT1::EgD9eS::Lip2, YAT1::EgD9eS-L35G::Pex20, FBAINm::EgD8M::Pex20, EXP1::EgD8M::Pex16, FBAIN::EgD8M::Lip1, YAT1::EgD8M::Pex20, GPD::EaD8S::Pex16 (2 카피), YAT1::E389D9eS/EgD8M::Lip1, YAT1::EgD9eS/EgD8M::Aco, FBAINm::EaD9eS/EaD8S::Lip2, GPDIN::Y1D9::Lip1, GPD::FmD12::Pex20, YAT1::FmD12::Oct, EXP1::FmD12S::Aco, GPDIN::FmD12::Pex16, FBAINm::PaD17::Aco, EXP1::PaD17::Pex16, YAT1::PaD17S::Lip1, YAT1::Y1CPT1::Aco, YAT1::MCS::Lip1, FBA::MCS::Lip1 (2 카피), EXP1::Y1PCT::Pex16, YAT1::MaLPAAT1S::Pex16 (2 카피), FBAINm::Y1LPAAT1::Lip1 (2 카피), YAT1::Y1PDAT::Lip1 (2 카피).

[0519]

균주 YOS9601(Leu-, Ura-), YOS9602(Leu-, Ura-), YOS9607(Ura-), YOS9608(Ura-) 및 Z5567U(Ura-) 대조군의 지방산 조성 및 오일 함량을 분석하기 위하여, 3벌의 플라스크 검정법을 상기 일반 방법에 기재된 바와 같이 수행하였다.

[0520]

표 18에는 총 DCW, TFA % DCW, TFA의 중량 백분율["% TFA"]로서의 각 지방산의 농도 및 EPA % DCW가 요약되어 있다. 지방산은 표 7에서와 같으며, 20:4(5,11,14,17)는 주니페론산을 지칭한다. 각 시료 중의 모든 지방산의 합은 총 100이다.

표 18

야로위아 균주 YOS9601, YOS9602, YOS9607, YOS9608 및 Z5567U 중의 지방산 조성 및 전체 지질 함량

균주 (델타-5 불포화화효소 유전자의 번호)	시료	% TFA													기타
		16:0	16:1	18:0	18:1	18:2	AL A	ED A	DGL A	AR A	EtrA	20:4 (5,11, 14,17)	ETA	EPA	
Z5567U (4 개의 유전자)	A	2.2	1.0	2.9	9.8	15.5	0.5	6.9	5.3	0.7	1.4	0.5	3.7	42.2	7.3
	B	2.2	1.0	3.0	9.9	15.6	0.6	6.9	5.2	0.7	1.4	0.5	3.7	42.0	7.3
	C	2.2	1.0	2.9	10.0	15.4	0.5	6.9	5.3	0.7	1.3	0.5	3.6	42.2	7.4
	평균	2.2	1.0	2.9	9.9	15.5	0.5	6.9	5.2	0.7	1.4	0.5	3.7	42.1	7.4
YOS9601(3 개의 델타-5 유전자)	A	4.0	1.5	4.2	14.0	17.3	0.5	7.6	5.5	0.5	1.5	0.3	4.0	33.0	5.9
	B	3.8	1.4	4.0	13.7	17.4	0.6	7.6	5.6	0.5	1.5	0.4	4.1	33.4	6.0
	C	3.8	1.4	4.1	13.7	17.4	0.6	7.5	5.6	0.5	1.6	0.4	4.1	33.4	5.9
	평균	3.9	1.5	4.1	13.8	17.4	0.6	7.5	5.6	0.5	1.5	0.4	4.1	33.2	6.0
YOS9602(1 개의 델타-5 유전자)	A	3.5	1.4	3.7	12.4	16.9	0.6	7.9	8.2	0.4	1.6	0.3	7.1	30.3	5.9
	B	3.4	1.4	3.7	12.2	16.9	0.5	8.0	8.3	0.5	1.6	0.2	7.1	30.4	5.9
	C	3.5	1.4	3.8	12.4	16.8	0.5	7.9	8.2	0.5	1.5	0.3	7.0	30.2	6.0
	평균	3.5	1.4	3.7	12.3	16.8	0.5	7.9	8.3	0.5	1.6	0.3	7.1	30.3	5.9
YOS9607(0 개의 델타-5 유전자)	A	3.0	1.7	2.6	11.4	14.3	0.5	8.5	15.2	0.2	1.6	0.4	36.6	0.0	4.1
	B	3.0	1.7	2.6	11.4	14.3	0.5	8.5	15.2	0.1	1.6	0.4	36.7	0.0	4.1
	C	3.0	1.7	2.6	11.4	14.3	0.5	8.6	15.2	0.1	1.6	0.4	36.6	0.0	4.1
	평균	3.0	1.7	2.6	11.4	14.3	0.5	8.5	15.2	0.1	1.6	0.4	36.6	0.0	4.1
YOS9608(0 개의 델타-5 유전자)	A	3.1	1.7	2.6	11.5	14.0	0.5	8.6	15.4	0.2	1.6	0.4	36.3	0.0	4.2
	B	3.1	1.7	2.6	11.5	14.1	0.5	8.7	15.3	0.2	1.6	0.4	36.4	0.0	4.1
	C	3.1	1.7	2.6	11.4	14.1	0.5	8.6	15.3	0.1	1.6	0.4	36.6	0.0	4.0
	평균	3.1	1.7	2.6	11.5	14.1	0.5	8.6	15.3	0.2	1.6	0.4	36.4	0.0	4.1

[0521]

[0522]

플라스크 실험의 데이터에 의해, 계놈 내에 3개의 델타-5 불포화화효소 유전자를 포함하는 균주 YOS9601(Leu-, Ura-)이 약 33 EPA % TFA를 생성하는 한편, 그의 계놈 내에 오직 1개의 델타-5 불포화화효소 유전자를 포함하는 균주 YOS9602(Leu- Ura-)는 약 30 EPA % TFA를 생성하는 것으로 입증되었다. 대조적으로, 균주 YOS9607(Ura-) 및 YOS9608(Ura-)은 어떠한 EPA도 생성할 수 없었으나, 약 36%의 ETA를 생성하였다. 균주 YOS9607(Ura-) 및 YOS9608(Ura-) 내의 델타-5 불포화화효소 활성의 결여는 상기 전체 지방산 분석에 의해 입증되었으며; 또한, PCR 분석에 의해, 델타-5 불포화화효소 유전자를 암호화하는 임의의 DNA 서열의 결여가 확인되었다. Z5567U(Ura-) 균주는 균주 Z5567과 비교하여, 플라스크 검정법에서 더 적은 EPA % DCW를 생성하였다.

[0523]

실시예 9

[0524]

적어도 약 49 TFA % DCW와 함께, 적어도 약 50 EPA % TFA를 생성하는 야로위아 리폴리티카 균주 Y8174, Y8184 및 Y8187의 생성

[0525]

본 실시예에는 플라스크 검정법에서 49 초과 TFA % DCW와 함께 약 50 초과 EPA % TFA를 생성할 수 있는 야로위아 리폴리티카 균주 YOS9607 및 YOS9608(실시예 8) 유래의 균주 Y8174, Y8184 및 Y8187의 구축이 기재되어 있다. 3개의 이중 돌연변이 델타-5 불포화화효소를 균주 YOS9607(Ura-) 및 YOS9608(Ura-)의 염색체 내로 통합 시킴으로써 이들 균주를 생성하여, 형질전환 균주가 EPA를 생성하는 능력을 복원시켰다.

[0526]

더욱 구체적으로, HPGG[서열 번호 181] 및 HDASH [서열 번호 183] 모티프(본 명세서에 참고로 포함되는 미국 가출원 제61/428,277호에 기재된 바와 같은[대리인 문서 번호 CL5267USPRV(출원일: 2010년 12월 30일)]) 둘 모두 내의 돌연변이를 포함하는 이중 돌연변이 델타-5 불포화화효소를 하기로 이루어진 균으로부터 선택하였다: EgD5S-36s157g(서열 번호 110; 실시예 11L), EaD5S-35a158g(서열 번호 112; 실시예 11M), EgD5M(즉, EgD5R*-34g158g; 서열 번호 106; 실시예 11I 및 11K) 및 EgD5M1(즉, EgD5R*-34g158g347s; 서열 번호 108; 실시예 11J 및 11K).

[0527]

3개의 키메라 돌연변이 델타-5 불포화화효소 유전자(즉, FBAIN::EgD5S-36s157g::Pex20, YAT1::EaD5S-35a158g::Oct 및 EXP1::EgD5M(EgD5R-34g158g)::Pex16)가 균주 YOS9607 및 YOS9608의 염색체 C의 1685392 내지 1687267의 영역 내로 통합되도록 구축물 pZR5AU-555(도 13a; 서열 번호 113)를 생성하여, 이에 의해 EPA의 생성을 가능하게 하였다.

[0528]

pZR5AU- 555 플라스미드에는 하기의 성분이 포함되어 있다:

표 19

플라스미드 pZR5AU-555(서열 번호 113)의 설명

서열 번호 113 내의 RE 부위 및 뉴클레오타드	단편 및 키메라 유전자 성분의 설명
<i>AscI/BsiWI</i> (7713-6820)	야로위아 염색체 C의 1685392 내지 1686281의 890 bp DNA 단편(도면에서 "R5-5' 영역"으로 표시)
<i>PacI/AscI</i> (11396-10436)	야로위아 염색체 C의 1686300 내지 1687260의 967 bp DNA 단편(도면에서 "R5-3' 영역"으로 표시)
<i>PmeI/Clal</i> (2476-1)	하기를 포함하는 YAT1::EaD5S-35a158g::Oct ● YAT1: 야로위아 리플리터카 YAT1 프로모터(도면에서 "YAT"로 표시; 미국 특허 출원 공개 제 2010-0068789-A1 호); ● EaD5S-35a158g: 유글레나 아나바에나 유래의 돌연변이 HaGG[서열 번호 428] 및 HDgSH[서열 번호 429] 모티프를 포함하는 합성 돌연변이 델타-5 불포화화효소(서열 번호 111); ● OCT: 야로위아 OCT 유전자의 OCT 종결자 서열(진뱅크 수탁 번호 X69988)
<i>EcoRII/BsiWI</i> (4127-6820)	하기를 포함하는 FBAIN::EgD5S-36s157g::Pex20: ● FBAIN: 야로위아 리플리터카 FBAIN 프로모터(도면에서 "FBA1+Intron"으로 표시; 미국 특허 제 7,202,356 호); ● EgD5S-36s157g: 유글레나 그라실리스 유래의 돌연변이 HPG[서열 번호 427] 및 HDgSH[서열 번호 429] 모티프를 포함하는 합성 돌연변이 델타-5 불포화화효소(서열 번호 109)(도면에서 "EgD5S"로 표시, 돌연변이 HPG 및 HDgSH 모티프가 나타나 있음); ● Pex20: 야로위아 Pex20 유전자로부터의 Pex20 종결자 서열(진뱅크 수탁 번호 AF054613)
<i>PacI/Clal</i> (11396-1)	하기를 포함하는 EXP1::EgD5M(EgD5R-34g158g)::Pex16: ● EXP1: 야로위아 리플리터카 수송 단백질(EXP1) 프로모터(도면에서 "EXP"로 표시; WO 2006/052870 호); ● EgD5M(EgD5R*-34g158g): 유글레나 그라실리스 유래의 돌연변이 HgGG[서열 번호 425] 및 HDAgH[서열 번호 432] 모티프를 포함하는 합성 돌연변이 델타-5 불포화화효소(서열 번호 105)(도면에서 "EgD5M"으로 표시, 돌연변이 HgGG 및 HDAgH 모티프가 나타나 있음); ● Pex16: 야로위아 Pex16 유전자로부터의 Pex16 종결자 서열(진뱅크 수탁 번호 U75433)
<i>PmeI/EcoRI</i> (2476-4127)	야로위아 Ura3 유전자(진뱅크 수탁 번호 AJ306421)

[0529]

[0530]

구축물 pZR5AU-555M(도 13b; 서열 번호 114)은 키메라 EXP1::EgD5M1(EgD5-34g158g347s)::Pex16 유전자를 pZR5AU-555의 키메라 EXP1::EgD5M(EgD5R-34g158g)::Pex16 유전자 대신에 사용하였던 것을 제외하고, pZR5AU-555와 동일하였다(다시 말하면, 여기서, EgD5-34g158g347s는 서열 번호 107에 기재되어 있다).

[0531]

pZR5AU-555 및 pZR5AU-555M 플라스미드를 따로 *AscI*로 분해한 다음, 일반 방법에 따라 개별적으로 균주 YOS9607 및 YOS9608의 형질전환을 위해 사용하였다. 형질전환 세포를 MM 플레이트 상에 플레이팅하고, 30℃에서 5일 동안 유지시켰다. 단일의 콜로니를 MM 플레이트 상으로 다시 스트리킹한 다음, 30℃에서 액체 MM으로 접종하고 250 rpm/분으로 2일 동안 셰이킹하였다. 세포를 원심분리에 의해 수집하고, HGM에 재현탁화시킨 다음, 250 rpm/분에서 5일 동안 셰이킹하였다. 세포를 일반 방법에 따라 지방산 분석으로 처리하였다.

[0532]

균주 YOS9607로의 96개의 pZR5AU-555m 형질전환체의 GC 분석에 의해, TFA의 50.8%의 EPA를 생성하였던 하나의 균주(즉, #86)를 동정하였으며; 이러한 균주는 균주 Z8184로 명명하였다. 유사하게, 균주 YOS9608으로의 96개의 pZR5AU-555 형질전환체의 스크리닝에 의해, TFA의 51.3%의 EPA를 생성하였던 하나의 균주(즉, #68)를 동정하였으며; 이러한 균주는 균주 Z8174로 명명하였다. 그리고, 균주 YOS9608으로의 96개 pZR5AU-555m 형질전환체의 GC 분석에 의해, TFA의 51.2%의 EPA를 생성하였던 하나의 균주(즉, #56)를 동정하였으며; 이러한 균주는 균주 Z8187로 명명하였다.

[0533]

이들 새로운 EPA 균주의 지방산 조성 및 오일 함량은 2벌의 플라스크 검정법을 행함으로써 결정하였다. 표 20에는 총 DCW, TFA % DCW, TFA의 중량 백분율["% TFA"]로서의 각 지방산의 농도 및 EPA % DCW가 요약되어 있다. 지방산은 표 7(상기)에서와 같이 확인되었으며, 20:4(5,11,14,17)는 주니페론산을 지칭한다. 각 시료 중의 모든 지방산의 합은 총 100이다.

[0534]

따라서, 모든 3개의 균주는 49 초과 TFA % DCW와 함께, 50 초과 EPA % TFA를 생성할 수 있었다.

표 20

야로위아 균주 Z8174, Z8184 및 Z8187의 지방산 조성 및 전체 지질 함량

균주	DCW(g/L)	TFA % DCW	% TFA										20:4 (5,11, 14,17)	ET A	EPA	EPA DCW	%
			16:0	16:1	18:0	18:1	18:2	AL A	ED A	DGL A	AR A	EtrA					
Z8174	4.4	49.7	2.1	0.8	2.2	5.8	14.8	0.7	3.6	6.2	0.8	0.6	0.5	4.1	50.9	25.3	
Z8184	4.5	51.6	2.1	0.8	2.1	5.8	14.9	0.7	3.7	6.6	0.7	0.5	0.2	4.1	50.8	26.2	
Z8187	4.3	50.8	2.1	0.8	2.1	5.6	14.8	0.7	3.9	6.6	0.8	0.6	0.3	4.2	50.7	25.7	

[0535]

[0536]

실시예 10

[0537]

리놀레산으로부터 에이코사다이엔산으로의 전환 효율이 향상된 돌연변이 델타-9 연장효소

[0538]

본 명세서에서 실시예 10A, 10B, 10C, 10D, 10E, 10F, 10G, 10H, 10I 및 10J로 나누어서 기재된 본 실시예는 서열 번호 1에 기재된 아미노산 서열을 포함하는 돌연변이 델타-9 연장효소 폴리펩티드의 설명을 지지하기 위한 실험 데이터를 기재하고 있으며, 여기서, 서열 번호 1은 적어도 하나의 아미노산 돌연변이에 의해 서열 번호 3과 상이하며, 상기 돌연변이(들)는 하기로 이루어진 군으로부터 선택된다: (i) L35F 돌연변이; (ii) L35M 돌연변이; (iii) L35G 돌연변이; (iv) L35G 돌연변이, 및 S9A, S9D, S9G, S9I, S9K, S9Q, Q12K, A21D, A21T, A21V, V32F, Y84C, Q107E, L108G, G127L, W132T, M143N, M143W, L161T, L161Y, W168G, I179M, I179R, C236N, Q244N, A254W 및 A254Y로 이루어진 군으로부터 선택되는 적어도 하나의 다른 돌연변이; (v) L35G, A21V, L108G 및 I179R 돌연변이; (vi) L35G, W132T 및 I179 돌연변이; (vii) L35G, S9D, Y84C 및 I179R 돌연변이; (viii) L35G, Y84C, I179R 및 Q244N 돌연변이; (ix) L35G, A21V, W132T, I179R 및 Q244N 돌연변이; (x) K58R 및 I257T 돌연변이; (xi) D98G 돌연변이; (xii) L130M 및 V243A 돌연변이; 및, (xiii) K58R, L35F, L35G, L35M, S9A, S9D, S9G, S9I, S9K, S9Q, Q12K, A21D, A21T, A21V, V32F, Y84C, D98G, Q107E, L108G, G127L, L130M, W132T, M143N, M143W, L161T, L161Y, W168G, I179M, I179R, C236N, V243A, Q244N, A254W, A254Y 및 I257T로 이루어진 군으로부터 선택되는 적어도 2개의 돌연변이를 포함하는 임의의 조합. 실시예 10A, 10B, 10C, 10D, 10E, 10F, 10G, 10H, 10I 및 10J는 또한 본 명세서에 전문이 참고로 포함되는 미국 가특허 출원 제61/377248호[대리인 문서 번호 CL4783USPRV(출원일: 2010년 8월 26일)]에도 기재되어 있다.

[0539]

실시예 10A: 야로위아 리폴리티카에서의 발현을 위해 코돈-최적화된 합성 델타-9 연장효소 유전자(유글레나 그라실리스로부터 유래) ["EgD9eS"]를 포함하는 야로위아 리폴리티카 발현 벡터 pZuFmEgD9ES의 구축

[0540]

키메라 FBAINm::EgD9eS::Pex20 유전자를 포함하는 야로위아 리폴리티카 벡터 pZuFmEgD9ES(도 14; 서열 번호 115)(여기서, EgD9eS는 유글레나 그라실리스 유래이며, 야로위아에서의 발현을 위해 코돈-최적화된 합성 델타-9 연장효소이다)의 구축이 본 명세서에서 참고로 포함되는 미국 특허 제7,645,604호의 실시예 8에 기재되어 있다. EgD9eS의 뉴클레오티드 서열(서열 번호 2)은 야생형 유글레나 그라실리스 델타-9 연장효소의 뉴클레오티드 서열("EgD9e"; 서열 번호 31)과 상이한데, 이는 번역 개시 부위의 변형에 더하여, 777 bp 암호화 영역 중 117 bp가 변형되고(15.1%) 106개 코돈이 최적화되었기 때문이다(40.9%)(그렇지만, 코돈-최적화된 유전자에 의해 암호화되는 단백질 서열[즉, 서열 번호 3]은 야생형 단백질 서열[즉, 서열 번호 32]과 동일하다).

[0541]

실시예 10B: 델타-9 연장효소 전환 효율이 증가된 돌연변이 델타-9 연장효소를 포함하는 야로위아 리폴리티카 형질전환체를 분석하기 위한 일반 방법

[0542]

본 실시예에는 오류 유발 중합효소 연쇄 반응["ePCR"] 라이브러리(실시예 10C), 부위-포화 라이브러리(실시예 10E), 슬로노맥스(등록 상표) 라이브러리(실시예 10G) 또는 조합 라이브러리(실시예 10I)(하기 기재)에서 생성된 다양한 돌연변이된 EgD9eS 유전자 또는 비-돌연변이 EgD9eS 유전자(서열 번호 2)("대조군" 또는 "야생형" 중 어느 하나로 지칭) 중 어느 하나를 발현하는 야로위아 리폴리티카 균주 Y2224(야생형 야로위아 균주 ATCC #20362의 Ura3 유전자의 자가 돌연변이로부터의 FOA 내성 돌연변이체[국제 출원 공개 WO 2008/073367호의 실시예 7에 기재된 분리])의 pZuFmEgD9ES 형질전환체 내의 지질 프로파일을 분석하기 위한 일반 방법이 기재되어 있다.

[0543]

돌연변이 라이브러리의 에스케리키아 콜라이 및 야로위아 리폴리티카로의 형질전환

[0544]

각각의 돌연변이 라이브러리로부터의 플라스미드를 전기천공법에 의해 에스케리키아 콜라이 탐 10 일렉트로-컴피턴트 세포(카달로그 번호 C404052, 미국 캘리포니아주 칼스배드 소재의 인비트로젠)로 형질전환시켰다. 형질전환된 세포를 100 mg/L의 앰피실린이 있는 루리아-베르타니["LB"] 아가 플레이트로 스프레딩하고, 37°C 인큐베

이터에서 하룻밤 성장시켰다. 플라스미드 DNA를 제조처의 프로토콜에 따라 쿼이프렙(등록 상표) 스핀 미니프렙 키트(미국 캘리포니아주 발렌시아 소재의 쿼아젠 인코포레이티드)를 사용하여 에스케리키아 콜라이 형질전환체로부터 추출하였다.

[0545] 이어서, DNA 분자를 일반 방법에 기재된 바와 같이 야로위아 리폴리티카 균주 Y2224로 형질전환시키고, 형질전환체를 MM 플레이트 상에서 선택하였다. 30℃에서의 2일의 성장 후에, MM 플레이트 상에서 선택한 형질전환체를 픽킹하고 새로운 MM 플레이트 상으로 다시 스트리킹하였다.

[0546] 퀵 스크린 플레이트 검정법(Quick Screen Plate Assay)

[0547] 퀵 스크린 "플레이트 검정법"을 각각의 돌연변이 라이브러리의 예비 기능 분석을 위해 사용하였다. 이러한 플레이트 검정법을 위하여, 상기 다시 스트리킹한 MM 플레이트로부터의 야로위아 리폴리티카 형질전환체를 배지 플레이트로부터 직접 분석하였다. FAME를 트라이메틸설포늄 하이드록시드["TMSH"]를 사용하여 제조하였다.

[0548] TMSH를 메탄올 중의 산화은과의 반응에 의한 하이드록시드 용액으로의 전환 후에 트라이메틸설포늄 아이오다이드["TMSI"]로부터 제조하였다. 구체적으로, 4.4 g의 TMSI를 100 mL의 MeOH 중에서 혼합하고, 50℃ 수조에서 1 시간 동안 인큐베이션되게 한 다음; 5 g의 Ag₂O를 용액에 첨가하고, 실온에서 4시간 동안 교반하였다. 최종 용액을 사용 전에 여과하였다. TMSH는 염기-촉매작용된 O-아실 지질(즉, TAG)의 에스테르 교환 및 유리 지방산의 에스테르화를 야기한다(문헌[A. H. El-Hamdy & W. W. Christie, *J. of Chromatography*, 630:438-441 (1993)]).

[0549] 1 μl 루프(loop)를 사용하여, 세포를 다시 스트리킹한 MM 플레이트로부터 직접 취하고, 0.35 mL 인서트가 있는 기체 크로마토그램["GC"] 바이얼에서 50 μl의 TMSH에 현탁화시켰다. 이어서, 헵탄(150 μl)을 바이얼 인서트에 첨가하고, 바이얼을 캡핑한 다음, 진탕시키면서 실온에서 20분 동안 인큐베이션시켰다. 이후에, 헵탄층으로부터의 1 μl를 FAME의 GC 분석을 위하여 오메가왁스(Omegawax) 320 융합 실리카 모세관 컬럼(미국 펜실베이니아주 벨폰테 소재의 수펠코 인코포레이티드(Supelco Inc.))이 장착된 휴렛-팩카드 7890 GC로 주입하였다. 정제 시간을 상용의 표준물질(표준물질 #461, 미국 미네소타주 엘리시안 소재의 누-체크 프렙 인코포레이티드(Nu-Chek Prep, Inc.))로부터의 메틸 에스테르에 대한 것과 비교하였다.

[0550] EgD9eS 돌연변이체를 포함하는 세포로부터 수득한 FAME 프로파일을 비-돌연변이 EgD9eS 대조군의 것과 비교하였다. 이러한 일차 스크린의 결과는 이차 확인 검정법으로 처리할 돌연변이체의 선택을 위한 기준으로 삼았다. 확인 검정법을 위한 돌연변이체를 선택하기 위해 사용되는 기준은 지질 프로파일, 특히 모든 적분된 피크의 합에 비한 백분율로서 해당하는 FAME의 GC 피크 영역으로 계산 시의 EDA의 농도["EDA % TFA"] 및/또는 LA로부터 EDA로의 전환 효율에 기초한 것이다. LA로부터의 EDA로의 전환 효율["전환%"]을 하기의 식에 따라 각 형질전환체에 대하여 계산하였다: $([\text{생성물}]/[\text{기질} + \text{생성물}]) \times 100$, 여기서, 생성물은 EDA % TFA였으며, 기질은 TFA의 영역 백분율로써의 LA의 농도["LA % TFA"]였다.

[0551] "확인" 검정법

[0552] 퀵 스크린 "플레이트 검정법"을 통하여 대조군에 비한 델타-9 연장효소 활성의 향상을 나타내는 EgD9eS 돌연변이체를 이후의 확인 검정법을 위해 선택하였다.

[0553] EgD9eS 돌연변이체를 포함하는 야로위아 리폴리티카 형질전환체를 먼저 MM 플레이트 상에 다시 스트리킹한 다음, 각각의 개별 형질전환체를 30℃에서 3 mL의 액체 MM의 3배의 배양물에 접종하고, 2일 동안 250 rpm/분으로 셰이킹하였다. 세포를 원심분리에 의해 수집하고, 지질을 추출하고, FAME를 소듐 메톡시드를 사용한 지질 추출물의 에스테르 교환에 의해 제조하고(문헌[Roughan, G., and Nishida I., *Arch. Biochem. Biophys.*, 276(1):38-46 (1990)]), 이후에 플레이트 검정법(상기)에 대해 기재된 바와 같이 GC에 의해 분석하였다.

[0554] 향상된 델타-9 연장효소 활성의 확인 후에, 각 돌연변이 pZUFmEgD9ES 플라스미드를 자이모프렙(Zymoprep)(상표명) 이스트 플라스미드 미니프렙 II 키트(카달로그 번호 D2004, 미국 캘리포니아주 오렌지 소재의 자이모 리서치)를 제조처에 의해 권고되는 바와 같이 사용하여 야로위아 리폴리티카 균주 Y2224 형질전환체로부터 회수하였다.

[0555] 복원된 플라스미드를 벡터 및 삽입물-특이적 프라이머를 사용하여 염료 종결(dye terminator) 기술(미국 특허 제5,366,860호; EP 272,007호)을 사용하여 ABI 오토매틱 시퀀서(ABI Automatic sequencer)에서 시퀀싱하였다. 서열의 비교를 당업자에게 잘 알려져 있는 표준 도구를 사용하여 달성하였다.

[0556] 실시예 10C: 2개의 EgD9eS 오류 유발 PCR 라이브러리의 구축

- [0557] 본 실시예는 2개의 델타-9 연장효소 오류 유발 중합효소 연쇄 반응["ePCR"] 라이브러리의 합성에 대해 기재하고 있다. 먼저 주형 내의 무작위 돌연변이를 포함하는 메가프라이머의 모음의 생성에 이어서, pZuFmEgD9ES로 점 돌연변이를 만들기 위한 이들 메가프라이머의 이용을 필요로 하는 2-단계 방법에서 2개의 ePCR 라이브러리를 생성하였다. 구축물 pZuFmEgD9ES(서열 번호 115)(실시예 10A)를 제1 10A ePCR 라이브러리를 위한 DNA 주형으로 사용하였다. 제2 ePCR 라이브러리는 DNA 주형으로서 제1 ePCR 라이브러리의 스크리닝으로부터의 히트(hit)를 사용하였다.
- [0558] 무작위 돌연변이유발 키트를 사용한 메가프라이머의 생성
- [0559] 진모르프 II 랜덤 뮤타지네시스 키트(GeneMorph II Random Mutagenesis Kit)(카달로그 번호 200550, 미국 캘리포니아주 라 졸라 소재의 스트라타진(Stratagene))를 사용하여 표적 단백질 내에 무작위 아미노산 치환을 생성하였다. 이는 A 및 T 대 G 및 C에서 동등한 돌연변이율을 갖는 덜 편향된 돌연변이 스펙트럼을 생성하도록 2개의 상이한 중합효소의 조합을 포함하는 신규한 오류 유발 PCR 효소 배합된 포플레이션을 사용하여 오류-유발 PCR 동안 표적 유전자 내로 돌연변이를 도입함으로써 가능하다. 높은 생성물 수율을 위해 최적화된 단일의 세트의 완충액 조건을 사용하여 kB당 1 내지 16개 돌연변이의 돌연변이율이 달성될 수 있는 것으로 언급되었다. 바람직한 돌연변이율은 단순히 반응 내의 처음의 주형 DNA의 양 및/또는 수행되는 증폭 사이클의 횟수를 다양하게 함으로써 조절할 수 있다.
- [0560] 상기 키트를 사용하여 제조처에 의해 권고되는 프로토콜을 사용하여 EgD9eS "메가프라이머"를 생성하였다. 이들 메가프라이머는 약 930 bp 길이였으며, EgD9eS를 암호화하는 777 bp(서열 번호 2)를 포함한다. 반응 혼합물은 제1 ePCR 라이브러리를 위하여 μl 당 16 ng의 DNA 주형 또는 제2 라이브러리를 위하여 μl 당 2.0 ng의 DNA 주형 중 어느 하나를 포함한다. 이는 또한 반응 완충액, dNTP(0.8 mM), 프라이머 pZUFm_6980_012208f(서열 번호 116)(2 μM), 프라이머 pZUFm_40_012208r(서열 번호 117)(2 μM) 및 뮤타자임(Mutazyme)(등록 상표) II DNA 중합효소(0.25 U/ μl)를 포함한다. PCR 반응을 마스터사이클러 구매 장비(Mastecycler gradient equipment)(미국 뉴욕주 웨스트버리 소재의 브린크만 인스트루먼츠, 인코포레이티드(Brinkmann Instruments, Inc.))에서 얇은 벽의 200 μl 튜브에서 수행하였다. PCR 증폭을 하기의 조건을 사용하여 수행하였다: 2분 동안 95°C에 이어서, 30초 동안 95°C에서의 변성, 30초 동안 55°C에서의 어닐링 및 90초 동안 72°C에서의 연장의 30 사이클. 72°C에서 4분 동안의 최종 연장 사이클에 이어서 4°C에서의 반응 종결을 수행하였다.
- [0561] PCR 생성물을 제조처에 의해 권고되는 바와 같이, DNA 클린 앤드 컨센트레이터(Clean & Concentrator)(상표명)-5 키트(카달로그 번호 D4003, 미국 캘리포니아주 오렌지 소재의 자이모 리서치)를 사용하여 정제하였다. 정제된 이중-가닥 PCR 생성물을 "메가프라이머"로 사용하였으며, 각각은 EgD9eS 내의 다양한 돌연변이를 함유한다.
- [0562] EgD9eS의 ePCR 돌연변이 유전자를 생성하기 위한 표준 클로닝 방법
- [0563] 제1 ePCR 라이브러리를 위하여, "메가프라이머"를 *NcoI* 및 *NotI* 제한 효소로 분해하였다. 그 다음, 겔 정제된 *NcoI/NotI* 유전자 단편을 실온에서 5시간 동안의 라이게이션 반응을 통하여 T4 DNA 리가아제(미국 위스콘신주 매디슨 소재의 프로메가)를 사용하여, 겔 정제된 *NcoI/NotI* pZUFmEgD9ES 벡터(서열 번호 115)로 직접 라이게이션시켰다.
- [0564] EgD9eS의 ePCR 돌연변이 유전자를 생성하기 위한 부위-지정 돌연변이유발
- [0565] 제2 ePCR 라이브러리를 생성하기 위하여, 상술된 "메가프라이머"를 "메가프라이머" 내의 EgD9eS 돌연변이를 pZuFmEgD9ES(도 14; 서열 번호 115)로 도입하도록 설계된 반응에서 사용하여, 비-돌연변이 EgD9eS 유전자를 다양한 돌연변이 EgD9eS 유전자로 대체하였다. 이는 퀵체인지(등록 상표) II XL 부위 지정 돌연변이유발 키트(카달로그 번호 200524, 미국 캘리포니아주 라 졸라 소재의 스트라타진)를 사용하여 달성하였다.
- [0566] 둘 모두의 플라스미드 가닥의 돌연변이 유발 프라이머-지시 복제를 위한 하이-피델리티(high-fidelity) *PfuUltra* DNA 중합효소를 사용하여 이중 가닥 벡터 내의 대상 삽입물 내에서 점 돌연변이를 생성하고, 아미노산을 대체하고, 단일/다수의 인접 아미노산을 결실시키거나 삽입하기 위하여, 퀵체인지(등록 상표) II 부위-지정 돌연변이유발 키트를 사용하였다. 키트는 특수 벡터, 독특한 제한 부위 또는 다수의 형질전환을 필요로 하지 않으며, 사실상 임의의 이중-가닥 플라스미드 내의 부위-특이적 돌연변이를 가능하게 한다. 기본적인 절차는 2개의 합성 올리고뉴클레오타이드 프라이머를 사용하며, 둘 모두는 목적하는 돌연변이를 함유하며, 벡터의 반대 가닥에 상보적이며, 이는 프라이머 대체 없이 하이-피델리티 DNA 중합효소에 의해 온도 사이클링 동안 연장된다. 올리고뉴클레오타이드 프라이머의 연장에 의해, 스테거드 닉(staggered nick)을 함유하는 돌연변이된 플라스미드를 생성하였으며, 이어서 이를 *Dpn I* 엔도뉴클레아제(endonuclease)로 처리하였다. 이러한 제한 효소는 메틸화

된 DNA 및 반-메틸화된 DNA에 대하여 특이적이며, 이에 의해, 모 DNA 주형의 분해 및 돌연변이를 함유하는 합성 DNA에 대한 선택이 가능하게 된다. 이어서, 목적 돌연변이를 함유하는 Nick 있는 벡터 DNA를 에스케리키아 콜라이 숙주에서 형질전환시키고 증식시켰다.

[0567] 그러나, 본 방법에서, 다양한 돌연변이 EgD9eS 유전자를 포함하는 이중 가닥 메가프라이머를 통상적인 합성 올리고뉴클레오타이드 프라이머 대신에 사용하였다. 구체적으로, 5.0 μ l의 키트-제공되는 10x 반응 완충액, 1.0 μ l의 50 ng/ μ l pZUFmEgD9ES 주형(서열 번호 115), 42 μ l의 메가프라이머, 1.0 μ l의 40 mM 키트-제공된 dNTP 믹스 및 1.0 μ l의 키트-제공된 Pfu-Ultra DNA 중합효소를 포함하는 50 μ l의 반응물을 제조하였다. 이러한 반응 혼합물을 얇은 벽 200 μ l-용량 PCR 튜브에 두고, 하기의 조건을 사용하여 PCR 증폭으로 처리하였다: 30초 동안 95°C에 이어서, 30초 동안 95°C에서의 변성, 1분 동안 55°C에서의 어닐링 및 6분 동안 68°C에서의 연장의 25 사이클. 8분 동안 68°C에서의 최종 연장 사이클을 수행하고, 4°C에서의 반응 종결로 이어졌다.

[0568] 키트-제공되는 DpnI 제한 효소(1.0 μ l)를 완성된 부위-지정 돌연변이유발 반응 혼합물에 직접 첨가하고, 효소 분해를 37°C에서 1시간 동안 수행하여, DNA 주형을 제거하였다. 분해된 생성물을 DNA 클리닝 키트(자이모 리서치)를 사용하여 정제하고, 용리시켜, pZUFmEgD9ES 벡터 백본에 포함되는 다양한 돌연변이 EgD9eS 유전자를 포함하는 10 μ l의 정제된 DNA를 생성하였다.

[0569] 실시예 10D: 델타-9 연장효소 전환 효율이 향상된 ePCR EgD9eS 라이브러리 돌연변이체의 동정

[0570] 본 실시예에는 1) 야생형 단백질 EgD9eS(서열 번호 3)과 비교하여 LA로부터 EDA로의 델타-9 연장효소 전환 효율이 향상된 EgD9eS ePCR 라이브러리 돌연변이체의 동정; 및 2) 이들 EgD9eS ePCR 라이브러리 돌연변이체의 서열 분석이 기재되어 있다.

[0571] EgD9eS ePCR 돌연변이체의 동정

[0572] 실시예 10C에서 제조된 ePCR 유전자 라이브러리 돌연변이체를 에스케리키아 콜라이 탐 10 일렉트로-컴피턴트 세포로 형질전환시키고, 정제하고, 이후에 실시예 10B에 기재된 바와 같이 야로위아 리폴리티카 균주 Y2224로 형질전환시켰다. 1,724개의 야로위아 형질전환체의 지방산 프로파일을 실시예 10B의 쿼 스크린 "플레이트 검정법"을 사용하여 스크리닝하였다. 대부분의 이들 돌연변이체는 대조군과 비교하여 감소된 활성을 나타내었다. 그러나, 5개의 돌연변이체는 실시예 10B의 확인 검정법에 기초하여, 대조군과 비교하여 향상된 델타-9 연장효소 활성을 나타내는 것으로 확인되었다.

[0573] 2개의 독립적인 확인 검정법으로부터의 데이터는 표 21 및 표 22에 제공되어 있으며, 개별 pZUFmEgD9ES 대조군 형질전환체의 FAME 프로파일을 ePCR 돌연변이체의 프로파일과 비교하였다. 더욱 구체적으로, 각각의 균주에 대한 모든 적분된 피크의 합에 비한 백분율["% TFA"]로서의 해당하는 FAME의 GC 피크 영역으로부터 계산된 각각의 지방산의 농도 및 LA로부터 EDA로의 전환%(실시예 10B에 기재된 바와 같이 결정)는 하기의 표 21 및 표 22에 나타나 있으며, 평균은 회색으로 강조표시하였고, "평균"으로 나타내었다. 지방산을 16:0(팔미트산), 16:1(팔미톨레산), 18:0(스테아르산), 18:1(올레산), LA 및 EDA로 동정하였다. EgD9eS 대조군에 대한 각각의 돌연변이체의 성능의 비교는 각각의 돌연변이체가 분석되는 특정 확인 검정법에서만 이루어져야 한다(즉, 검정법 #1과 검정법 #2 간의 비교는 이루어질 수 없다).

표 21

확인 검정법 #1: EgD9eS 또는 그의 ePCR 라이브러리 돌연변이 변이체를 발현하는 형질전환체
야로위아 리폴리티카 균주 Y2224에서의 지질 조성

균주	반복검정 번호	% TFA						LA로부터 EDA로의 전환%
		16:0	16:1	18:0	18:1	LA	EDA	
EgD9eS 대조군-1	1	12.8	12.9	3.0	46.6	17.3	3.0	14.8
	2	12.9	12.6	3.0	45.5	17.2	3.1	15.2
	3	12.6	12.5	2.9	47.0	17.1	3.1	15.3
EgD9eS 대조군-2	1	14.0	12.4	3.6	45.6	16.2	3.2	16.3
	2	12.2	12.5	2.6	47.4	17.3	3.1	15.1
	3	13.2	13.1	3.1	45.4	17.2	3.1	15.3
EgD9eS 대조군-3	1	13.3	13.2	3.1	45.0	17.2	3.2	15.7
	2	12.8	12.6	2.9	46.5	17.3	3.1	15.0
	3	13.2	13.0	3.1	45.3	17.2	3.2	15.7
EgD9eS 대조군-4	1	13.7	11.9	3.7	46.2	15.8	3.4	17.6
	2	12.6	13.0	2.7	45.5	17.9	3.2	15.3
	3	12.7	12.9	2.9	45.6	17.6	3.2	15.5
EgD9eS 대조군-5	1	12.9	12.6	3.0	45.7	17.7	3.1	14.9
	2	12.1	12.1	2.7	47.9	17.3	3.1	15.2
대조군 평균	--	12.9	12.7	3.0	46.1	17.2	3.1	15.5
1.2ep-8	1	11.8	12.6	2.4	47.7	17.1	3.6	17.3
	2	12.1	12.9	2.5	47.0	16.9	3.7	17.9
	3	12.7	12.8	2.9	45.9	16.9	3.7	18.0
	평균	12.2	12.8	2.6	46.9	17.0	3.7	17.8
1.9ep-63	1	12.5	12.9	2.7	46.1	17.5	3.3	15.9
	2	12.6	12.7	2.8	46.2	17.3	3.3	16.0
	3	13.0	12.6	3.2	45.7	16.9	3.4	16.8
	평균	12.7	12.7	2.9	46.0	17.2	3.3	16.3
1.4ep-161	1	13.7	12.3	3.5	45.8	16.2	3.3	17.0
	2	12.4	12.7	2.9	46.9	16.8	3.2	16.0
	3	12.5	12.4	3.0	46.8	16.9	3.3	16.3
	평균	12.9	12.5	3.1	46.5	16.6	3.3	16.4

[0574]

표 22

확인 검정법 #2: EgD9eS 또는 그의 ePCR 라이브러리 돌연변이 변이체를 발현하는 형질전환체
야로위아 리폴리티카 균주 Y2224에서의 지질 조성

균주	반복검정 번호	% TFA						LA로부터 EDA로의 전환%
		16:0	16:1	18:0	18:1	LA	EDA	
EgD9eS 대조군-2	1	12.0	12.1	3.0	50.2	14.0	2.9	16.9
	2	12.0	11.6	3.1	50.4	14.0	2.7	16.0
	3	11.8	12.1	3.0	51.1	14.4	2.8	16.3
EgD9eS 대조군-5	1	11.8	12.1	3.2	50.6	13.7	3.2	18.9
	2	11.9	12.2	3.2	51.0	13.9	3.3	19.2
	3	11.8	11.8	3.2	51.8	13.9	2.9	17.1
EgD9eS 대조군-6	1	11.8	11.7	3.2	51.1	14.0	2.8	16.6
	2	11.8	11.9	3.3	51.0	14.5	2.8	16.4
	3	11.6	12.2	2.8	51.3	14.9	2.8	15.8
EgD9eS 대조군-7	1	11.9	11.8	3.4	51.1	14.3	2.8	16.2
	2	11.8	12.0	3.2	51.1	14.2	2.8	16.6
	3	12.0	12.0	3.2	50.8	14.1	2.8	16.5
대조군 평균	--	11.9	11.9	3.2	51.0	14.1	2.9	16.9
2.1ep-94	1	11.7	11.1	2.8	50.0	14.9	3.5	19.2
	2	10.8	11.8	2.0	50.5	15.8	3.9	19.9
	3	11.1	11.5	2.0	51.2	15.3	3.9	20.2
	평균	11.2	11.5	2.3	50.5	15.3	3.8	19.8
2.1ep-95	1	11.8	10.9	2.7	50.4	15.1	3.5	18.9
	2	11.8	11.0	2.6	50.6	15.5	3.5	18.6
	3	12.0	11.0	3.2	50.1	15.1	3.5	18.7
	평균	11.9	11.0	2.8	50.4	15.2	3.5	18.8

[0575]

[0576]

확인 검정법 #1에서의 상술된 데이터의 요약에 있어서, 비-돌연변이 코돈-최적화된 EgD9eS 유전자를 포함하는 pZuFmEgD9ES로 형질전환된 야로위아 리폴리티카 균주 Y2224의 클론이 평균 3.1 EDA % TFA를 생성하였으며, 여기서, 이들 5개의 클론에서의 LA로부터 EDA로의 평균 전환 효율["전환%"]은 약 15.5%인 것으로 결정되었다. 대조적으로, 돌연변이 균주 1.2ep-8에 대한 LA로부터 EDA로의 평균 전환율은 17.8%(또는 대조군과 비교해서는 115%)였으며, 돌연변이 균주 1.9ep-63에 대한 평균 전환율은 16.3%(또는 대조군과 비교해서는 105%)였으며; 돌연변이 균주 1.4ep-161에 대한 평균 전환율은 16.4%(또는 대조군과 비교해서는 106%)였다.

[0577] 확인 검정법 #2에서, pZuFmEgD9ES로 형질전환된 야로위아 리폴리티카 균주 Y2224의 클론을 2.9 EDA % TFA를 생성하였으며, 이들 4개의 균주에서의 LA로부터 EDA로의 평균 전환%는 약 16.9%인 것으로 결정되었다. 돌연변이 균주 2.1ep-94에 대한 LA로부터 EDA로의 평균 전환%는 19.8%(또는 대조군과 비교해서는 117%)였고; 돌연변이 균주 2.1ep-95에 대한 평균 전환%는 18.8%(또는 대조군과 비교해서는 111%)였다.

[0578] 따라서, 이들 실험에 의해, EgD9eS ePCR 돌연변이체 1.2ep-8, 1.9ep-63, 1.4ep-161, 2.1ep-94 및 2.1ep-95에 의해 향상된 델타-9 연장효소 전환 효율이 나타나는 것이 확인되었다.

[0579] EgD9eS ePCR 돌연변이체의 서열

[0580] 돌연변이체 1.2ep-8, 1.9ep-63, 1.4ep-161, 2.1ep-94 및 2.1ep-95로부터 복원된 플라스미드를 DNA 시퀀싱에 의해 특성화하였으며, 분석에 의해, 표 23에 나타난 바와 같이, 돌연변이 EgD9eS 유전자 내에서 다양한 뉴클레오티드 치환이 드러났고, 아미노산 치환이 발현되었다. 아미노산 치환을 나타내는 명칭은 각 돌연변이 EgD9eS 유전자 및 돌연변이 EgD9eS 유전자를 포함하는 각 돌연변이 pZuFmEgD9ES 플라스미드에 제공하였다. 열거된 각각의 치환(즉, L35G)에 대하여, 처음의 문자는 비-돌연변이 EgD9eS(즉, 서열 번호 3) 내의 아미노산에 해당하며, 두번째 문자는 돌연변이체 내의 동일한 위치에서 관찰되는 아미노산에 해당하며, 다시 말하면, L35G는 위치 35에서의 EgD9eS의 Leu로부터 EgD9eS 돌연변이체의 Gly로의 변화를 나타낸다.

표 23

시퀀싱된 EgD9eS ePCR 라이브러리 돌연변이체의 요약

ePCR 돌연변이체	뉴클레오티드 치환	생성된 아미노산 치환(침묵 돌연변이)	돌연변이 유전자에 대한 명칭	돌연변이 플라스미드에 대한 명칭
1.2ep-8	C103T 및 A654G	L35F 및 (G218G)	"EgD9eS-L35F"(서열 번호 118)	pZuFmEgD9ES-L35F(서열 번호 120)
1.9ep-63	A173G, T234G, G402A 및 T770C	K58R, (S78S), (Q134Q) 및 I257T	"EgD9eS-K58R/I257T"(서열 번호 121)	pZuFmEgD9eS-K58R/I257T(서열 번호 123)
1.4ep-161	C388A, C450T 및 T728C	L130M, (N150N) 및 V243A	"EgD9eS-L130M/V243A ₁ "(서열 번호 124)	pZuFmEgD9ES-L130M/V243A ₁ (서열 번호 126)
2.1ep-95	A293G	D98G	"EgD9eS-D98G"(서열 번호 127)	pZuFmEgD9ES-D98G(서열 번호 129)
2.1ep-94	C21T, C388A, C450T 및 T728C	(I71), L130M, (N150N) 및 V243A	"EgD9eS-L130M/V243A ₂ "(서열 번호 130)	pZuFmEgD9ES-L130M/V243A ₂ (서열 번호 132)

[0581]

[0582] 따라서, 예를 들어, 돌연변이 1.2ep-8로부터 복원된 플라스미드는 2개의 뉴클레오티드 치환(즉, C103T 및 A654G)을 포함한다. 이들 2개의 뉴클레오티드 치환은 하나의 발현되는 아미노산 치환(즉, L35F) 및 하나의 침묵 아미노산 돌연변이에 해당한다(즉, G218G; GGA 및 GGG 둘 모두가 Gly를 암호화하기 때문에, 아미노산은 A654G 뉴클레오티드 치환의 결과로서 돌연변이 단백질에서 변경되지 않는다). 아미노산 변화 L35F를 야기하는 C103T 및 A654G 돌연변이를 포함하는 플라스미드는 pZuFmEgD9ES-L35F(서열 번호 120)로 명명하는 한편, 거기에서 돌연변이 델타-9 연장효소의 뉴클레오티드 서열은 "EgD9eS-L35F"(서열 번호 118)로 명명하고, 이는 서열 번호 119에 기재된 단백질 서열을 갖는다.

[0583] 실시예 10E: 2-부위 포화 EgD9eS 유전자 라이브러리의 구축

[0584] 본 실시예는 EgD9eS(서열 번호 3) 내의 아미노산 위치 35 및 107을 표적화함으로써 만들어지는 부위-포화[SS] 라이브러리의 합성을 기재하고 있다. 위치 35를 표적화하는 이유는 실시예 10D의 결과에 기초한 것이며, 위치 107을 표적화하는 이유는 후술된다. 먼저 주형 내의 표적화된 돌연변이를 포함하는 메가프라이머의 생성에 있어서, pZuFmEgD9ES로 점 돌연변이를 만들기 위한 이들 메가프라이머의 이용을 필요로 하는 2-단계 방법에서 SS 라이브러리를 생성하였다.

[0585] EgD9eS의 위치 107을 표적화하는 이유

[0586] 먼저, 하기 표 24에 기재된 바와 같은 17개 지방산 연장효소의 아미노산 서열을 클러스터링을 정렬 방법을 사

용하여 정렬하였다.

표 24

보존 패턴 분석으로 처리되는 지방산 연장효소

연장효소 약어	유기체	참조	서열 번호
Ci_elo	키오나 인테스티날리스	진뱅크 수탁 번호 AAV67802	133
Om_elo	온코린처스 마이키스	진뱅크 수탁 번호 AAV67803	134
Mp_elo1	마르잔티아 폴리모르파	진뱅크 수탁 번호 AAT85662	135
Pp_elo1	페스코미트렐라 파렌스	진뱅크 수탁 번호 AAL84174	136
Mp_d5e	마르잔티아 폴리모르파	진뱅크 수탁 번호 BAE71130	137
Ot_elo1	오스트레오코커스 타우리	진뱅크 수탁 번호 AAV67797	138
Pav_elo2	파블로바 중 CCMP459	진뱅크 수탁 번호 AAV33630	139
Ps_elo2	파블로바 살리나	진뱅크 수탁 번호 AAY15135	140
Ot_elo2	오스트레오코커스 타우리	진뱅크 수탁 번호 AAV67798	141
Ea_d9e	유글레나 아나바에나	미국 특허 제 7,794,701 호	34
Eg_d9e	유글레나 그라실리스	미국 특허 제 7,645,604 호	32
E398_d9e	유타램피렐라 중 CCMP389	미국 특허 제 7,645,604 호	38
Ig_d9e	아이소크라이시스 갈바나	PCT 공개 WO 2002/077213 호, WO 2005/083093 호, WO 2005/012316 호 및 WO 2004/057001 호; 진뱅크 수탁 번호 AAL37626	42
Tp_elo2	탈라시오시라 슈도나나	진뱅크 수탁 번호 AAV67800	142
Tp_elo1	탈라시오시라 슈도나나	진뱅크 수탁 번호 AAV67799	143
Ma_d6e	모르티에렐라 알피나	진뱅크 수탁 번호 AAF70417	144
Th_elo2	트라우스토키트리움 중 FJN-10	진뱅크 수탁 번호 ABC18314	145

[0587]

[0588]

톱슨 등에 의해 기재된 클러스털 더블유 정렬 방법(문헌[Nucleic Acids Res. 22:4673-4680 (1994)])을 디폴트 파라미터(즉, 단백질 중량 매트릭스 = 고넷 250, 갭 개방 페널티(gap opening penalty) = 10, 갭 연장 페널티 = 0.2 및 전체 정렬 알고리즘)와 함께 클러스털더블유 패키지(버전 1.83)를 사용하여 수행하였다. 정렬의 결과는 도 15(도 15a, 15b, 15c, 15d, 15e, 15f, 15g 및 15h를 포함)에 나타나 있다. "Trace_1", "Trace_2", "Trace_3" 및 "Trace_4"는 하기에 정의되는 바와 같은 기능성 그룹 I, 그룹 II, 그룹 III 및 그룹 IV에 대한 각각의 열의 공통서열을 나타내며, 다시 말하면, Trace 1은 Ci_elo, Om_elo, Mp_elo1, Pp_elo1, Mp_d5e 및 Ot_elo1을 포함하는 그룹 I에서의 단백질 공통 서열을 나타낸다. 각 열의 공통서열은 하기에 정의되어 있다. 구체적으로, 열이 완전히 보존되었다면, 공통서열은 대문자로 나타낸 보존된 아미노산으로 표시하였다. 열이 물리-화학적 특성의 면에서 보존되었다면, 공통서열은 소문자로 표시하였으며, 여기서, "k"는 아미노산 D 및 E (음으로 하전)를 나타내며, "q"는 아미노산 H, K 및 R(양으로 하전)을 나타내고, "p"는 아미노산 N 및 Q(극성)를 나타내며, "a"는 아미노산 I, L 및 V(지방족)를 나타내고, "d"는 아미노산 F, W 및 Y(방향족)를 나타내며, "h"는 아미노산 A 및 G(아주 작은)를 나타내고, "s"는 아미노산 D, E, N, Q, H, K, R, S 및 T(친수성)를 나타내며, "f"는 아미노산 I, L, V, F, W, Y, C 및 M(소수성)을 나타낸다. 열이 보존되지 않았다면, 공통서열은 대문자 "X"로 나타내었다.

[0589]

이웃-연결수(neighbor-joining tree)를 클러스털 더블유 정렬로부터 생성하였다. 이웃-연결수 토폴로지에 기초하여, 17개의 서열을 4개의 그룹으로 분배하였으며, 이들 그룹은 상이한 기질 특이성의 기능성 그룹에 해당하는 것으로 가정된다: 그룹 I은 Ci_elo, Om_elo, Mp_elo1, Pp_elo1, Mp_d5e 및 Ot_elo1을 포함하며; 그룹 II는 Pav_elo2, Ps_elo2 및 Ot_elo2를 포함하고; 그룹 III은 Ea_d9e, Eg_d9e, E398_d9e 및 Ig_d9e를 포함하며; 그룹 IV는 Tp_elo2, Tp_elo1, Ma_d6e 및 Th_elo2를 포함한다.

[0590]

도 15의 정렬 및 이웃-연결수의 그룹화를 고려하여, 하기의 결론을 도출하였다. 먼저, 몇몇 위치는 그룹 I, II, III 및 IV 내의 모든 17개 서열에 걸쳐 완전히 보존된다. 이들 위치는 연장효소의 촉매적 활성에 필수적인 것 같다고 여겨졌으며, 이에 따라 돌연변이에 대한 표적으로 배제하였다. 몇몇 위치는 그룹 I, II, III 및 IV 내의 오직 일부의 서열에서만 보존되었다(즉, 완전하게 보존되지 않았다). 이들 위치는 그룹 I, II, III 또는 IV의 기능성 그룹 내에서 연장효소에 의해 나타나는 기질 특이성에 중요할 것 같다고 여겨졌다. 일부 위치가 그룹 III(모든 4개의 공지되어 있는 델타-9 연장효소 포함)에서 상대적으로 보존되었으나, 변이도 또한 존재하였으며; EgD9e의 넘버링에 기초하여, 아미노산 위치 22, 47, 54, 101, 107, 111, 115, 161, 182, 192 및 242를 참조한다. 이들 위치는 델타-9 연장효소의 활성에 중요할 것 같다고 여겨졌으며, Ea_d9e(서열 번호 34),

Eg_d9e(서열 번호 32), E398_d9e(서열 번호 38) 및 Ig_d9e(서열 번호 42)의 기질 특이성의 차이를 조절하는 것으로 가정되었다.

[0591] EgD9eS 내의 막관통["TM"] 도메인의 분석을 TMHMM 프로그램("단백질 내의 막관통 헬릭스의 예측(Prediction of transmembrane helices in proteins"; TMHMM Server v. 2.0, 덴마크 DK-2800 링비 소재의 덴마크 공과 대학 바이오센트럼(BioCentrum)-DTU 생물 서열 분석 센터(Center for Biological Sequence Analysis))을 사용하여 수행하였다. 예측에 의해, N- 및 C-말단 둘 모두가 세포질 측에 위치하는 6개의 막-스패닝 헬릭스(아미노산 잔기 32-51, 66-88, 114-136, 156-175, 188-206, 221-243에 해당)가 나타났다. Ot_elo2, Ig_elo1, Pav_elo2 및 Tp_elo2를 TMHMM 프로그램을 사용하여 유사하게 분석하는 경우, 막-스패닝 헬릭스의 개수는 4 내지 8로 달라졌다. 따라서, 이들 달라지는 예측을 확고히 하기 위하여, 하기의 기능 정보의 부분을 사용하였다.

[0592] 1. 고도로 보존된 히스티딘-풍부 모티프 [Q/H]xxHH("His-box")는 Ig_d9e(서열 번호 42)의 최적의 효소 활성에 필수적인 것으로 관찰되었으나, 기질 특이성의 직접적인 원인인 것은 아니다(문헌[Qi et al., *FEBS Letters*, 547:137-139 (2003)]). 따라서, His-box(EgD9eS 내의 아미노산 잔기 134-138에 해당)가 활성 부위에 포함되는 것이 강력하게 시사되었으며; 이는 기질이 활성 부위에 접근할 수 있도록 폴딩된 단백질의 세포질 측 내에 또는 그 근처에 위치해야 한다.

[0593] 2. 하전된 잔기를 갖는 몇몇의 고도로 보존된 위치는 EgD9eS의 C-말단에 존재한다. 그들은 활성에 관련이 있을 것 같으며, 이에 따라, C-말단은 아마도 폴딩된 단백질의 세포질 측에 위치할 것이다.

[0594] 아미노산 잔기 114-136 및 아미노산 잔기 156-175에서 막-스패닝 헬릭스를 예측한 TMHMM 결과와 대조적으로, 상기 고려사항에 의해 잔기 114-136의 서열 영역이 막을 스패닝하지 않는 것이 나타났는데, 이는 His-box가 막의 외면에 위치할 수 없기 때문이다. C-말단이 세포질 측에 위치한다면, 156 내지 175의 예측된 TM 도메인도 또한 막을 스패닝하지 않는다. 연장효소에 대한 기질이 고도로 소수성이기 때문에, 그는 아마도 지질 이중층으로 분배될 것이다. 활성 부위(His-box 포함)는 막 표면에서 또는 그의 바로 근처에서 나타날 수 있다.

[0595] 따라서, 이들 2개의 소수성 영역(즉, 아미노산 잔기 114-136 및 아미노산 잔기 156-175에 해당)이 내막 리플릿 내에 또는 그 근처에 놓여 있어, 활성 부위가 막 근처에서 자리잡게 하는 것으로 본 명세서에서 예측되었다. EgD9eS에 대해 예측된 최종 막 토폴로지 모델은 도 16a에 나타나 있다. 구체적으로, 각각의 수직의 원통은 막-스패닝 세그먼트를 나타내는 한편, 각각의 수평의 원통은 내막 리플릿 내에 또는 그 근처에 놓여 있는 소수성 스트레치를 나타낸다. His-box 내의 보존된 Gln[Q] 및 His[H] 잔기(즉, 아미노산 잔기 134-138에 해당)는 작은 원으로 표시되어 있다. 마지막으로, "내부"는 세포질 공간에 해당하고, "외부"는 주변세포질 공간에 해당한다.

[0596] 보존 패턴 분석에 의해, 효소 활성에 영향을 미치는 것으로 예측되었던 그룹 III 델타-9 지방산 연장효소(즉, Ea_d9e[서열 번호 34], Eg_d9e[서열 번호 32], E398_d9e[서열 번호 38] 및 Ig_d9e[서열 번호 42]) 내에서 11개의 상이한 아미노산 잔기가 동정되었으나, 예측되는 토폴로지 모델로부터의 결과에 의해 후보 잔기가 추가로 제한되었다. 구체적으로, 효소 활성에 중요한 위치는 활성 부위가 놓여 있는 세포질 측에 또는 그 근처에 존재하여야 하는 것으로 추론되었다. 아미노산 잔기 47, 54 및 192는 이러한 기준을 만족시키지 못하였으며, 이에 따라, 그들이 델타-9 연장효소의 활성을 조절하는데 중요할 리가 없는 것으로 추정되었다.

[0597] 상기 이유에 기초하여, EgD9eS의 델타-9 연장효소 활성에 유의미하게 영향을 미칠 것 같은 후보 잔기는 서열 번호 3의 전장 단백질 내의 258개 잔기로부터 위치 22, 101, 107, 111, 115, 161, 182 및 242에 해당하는 오직 8개의 잔기로 줄었다. 이들 8개의 위치는 EgD9eS의 기질 전환율을 향상시키기 위한 부위-지정 돌연변이유발에 대한 표적으로 권고되었다. 하기의 실험 데이터는 위치 107을 표적화한 것이다.

[0598] 부위-포화 라이브러리의 구축을 위한 메가프라이머의 생성

[0599] 올리고뉴클레오타이드 EgD9E_102_053008f(서열 번호 146) 및 EgD9E_760_053008r(서열 번호 147)을 각각 EgD9eS(서열 번호 3)의 아미노산 잔기 35 및 107을 표적화하도록 설계하였다. 이들 올리고뉴클레오타이드의 상업적 합성 후에, 그들을 PCR 반응에서 사용하여 SS 라이브러리의 구축에 사용하기에 적절한 메가프라이머를 생성하였다. 구체적으로, 50 μ l의 반응 혼합물은 Pfu-Ultra 중합효소(스트라타진)가 제공된 5.0 μ l의 10x 반응 완충액, 1.0 μ l의 50 ng/ μ l EgD9eS(서열 번호 2), 1.0 μ l의 10 pmol/ μ l 프라이머 EgD9E_102_053008f(서열 번호 146), 1.0 μ l의 10 pmol/ μ l 프라이머 EgD9E_760_053008r(서열 번호 147), 1.0 μ l의 40 mM dNTP 믹스(미국 위스콘신주 매디슨 소재의 프로메가), 1.0 μ l의 하이 피델리티 Pfu-Ultra DNA 중합효소(스트라타진) 및 40 μ l의 물을 함유하도록 제조하였다. 혼합물을 마스터사이클러 구매 장비(미국 뉴욕주 웨스트버리 소재의 브린크만 인스트루먼트, 인코포레이티드)에서 PCR 반응을 위하여 얇은 벽의 200 μ l 튜브에 배치하였다. PCR 증폭을 하기의

조건을 사용하여 수행하였다: 30초 동안 95℃에 이어서, 30초 동안 95℃에서의 변성, 1분 동안 54℃에서의 어닐링 및 2분 동안 72℃에서의 연장의 30 사이클. 4분 동안 72℃에서의 최종 연장 사이클을 수행하고, 4℃에서의 반응 종결로 이어졌다.

- [0600] PCR 생성물을 제조처에 의해 권고되는 바와 같이 DNA 클린 앤드 컨센트레이터(상표명)-5 키트(카달로그 번호 D4003, 미국 캘리포니아주 오렌지 소재의 자이모 리서치)를 사용하여 정제하였다. 정제된 이중-가닥 PCR 생성물을 "메가프라이머"로 사용하였으며, 각각은 EgD9eS 내의 다양한 돌연변이를 함유한다.
- [0601] EgD9eS의 부위-포화 돌연변이 유전자를 생성하기 위한 부위-지정 돌연변이유발
- [0602] 이어서, 상술된 "메가프라이머"를 "메가프라이머" 내의 EgD9eS 돌연변이를 pZuFmEgD9ES(도 14; 서열 번호 115)로 도입하도록 설계된 반응에서 사용하여 비-돌연변이 EgD9eS 유전자를 다양한 돌연변이 EgD9eS 유전자로 대체하였다. 이는 실시예 10C에 기재된 바와 같이 퀵체인지(등록 상표) II XL 부위 지정 돌연변이유발 키트(카달로그 번호 200524, 미국 캘리포니아주 라 줄라 소재의 스트라타진)를 사용하여 달성하였다. 구체적으로, 부위 지정 돌연변이유발 반응의 조성 및 증폭 조건은 *DpnI* 제한효소 및 DNA 클린-업 방법과 같이, 실시예 10C에 기재된 것과 동일하였다.
- [0603] 실시예 10F: 델타-9 연장효소 전환 효율이 향상된 EgD9eS 부위-포화 라이브러리 돌연변이체의 동정
- [0604] 본 실시예에는 1) 야생형 단백질 EgD9eS(서열 번호 3)와 비교하여 LA로부터 EDA로의 델타-9 연장효소 전환 효율이 향상된 EgD9eS 돌연변이체의 동정; 및 2) 이들 EgD9eS 돌연변이체의 서열 분석이 기재되어 있다.
- [0605] EgD9eS 부위-포화 돌연변이체의 동정
- [0606] 실시예 10E에서 제조된 SS 라이브러리를 에스케리키아 콜라이 탐 10 일렉트로-컴피턴트 세포로 형질전환시키고, 정제하고, 이후에 실시예 10B에 기재된 바와 같이 야로위아 리폴리티카 균주 Y2224로 형질전환시켰다. SS 라이브러리로부터의 구축물을 갖는 510개의 야로위아 형질전환체의 지방산 프로파일을 실시예 10B의 퀵 스크린 "플레이트 검정법"을 사용하여 분석하였다. 3개의 형질전환체는 실시예 10B의 확인 검정법에 기초하여 대조군과 비교하여 향상된 델타-9 연장효소 활성을 나타내는 것으로 확인되었다.
- [0607] 확인 검정법으로부터의 데이터는 표 25에 제시되어 있으며, 개별 pZuFmEgD9ES 대조군 형질전환체의 FAME 프로파일을 SS 라이브러리 돌연변이체의 FAME 프로파일과 비교하였다. 더욱 구체적으로, 각각의 균주에 대한 TFA의 영역 백분율["% TFA"]로서의 각각의 지방산의 농도 및 LA로부터 EDA로의 전환%(실시예 10B에 기재된 바와 같이 결정)는 하기의 표 25에 나타내었으며, 평균은 회색으로 강조표시하였고, "평균"으로 나타내었다. 지방산은 실시예 10D의 약어에 기초하여 나타내었다.

표 25

확인 검정법: EgD9eS 또는 그의 SS 라이브러리 돌연변이 변이체를 발현하는 형질전환체 야로위아 리폴리티카 균주 Y2224에서의 지질 조성

균주	반복검정 번호	% TFA						LA로부터 EDA로의 전환%
		16:0	16:1	18:0	18:1	LA	EDA	
EgD9eS 대조군-1	1	12.6	11.5	5.2	47.0	15.0	3.4	18.3
	2	12.5	11.8	4.9	47.1	15.1	3.4	18.4
EgD9eS 대조군-2	1	12.7	11.5	4.2	48.5	15.4	3.4	18.0
	2	12.0	12.0	4.1	47.1	16.9	3.8	18.2
	3	12.9	11.3	4.4	48.3	15.4	3.4	18.0
EgD9eS 대조군-3	1	12.5	11.7	3.7	49.4	15.5	3.5	18.6
	2	12.4	11.6	5.1	47.8	15.0	3.4	18.6
	3	12.1	11.8	5.0	48.3	15.4	3.5	18.6
EgD9eS 대조군-4	1	12.3	11.5	5.2	47.7	15.0	3.7	19.8
	2	12.4	11.8	4.7	48.1	15.0	3.5	19.1
	3	12.7	11.7	3.8	48.7	15.1	3.7	19.8
대조군 평균	--	12.5	11.6	4.6	48.0	15.3	3.5	18.7
2.4sd-24	1	12.6	11.8	4.0	48.6	13.3	4.9	27.0
	2	12.6	11.6	4.6	48.6	13.0	4.9	27.5
	3	12.5	11.8	3.9	49.6	13.2	4.9	27.0
	평균	12.6	11.7	4.1	48.9	13.2	4.9	27.2
2.4sd-52	1	12.6	11.5	3.8	50.3	13.2	4.7	26.4
	2	12.5	11.2	4.3	49.4	13.2	4.7	26.2
	3	12.6	11.2	5.0	48.7	12.8	4.8	27.2
	평균	12.6	11.3	4.4	49.4	13.1	4.7	26.6
2.4sd-53	1	12.6	12.0	3.6	50.1	13.8	4.5	24.8
	2	12.3	12.0	3.8	49.1	14.4	4.6	24.3
	3	12.5	12.4	3.6	49.2	13.6	4.4	24.6
	평균	12.5	12.1	3.7	49.5	13.9	4.5	24.6

[0608]

[0609]

확인 검정법에서, 비-돌연변이 코돈-최적화된 EgD9eS 유전자를 포함하는 pZuFmEgD9ES로 형질전환된 야로위아 리폴리티카 균주 Y2224의 클론이 평균 3.5 EDA % TFA를 생성하였으며, 여기서, 이들 4개의 균주에서의 LA로부터 EDA로의 평균 전환 효율["전환%"]은 약 18.7%인 것으로 결정되었다. 그와 비교하여, 돌연변이 균주 2.4sd2-24에 대한 LA로부터 EDA로의 평균 전환%는 27.2%(또는 대조군과 비교해서는 145%)였으며; 돌연변이 균주 2.4sd2-52에 대한 평균 전환%는 26.6%(또는 대조군과 비교해서는 142%)였고; 돌연변이 균주 2.4sd2-53에 대한 평균 전환%는 24.6%(또는 대조군과 비교해서는 132%)였다. 따라서 이러한 검정법에 의해, 부위-포화 돌연변이체 2.4sd2-24, 2.4sd2-52 및 2.4sd2-53에 의해 향상된 델타-9 연장효소 전환 효율이 나타나는 것이 확인되었다.

[0610]

EgD9eS 부위-포화 돌연변이체의 서열

[0611]

돌연변이체 2.4sd-24, 2.4sd-52 및 2.4sd-53으로부터 복원된 플라스미드를 DNA 시퀀싱에 의해 특성화하였으며, 분석에 의해, 표 26에 나타난 바와 같이 돌연변이 EgD9eS 유전자 내에서 다양한 뉴클레오티드 치환이 드러났고, 아미노산 치환이 발현되었다. 아미노산 치환을 나타내는 명칭은 각 돌연변이 EgD9eS 유전자 및 돌연변이 EgD9eS 유전자를 포함하는 각각의 돌연변이 pZuFmEgD9ES 플라스미드에 제공하였다.

표 26

시퀀싱된 EgD9eS SS 라이브러리 돌연변이체의 요약

부위-포화 돌연변이체	뉴클레오티드 치환	생성된 아미노산 치환	돌연변이 유전자에 대한 명칭	돌연변이 플라스미드에 대한 명칭
2.4sd-24	C103G, T104G, C105G	L35G	"EgD9eS-L35G" (서열 번호 43)	pZuFmEgD9ES-L35G (서열 번호 148)
2.4sd-52	C103G, T104G, C105G	L35G		
2.4sd-53	C103A, C105G, C319G	L35M 및 Q107E	"EgD9eS- L35M/Q107E" (서열 번호 149)	pZuFmEgD9ES- L35M/Q107E (서열 번호 151)

[0612]

[0613]

당업자에게 명백할 것처럼, 본 출원인은 다양한 뉴클레오티드 서열이 유전자 코드의 축퇴성에 기초하여, 예를 들어 EgD9eS-L35G에 기재된 단백질을 암호화할 수 있음을 인식한다. 따라서, 예를 들어, 아미노산 잔기 위치 35에서 서열 번호 44에 기재된 돌연변이 단백질에서 암호화되는 Gly는 GGG(서열 번호 43에 기재된 델타-9 연장

효소 오픈 리딩 프레임["ORF"]에서와 같이), GGA(서열 번호 152에 기재된 델타-9 연장효소 ORF에서와 같이), GGC(서열 번호 153에 기재된 델타-9 연장효소 ORF에서와 같이) 및 GGT(서열 번호 154에 기재된 델타-9 연장효소 ORF에서와 같이)에 의해 암호화될 수 있다. 암호화되는 단백질에서 침묵 돌연변이를 야기하는 다양한 다른 뉴클레오티드 치환도 또한 고려되며, 이에 따라, EgD9eS-L35G를 암호화하는 본 명세서에 제공되는 뉴클레오티드 서열(서열 번호 44)은 본 개시물에 대한 제한으로 간주되어서는 안된다. 본 발명의 돌연변이 단백질을 암호화하며 델타-9 연장효소 활성을 갖는 본 명세서에 기재된 임의의 뉴클레오티드 서열 내의 유사한 변이가 고려된다.

[0614] 실시예 10G: EgD9eS-L35G 슬로노맥스(등록 상표) 라이브러리의 생성

[0615] 본 실시예에는 EgD9eS-L35G 돌연변이체(서열 번호 44; 실시예 10F) 내의 50개의 별개의 아미노산 위치를 표적화함으로써 제조되며, 모 효소와 비교하여 LA로부터 EDA로의 전환 효율이 42 내지 45% 향상된 것으로 증명된 슬로노맥스(등록 상표) 라이브러리의 합성이 기재되어 있다. 따라서, 본 실시예는 L35 돌연변이를 포함하는 EgD9eS 돌연변이체로 "스테킹"될 수 있는 추가의 유익한 돌연변이를 동정하고자 한다.

[0616] 하기에 추가로 상세히 기재되는 자동화 로봇식 플랫폼인 슬로노믹스(등록 상표)에 의해, 슬로노맥스(등록 상표) 라이브러리를 생성하였으며, 여기서, 서열 위치마다의 돌연변이체의 개수 및 그들의 비는 매우 정밀하게 조절될 수 있다. 따라서, 자동화 과정은 부위-포화 라이브러리의 생성 시에 고려되는 제한되는 잔기와 대조적으로(실시예 10E), 델타-9 연장효소 활성에서의 그들의 영향을 결정하기 위하여 실험적으로 시험할 수 있는 후보 잔기의 개수가 크게 증가될 수 있다는 점에서 이점을 제공한다.

[0617] 기능적 부위 평가를 위하여 EgD9eS 내의 50개의 별개의 잔기를 표적화하는 이유

[0618] 델타-9 연장효소를 아이소크라이시스 갈바나 ["IgD9e"](서열 번호 42; PCT 공개 WO 2002/077213호, WO 2005/083093호, WO 2005/012316호 및 WO 2004/057001호; 진뱅크 수탁 번호 AAL37626), 유트랩티엘라 종 CCMP389["E389D9e"](서열 번호 38; 미국 특허 제7,645,604호), 유글레나 그라실리스["EgD9e"](서열 번호 32; 미국 특허 제7,645,604호) 및 유글레나 아나바에나["EaD9e"](서열 번호 34; 미국 특허 제7,794,701호)로부터 동정하고 기능적으로 특성화하였다. 각각의 이들 연장효소는 LA를 EDA로 전환시킬 수 있는 것으로 보인다. EgD9e, EaD9e 및 E389D9e는 서로 60% 초과와 서열 유사성을 공유하는 한편, IgD9e는 EgD9e, EaD9e 및 E389D9e 중 임의의 것과 오직 약 35%의 서열 유사성만을 공유한다(디폴트 파라미터(즉, 단백질 중량 매트릭스 = 고넷 250, 갭 개방 페널티 = 10, 갭 연장 페널티 = 0.2 및 전체 정렬 알고리즘)를 사용하는 클러스털더블유(버전 1.83) 분석에 기초하여).

[0619] 델타-9 연장효소 전환 효율이 향상된 돌연변이체를 야기하는 위치(예컨대, D98G[실시예 10D] 및 L35G[실시예 10F])가 중등의 서열 보존성을 갖는 것으로 관찰되었다. IgD9e, EgD9e, EaD9e 및 E389D9e의 아미노산 서열 정렬을 생성하여, 벡터 엔티아이(등록 상표)의 얼라인엑스 프로그램(미국 캘리포니아주 칼스배드 소재의 인비트로젠 코포레이션)의 디폴트 파라미터를 사용하여, 다른 중등으로 보존된 잔기를 동정하였다(도 17). 모든 정렬된 서열 간에 보존된 미국 특허 제7,645,604호의 델타-9 연장효소 모티프는 공통 서열 내의 밑줄이 그어진 굵은 문자로 도면에 나타나 있다. 서열 번호 32의 EgD9e 서열 내의 굵은 문자의 잔기(서열이 서열 번호 3에 기재된 EgD9eS의 서열과 동일)는 돌연변이되어 델타-9 연장효소 활성이 향상된 돌연변이 연장효소를 야기하는 잔기를 나타낸다. 또한, 이들 돌연변이의 위치는 각각의 정렬의 열에 걸쳐 별표로 강조표시되어 있다.

[0620] 이들 중등으로 보존된 잔기가 아미노산 치환에 대한 표적으로서의 우수한 후보일 수 있어, 비-돌연변이 EgD9eS 대조군과 비교하여 활성이 향상된 제2 세대의 돌연변이 효소를 제공할 가능성이 있는 것으로 가정되었다.

[0621] 이들 4개의 상동성 효소의 서열을 비교하여, 258개 아미노산 위치 중 58개가 모든 4개의 연장효소 중에 보존되는 것으로 결정되었으며; 이에 따라, 이들 잔기를 고려대상으로부터 배제하였다. 추가로, 92개의 위치가 EgD9e, EaD9e 및 E389D9e 간에 보존된 것으로 결정되었으며; 이들 위치도 또한 고려대상으로부터 배제하였다. 마지막으로, 상동체 간에 무작위 아미노산 변화를 갖는 위치는 통상 단백질 기능에 유의미한 역할을 수행하지 않기 때문에, 이에 따라 모든 4개의 연장효소 간에 4가지 상이한 아미노산 잔기를 갖는 것으로 결정된 추가의 22개 위치를 기능 평가를 위한 표적화된 위치로서의 고려대상으로부터 배제하였다.

[0622] 서열 번호 32 내의 나머지 86개 위치(즉, 위치 1, 3, 4, 5, 9, 12, 21, 22, 27, 28, 29, 32, 35, 37, 41, 42, 45, 47, 48, 51, 52, 53, 54, 57, 58, 60, 62, 63, 66, 67, 70, 71, 73, 74, 80, 83, 84, 85, 89, 94, 98, 101, 104, 105, 107, 108, 111, 115, 127, 131, 132, 143, 149, 152, 153, 155, 156, 161, 168, 169, 179, 181, 182, 192, 196, 204, 207, 209, 210, 211, 216, 218, 222, 223, 225, 229, 236, 239, 242, 244, 245,

247, 250, 254, 257 및 258)를 기능적 부위 평가에 대한 유력한 표적으로 고려하였다. EgD9e(서열 번호 32), EaD9e(서열 번호 34) 및 E389E9e(서열 번호 38) 내의 이들 각각의 하나의 위치에서 암호화된 아미노산 잔기의 비교를 하기 표 27에 나타내었다.

표 27

기능적 부위 평가를 위한 위치

위치*	EgD9e	EaD9e	389D9e	위치*	EgD9e	EaD9e	389D9e
1	M	M	I	105	T	T	A
3	V	A	V	107	Q	Q	K
4	V	A	A	108	L	L	V
5	N	K	N	111	L	L	Y
9	S	S	A	115	L	V	L
12	Q	Q	A	127	G	D	A
21	A	A	Q	131	T	S	S
22	Q	Q	R	132	W	F	F
27	A	A	I	143	M	I	M
28	S	S	Y	149	Y	Y	V
29	H	S	S	152	R	R	S
32	V	V	L	153	N	N	G
35	L	L	F	155	A	G	S
37	I	V	I	156	V	V	I
41	I	A	I	161	L	L	F
42	L	I	L	168	W	W	F
45	T	M	T	169	I	I	V
47	G	R	G	179	I	I	M
48	P	P	E	181	L	L	F
51	P	L	D	182	K	N	N
52	K	K	S	192	S	S	A
53	G	R	G	196	I	I	T
54	Q	Q	K	204	I	I	L
57	M	L	L	207	K	K	W
58	K	K	R	209	R	R	K
60	V	L	L	210	N	N	D
62	T	T	K	211	I	V	I
63	N	A	W	216	Q	Q	K
66	L	F	L	218	G	G	P
67	L	L	F	222	F	F	L
70	I	I	V	223	G	A	A
71	Y	Y	F	225	F	I	I
73	L	F	L	229	F	W	W
74	G	G	V	236	C	L	L
80	A	A	G	239	L	L	I
83	M	L	I	242	Y	Y	F
84	Y	S	Y	244	Q	Q	K
85	T	V	T	245	T	T	S
89	M	L	Y	247	I	I	V
94	E	E	D	250	K	P	K
98	D	N	D	254	A	R	A
101	V	V	L	257	I	K	K
104	I	I	F	258	Q	E	V

*위치는 EgD9eS(서열 번호 3)의 것과 동일한 서열을 갖는 EgD9e(서열 번호 32)에 대한 정렬을 기초로 한 것이다.

[0623]

[0624]

표 27에서 상기 동정된 86개 위치 중에, 도 16a의 막 토폴로지 모델에 기초하여, 주변세포질 공간에 가장 근접한 부위는 추가의 고려대상으로부터 배제하였다(즉, 위치 45, 47, 48, 51, 52, 53, 54, 57, 58, 60, 62, 63, 66, 67, 70, 71, 73, 74, 204, 207, 209, 210, 211, 216, 218, 222, 223, 225 및 229). 회색의 굵은 문자로 강조표시된 부위(즉, EgD9eS의 위치 3, 5, 9, 12, 21, 22, 27, 28, 32, 37, 41, 42, 80, 84, 85, 94, 98, 101, 104, 105, 107, 108, 111, 115, 127, 131, 132, 143, 149, 152, 153, 156, 161, 168, 169, 179, 181, 182, 192, 196, 236, 239, 242, 244, 245, 247, 250, 254, 257 및 258)를 추가의 실험적 평가를 위해 선택하였다.

[0625]

EgD9eS-L35G의 슬로노맥스(등록 상표) 돌연변이 유전자를 생성하기 위한 슬로노믹스(등록 상표)

[0626]

슬로노믹스(등록 상표)(미국 특허 제7,695,906호)는 "한번에 하나의 코돈"의 조합 라이브러리 합성을 위한 보편적인 빌딩 블록(universal building block)으로서 일련의 이중 가닥의 DNA 트리플렛(triplet)을 사용한다(독일 푸하임 소재의 슬로닝 바이오테크놀로지(Sloning BioTechnology)). 라이브러리 생성을 위하여, 임의의 목적하는 서열 위치에서 동시에 다수의 코돈을 도입할 수 있다. 임의의 비로 최대 20개의 코돈의 전달을 선택하고 정확하게 조절하는 능력 및 기능적 편향의 부재에 의해 완전한 목적하는 돌연변이체의 세트를 함유하는 특별히 고 품질의 라이브러리가 야기된다.

- [0627] 이에 따라, 슬로노맥스(등록 상표) 유전자 라이브러리(총 50개)를 주형으로서 pZuFmEgD9ES-L35G(서열 번호 148)를 사용하여 슬로닝 바이오테크놀로지에 의해 생성하였으며, 각각의 유전자 라이브러리는 표적화된 위치(다시 말하면, EgD9eS의 위치 3, 5, 9, 12, 21, 22, 27, 28, 32, 37, 41, 42, 80, 84, 85, 94, 98, 101, 104, 105, 107, 108, 111, 115, 127, 131, 132, 143, 149, 152, 153, 156, 161, 168, 169, 179, 181, 182, 192, 196, 236, 239, 242, 244, 245, 247, 250, 254, 257 또는 258)에서 적어도 16개의 독립적이고 독특한 서열 돌연변이를 갖는다.
- [0628] 실시예 10B에 기재된 바와 같이, 모든 EgD9eS-L35G 돌연변이체를 pZuFmEgD9ES-L35G에 의해 제공되는 벡터 백본으로 클로닝하고, 이후에 야로위아 리폴리티카 균주 Y2224로 형질전환시키고, 배양하였다. 형질전환된 세포(동결된 글리세롤 스톡(stock)으로 제공) 및 DNA를 슬로닝 바이오테크놀로지로부터 획득하였다. 약간의 형질전환된 세포 및 DNA를 시퀀싱하고 확인하였다.
- [0629] 실시예 10H: 텔타-9 연장효소 전환 효율이 향상된 EgD9eS-L35G 슬로노맥스(등록 상표) 라이브러리 돌연변이체의 동정
- [0630] 본 실시예는 실시예 10F에서 동정된 변이체 단백질 EgD9eS-L35G(서열 번호 44)의 LA로부터 EDA로의 텔타-9 연장효소 전환 효율과 비교하여 전환 효율이 향상된 EgD9eS-L35G 슬로노맥스(등록 상표) 돌연변이체의 동정을 기재하고 있다.
- [0631] 슬로노맥스(등록 상표) 라이브러리로부터의 구축물을 갖는 807개의 야로위아 형질전환체의 지방산 프로파일을 실시예 10B의 "확인 검정" 방법을 사용하여 스크리닝하여, 다시-스트리킹한 새로운 MM 플레이트에서 성장한 세포를 사용하여 3 mL 액체 MM을 포함하는 3벌의 배양물을 개별 접종하였다. 807개의 돌연변이체에 더하여, pZuFmEgD9ES-L35G(서열 번호 148)를 포함하는 야로위아 균주 Y2224 형질전환체를 실험 대조군으로서 3벌로 접종하였다.
- [0632] 확인 검정법에서 선택된 돌연변이체로부터의 데이터는 표 28에 제시되어 있으며, 3개의 대표적인 EgD9eS-L35G 대조군의 FAME 프로파일을 슬로노맥스(등록 상표) 라이브러리 돌연변이체의 것과 비교하여, LA로부터 EDA로의 평균 전환%의 증가가 나타났다. 더욱 구체적으로 각각의 균주에 대한 TFA의 영역 백분율["% TFA"]로서의 각각의 지방산의 평균("평균"으로 나타냄) 농도 및 LA로부터 EDA로의 평균 전환%(실시예 10B에 기재된 바와 같이 결정)는 하기 표 28에 나타나 있다. 지방산은 실시예 10D의 약어에 기초하여 나타내었다. 각각의 균주 설명은 돌연변이 EgD9eS 유전자 각각에 존재하는 특정 아미노산 치환을 나타낸다. 따라서, 균주 EgD9eS-L35G/S9A는 서열 번호 3에 기재된 EgD9eS의 서열과 비교하여 L35G 돌연변이 및 S9A 돌연변이를 갖는 돌연변이 EgD9eS 유전자를 포함하는 돌연변이 pZuFmEgD9ES 플라스미드를 포함한다.

표 28

확인 검정법: EgD9eS-L35G 또는 그의 슬로노맥스(등록 상표) 돌연변이 변이체를 발현하는 형질전환체 야로위아 리플리타 균주 Y2224에서의 지질 조성

균주	반복검정 번호	% TFA						LA로부터 EDA로의 전환%
		16:0	16:1	18:0	18:1	18:2	20:2	
EgD9eS-L35G 대조군-1	1	13.7	12.0	2.9	40.4	19.0	4.8	20.2
	2	14.2	12.7	2.9	41.7	19.5	4.4	18.5
	3	13.8	12.2	3.0	41.1	19.1	4.9	20.4
EgD9eS-L35G 대조군-2	1	13.8	12.5	2.8	40.7	19.8	4.5	18.4
	2	14.0	12.5	2.8	41.1	19.8	3.7	15.6
	3	13.8	12.3	2.9	41.0	19.7	4.5	18.6
EgD9eS-L35G 대조군-3	1	13.9	12.8	2.7	41.7	18.5	3.9	17.5
	2	13.7	12.7	2.7	42.6	19.0	3.7	16.2
	3	14.1	12.9	2.8	41.7	19.1	4.1	17.5
대조군 평균	--	13.9	12.5	2.8	41.3	19.3	4.3	18.1
EgD9eS-L35G/S9A	3 가지의 평균	12.3	12.8	3.1	49.9	12.9	3.8	22.8
EgD9eS-L35G/S9D	3 가지의 평균	12.3	12.3	3.2	48.6	12.2	4.3	25.6
EgD9eS-L35G/S9G	3 가지의 평균	12.6	12.7	3.1	51.1	12.2	3.8	23.3
EgD9eS-L35G/S9I	3 가지의 평균	13.0	12.2	2.9	52.5	12.2	3.1	20.4
EgD9eS-L35G/S9K	3 가지의 평균	12.4	12.4	2.9	52.3	12.1	3.5	22.1
EgD9eS-L35G/S9Q	3 가지의 평균	12.5	13.1	2.8	52.0	12.4	3.1	20.1
EgD9eS-L35G/Q12K	3 가지의 평균	12.5	14.1	2.6	51.4	11.6	3.3	22.3
EgD9eS-L35G/A21D	3 가지의 평균	12.4	14.2	2.7	49.7	12.1	3.3	21.4
EgD9eS-L35G/A21T	3 가지의 평균	12.3	13.3	2.9	50.3	13.2	3.3	20.0
EgD9eS-L35G/A21V	3 가지의 평균	12.7	15.1	2.3	49.1	13.4	3.6	21.3
EgD9eS-L35G/V32F	3 가지의 평균	12.2	14.9	2.4	49.4	14.0	3.2	18.8
EgD9eS-L35G/Y84C	3 가지의 평균	11.9	13.3	2.6	51.5	12.8	4.5	26.1
EgD9eS-L35G/L108G	3 가지의 평균	13.0	13.4	3.0	48.4	14.8	3.4	18.8
EgD9eS-L35G/G127L	3 가지의 평균	9.7	11.0	1.9	36.6	24.5	5.7	18.8

[0633]

EgD9eS-L35G/W132T	3 가지의 평균	13.8	12.8	3.0	43.7	18.2	4.0	18.1
EgD9eS-L35G/M143N	3 가지의 평균	10.1	11.1	2.0	39.6	21.4	4.5	17.4
EgD9eS-L35G/M143W	3 가지의 평균	11.4	12.2	2.3	43.8	18.4	4.4	19.1
EgD9eS-L35G/L161T	3 가지의 평균	11.1	12.1	2.7	41.3	17.8	5.6	23.7
EgD9eS-L35G/L161Y	3 가지의 평균	9.9	11.8	2.6	37.4	22.3	6.1	21.5
EgD9eS-L35G/W168G	3 가지의 평균	11.5	12.3	2.5	44.0	17.6	4.7	20.8
EgD9eS-L35G/I179M	3 가지의 평균	13.8	12.5	3.0	41.5	18.5	4.3	18.8
EgD9eS-L35G/I179R	3 가지의 평균	10.2	11.9	2.2	40.5	18.4	6.3	25.5
EgD9eS-L35G/C236N	3 가지의 평균	13.3	13.4	2.8	45.5	16.9	3.8	18.5
EgD9eS-L35G/Q244N	3 가지의 평균	10.2	12.4	2.2	38.2	17.9	5.7	24.2
EgD9eS-L35G/A254W	3 가지의 평균	11.7	16.8	2.0	48.8	14.8	3.7	20.2
EgD9eS-L35G/A254Y	3 가지의 평균	13.1	16.2	2.5	48.4	12.9	3.4	21.0

[0634]

- [0635] 표 28에 제시된 반복 EgD9eS-L35G 대조군의 지방산 프로파일 및 LA로부터 EDA로의 전환%가 이전에 제시된 EgD9eS-L35G 프로파일과 다소 상이한 것은 주목할 만한 것이다. 본 실험 세트에서, EgD9eS-L35G 대조군은 이전 분석과 비교하여 "저조하였다"(즉, 상기에서 LA로부터 EDA로의 평균 전환%는 약 18.1%인 것으로 결정되었으며, 실시예 10F, 표 25에서 LA로부터 EDA로의 평균 전환%는 약 26.6% 및 27.2%인 것으로 결정되었다). 그러나, EgD9eS-L35G가 있는 형질전환체는 4.3 EDA % TFA(상기의 평균)를 생성하였으며, 이는 EgD9eS가 있는 형질전환체에서 생성되는 것보다 유의미하게 더 컸다(즉, 3.1 EDA % TFA[실시예 10D, 표 21], 2.9 EDA % TFA[실시예 10D, 표 22] 및 3.5 EDA % TFA[실시예 10F, 표 25]). 이러한 이유로, EgD9eS-L35G의 기능 분석을 반복한 이전의 실험으로부터의 성능(데이터 미도시)을 표 28에 제시되는 EgD9eS 부위-평가 라이브러리로부터의 돌연변이체의 비교를 위한 기준으로서 본 실험에서 EgD9eS-L35G 성능에 더하여 사용하였다.
- [0636] 표 28에 제시된 26개의 선택된 연장효소 변이체 중에, 표 28의 대조군 데이터와 비교하여 유사하거나 유의미하게 향상된 델타-9 연장효소 전환 활성을 나타내는 11개를 동정하였다(굵은 문자로 강조표시). 이들 돌연변이체에는 EgD9eS-L35G/S9D(141% 증가), EgD9eS-L35G/A21V(118% 증가), EgD9eS-L35G/V32F(104% 증가), EgD9eS-L35G/Y84C(144% 증가), EgD9eS-L35G/L108G(104% 증가), EgD9eS-L35G/W132T(100% 증가), EgD9eS-L35G/M143N(96% 증가), EgD9eS-L35G/L161T(131% 증가), EgD9eS-L35G/I179R(141% 증가), EgD9eS-L35G/C236N(102% 증가) 및 EgD9eS-L35G/Q244N(134% 증가)이 포함되며, EgD9eS에 대한 델타-9 연장효소 전환 활성은 괄호 안에 나타나 있다. 따라서, 최대 44%의 LA로부터 EDA로의 전환 효율의 향상이 입증되었다.
- [0637] 실시예 10I: EgD9eS-L35G/ S9D/ A21V/ V32F/ Y84C/ L108G/ W132T/ M143N/ L161T/ I179R/ C236N/ Q244N 조합 라이브러리의 생성
- [0638] 본 실시예는 돌연변이 EgD9eS 조합 라이브러리의 합성을 기재하고 있으며, 여기서, 상기 실시예 10H에서 동정된 유익한 돌연변이의 다양한 조합(즉, L35G, S9D, A21V, V32F, Y84C, L108G, W132T, M143N, L161T, I179R, C236N 및 Q244N)을 L35G 돌연변이를 포함하는 EgD9eS 돌연변이체로 함께 "스태킹"하였다.
- [0639] 조합 라이브러리의 구축을 위한 합성 프라이머의 생성
- [0640] 11쌍의 프라이머를 서열 번호 155 내지 176에 기재된 바와 같이 상업적으로 합성하였다(하기 표 29 참조). 하기의 돌연변이 중 하나를 EgD9eS-L35G 유전자로 도입하도록 각 프라이머 쌍을 설계하였다: S9D, A21V, V32F, Y84C, L108G, W132T, M143N, L161T, I179R, C236N 및 Q244N.
- [0641] 프라이머를 T4 폴리뉴클레오티드 키나제["PNK"] (카탈로그 번호 70031Z, 유에스비 코퍼레이션)를 사용하여 37°C에서 60분 동안 인산화시킨 다음, 65°C에서 20분 동안 불활성화시켰다. 각 20 μ l 인산화 반응 혼합물에는 2.0 μ l의 10x T4 PNK 완충액, 15.0 μ l의 프라이머 DNA(약 7 μ m), 0.6 μ l의 100 mM ATP, 0.4 μ l의 T4 PNK 및 2.0 μ l의 물이 포함되어 있다.
- [0642] EgD9eS-L35G의 조합 돌연변이 유전자를 생성하기 위한 다수의 돌연변이 부위 돌연변이유발
- [0643] 체인지(Change)-IT(상표명) 다중 돌연변이 부위 지정 돌연변이유발 키트(카탈로그 번호 78480, 미국 오하이오주 클리블랜드 소재의 유에스비 코퍼레이션)를 사용하여 일련의 6개의 반응에서 S9D, A21V, V32F, Y84C, L108G, W132T, M143N, L161T, I179R, C236N 및 Q244N 돌연변이를 EgD9eS-L35G 내로 도입하고, 각 반응(최종 반응을 제외하고)에서 프라이머 세트 "A"의 정방향 프라이머 및 역방향 프라이머, 및 프라이머 세트 "B"의 정방향 프라이머 및 역방향 프라이머의 포함에 기초하여 2개의 새로운 돌연변이를 도입하였다(표 29). 일련의 반응에서의 초기의 주형이 EgD9eS-L35G였지만, 체인지-IT(상표명) 반응 1의 생성물은 체인지-IT(상표명) 반응 2의 주형으로 사용되는 등등이었다.

표 29

체인지-IT(상표명) 다수 돌연변이 부위 직접 돌연변이유발 반응의 요약						
체인지-IT(상표명) 반응 번호	주형	생성물	프라이머 세트 "A"		프라이머 세트 "B"	
			정방향 프라이머	역방향 프라이머	정방향 프라이머	역방향 프라이머
#1	EgD9eS-L35G	체인지-IT(상표명) 반응 번호 1	Eg_9D_122709f (서열 번호 155)	Eg_84C_122709r (서열 번호 156)	Eg_84C_122709f (서열 번호 157)	Eg_9D_122709r (서열 번호 158)
#2	체인지-IT(상표명) 반응 번호 1	체인지-IT(상표명) 반응 번호 2	Eg_161T_122709f (서열 번호 159)	Eg_179R_122709r (서열 번호 160)	Eg_179R_122709f (서열 번호 161)	Eg_161T_122709r (서열 번호 162)
#3	체인지-IT(상표명) 반응 번호 2	체인지-IT(상표명) 반응 번호 3	Eg_244N_122709f (서열 번호 163)	Eg_21V_010710r (서열 번호 164)	Eg_21V_010710f (서열 번호 165)	Eg_244N_122709r (서열 번호 166)
#4	체인지-IT(상표명) 반응 번호 3	체인지-IT(상표명) 반응 번호 4	Eg_32F_010710f (서열 번호 167)	Eg_108G_010710r (서열 번호 168)	Eg_108G_010710f (서열 번호 169)	Eg_32F_010710r (서열 번호 170)
#5	체인지-IT(상표명) 반응 번호 4	체인지-IT(상표명) 반응 번호 5	Eg_132T_010710f (서열 번호 171)	Eg_143N_010710r (서열 번호 172)	Eg_143N_010710f (서열 번호 173)	Eg_132T_010710r (서열 번호 174)
#6	체인지-IT(상표명) 반응 번호 5	체인지-IT(상표명) 반응 번호 6	Eg_236N_010710f (서열 번호 175)	Eg_236N_010710r (서열 번호 176)	--	--

[0644]

[0645]

더욱 구체적으로, 2개의 25 μ l PCR 반응 혼합물을 준비하였으며, 각각의 것은 2.5 μ l의 10x 체인지-IT(상표명) 완충액, 2.5 μ l의 인산화된 정방향 프라이머, 2.5 μ l의 인산화된 역방향 프라이머, 1.0 μ l의 주형(50 ng/ μ l), 15.5 μ l의 뉴클레아제-미함유 물 및 1.0 μ l의 체인지-IT(상표명) 피델리티택 효소를 포함한다. 제1 반응에서는 프라이머 세트 "A"로부터의 프라이머를 사용하는 한편, 제2 반응에서는 프라이머 세트 "B" 프라이머를 사용하였다. PCR 증폭을 하기의 조건을 사용하여 수행하였다: 2분 동안 95℃에 이어서, 30초 동안 95℃에서의 변성, 30초 동안 55℃에서의 어닐링 및 25분 동안 68℃에서의 연장/라이게이션의 30 사이클. 30분 동안 68℃에서 최종 연장/라이게이션 사이클을 수행하고, 4℃에서의 반응 종결로 이어졌다.

[0646]

증폭 후에, *DpnI* 효소를 첨가함으로써 주형을 제거하고, 37℃에서 3시간 동안 분해를 수행하였다. PCR DNA를 사용하여 에스케리키아 콜라이 탐 10 일렉트로-컴피턴트 세포(카탈로그 번호 C404052, 미국 캘리포니아주 칼스배드 소재의 인비트로젠)를 형질전환시켰다. 형질전환된 세포를 100 mg/L 앰피실린이 있는 LB 아가 플레이트로 스프레딩하고, 37℃ 인큐베이터에서 하룻밤 성장시켰다. 플라스미드 DNA를 제조처의 프로토콜에 따라 쿼이프랩(등록 상표) 스핀 미니프랩 키트(미국 캘리포니아주 발렌시아 소재의 쿼아젠 인코포레이티드)를 사용하여 형질전환체 에스케리키아 콜라이 세포로부터 추출하였다. 그 다음, 정제된 DNA를 다음의 체인지-IT(상표명) 반응에서 주형으로 사용하였다. 11개의 돌연변이 중 마지막을 원래의 EgD9eS-L35G 주형에 도입하는 6번째 반응 후에, 상술된 바와 같이 DNA를 형질전환체 에스케리키아 콜라이 세포로부터 정제하였다. 이어서, DNA를 야로위아 리폴리티카 균주 Y2224(상기 실시예 10B)로 형질전환시켰다.

[0647]

실시예 10J: 델타-9 연장효소 전환 효율이 향상된 EgD9eS-L35G/ S9D/ A21V/ V32F/ Y84C/ L108G/ W132T/ M143N/ L161T/ I179R/ C236N/ Q244N 조합 라이브러리 돌연변이체의 동정

[0648]

본 실시예에는 1) 야생형 단백질 EgD9eS(서열 번호 3)와 비교하여 LA로부터 EDA로의 델타-9 연장효소 전환 효율이 향상된 EgD9eS-L35G/ S9D/ A21V/ V32F/ Y84C/ L108G/ W132T/ M143N/ L161T/ I179R/ C236N/ Q244N 조합 라이브러리 돌연변이체의 동정; 2) 이들 EgD9eS 돌연변이체의 서열 분석; 및 3) 시퀀싱된 EgD9eS 돌연변이체의 재생성에 의한 델타-9 연장효소 전환 효율의 향상 확인이 기재되어 있다.

[0649]

EgD9eS-L35G/ S9D/ A21V/ V32F/ Y84C/ L108G/ W132T/ M143N/ L161T/ I179R/ C236N/ Q244N 조합 라이브러리 돌연변이체의 동정

[0650]

조합 라이브러리(실시예 10I)로부터의 구축물을 갖는 2,388개의 야로위아 형질전환체의 지방산 프로파일을 실시예 10B의 퀵 스크린 "플레이트 검정법"을 사용하여 스크리닝하였다. 이들 대부분의 돌연변이체는 야생형 대조군, EgD9eS(서열 번호 3)와 비교하여 감소된 LA로부터 EDA로의 전환을 나타내었다. 그러나, 5개의 돌연변이체는 실시예 10B의 확인 검정법에 기초하여, 대조군과 비교하여 향상된 델타-9 연장효소 활성을 나타내는 것으로 확인되었다.

[0651]

돌연변이 EgD9eS 유전자의 DNA 서열을 콜로니 PCR을 사용하여 결정하였다. 약술하면, 소량의 효모 세포를 멸균 피펫 팁을 사용하여, 스트리킹한 새로운 플레이트로부터 시료 추출하고, 세포를 20 μ l의 분자 등급의 물에서 현탁화시켰다. 세포 현탁액(2 μ l)을 제조처(타카라 바이오테크놀로지 컴퍼니, 리미티드(Takara Biotechnology Co., LTD.))의 권고에 따라 제조한 타카라(TaKaRa) Ex Taq PCR 믹스로 옮겼다. 콜로니 PCR을 위해 사용한 프

라이머는 정방향 프라이머 FBAIN-F(서열 번호 366) 및 역방향 프라이머 Y1026(서열 번호 367)이었다. 썬열 사이클러(thermal cycler) 프로그램에는 5분 동안 94℃에서의 주형의 초기 변성에 이은, 30초 동안 94℃에서의 변성, 30초 동안 56℃에서의 어닐링 및 3분 동안 72℃에서의 연장의 40 사이클이 포함되었다. 6분 동안 72℃에서의 최종 연장을 수행하였다.

[0652] PCR 생성물을 프라이머 FBAIN-F(서열 번호 366) 및 Y1026(서열 번호 367)을 사용하여 시퀀싱하였다. DNA 서열 데이터의 분석에 의해, 돌연변이 EgD9eS 유전자 내에서 뉴클레오티드 치환이 드러났으며, 아미노산 치환이 발견되었다. 아미노산 치환을 나타내는 명칭을 표 30에 나타낸 바와 같이, 돌연변이 EgD9eS 유전자 및 돌연변이 EgD9eS 유전자를 포함하는 돌연변이 pZuFmEgD9ES 플라스미드에 제공하였다.

표 30

시퀀싱된 EgD9eS 조합 라이브러리 돌연변이체의 요약

조합 돌연변이체	뉴클레오티드 치환	생성된 아미노산 치환	돌연변이 유전자에 대한 명칭	돌연변이 플라스미드에 대한 명칭
EgD9EN-427	C103G, T104G 및 C105G	L35G	EgD9eS-L35G/W132T/I179R (서열 번호 368)	pZuFmEgD9ES-L35G/W132T/I179R (서열 번호 370)
	T394A, G395C, G396C	W132T		
	A535C, T536G, C537A	I179R		
EgD9EN-1043	T25G, C26A	S9D	EgD9eS-S9D/L35G/Y84C/I179R (서열 번호 371)	pZuFmEgD9ES-S9D/L35G/Y84C/I179R (서열 번호 373)
	C103G, T104G 및 C105G	L35G		
	A251G, C252T	Y84C		
	A535C, T536G, C537A	I179R		
EgD9EN-1534	C62T 및 T63G	A21V	EgD9eS-A21V/L35G/L108G/I179R (서열 번호 177)	pZuFmEgD9ES-A21V/L35G/L108G/I179R(서열 번호 179)
	C103G, T104G 및 C105G	L35G		
	C322G, T323G 및 G324T	L108G		
	A535C, T536G 및 C537A	I179R		
EgD9EN-1635	C103G, T104G 및 C105G	L35G	EgD9eS-L35G/Y84C/I179R/Q244N (서열 번호 374)	pZuFmEgD9ES-L35G/Y84C/I179R/Q244N (서열 번호 376)
	A251G, C252T	Y84C		
	A535C, T536G, C537A	I179R		
	C730A, G732C	Q244N		
EgD9EN-1734	C62T, T63G	A21V	EgD9eS-A21V/L35G/ W132T/I179R/Q244N (서열 번호 377)	pZuFmEgD9ES-A21V/L35G/W132T/I179R/Q244N (서열 번호 379)
	C103G, T104G 및 C105G	L35G		
	T394A, G395C, G396C	W132T		
	A535C, T536G, C537A	I179R		
	C730A, G732C	Q244N		

[0653]

[0654] 부위-지정 돌연변이유발을 위한 새로운 프라이머를 표 30의 아미노산 치환에 기초하여 설계하였다. 이어서, 이들 프라이머를 "메가프라이머" 내의 EgD9eS 돌연변이를 pZuFmEgD9ES(도 2; 서열 번호 115)로 도입하도록 설계된 반응에서 사용하고, 이에 의해, 비-돌연변이 EgD9eS 유전자를 표 30에 열거된 다양한 돌연변이 EgD9eS 유전자로 대체하였다. 이는 실시예 10C에 기재된 바와 같이, 콕체인지(등록 상표) II XL 부위 지정 돌연변이유발 키트 (카달로그 번호 200524, 미국 캘리포니아주 라 졸라 소재의 스트라타진)를 사용하여 달성하였다. 이들 돌연변이 유전자를 실시예 10B에 기재된 바와 같이 에스케리키아 콜라이 탐 10 일렉트로-컴피턴트 세포로 형질전환시키고, 정제하고, 시퀀싱하고, 이후에 야로위아 리폴리티카 균주 Y2224로 형질전환시켰다. 이러한 방식으로, 표 30에 나타낸 돌연변이 EgD9eS 유전자를 플라스미드 상에 재생성하였으며, 균주 Y2224 내로 다시 도입하여, EgD9eS 조합 돌연변이체에 의해 나타나는 향상된 델타-9 연장효소 전환 효율이 동정된 아미노산 치환에 기인함을 확인하였다.

[0655] 이들 확인 검정법으로부터의 데이터는 표 31에 제시되어 있으며, 개별 pZuFmEgD9ES 대조군 형질전환체의 FAME 프로파일을 조합 라이브러리의 돌연변이체와 비교하였다. 보존성 비교를 위하여, 각 균주에 대해 나타낸 데이

터는 각 균주에 대한 LA로부터 EDA로의 가장 높은 전환%를 갖는 3개의 분리주에 대한 FAME 프로파일을 나타낸다. 더욱 구체적으로, 각각의 균주에 대한 TFA의 영역 백분율["% TFA"]로서의 각각의 지방산의 농도 및 LA로부터 EDA로의 전환%(실시에 10B에 기재된 바와 같이 결정)는 하기에 나타내었으며, 평균은 회색으로 강조표시하였고, "평균"으로 나타내었다. 지방산은 실시에 10D의 약어에 기초하여 나타내었다.

표 31

확인 검정법: EgD9eS 또는 그의 조합 돌연변이 변이체를 발현하는 형질전환체 야로위아 리폴리티카 균주 Y2224에서의 지질 조성

균주	반복검정 번호	% TFA						LA로부터 EDA로의 전환%
		16:0	16:1	18:0	18:1	18:2	20:2	
EgD9eS-대조군	1	12.5	12.6	2.8	50.6	13.1	2.6	16.7
	2	12.1	13.1	2.5	52.1	13.4	2.5	15.8
	3	12.5	13.1	2.8	51.0	13.3	2.5	15.8
	평균	12.4	12.9	2.7	51.2	13.2	2.5	16.1
EgD9EN-427	1	11.6	13.2	2.6	49.7	14.2	3.2	18.1
	2	12.2	12.7	2.6	51.5	13.0	2.9	18.0
	3	11.8	12.7	2.5	52.4	13.0	2.7	17.2
	평균	11.9	12.9	2.6	51.2	13.4	2.9	17.8
EgD9EN-1043	1	11.7	13.3	2.4	52.2	13.3	2.8	17.7
	2	11.8	12.8	2.5	51.9	12.8	2.8	17.9
	3	11.8	12.6	2.5	51.6	12.8	2.6	16.9
	평균	11.8	12.9	2.4	51.9	13.0	2.8	17.5
EgD9EN-1534	1	11.4	11.8	2.4	48.4	12.5	2.6	17.2
	2	12.0	12.4	2.5	49.8	13.3	2.6	16.6
	3	12.2	12.4	2.6	50.8	13.1	2.6	16.6
	평균	11.9	12.2	2.5	49.7	13.0	2.6	16.8
EgD9EN-1635	1	11.1	12.9	2.4	51.2	13.1	3.0	18.8
	2	11.5	13.8	2.5	49.4	14.1	3.1	18.1
	3	11.9	13.5	2.6	50.1	14.1	2.9	17.1
	평균	11.5	13.4	2.5	50.2	13.8	3.0	18.0
EgD9EN-1734	1	11.5	12.7	2.4	51.7	12.4	3.3	21.2
	2	11.3	12.5	2.2	51.3	12.1	3.2	20.7
	3	11.8	12.8	2.4	52.8	13.0	2.8	18.0
	평균	11.5	12.7	2.3	51.9	12.5	3.1	20.0

[0656]

[0657]

서열 번호 2(비-돌연변이체)의 코돈-최적화된 EgD9eS 유전자를 포함하는 pZuFmEgD9ES로 형질전환된 야로위아 리폴리티카 균주 Y2224의 클론은 평균 2.5 EDA % TFA를 생성하였으며, 여기서, 이들 3개의 클론에서의 LA로부터 EDA로의 평균 전환 효율["전환%"]은 약 16.1%인 것으로 결정되었다. 대조적으로, 돌연변이 균주 EgD9EN-427에 대한 LA로부터 EDA로의 평균 전환%는 17.8%(또는 대조군과 비교해서는 110%)였다. 유사하게, 돌연변이 균주 EgD9EN-1043에 대한 LA로부터 EDA로의 평균 전환%는 17.5%(또는 대조군과 비교해서는 108%)였다. 돌연변이 균주 EgD9EN-1534에 대한 LA로부터 EDA로의 평균 전환%는 16.8%(또는 대조군과 비교해서는 104%)였으며; 돌연변이 균주 EgD9EN-1635에 대한 평균 전환%는 18.0%(또는 대조군과 비교해서는 111%)였으며; 돌연변이 균주 EgD9EN-1734에 대한 평균 전환%는 20.0%(또는 대조군과 비교해서는 123%)였다.

[0658]

따라서, 이들 실험에 의해, EgD9eS 조합 라이브러리 돌연변이체 EgD9EN-427, EgD9EN-1043, EgD9EN-1534, EgD9EN-1635 및 EgD9EN-1734에 의해 향상된 델타-9 연장효소 전환 효율이 나타나는 것으로 확인되었으며, 여기서, 향상은 4 내지 23% 범위였다.

[0659]

실시에 11

[0660]

돌연변이 HPGG(서열 번호 181) 모티프 및 HDASH(서열 번호 183) 모티프 델타-5 불포화화효소

[0661]

본 명세서에서 실시에 11A, 11B, 11C, 11D, 11E, 11F, 11G, 11H, 11I, 11J, 11K, 11L 및 11M으로 나누어 기재되어 있는 본 실시에에는 (a) 서열 번호 180[HxGx]에 기재된 아미노산 모티프 - 서열 번호 180[HxGx]은 서열 번호 181[HPGG]과 동일하지 않음 - ; 및 (b) 서열 번호 182[HxxxH]에 기재된 아미노산 모티프 - 서열 번호

182[HxxxH]는 서열 번호 183[HDASH]과 동일하지 않음 - 를 포함하는 델타-5 불포화화효소 활성을 갖는 돌연변이 폴리펩티드의 설명을 뒷받침하는 실험 데이터가 기재되어 있다.

- [0662] 더욱 구체적으로, 델타-5 불포화화효소 활성을 가지며, 서열 번호 110[EgD5S-36s157g 또는 EgD5S-HPGs_HDgSH]; 서열 번호 112[EaD5S-35a158g 또는 EaD5S-HaGG_HDgSH]; 서열 번호 106[EgD5R*-34g158g 또는 EgD5R*-HgGG_HDAgH]; 및 서열 번호 108[EgD5R*-34g158g347s 또는 EgD5R*-HgGG_HDAgH_347s]로 이루어진 군으로부터 선택되는 아미노산 서열을 갖는 돌연변이 폴리펩티드의 설명이 후술된다.
- [0663] 실시예 11A, 11C, 11D, 11E, 11F, 11G, 11H, 11I, 11J, 11K, 11L 및 11M은 본 명세서에 전문이 참고로 포함되는 미국 특허 출원 제61/428,277호[대리인 문서 번호 CL5267USPRV(출원일: 2010년 12월 30일)]에도 또한 기재되어 있다.
- [0664] 국제 출원 공개 WO 2008/073367호에 기재된 야로위아 리폴리티카 균주 Y4036U(*Leu-*, *Ura-*)를 아래의 실시예 11D, 11E, 11F, 11H, 11I 및 11K에서 숙주로 사용하였다.
- [0665] 균주 Y2224(*Ura3-*, *Ura3* 유전자의 자가 돌연변이로부터의 FOA 내성 돌연변이체), 균주 Y4001(*Leu-* 표현형과 함께 17% EDA를 생성), 균주 Y4001U1(*Leu-* 및 *Ura-*) 및 균주 Y4036(*Leu-* 표현형과 함께 18% DGLA를 생성)의 구축을 통하여 야로위아 리폴리티카 ATCC #20362로부터 균주 Y4036U를 유도하였다.
- [0666] 야생형 야로위아 리폴리티카 ATCC #20362에 대한 균주 Y4036U의 최종 유전형은 *Ura3-*, YAT1::ME3S::Pex16, EXP1::EgD9eS::Lip1, FBAINm::EgD9eS::Lip2, GPAT::EgD9e::Lip2, FBAINm::EgD8M::Pex20, EXP1::EgD8M::Pex16, GPD::FmD12::Pex20, YAT1::FmD12::OCT였다.
- [0667] 실시예 11A: 유글레나 그라실리스 델타-5 불포화화효소["EgD5"]를 위한 토폴로지 모델의 개발
- [0668] 유글레나 그라실리스로부터의 델타-5 불포화화효소["EgD5"; 미국 특허 제7,678,560호; 서열 번호 184 및 185] 내의 HDASH 모티프의 가능한 중요성을 더 잘 예측하기 위하여, 토폴로지 모델(도 18)을 하기의 논리와 분석에 기초하여 개발하였다.
- [0669] 먼저, EgD5의 막관통 도메인의 분석을 TMHMM 프로그램("단백질 내의 막관통 헬릭스의 예측"; TMHMM Server v. 2.0, 덴마크 DK-2800 링비 소재의 덴마크 공과 대학 바이오센트럼-DTU 생물 서열 분석 센터)을 사용하여 수행하였다. 예측에 의해, N- 및 C-말단 둘 모두가 막의 세포질 측에 위치하는 6개의 막-스패닝 헬릭스(아미노산 잔기 103-125, 130-152, 165-187, 234-256, 280-302 및 306-328)가 나타났다.
- [0670] 유사한 TMHMM 분석을 하기의 EgD5의 상동체를 사용하여 수행하였다: 진뱅크 수탁 번호 AAT09160[니츠시아 클로스테리움 예프. 미누티시마(*Nitzschia closterium f. minutissima*)], 진뱅크 수탁 번호 BAG71007[오블론기치트리움(*Oblongichytrium*) 종 SEK 347] 및 진뱅크 수탁 번호 AAL92562[파에오다크틸룸 트리코르누툼(*Phaeodactylum tricornutum*)]. 각각의 상동체에 대하여 4개의 막관통 세그먼트가 예측되었으며, 이는 EgD5에 대하여 예측된 처음의 2개 및 마지막 2개의 막관통 도메인에 해당한다.
- [0671] 막-결합 지방산 불포화화효소는 3개의 히스티딘-풍부(His-rich) 모티프를 특징으로 하는 막 이철 단백질의 슈퍼패밀리에 속한다: HX₍₃₋₄₎H(서열 번호 186 및 187), HX₍₂₋₃₎HH(서열 번호 188 및 189) 및 (H/Q)X₍₂₋₃₎HH(서열 번호 190 및 191). 이들 His-풍부 잔기는 막의 세포질 측면에 위치하는 것으로 예측되었으며, 효소 활성에 중요한 것으로 보였다(문헌[Shanklin, J. et al., *Biochemistry*, 33:12787-12794 (1994)]; 문헌[Shanklin, J., and Cahoon, E. B., *Annu. Rev. Plant Physiol. Plant Mol. Biol.*, 49:611-641(1998)]). EgD5 내에서, 제1 His-풍부 영역(HDASH[서열 번호 183])은 아미노산 잔기 165 내지 187을 스패닝하는 예측된 제3 막관통 세그먼트 전에 위치하는 한편, 제2 His-풍부 영역(HIMRHH[서열 번호 189])은 이러한 막관통 세그먼트 후에 위치한다. 제3 막관통 세그먼트가 실제로 막을 스패닝한다면, 제2 His-풍부 영역은 주변세포질 공간에 위치할 것이며---이에 따라, 철-활성 부위에서의 그의 참여가 방지된다. 결과적으로, 제3 막관통 세그먼트(아미노산 잔기 165-187) 또는 제4 막관통 세그먼트(아미노산 잔기 234-256) 중 어느 것도 막을 스패닝하지 않는 것으로 가정되었다. 이는 3개의 델타-5 불포화화효소 상동체(즉, 진뱅크 수탁 번호 AAT09160, BAG71007 및 AAL92562)에 대한 TMHMM 예측과 일치하였다.
- [0672] 델타-5 불포화화효소 기질(즉, DGLA, ETA)이 고도로 소수성이기 때문에, 이는 아마도 지질 이중층으로 분배될 것으로 추정된다. 유사하게, 3개의 His-풍부 클러스터로부터 조립된 활성 부위는 아마도 막 표면에서 또는 그에 매우 근접하여 발생할 것으로 추정되었다. 따라서, 원래 TMHMM에 의해 막을 스패닝하는 것으로 예측되었던 잔기 165 내지 187 및 234 내지 256에서 관찰되는 각각 제3 및 제4 막관통 세그먼트는 대신에 활성 부위가 막에

근접하게 위치하는 것을 보장하도록 막 표면 근처에 놓여 있는 것으로 예측되었다. 아미노산 잔기 103-125, 130-152, 280-302 및 306-328에서의 막관통 영역은 TMHMM에 의해 예측되는 바와 같이 남아 있었다.

[0673] 따라서, EgD5에 대해 예측된 최종 토폴로지 모델은 도 18에 나타나 있다. 수직의 원통은 막 스페닝 도메인을 나타내는 한편, 수평의 원통은 막을 스페닝하지 않으나 내막 표면 근처에 놓여 있는 2개의 고도의 소수성 영역을 나타낸다. 원은 아마도 활성 부위 내에 포함될 것 같은 His 잔기에 해당한다. 또한, HPGG(서열 번호 181) 모티프 및 HDASH(서열 번호 183) 모티프의 위치도 동정되었다. 마지막으로, "내부"는 세포질 공간에 해당하는 한편, "외부"는 주변세포질 공간에 해당한다.

[0674] 실시예 11B: 불포화화효소 내의 천연의 HDASH(서열 번호 183) 모티프 변이의 결정

[0675] 선택된 불포화화효소 단백질 서열을 천연의 변이가 HDASH(서열 번호 183) 모티프 내에서 발생하는지 여부를 결정하기 위하여 시험하였다. 구체적으로, 불포화화효소 단백질은 유글레나 그라실리스 델타-5 불포화화효소 ["EgD5"; 미국 특허 제7,678,560호], 모르티에렐라 알피나 델타-5 불포화화효소["MaD5"; 미국 특허 제 5,972,664호], 및 다른 공지되어 있는 델타-5 불포화화효소 및/또는 델타-5 불포화화효소와 밀접하게 관련되어 있는 것으로 공지되어 있는 델타-6 불포화화효소에 대한 BLAST 히트를 포함한다. 선택된 서열은 레이저진 바이오인포매틱스 컴퓨팅 스위트(미국 위스콘신주 메디슨 소재의 디엔에이스타 인코포레이티드(DNASTAR Inc.))의 맥얼라인(상표명) 프로그램을 사용하여 정렬하였으며, HDASH 모티프(또는 그의 변이체)는 하기 표 32에 약술되어 있다.

표 32

HDASH(서열 번호 183) 모티프의 천연 변이체

진뱅크 수탁 번호 또는 특허 번호	유기체	변이체 HDASH 모티프	서열 번호
CBL59059.1(gi_295016816)	모르티에렐라 알피나	HDASH	183
CAL49887.1(gi_116001271)	피토프토라 소자예(<i>Phytophthora sojae</i>)	HDASH	183
CBL59057.1(gi_295016812)	피스코미트렐라 파텐스	HDgnH	380
CAT16395.1(gi_218101624)	유글레나 그라실리스	HDASH	183
CBL59055.1(gi_295016808)	파에오다크틸룸 트리코르누툼	HDAnH	381
CBL59102.1(gi_295016902)	탈라시오시라 슈도나나	HDAnH	381
CAM55833.1(gi_126633754)	탈라시오시라 슈도나나	HDAnH	381
AAL13311.1(gi_16033740)	피티움 이레귤라레(<i>Pythium irregulare</i>)	HDsSH	430
CAD53323.1(gi_23894018)	피토프토라 메가스퍼마(<i>Phytophthora megasperma</i>)	HDASH	183
BAD95486.1(gi_62484905)	모르티에렐라 알피나	HDASH	183
NP_501751.1(gi_17542396)	카에노라브디티스 엘레간스(<i>Caenorhabditis elegans</i>)	HefaH	382
CAE65324.1(gi_39585564)	카에노라브디티스 브리그사예(<i>Caenorhabditis briggsae</i>)	HefH	383
AAM09687.1(gi_20069123)	트라우스토키트리움 종 ATCC 21685	HemgH	384
CAJ07076.1(gi_68124314)	리슈마니아(<i>Leishmania</i>) 주요 균주 프리에틀린(<i>Friedlin</i>)	HeAgH	385
AAH26831.1(gi_20070924)	무스 무스쿨루스(<i>Mus musculus</i>)	HDfGH	386
NP_571720.2(gi_42476248)	다니오 레리오(<i>Danio rerio</i>)	HDfGH	386
AAL82631.2(gi_55846441)	살모 살라(<i>Salmo salar</i>)	HDyGH	387
AAL92562.1(gi_19879687)	파에오다크틸룸 트리코르누툼	HDAnH	381
AAX14502.1(gi_60172920)	탈라시오시라 슈도나나	HDAnH	381
AAT09160.1(gi_47028617)	니즈시아 클로스테리움 에프. 미누티시마	HDAnH	381
AAT85663.1(gi_50882495)	마르찬티아 폴리모르파	HDgnH	380
XP_638329.1(gi_66809213)	딕티오스텔리움 디스코이테움(<i>Dictyostelium discoideum</i>) AX4	HDscH	388
XP_640331.1(gi_66812304)	딕티오스텔리움 디스코이테움 AX4	HDAcH	389
미국 특허 제 7,678,560 호	유글레나 그라실리스	HDASH	183
미국 특허 제 5,972,664 호	모르티에렐라 알피나	HDASH	183

[0676]

- [0677] 상기 분석에 기초하여, HDASH(서열 번호 183) 모티프의 Asp["D"] 잔기가 아마도 Glu 잔기["E"]로 치환될 수 있으며, Ala["A"] 잔기가 아마도 Gly["G"], Ser["S"], Phe["F"], Tyr["Y"] 또는 Met["M"] 잔기로 치환될 수 있고 /거나 HDASH(서열 번호 183) 모티프의 Ser["S"] 잔기가 아마도 Cys["C"], Asn["N"], Gly["G"], Ala["A"] 또는 Thr["T"] 잔기로 치환될 수 있는 것으로 보였다.
- [0678] 실시예 11C: 야생형 유글레나 그라실리스 델타-5 불포화화효소["EgD5"]의 서열
- [0679] 미국 특허 제7,678,560호에는 유글레나 그라실리스로부터의 델타-5 불포화화효소(즉, EgD5, 서열 번호 185)의 분리 및 클로닝이 기재되어 있다. 최근에, 상기 문헌에서 클로닝된 EgD5의 더욱 상세한 분석에 의해, 하나 이상의 변이체 "야생형" 유글레나 그라실리스 델타-5 불포화화효소 서열이 동정되었으며, 이를 EgD5R로 명명하고, 본 명세서에 서열 번호 192 및 193으로 기재하였고, 이는 이전에는 인식되지 않았던 것이다. 미국 특허 제 7,678,560호에 기재된 EgD5의 위치 347의 Ser 잔기 대신에, EgD5R(서열 번호 193)은 위치 347의 Arg 잔기를 포함한다. 이러한 차이는 PCR 또는 cDNA 생성 방법의 결과로 야기되는 것으로 가정된다.
- [0680] 구체적으로, EgD5(서열 번호 184, 미국 특허 제7,678,560호의 서열 번호 1에 해당)를 주형으로서 유글레나 그라실리스의 이중-가닥 cDNA를 사용하여 5'- 및 3'-RACE 기술을 사용하여 획득하였다(미국 특허 제7,678,560호의 실시예 4 및 5). 이어서, 유글레나 그라실리스 델타-5 불포화화효소를 암호화하는 ORF는 주형으로서 유글레나 그라실리스 cDNA를 사용하여 PCR에 의해 증폭시키고, 정제하고, 제한효소 분해로 처리한 다음, 적절한 벡터 내로 직접 라이게이션시켜, pDMW367(미국 특허 제7,678,560호의 실시예 6)을 얻었다. pDMW367의 서열은 미국 특허 제7,678,560호에서 서열 번호 23으로 제공되어 있다(본 명세서에서의 서열 번호 194에 해당). 미국 특허 제 7,678,560호에서, pDMW367이 키메라 FBAIN::EgD5::Pex20 유전자를 포함하는 것으로 보고되어 있지만, 이러한 키메라 유전자 내의 델타-5 불포화화효소 서열이 실제로 EgD5R(서열 번호 192)의 뉴클레오티드 서열이었음이 이제 인식되었다.
- [0681] EgD5(서열 번호 184) 및 EgD5R(서열 번호 192)(도 19a 및 19b)의 정렬에 의해, 4개의 뉴클레오티드 차이가 나타났다. 서열 번호 184에 대한 돌연변이는 G819GA, T948C, C1041A 및 G1349A이다. G1349A 돌연변이는 pDMW367 내로의 클로닝을 위해 EgD5를 증폭시키기 위해 사용되는 특이적 프라이머 서열에 기인한다. EgD5(서열 번호 185) 및 EgD5R(즉, 서열 번호 193)의 번역된 생성물의 정렬에 의해, 단일의 아미노산 차이, 즉, S347R 돌연변이가 드러났다.
- [0682] 또한, 미국 특허 제7,678,560호, 실시예 9는 야로위아 리폴리티카에서의 발현을 위해 코돈-최적화된 EgD5 유래의 합성 델타-5 불포화화효소(즉, EgD5S; 서열 번호 195 및 196)의 생성을 기재하고 있다. EgD5의 코돈-최적화에 의해, 1350 bp 암호화 영역 중 196 bp(14.5%)의 변형 및 전체 449개 코돈 중 189개 코돈(42%)의 최적화가 야기되었다. 코돈-최적화된 EgD5S 유전자에 의해 암호화되는 단백질 서열(즉, 서열 번호 196)은 야생형 단백질 서열(즉, 서열 번호 185)의 것과 동일하였으며, 여기서, 347 위치에서의 아미노산은 Ser이다.
- [0683] 실시예 11D: 4개의 제한 엔도뉴클레아제 부위가 제거된 야생형 EgD5R["EgD5R*"]을 포함하는 구축물 pDMW367-M4의 생성
- [0684] 본 실시예에는 키메라 FBAIN::EgD5R::Pex20 유전자를 포함하는 플라스미드 pDMW367-M4(도 20a, 20b 및 20c)의 구축이 기재되어 있다. EgD5R*(서열 번호 197)은 이후의 클로닝 절차를 용이하게 하기 위해 생성된 변형된 야생형 EgD5R의 변이체(서열 번호 192)이었으며, 여기서, 이러한 변형에 의해, 야생형 EgD5R 암호화 영역으로부터 4개의 제한 효소 부위(즉, *EcoRI*, *HindIII*, *BglII* 및 *NcoI*)의 제거가 야기되었다. EgD5R(서열 번호 193) 및 EgD5R*(서열 번호 198)의 아미노산 서열은 동일하다.
- [0685] 구체적으로, 플라스미드 pDMW367-M4(서열 번호 199; 도 20c)를 pDMW367(서열 번호 194, 실시예 11C; 도 20a)로부터 유도하였다. 천연의 *EcoRI*, *HindIII*, *BglII* 및 *NcoI* 제한 효소 부위를 이후에 EgD5R 암호화 영역으로부터 제거하여, pDMW367-M4를 생성하였다. 먼저, *EcoRI* 및 *BglII* 부위를 주형으로서 pDMW367(서열 번호 194)을 사용하고, 프라이머로서 2쌍의 올리고뉴클레오티드를 사용하여, 시험관 내 돌연변이유발에 의해 제거하였다. 프라이머 쌍 YL813(서열 번호 200) 및 YL814(서열 번호 201)는 *EcoI* 부위의 돌연변이를 가능하게 하는 한편, 프라이머 쌍 YL815(서열 번호 202) 및 YL816(서열 번호 203)은 *BglII* 부위의 돌연변이를 가능하게 하였다. 이들 반응에 의해, 구축물 pDMW367-M2(도 20b; 서열 번호 204)가 생성되었다. 서열 분석에 의해, pDMW367-M2 내의 변이체 EgD5R의 아미노산 서열이 pDMW367 내의 EgD5R의 아미노산 서열과 동일한 것을 확인하였다.
- [0686] 이어서, *HindIII* 및 *NcoI* 부위를 주형으로서 pDMW367-M2를 사용하고, 프라이머로서 2쌍의 올리고뉴클레오티드를 사용하여 시험관 내 돌연변이유발에 의해 제거하였다. 프라이머 쌍 YL829(서열 번호 205) 및 YL830(서열 번호

206)은 *Hind*III 부위의 돌연변이를 가능하게 하는 한편, 프라이머 쌍 YL831(서열 번호 207) 및 YL832(서열 번호 208)는 *Nco*I 부위의 돌연변이를 가능하게 하였다. 이에 의해 pDMW367-M4의 생성이 야기되었다. 다시 말하면, 서열 분석에 의해, pDMW367-M4 내의 변이체 EgD5(즉, EgD5R*)의 아미노산 서열이 pDMW367 내의 EgD5R의 아미노산 서열과 동일한 것으로 확인되었다.

[0687] 이후의 실시예에 대하여, 야생형 EgD5에 대한 참조는 다르게 특정되지 않는 한, 사실상 EgD5R(서열 번호 192 및 193) 및 EgD5R*(서열 번호 197 및 198)에 대한 참조를 포함할 것이다.

[0688] 실시예 11E: EgD5R*의 델타-5 불포화화효소 활성에 대하여 유사한 델타-5 불포화화효소 활성을 야기하는 HDxSH(서열 번호 434) 돌연변이의 동정

[0689] HDASH(서열 번호 183) 모티프는 EgD5R*(서열 번호 198)의 아미노산 잔기 155 내지 159를 스패닝한다. 단일의 아미노산 돌연변이를 부위-지정 돌연변이유발(퀵체인지 키트, 미국 캘리포니아주 소재의 스트라타진)에 의해, 주형으로서 pDMW367-M4(실시예 11D)를 사용하고, EgD5R*의 HDASH(서열 번호 183) 모티프의 Ala 잔기를 개별적으로 돌연변이시키기 위한 프라이머로서 19쌍의 올리고뉴클레오타이드(서열 번호 209-246; 하기 표 33)를 사용해 수행하여, 이에 의해, 가능한 모든 아미노산 치환(즉, HDxSH[서열 번호 434] 돌연변이체)을 생성하였다. 각각의 돌연변이로부터의 플라스미드를 에스케리키아 콜라이 XL2Blue 세포로 형질전환시켰다. 각 19개의 형질전환으로부터의 3개의 콜로니를 피킹하고, 개별적으로 37°C에서 하룻밤 액체 배지에서 성장시켰다. 플라스미드(즉, 총 57개)를 이들 배양물로부터 분리하고, 개별적으로 시퀀싱하여, 돌연변이를 확인하였다.

[0690] 야생형 pDMW367-M4 플라스미드 및 분리된 돌연변이 플라스미드를 일반 방법에 기재된 바와 같이, 각각 야로위아 리폴리티카 균주 Y4036U1 내로 형질전환시켰다. 형질전환체를 MMLeu 플레이트 상에서 선택하였다. 30°C에서 2일의 성장 후에, 각 형질전환 반응으로부터의 3개의 형질전환체를 새로운 MMLeu 플레이트로 스트리킹하고, 30°C에서 추가 2일 동안 인큐베이션시켰다. 이어서, 콜로니를 사용하여 24 웰 쿼젠텐 블록(block) 내의 3 mL의 MMLeu에 접종하였다. 블록을 200 rpm에서 셰이킹하는 30°C 인큐베이터에서 인큐베이션시켰다. 배양물을 2일 동안 인큐베이션시킨 후에, 블록을 원심분리하고, 상층액을 제거하고, 3 mL의 HGM을 첨가하였다. 블록을 추가 5일 동안 200 rpm에서 셰이킹하는 30°C 인큐베이터에 다시 두었다. 세포를 원심분리에 의해 수집하고, 지질을 추출하고, 지방산 메틸 에스테르["FAME"]를 에스테르 교환에 의해 제조하고, 이후에 휴렛-팩카드 6890 GC로 분석하였다.

[0691] 각각의 돌연변이 HDASH(서열 번호 183) 모티프가 원인이 되는 델타-5 불포화화효소 활성(3개의 형질전환체의 평균)은 하기 표 33에 요약되어 있다. 돌연변이 pDMW367M4 구축물 - 여기서, 돌연변이 구축물은 EgD5R* 돌연변이체를 포함한다 - 을 포함하는 형질전환체를 EgD5R* 내의 위치 157의 Ala 잔기에 대해 발생하는 아미노산 치환에 따라 명명한다(즉, 형질전환체 pDMW367M4-157c는 위치 157의 Ala에 대한 Cys로의 치환을 가져 HDcSH[서열 번호 390] 모티프를 제공하는 EgD5R*-157c로 명명된 돌연변이 델타-5 불포화화효소를 포함하며; 형질전환체 pDMW367M4-157g는 Ala에 대한 Gly로의 치환을 가져 HDgSH[서열 번호 429] 모티프를 제공하는 EgD5R*-157g로 명명된 돌연변이 델타-5 불포화화효소를 포함하는 등등이다). 전환 효율("평균 전환 효율")은 하기의 식에 따라 측정된다: $([\text{생성물}]/[\text{기질}+\text{생성물}]) \times 100$, 여기서, '생성물'은 중간체 생성물 및 그로부터 유래되는 경로에서의 모든 생성물을 포함한다. 결과를 플라스미드 pDMW367-M4 내의 야생형 EgD5R*(서열 번호 198)의 결과와 비교하고, GC 분석에 의해, 10.8%의 DGLA가 결정되었으며, 전체 지질의 3.6%의 ARA가 형질전환체에 의해 생성되었다(즉, 평균 전환 효율은 24.8%였다).

표 33

EgD5R* 및 HDxSH(서열 번호 434) 모티프 돌연변이체에서의 델타-5 불포화화효소 활성

Y4036U1 형질전환체*	프라이머의 서열 번호	돌연변이 모티프의 서열	HDASH	평균 전환 효율	활성 백분율**
pDMW367-M4	--	HDASH (서열 번호 183)		24.8%	100
pDMW367M4-157c	209 및 210	HDcSH (서열 번호 390)		10.7%	43.1%
pDMW367M4-157d	211 및 212	HDdSH (서열 번호 391)		1.0%	4.0%
pDMW367M4-157e	213 및 214	HDeSH (서열 번호 392)		0.9%	3.6%
pDMW367M4-157f	215 및 216	HDfSH (서열 번호 393)		1.0%	4.0%
pDMW367M4-157g	217 및 218	HDgSH (서열 번호 429)		23.8%	96%
pDMW367M4-157h	219 및 220	HDhSH (서열 번호 394)		1.0%	4.0%
pDMW367M4-157i	221 및 222	HDiSH (서열 번호 395)		0.9%	3.6%
pDMW367M4-157k	223 및 224	HDkSH (서열 번호 396)		1.0%	4.0%
pDMW367M4-157l	225 및 226	HDlSH (서열 번호 397)		1.1%	4.4%
pDMW367M4-157m	227 및 228	HDmSH (서열 번호 398)		1.0%	4.0%
pDMW367M4-157n	229 및 230	HDnSH (서열 번호 399)		1.1%	4.4%
pDMW367M4-157p	231 및 232	HDpSH (서열 번호 400)		2.3%	9.3%
pDMW367M4-157q	233 및 234	HDqSH (서열 번호 401)		0.6%	2.4%
pDMW367M4-157r	235 및 236	HDrSH (서열 번호 402)		0.8%	3.2%
pDMW367M4-157s	237 및 238	HDsSH (서열 번호 430)		23.3%	94%
pDMW367M4-157t	239 및 240	HDtSH (서열 번호 403)		1.0%	4.0%
pDMW367M4-157v	241 및 242	HDvSH (서열 번호 404)		0.3%	1.2%
pDMW367M4-157w	243 및 244	HDwSH (서열 번호 405)		0.9%	3.6%
pDMW367M4-157y	245 및 246	HDySH (서열 번호 406)		0.7%	2.8%

* 각 EgD5R* 유전자(돌연변이체 또는 유전형)는 pDMW367-M4 내에서 발현되었다.

**활성 백분율은 EgD5R*에 대한 것이다.

[0692]

[0693]

상기에 기초하여, HDASH(서열 번호 183) 모티프 내의 Ala 잔기가 EgD5R*의 델타-5 불포화화효소 활성에 실질적으로 영향을 주지 않고, Gly 또는 Ser 중 어느 하나로 치환될 수 있는 것이 명백하다. 구체적으로, pDMW367M4-157g 형질전환체에서의 EgD5R*-157g(서열 번호 247)는 23.8% 전환 효율로 DGLA를 ARA로 전환시킬 수 있는 한편, pDMW367M4-157s 형질전환체에서의 EgD5R*-157s(서열 번호 248)는 23.3% 전환 효율로 DGLA를 ARA로 전환시킬 수 있었다.

[0694]

실시에 11F: EgD5R*의 델타-5 불포화화효소 활성에 대하여 유사한 델타-5 불포화화효소 활성을 야기하는 HDxH(서열 번호 435) 돌연변이의 동정

[0695]

단일의 아미노산 돌연변이를 부위-지정 돌연변이유발(퀵체인지 키트, 미국 캘리포니아주 소재의 스트라타진)에 의해, 주형으로서 pDMW367-M4(실시에 11D)를 사용하고, EgD5R*(서열 번호 198)의 HDASH(서열 번호 183) 모티프의 Ser 잔기를 개별적으로 돌연변이시키기 위한 프라이머로서 19쌍의 올리고뉴클레오타이드(서열 번호 249-286; 하기 표 34)를 사용하여 수행하여, 이에 의해, 가능한 모든 아미노산 치환(즉, HDxH(서열 번호 435) 돌연변이체)을 생성하였다. 돌연변이유발 후에, 실시에 11E에 기재된 바와 같이, 플라스미드를 야로위아 리폴리티카 Y4036U1 내로 형질전환시키고, 형질전환체를 선택하고, MMLeu 및 HGM에서 성장시키고, FAME를 제조하고, GC에 의해 분석하였다.

[0696]

HDASH(서열 번호 183) 모티프 내의 각각의 돌연변이가 원인이 되는 델타-5 불포화화효소 활성(3개의 형질전환체의 평균)은 하기 표 34에 요약되어 있다. 돌연변이 pDMW367M4 구축물을 포함하는 형질전환체 - 여기서, 돌연변이

이 구축물은 EgD5R* 돌연변이체를 포함한다 - 를 EgD5R* 내의 위치 158의 Ser 잔기에 대해 발생하는 아미노산 치환에 따라 명명하였다(즉, 형질전환체 pDMW367M4-158a는 위치 158에서 Ser에 대한 Ala로의 치환을 가져, HDAaH [서열 번호 431] 모티프를 제공하는 EgD5R*-158a로 명명된 돌연변이 델타-5 불포화화효소를 포함하며; 형질전환체 pDMW367M4-158r은 Ser에 대한 Arg로의 치환을 가져, HDArH[서열 번호 419] 모티프를 제공하는 EgD5R*-158r로 명명된 돌연변이 델타-5 불포화화효소를 포함하는 등등이다). 전환 효율은 실시예 11E에 기재된 식에 따라 측정하였다. 결과를 플라스미드 pDMW367-M4 내의 야생형 EgD5R*(서열 번호 198)의 결과와 비교하고, GC 분석에 의해, 11.3%의 DGLA가 결정되었으며, 전체 지질의 3.4%의 ARA가 형질전환체에 의해 생성되었다(즉, 평균 전환 효율은 23.3%였다).

표 34

EgD5R* 및 HDAxH(서열 번호 435) 모티프 돌연변이체 내의 델타-5 불포화화효소 활성

Y4036U1 형질전환체*	프라이머의 서열 번호	돌연변이 HDASH 모티프의 서열	평균 전환 효율	활성 백분율**
pDMW367-M4	--	HDASH (서열 번호 183)	23.3%	100%
pDMW367M4- 158a	249 및 250	HDAaH (서열 번호 431)	23.5%	100.9%
pDMW367M4- 158c	251 및 252	HDAcH (서열 번호 407)	17.9%	76.8%
pDMW367M4-158d	253 및 254	HDAAdH (서열 번호 408)	2.8%	12.0%
pDMW367M4-158e	255 및 256	HDAeH (서열 번호 409)	1.9%	8.2%
pDMW367M4-158f	257 및 258	HDAfH (서열 번호 410)	1%	4.3%
pDMW367M4- 158g	259 및 260	HDAgH (서열 번호 432)	25.1%	107.7%
pDMW367M4-158h	261 및 262	HDAhH (서열 번호 411)	1.6%	6.9%
pDMW367M4-158i	263 및 264	HDAiH (서열 번호 412)	1.1%	4.7%
pDMW367M4-158k	265 및 266	HDAkH (서열 번호 413)	1%	4.3%
pDMW367M4-158l	267 및 268	HDAIH (서열 번호 414)	1.1%	4.7%
pDMW367M4- 158m	269 및 270	HDAmH (서열 번호 415)	2.3%	9.9%
pDMW367M4- 158n	271 및 272	HDAH (서열 번호 416)	16.5%	70.8%
pDMW367M4-158p	273 및 274	HDApH (서열 번호 417)	1.2%	5.2%
pDMW367M4-158q	275 및 276	HDAqH (서열 번호 418)	10.4%	44.6%
pDMW367M4-158r	277 및 278	HDArH (서열 번호 419)	10.0%	42.9%
pDMW367M4-158t	279 및 280	HDAtH (서열 번호 420)	9.6%	41.2%
pDMW367M4-158v	281 및 282	HDAvH (서열 번호 421)	1.5%	6.4%
pDMW367M4- 158w	283 및 284	HDAwH (서열 번호 422)	9.3%	40.0%
pDMW367M4-158y	285 및 286	HDAyH (서열 번호 423)	1.1%	4.7%

각 EgD5R 유전자(돌연변이체 또는 야생형)는 pDMW367-M4 내에서 발현되었다.

**활성 백분율은 EgD5R*에 대한 것이다.

[0697]

[0698]

결과에 의해, HDASH(서열 번호 183) 모티프 내의 Ser 잔기가 EgD5R*의 델타-5 불포화화효소 활성에 실질적으로 영향을 주지 않으면서, Ala 또는 Gly 중 어느 하나로 치환될 수 있음이 증명되었다. 구체적으로, pDMW367M4-158a 형질전환체 내의 EgD5R*-158a(서열 번호 287)는 23.5% 전환 효율로 DGLA를 ARA로 전환시킬 수 있는 한편, pDMW367M4-158g 형질전환체 내의 EgD5R*-158g(서열 번호 288)는 25.1% 전환 효율로 DGLA를 ARA로 전환시킬 수 있었다.

[0699]

실시예 11G: EgD5R*의 델타-5 불포화화효소 활성에 대하여 유사한 델타-5 불포화화효소 활성을 야기하는 HxGx

(서열 번호 180) 및 HDxxH(서열 번호 424) 돌연변이의 동정

- [0700] 미국 특허 공개 제2010-0075386-A1호에는 시토크롬 b_5 -유사 도메인의 HPGG(서열 번호 181) 모티프 내의 적어도 하나의 돌연변이(즉, HxGx[서열 번호 180] 돌연변이)를 갖는 돌연변이 델타-5 불포화화효소가 기재되어 있다. HPGG(서열 번호 181) 모티프는 EgD5R*(서열 번호 198)의 아미노산 잔기 33 내지 36을 스패닝한다.
- [0701] 본 실시예에서는 EgD5R*-157g(실시예 11E, 서열 번호 247), EgD5R*-158a(실시예 11F, 서열 번호 287) 및 EgD5R*-158g(실시예 11F, 서열 번호 288)의 HPGG(서열 번호 181) 모티프 내에 돌연변이를 도입하여, HPGG(서열 번호 181) 및 HDASH(서열 번호 183) 도메인 내의 이중 돌연변이의 효과를 확인하였다.
- [0702] 단일의 아미노산 돌연변이를 부위-지정 돌연변이유발(퀵체인지 키트, 미국 캘리포니아주 소재의 스트라타진)에 의해, 주형으로서 pDMW367M4-157g(실시예 11E, 서열 번호 289), pDMW367M4-158a(실시예 11F, 서열 번호 290) 및 pDMW367-158g(실시예 11F, 서열 번호 291)를 사용하고, 돌연변이 델타-5 불포화화효소 유전자의 HPGG(서열 번호 181) 모티프의 Pro 잔기 또는 제2 Gly 잔기 중 어느 하나를 개별적으로 돌연변이시키기 위한 프라이머로서 몇쌍의 올리고뉴클레오타이드(서열 번호 292-297; 표 35)를 사용해 행하여, 이에 의해 HPGG(서열 번호 181) 및 HDASH(서열 번호 183) 모티프 내에 이중 돌연변이를 생성하였다. 돌연변이유발 후에, 플라스미드를 실시예 11E에 기재된 바와 같이, 야로위아 리폴리티카 균주 Y4036U1 내로 형질전환시키고, 형질전환체를 선택하고 MMLeu 및 HGM에서 성장시키고, FAME를 제조하고 GC에 의해 분석하였다.
- [0703] HxGx(서열 번호 180) 및 HDxxH(서열 번호 424) 돌연변이 둘 모두가 있는 돌연변이 델타-5 불포화화효소의 델타-5 불포화화효소 활성은 하기 표 35에 요약되어 있다. 돌연변이 pDMW367M4 구축물을 포함하는 형질전환체 - 여기서, 돌연변이 구축물은 EgD5R* 돌연변이체를 포함한다 - 는 EgD5R*의 HDASH(서열 번호 183) 모티프 내의 Ala 잔기 또는 Ser 잔기에 대한 아미노산 치환과 조합된, EgD5R*의 HPGG(서열 번호 181) 모티프 내의 Pro 잔기 또는 제2 Gly 잔기에 대한 아미노산 치환에 따라 명명된다. 다시 말하면, 예컨대, 형질전환체 pDMW367-34g158g는 위치 34에서 Pro에 대한 Gly로의 치환을 가지며(이에 의해, HgGG[서열 번호 425] 모티프를 생성하는) 위치 158에서 Ser에 대한 Gly로의 치환을 갖는(이에 의해, HDAGH[서열 번호 432] 모티프를 생성하는) EgD5R*-34g158g로 명명된 돌연변이 델타-5 불포화화효소를 포함하는 등등이다. 전환 효율은 실시예 11E에 기재된 식에 따라 측정하였다. 결과를 플라스미드 pDMW367-M4 내의 야생형 EgD5R*의 결과와 비교하고, GC 분석에 의해, 11.7%의 DGLA가 결정되었으며, 전체 지질의 4.4%의 ARA가 형질전환체에 의해 생성되었다(즉, 평균 전환 효율은 27.5%였다).

표 35

HxGx(서열 번호 180) 및 HDxxH(서열 번호 424) 모티프를 동시에 포함하는 EgD5R* 돌연변이체 내의 델타-5 불포화화효소 활성						
Y4036U1 형질전환체	돌연변이 유전자	프라이머의 서열 번호	돌연변이 모티프의 서열	돌연변이 HPGG 모티프의 서열	돌연변이 HDASH 모티프의 서열	평균 전환 효율
pDMW367-M4	--	--	HPGG (서열 번호 181)	HDASH (서열 번호 183)		27.5%
pDMW367- 34g157g	EgD5R*-34g157g (서열 번호 298)	292 및 293	HgGG (서열 번호 425)	HDgSH (서열 번호 429)		22.9%
pDMW367- 34g158a	EgD5R*-34g158a (서열 번호 300)	292 및 293	HgGG (서열 번호 425)	HDAAH (서열 번호 431)		24.3%
pDMW367- 34g158g	EgD5R*-34g158g (서열 번호 302)	292 및 293	HgGG (서열 번호 425)	HDAGH (서열 번호 432)		26.8%
pDMW367- 34h158a	EgD5R*-34h158a	294 및 295	HhGG (서열 번호 426)	HDAAH (서열 번호 431)		18.7%
pDMW367- 34h158g	EgD5R*-34h158g	294 및 295	HhGG (서열 번호 426)	HDAGH (서열 번호 432)		22%
pDMW367- 36s158a	EgD5R*-34s158a	296 및 297	HPGs (서열 번호 427)	HDAAH (서열 번호 431)		17.5%
pDMW367- 36s158g	EgD5R*-34s158g	296 및 297	HPGs (서열 번호 427)	HDAGH (서열 번호 432)		18.9%

가 EgD5R 유전자(돌연변이체 또는 야생형)는 pDMW367-M4 내에서 발현되었다.

- [0704]
- [0705] 결과에 의해, HPGG(서열 번호 181) 모티프 및 HDASH(서열 번호 183) 모티프가 델타-5 불포화화효소 활성에 중요하지만, 야생형에 비해 적어도 64%의 델타-5 불포화화효소 활성을 보유하는 HxGx(서열 번호 180) 및 HDxxH(서열 번호 424) 모티프를 갖는 불포화화효소를 구축할 수 있는 것이 증명되었다. 구체적으로, HPGG(서열 번호 181) 모티프 내의 Pro 잔기가 Gly으로 치환됨과 동시에, 1) HDASH(서열 번호 183) 모티프 내의 Ala 잔기가 Gly로 치환되거나; 또는 2) HDASH(서열 번호 183) 모티프 내의 Ser 잔기가 Ala 또는 Gly로 치환될 수 있다. 또한, HPGG(서열 번호 181) 모티프 내의 Pro 잔기가 His로 치환됨과 동시에, HDASH(서열 번호 183) 모티프 내의 Ser 잔기가 Ala 또는 Gly 중 어느 하나로 치환될 수 있다. 그리고, HPGG(서열 번호 181) 모티프 내의 두번째 Gly 잔기가 Ser으로 치환됨과 동시에, HDASH(서열 번호 183) 모티프 내의 Ser은 Ala 또는 Gly 중 어느 하나로 치환

될 수 있다.

[0706] 바람직한 이중 돌연변이체는 EgD5R*-34g157g(서열 번호 298 및 299; pDMW367-34g157g 형질전환체에서 22.9% 전환 효율로 DGLA를 ARA로 전환시킬 수 있음), EgD5R*-34g158a(서열 번호 300 및 301; pDMW367-34g158a 형질전환체에서 24.3% 전환 효율로 DGLA를 ARA로 전환시킬 수 있음) 및 EgD5R*-34g158g(서열 번호 302 및 303; pDMW367-34g158g 형질전환체에서 26.8% 전환 효율로 DGLA를 ARA로 전환시킬 수 있음)였다.

[0707] 실시예 11H: EgD5R*-34g158g 유래의, 야로위아 리폴리티카에서의 발현을 위한 N-말단 코돈-최적화된 돌연변이 델타-5 불포화화효소 유전자("EgD5M")의 합성

[0708] EgD5R*-34g158g(서열 번호 302, 실시예 11G)의 5' 부분의 코돈 사용을 미국 특허 제7,125,672호에 기재된 것과 유사한 방식으로 야로위아 리폴리티카에서의 발현을 위해 최적화하였다. 구체적으로, EgD5R*-34g158g의 처음의 204 bp를 코돈-최적화시켜, "EgD5M"으로 명명된 코돈-최적화된 델타-5 불포화화효소 유전자의 합성을 야기하였다(서열 번호 105 및 106). EgD5M을 EgD5R*-34g158g의 델타-5 불포화화효소 유전자의 암호화 서열에 기초하여, 야로위아 코돈 사용 패턴(미국 특허 제7,125,672호), 'ATG' 번역 개시 코돈 주위의 공통 서열 및 RNA 안정성의 일반적인 규칙(문헌[Guhaniyogi, G. and J. Brewer, *Gene*, 265(1-2):11-23 (2001)])에 따라 설계하였다. 번역 개시 부위의 변형에 더하여, 암호화 영역의 N-말단 내의 204 bp 중 52 bp를 변형시켰으며(25.5%; 도 21), 불포화화효소 단백질의 N-말단 내의 68개 아미노산 중 45개 코돈을 최적화시켰다(66.2%). *NcoI* 부위 및 *NotI* 부위를 각각 번역 개시 코돈 주위 및 EgD5M의 정지 코돈 후에 혼입시켰다. 코돈-최적화된 EgD5M 유전자(즉, 서열 번호 106)에 의해 암호화되는 단백질 서열은 야생형 EgD5R*-34g158g 단백질 서열(즉, 서열 번호 303)의 것과 동일하다. 설계된 EgD5M 유전자(서열 번호 105)를 진스크립트 코포레이션(미국 뉴저지주 피스카타웨이 소재)에 의해 합성하고, pUC57(진뱅크 수탁 번호 Y14837) 내로 클로닝하여, pEgD5M(도 22a; 서열 번호 304)을 생성하였다.

[0709] 실시예 11I: EgD5M을 포함하는 구축물 pDMW367-5M의 생성

[0710] 본 실시예에는 키메라 FBAIN::EgD5M::Pex20 유전자를 포함하는 플라스미드 pDMW367-5M의 구축이 기재되어 있다. 플라스미드 pDMW367-5M(도 22b; 서열 번호 305)을 pDMW367-M4(도 20c; 서열 번호 199)의 *NcoI*/*NotI* EgD5R* 단편을 pEgD5M(도 22a; 서열 번호 304)으로부터의 *NcoI*/*NotI* EgD5M 단편으로 대체함으로써 구축하였다. 이러한 라이게이션의 생성물은 pDMW367-5M이었으며, 이는 이에 의해 하기의 성분을 포함하였다:

표 36

플라스미드 pDMW367-5M(서열 번호 305)의 성분

서열 번호 305 내의 RE 부위 및 뉴클레오티드	단편 및 키메라 유전자 성분의 설명
<i>EcoR I</i> / <i>BstW I</i> (6063-318)	하기를 포함하는 FBAIN::EgD5M::Pex20: ● FBAIN: 야로위아 리폴리티카 FBAIN 프로모터(미국 특허 제 7,202,356 호); ● EgD5M: 유칼레나 그라실리스 유래의 합성 돌연변이 N-말단 코돈-최적화된 EgD5R*-34g158g ["EgD5M"] 델타-5 불포화화효소(서열 번호 105); ● Pex20: 야로위아 Pex20 유전자(진뱅크 수탁 번호 AF054613)의 Pex20 종결자 서열
1354-474	<i>ColE1</i> 플라스미드 복제 원점
2284-1424	에스케리키아 콜라이에서의 선택을 위한 앰피실린-내성 유전자(<i>Amp^R</i>)
3183-4476	야로위아 자가 복제 서열(ARS18; 진뱅크 수탁 번호 A17608)
6020-4533	야로위아 <i>Ura 3</i> 유전자(진뱅크 수탁 번호 AJ306421)

[0711]

[0712] 실시예 11J: N-말단 코돈-최적화된 돌연변이 델타-5 불포화화효소 유전자의 변이체 "EgD5M1"을 포함하는 구축물 pDMW367-5M1의 생성

[0713] 본 실시예에는 키메라 FBAIN::EgD5M1::Pex20 유전자를 포함하는 플라스미드 pDMW367-5M1(서열 번호 307)의 구축이 기재되어 있다. EgD5M1(서열 번호 107)의 뉴클레오티드 서열은 EgD5M에서의 위치 347의 Arg에 대한 CGA 코돈이 EgD5M에서 Ser에 대한 AGC 코돈을 암호화하도록 변경되었다는 것을 제외하고는 EgD5M(서열 번호 105)의 것과 동일하다. 이러한 변형을 설계하여, 델타-5 불포화화효소 활성에 대한 R347S 돌연변이(실시예 11C에 기재)의 영향을 분석하였다.

- [0714] 설계된 EgD5M1 유전자("EgD5R*-34g158g347s"로도 지칭; 서열 번호 107)를 진스크립트 코포레이션(미국 뉴저지주 피스카타웨이 소재)에 의해 합성하고, pUC57(진뱅크 수탁 번호 Y14837) 내로 클로닝하여, pEgD5M1(서열 번호 306)을 생성하였다.
- [0715] 플라스미드 pDMW367-5M1(서열 번호 307)을 pDMW367-M4(도 20c; 서열 번호 199)의 *NcoI/NotI* EgD5R* 단편을 pEgD5M1(서열 번호 306)로부터의 *NcoI/NotI* EgD5M1 단편으로 대체함으로써 구축하였다. 이러한 라이게이션의 생성물은 키메라 FBAIN::EgD5M1::Pex20 유전자를 포함하는 pDMW367-5M1이었다.
- [0716] 실시예 11K: 야로위아 리폴리티카 균주 Y4036U1에서 EgD5M 및 EgD5M1 델타-5 불포화화효소의 기능 분석
- [0717] 대조군 플라스미드 pDMW367-M4(서열 번호 199; 실시예 11D) 및 플라스미드 pDMW367-5M(서열 번호 305; 실시예 11I) 및 pDMW367-5M1(서열 번호 307; 실시예 11J)을 각각 야로위아 리폴리티카 균주 Y4036U1 내로 따로 형질전환시켰다. 형질전환체를 실시예 11E에 기재된 바와 같이 선택하고, MMLeu 및 HGM에서 성장시키고, FAME를 제조하고, GC에 의해 분석하였다.
- [0718] EgD5R*, EgD5M 및 EgD5M1의 델타-5 불포화화효소 활성(3개 형질전환체의 평균)은 하기 표 37에 요약되어 있다. 전환 효율("전환 효율")을 실시예 11E에 기재된 식에 따라 측정하였다. 결과는 플라스미드 pDMW367-M4 내의 야생형 EgD5R*(서열 번호 198)의 결과와 비교하고, GC 분석에 의해, 10.8%의 DGLA가 결정되었고, 전체 지질의 3.6%의 ARA가 형질전환체에 의해 생성되었다(즉, 평균 전환 효율은 24.8%였다).

표 37

EgD5R*, EgD5M 및 EgD5M1에서의 델타-5 불포화화효소 활성

Y4036U1 내로 형질전환된 플라스미드	델타-5 불포화화효소	돌연변이 HPGG 및 HDASH 모티프의 서열	잔기 347의 아미노산	전환 효율
pDMW367-M4	EgD5R* (서열 번호 197 및 198)	HPGG(서열 번호 181), HDASH(서열 번호 183)	R	24.8%
pDMW367-5M	EgD5M (서열 번호 105 및 106)	HgGG(서열 번호 425), HDAGH(서열 번호 432)	R	26.5%
pDMW367-5M1	EgD5M1 (서열 번호 107 및 108)	HgGG(서열 번호 425), HDAGH(서열 번호 432)	S	27.6%

- [0719]
- [0720] 결과에 의해, EgD5M(서열 번호 106) 및 EgD5M1(서열 번호 108) 둘 모두가 야생형 EgD5R*(서열 번호 198)보다 델타-5 불포화화효소 활성이 더 큰 것으로 입증되었다. EgD5M에 비해 향상된 EgD5M1의 델타-5 불포화화효소 활성에 의해, 아미노산 잔기 347이 단백질의 델타-5 불포화화효소 활성에 영향을 주며, Arg가 아니라 Ser이 바람직한 것으로 입증되었다.
- [0721] 실시예 11L: 야로위아 리폴리티카에서의 발현을 위해 코돈-최적화된 유글레나 그라실리스 유래의 합성 델타-5 불포화화효소 유전자("EgD5S") 내의 HPGs(서열 번호 427) 및 HxxxH(서열 번호 186) 돌연변이의 동정
- [0722] 본 실시예에서는 돌연변이 EgD5S-36s(또는 "EgD5S-HPGs") 유전자의 HDASH(서열 번호 183) 모티프 내에 돌연변이를 도입하여, HPGG(서열 번호 181) 및 HDASH(서열 번호 183) 보존된 도메인 내의 이중 돌연변이의 영향을 결정한다.
- [0723] EgD5S(서열 번호 195 및 196)는 야로위아 리폴리티카에서의 발현을 위해 코돈-최적화된 EgD5(실시예 11C) 유래의 합성 델타-5 불포화화효소이다(미국 특허 제7,678,560호). EgD5S의 아미노산 서열이 EgD5와 동일하지만, 뉴클레오타이드 서열은 상이하며, 구체적으로, 번역 개시 부위의 변형에 더하여, 1350 bp 암호화 영역 중 196 bp가 변형되고(14.5%) 189개 코돈이 최적화되었다(42%). GC 함량은 야생형 유전자(즉, EgD5) 내에서의 55.5%로부터 합성 유전자(즉, EgD5S) 내에서의 54.4%로 감소하였다. 그리고, *NcoI* 부위 및 *NotI* 부위를 각각 번역 개시 코돈 주위에, 그리고 EgD5S의 정지 코돈 후에 혼입시켰다.
- [0724] 미국 특허 공개 제2010-0075386-A1호의 실시예 1 내지 4에는 사이토크롬 *b₅*-유사 도메인의 아미노산 잔기 33 내지 36을 스패닝하는 EgD5S의 HPGG(서열 번호 181) 모티프의 두번째 Gly 잔기를 변형시키도록(즉, HPGx[서열 번호 436] 돌연변이) 표적화된 부위-지정 돌연변이유발 반응에서 주형으로서 EgD5S를 사용한 돌연변이 EgD5S-36s(서열 번호 308)의 동정이 기재되어 있다. 따라서, 돌연변이체 EgD5S-36s는 HPGs(서열 번호 427) 모티프를

포함하며, 여기서, HPGG(서열 번호 181) 모티프의 두번째 Gly 잔기는 주형으로서 EgD5S(서열 번호 196)를 사용하여 Ser로 치환시켰다. EgD5S-36s의 델타-5 불포화화효소 활성(미국 특허 공개 제2010-0075386-A1호)은 EgD5S의 델타-5 불포화화효소 활성의 약 106.9%였다. 플라스미드 pDMW369S(서열 번호 309)는 돌연변이 EgD5S-36s 유전자를 포함하며; 벡터 성분은 EgD5M 유전자 대신에 돌연변이 EgD5S-36s 유전자가 있다는 것을 제외하고, pDMW367-5M(본 명세서에서 도 22b)의 성분과 유사하다.

[0725] 실시예 11G에서 성공적인 이중 EgD5R* 돌연변이(즉, 돌연변이 HPGG[서열 번호 181] 및 돌연변이 HDASH[서열 번호 183] 모티프를 동시에 포함하는)의 생성에 기초하여, EgD5S-36s 내로 도입되는 경우 유사한 HxxxH(서열 번호 186) 돌연변이가 용인될 것으로 예상된다. 구체적으로, 단일의 아미노산 돌연변이를 부위-지정 돌연변이유발(퀵체인지 키트, 미국 캘리포니아주 소재의 스트라타진)에 의해, 주형으로서 pDMW369S(키메라 FBAIN::EgD5S-36S::Pex20 유전자를 포함하는)를 사용하고, EgD5S-36s(서열 번호 308)의 HDASH(서열 번호 183) 모티프 내의 Asp, Ala 또는 Ser 잔기 중 어느 하나를 개별적으로 돌연변이시키기 위한 프라이머로서 9쌍의 올리고뉴클레오타이드(서열 번호 310-327; 표 38)를 사용하여 수행하여, 이에 의해 9개의 선택된 아미노산 치환을 생성하였다. 돌연변이유발 후에, 플라스미드를 실시예 11E에 기재된 바와 같이, 야로위아 리폴리티카 균주 Y4036U1 내로 형질 전환시키고, 형질전환체를 선택하고 MMLeu 및 HGM에서 성장시키고, FAME를 제조하고, GC에 의해 분석하였다.

[0726] HPGs(서열 번호 427) 및 HxxxH(서열 번호 186) 돌연변이 둘 모두를 갖는 돌연변이 델타-5 불포화화효소의 델타-5 불포화화효소 활성(3개의 형질전환체의 평균)은 하기의 표 38에 요약되어 있다. 돌연변이 pDMW369S 구축물을 포함하는 형질전환체 - 여기서, 돌연변이 구축물은 EgD5S-36s의 돌연변이를 포함한다 - 를 HDASH(서열 번호 183) 모티프 내의 Asp, Ala 또는 Ser 잔기에 대하여 발생한 아미노산 치환에 따라 명명하였다(즉, 형질전환체 pDMW369S-156e는 위치 156의 Asp에 대한 Glu로의 치환을 가져, HeASH[서열 번호 433] 모티프를 제공하며 EgD5S-36s156e로 명명된 돌연변이 델타-5 불포화화효소를 포함하며, 형질전환체 pDMW369S-157g는 Ala에 대한 Gly로의 치환을 가져, HDgSH[서열 번호 429] 모티프를 제공하는 EgD5S-36s157g로 명명된 돌연변이 델타-5 불포화화효소를 포함하는 등등이다). 전환 효율은 실시예 11E에 기재된 식에 따라 측정하였다. 결과를 플라스미드 pDMW369S 내의 EgD5S-36S(서열 번호 308)와 비교하고, GC 분석에 의해, 8.1%의 DGLA가 결정되었으며, 전체 지질의 6.8%의 ARA가 형질전환체에 의해 생성되었다(즉, 평균 전환 효율은 45.8%였다).

표 38

HPGs(서열 번호 427) 및 HxxxH(서열 번호 186) 모티프를 동시에 포함하는 EgD5S 돌연변이체 내의 델타-5 불포화화효소 활성					
Y4036U1 형질전환체	돌연변이 유전자	프라이머의 서열 번호	돌연변이 HDASH 모티프의 서열	평균 전환 효율	EgD5S-36s에 대한 활성 백분율
pDMW369S	EgD5S-36s (서열 번호 308)	--	HDASH (서열 번호 183)	45.8%	100%
pDMW369S-157f	EgD5S-36s157f	310 및 311	HDfSH (서열 번호 393)	3.4%	7.4%
pDMW369S-157m	EgD5S-36s157m	312 및 313	HDmSH (서열 번호 398)	2.4%	5.2%
pDMW369S-157g	EgD5S-36s157g (서열 번호 109)	314 및 315	HDgSH (서열 번호 429)	36.6%	79.9%
pDMW369S-157s	EgD5S-36s157s	316 및 317	HDsSH (서열 번호 430)	17.9%	39.1%
pDMW369S-158a	EgD5S-36s158a (서열 번호 330)	318 및 319	HDaSH (서열 번호 431)	39.1%	85.4%
pDMW369S-158n	EgD5S-36s158n	320 및 321	HDnSH (서열 번호 399)	13.0%	28.4%
pDMW369S-158t	EgD5S-36s158t	322 및 323	HDtSH (서열 번호 403)	4.5%	9.8%
pDMW369S-158g	EgD5S-36s158g (서열 번호 332)	324 및 325	HDgSH (서열 번호 432)	34.3%	74.9%
pDMW369S-156e	EgD5S-36s156e (서열 번호 328)	326 및 327	HeASH (서열 번호 433)	36.2%	79.0%

[0727]

[0728] 결과에 의해, 코돈-최적화된 EgD5S 델타-5 불포화화효소가 돌연변이 HPGG(서열 번호 181) 및 돌연변이 HDASH(서열 번호 183) 모티프 둘 모두를 포함하도록 변경될 수 있는 한편, 오직 하나의 돌연변이 HPGG 모티프(즉, HPGs [서열 번호 427])를 갖는 돌연변이 EgD5S-36s와 비교하여, 여전히 적절한 델타-5 불포화화효소 활성을 보유하는 것이 증명되었다. 바람직한 이중 돌연변이체는 EgD5S-36s156e(서열 번호 328 및 329; pDMW369S-156e 형질전환체에서 36.2% 전환 효율로 DGLA를 ARA로 전환시킬 수 있는), EgD5S-36s157g(서열 번호 109 및 110; pDMW369S-157g 형질전환체에서 36.6% 전환 효율로 DGLA를 ARA로 전환시킬 수 있는), EgD5S-36s158a(서열 번호 330 및 331; pDMW369S-158a 형질전환체에서 39.1% 전환 효율로 DGLA를 ARA로 전환시킬 수 있는) 및 EgD5S-36s158g(서열 번호 332 및 333; pDMW369S-158g 형질전환체에서 34.3% 전환 효율로 DGLA를 ARA로 전환시킬 수 있는)였다.

- [0729] 실시예 11M: 야로위아 리폴리티카에서의 발현을 위해 코돈-최적화된 유글레나 아나바에나 유래의 합성 델타-5 불포화화효소 유전자("EaD5S") 내의 HaGG(서열 번호 428) 및 HxxxH(서열 번호 186) 돌연변이의 동정
- [0730] 본 실시예에서는 돌연변이 EaD5S-35a(또는 "EaD5S-HaGG") 유전자의 HDASH(서열 번호 183) 모티프 내에 돌연변이를 도입하여, HPGG(서열 번호 181) 및 HDASH(서열 번호 183) 보존된 도메인 내의 이중 돌연변이의 영향을 결정한다.
- [0731] 미국 특허 제7,943,365호에는 유글레나 아나바에나로부터의 델타-5 불포화화효소의 분리 및 클로닝이 기재되어 있다(즉, EaD5; 서열 번호 335 및 336). 이어서, 이러한 유전자를 야로위아 리폴리티카에서의 발현을 위해 코돈-최적화시켜, 합성 델타-5 불포화화효소 EaD5S(서열 번호 337 및 338)를 초래하였다. EaD5S의 아미노산 서열이 EaD5와 동일하지만, 뉴클레오티드 서열은 상이하며, 구체적으로 번역 개시 부위의 변형에 더하여, 1362 bp 암호화 영역 중 183 bp가 변형되고(13.4%) 174개 코돈이 최적화되었다(38.3%). GC 함량은 야생형 유전자(즉, EaD5; 서열 번호 335) 내에서의 57.6%로부터 합성 유전자(즉, EaD5S; 서열 번호 337) 내에서의 54.6%로 감소하였다. 그리고, *NcoI* 부위 및 *NotI* 부위를 각각 번역 개시 코돈 주위 및 EaD5S의 정지 코돈 후에 혼입시켰다.
- [0732] 미국 특허 공개 제2010-0075386-A1호의 실시예 6에는 사이토크롬 *b₅*-유사 도메인의 아미노산 잔기 34 내지 37을 스패닝하는 EaD5S의 HPGG(서열 번호 181) 모티프의 Pro 잔기를 변형시키도록(즉, HxGG[서열 번호 437] 돌연변이) 표적화된 부위-지정 돌연변이유발 반응에서 주형으로서 EaD5S를 사용한 돌연변이 EaD5S-35a(서열 번호 334)의 동정이 기재되어 있다. 따라서, 돌연변이 EaD5S-35a(서열 번호 334)는 HaGG(서열 번호 428) 모티프를 포함하며, 여기서, HPGG(서열 번호 181) 모티프의 Pro 잔기는 주형으로서 EaD5S(서열 번호 338)를 사용하여 Ala로 치환되었다. EaD5S-35a의 델타-5 불포화화효소 활성(미국 특허 공개 제2010-0075386-A1호)은 EaD5S의 델타-5 불포화화효소 활성의 약 99.2%였다. 플라스미드 pZuFmEaD5S-A(S)(서열 번호 339)는 돌연변이 EaD5S-35a 유전자를 함유하며; 벡터 성분은 EgD5M 유전자 대신에 돌연변이 EaD5S-35a 유전자가 있다는 것을 제외하고 pDMW367-5M(본 명세서에서 도 22b; 서열 번호 305)의 성분과 동일하다.
- [0733] 실시예 11G에서의 성공적인 이중 EgD5R* 돌연변이 및 실시예 11L에서의 이중 EgD5S 돌연변이(즉, 돌연변이 HPGG[서열 번호 181] 및 돌연변이 HDASH[서열 번호 183] 모티프를 동시에 포함하는)의 생성에 기초하여, EaD5S-35a 내로 도입되는 경우 유사한 HxxxH(서열 번호 186) 돌연변이가 용인될 것으로 예상된다. HDASH(서열 번호 183) 모티프는 EaD5S 및 EaD5S-35a의 아미노산 잔기 156 내지 160을 스패닝한다.
- [0734] 단일의 아미노산 돌연변이를 부위-지정 돌연변이유발(퀵체인지 키트, 미국 캘리포니아주 소재의 스트라타진)에 의해, 주형으로서 pZuFmEaD5S-A(S)(키메라 FBAIN::EaD5S-35a::Pex20 유전자를 포함하는)를 사용하고, EaD5S-35a(서열 번호 334)의 HDASH(서열 번호 183) 모티프 내의 Asp, Ala 또는 Ser을 개별적으로 돌연변이시키기 위한 프라이머로서 9쌍의 올리고뉴클레오티드(서열 번호 340-361; 표 39)를 사용하여 수행하여 이에 의해 9개의 선택된 아미노산 치환을 생성하였다. 돌연변이유발 후에, 플라스미드를 실시예 11E에 기재된 바와 같이, 야로위아 리폴리티카 균주 Y4036U1 내로 형질전환시키고, 형질전환체를 선택하고 MMLeu 및 HGM에서 성장시키고, FAME를 제조하고, GC에 의해 분석하였다.
- [0735] HaGG(서열 번호 428) 및 HxxxH(서열 번호 186) 돌연변이를 포함하는 돌연변이 델타-5 불포화화효소의 델타-5 불포화화효소 활성(3개의 형질전환체의 평균)은 하기의 표 39에 요약되어 있다. 돌연변이 pZuFmEaD5S-A(S) 구축물을 포함하는 형질전환체 - 여기서, 돌연변이 구축물은 EaD5S-35a의 돌연변이체를 포함한다 - 를 HDASH(서열 번호 183) 모티프 내의 Asp, Ala 또는 Ser 잔기에 대하여 발생하는 아미노산 치환에 따라 명명한다. 다시 말하면, 예를 들어, 형질전환체 pZuFmEaD5S-A(S)-157e는 위치 157에서 Asp에 대한 Glu로의 치환을 가져, HeASH(서열 번호 433) 모티프를 제공하는 EaD5S-35a157e로 명명된 돌연변이 델타-5 불포화화효소를 포함하며; 형질전환체 pZuFmEaD5S-A(S)-158g는 Ala에 대한 Gly로의 치환을 가져, HDgSH(서열 번호 429) 모티프를 제공하는 EaD5S-35a158g로 명명된 돌연변이 델타-5 불포화화효소를 포함하는 등등이다. 전환 효율을 실시예 11E에 기재된 식에 따라 측정하였다. 결과를 플라스미드 pZuFmEaD5S-A(S) 내의 EaD5S-35a(서열 번호 334)의 결과와 비교하고, GC 분석에 의해, 8.6%의 DGLA가 결정되었으며, 전체 지질의 5.1%의 ARA는 형질전환체에 의해 생성되었다(즉, 평균 전환 효율은 37.2%였다).

표 39

HGG(서열 번호 428) 및 돌연변이 HxxxH(서열 번호 186) 모터프를 동시에 포함하는 EaD5S 돌연변이체 네의 델타-5 불포화화효소 활성					
Y4036U1 형질전환체	돌연변이 유전자	프라이머의 서열 번호	돌연변이 HDASH 모터프의 서열	평균 전환 효율	EaD5S-35a에 대한 활성 백분율
pZuFmEaD5S-A(S)	EaD5S-35a (서열 번호 334)	--	HDASH (서열 번호 183)	37.2%	100%
pZuFmEaD5S-A(S)-157e	EaD5S-35a157e	340 및 341	HeASH (서열 번호 433)	14.0%	37.6%
pZuFmEaD5S-A(S)-158f	EaD5S-35a157f	342 및 343	HDSSH (서열 번호 393)	2.1%	5.6%
pZuFmEaD5S-A(S)-158g	EaD5S-35a158g (서열 번호 111)	344 및 345	HDgSH (서열 번호 429)	28.4%	76.3%
pZuFmEaD5S-A(S)-158m	EaD5S-35a158m	346 및 347	HDmSH (서열 번호 398)	1.8%	4.8%
pZuFmEaD5S-A(S)-158s	EaD5S-35a158s (서열 번호 362)	348 및 349	HDsSH (서열 번호 430)	27.4%	73.7%
pZuFmEaD5S-A(S)-158y	EaD5S-35a158y	350 및 351	HDySH (서열 번호 406)	1.9%	5.1%
pZuFmEaD5S-A(S)-159a	EaD5S-35a159a	352 및 353	HDAaH (서열 번호 431)	2.0%	5.4%
pZuFmEaD5S-A(S)-159c	EaD5S-35a159c	354 및 355	HDAcH (서열 번호 407)	14.2%	38.2%
pZuFmEaD5S-A(S)-159g	EaD5S-35a159g (서열 번호 364)	356 및 357	HDAGH (서열 번호 432)	26.5%	71.2%
pZuFmEaD5S-A(S)-159n	EaD5S-35a159n	358 및 359	HDAnH (서열 번호 416)	4.2%	11.3%
pZuFmEaD5S-A(S)-159t	EaD5S-35a159t	360 및 361	HDAtH (서열 번호 420)	9.8%	26.3%

[0736]

[0737]

결과에 의해, 코돈-최적화된 EaD5S 델타-5 불포화화효소가 돌연변이 HPGG(서열 번호 181) 및 돌연변이 HDASH(서열 번호 183) 모터프 둘 모두를 포함하도록 변경될 수 있는 한편, 오직 하나의 돌연변이 HPGG 모터프(즉, [서열 번호 428])를 갖는 돌연변이 EaD5S-35a와 비교하여, 여전히 적절한 델타-5 불포화화효소 활성을 보유하는 것이 증명되었다. 바람직한 이중 돌연변이체는 EaD5S-35a158g(서열 번호 111 및 112; pZuFmEaD5S-A(S)-158g 형질전환체에서 28.4% 전환 효율로 DGLA를 ARA로 전환시킬 수 있는), EaD5S-35a158s(서열 번호 362 및 363; pZuFmEaD5S-A(S)-158s 형질전환체에서 27.4% 전환 효율로 DGLA를 ARA로 전환시킬 수 있는) 및 EaD5S-35a159g(서열 번호 364 및 365; pZuFmEaD5S-A(S)-159g 형질전환체에서 26.5% 전환 효율로 DGLA를 ARA로 전환시킬 수 있는)였다.

수탁번호

[0738]

기탁기관명 : ATCC

수탁번호 : PTA-10025

수탁일자 : 20090514

기탁기관명 : ATCC

수탁번호 : PTA-10026

수탁일자 : 20090514

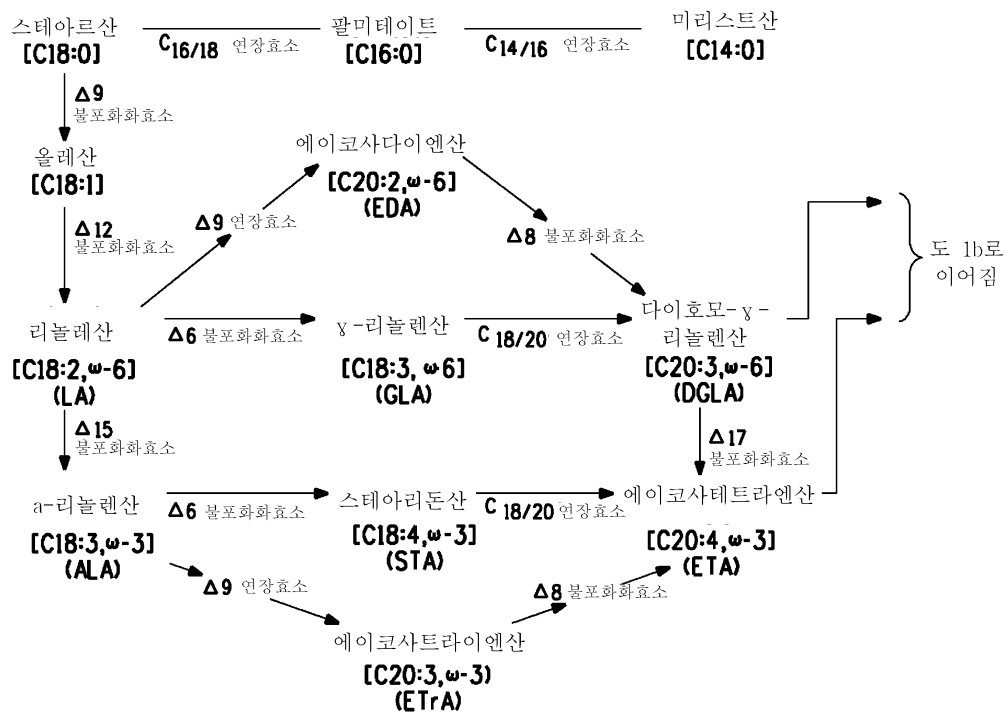
기탁기관명 : ATCC

수탁번호 : PTA-10027

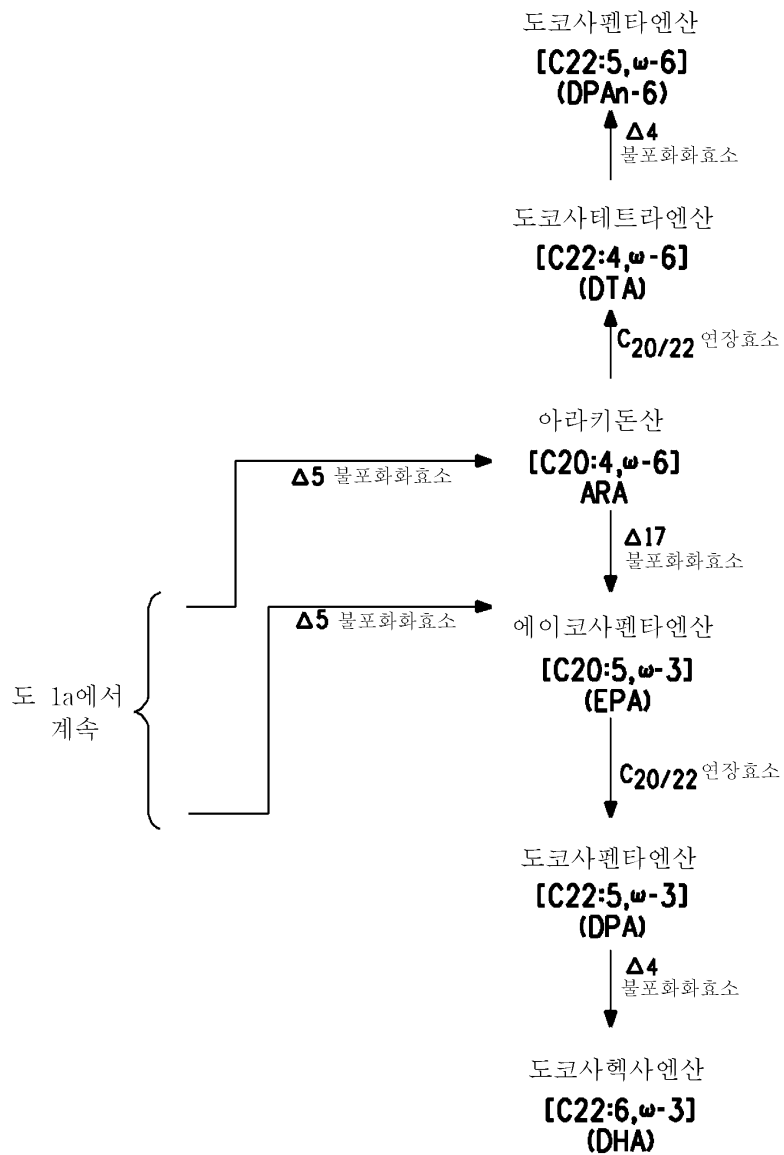
수탁일자 : 20090514

도면

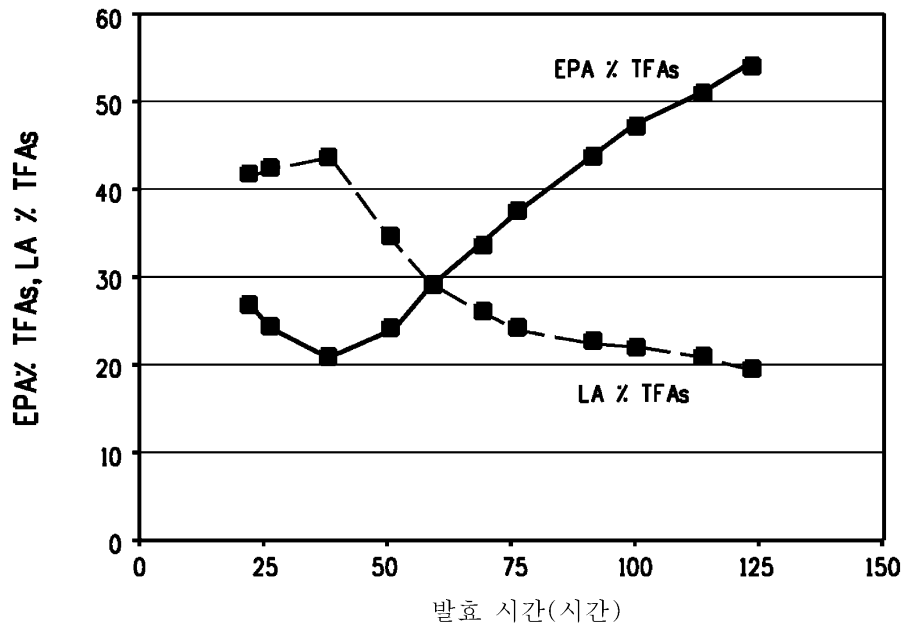
도면1a



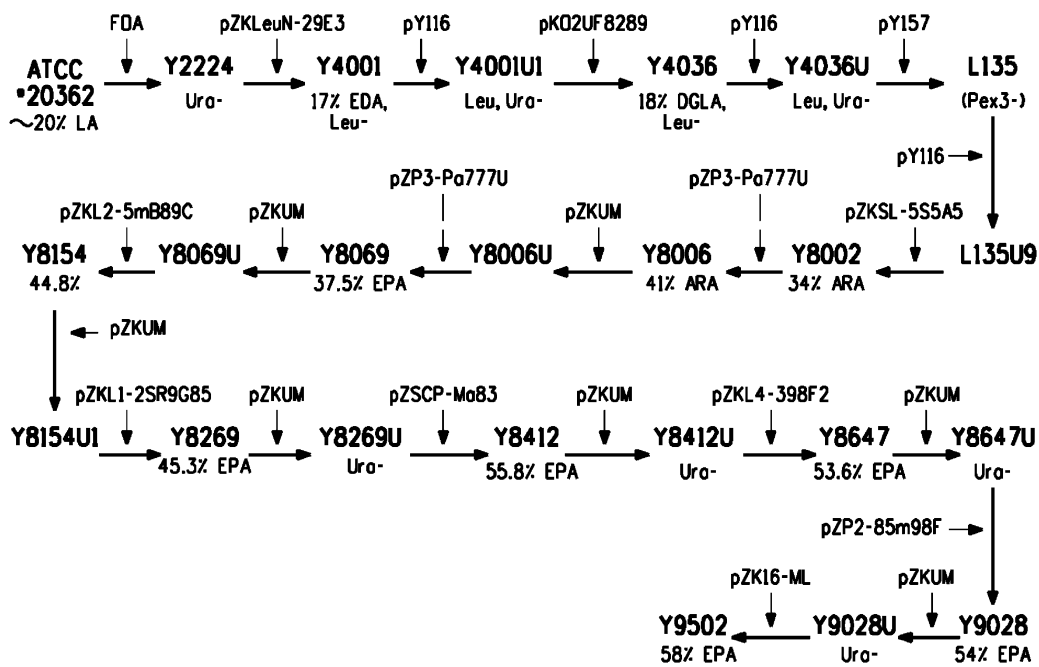
도면1b



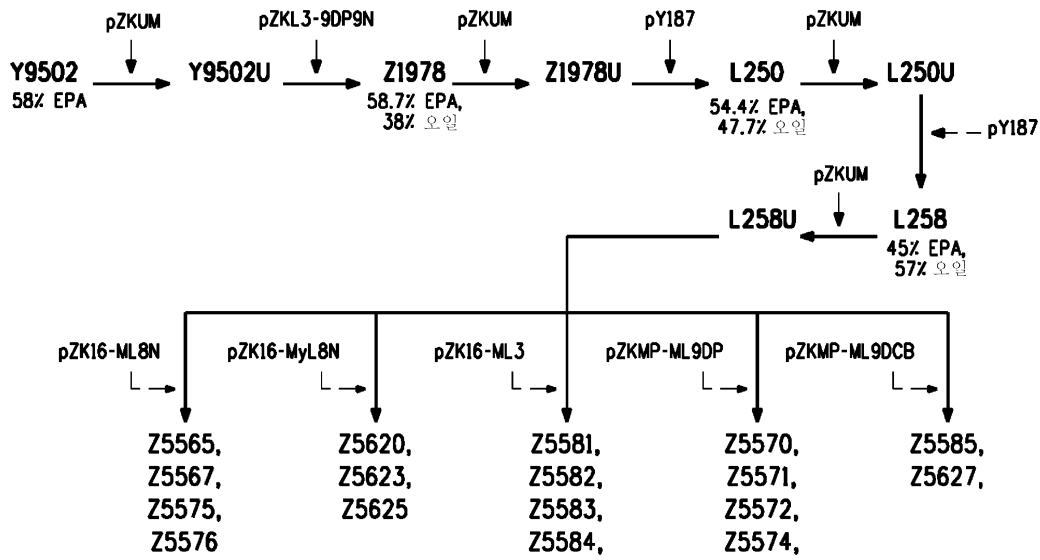
도면2



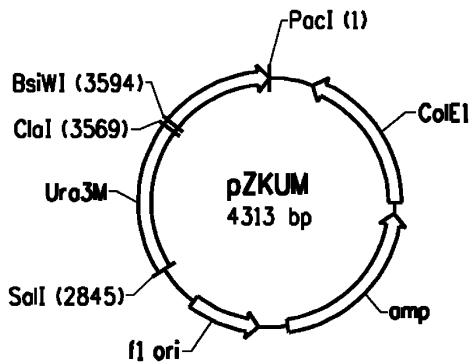
도면3a



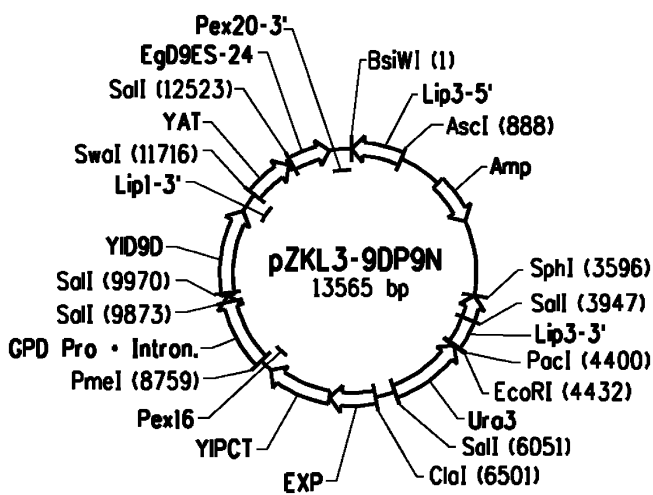
도면3b



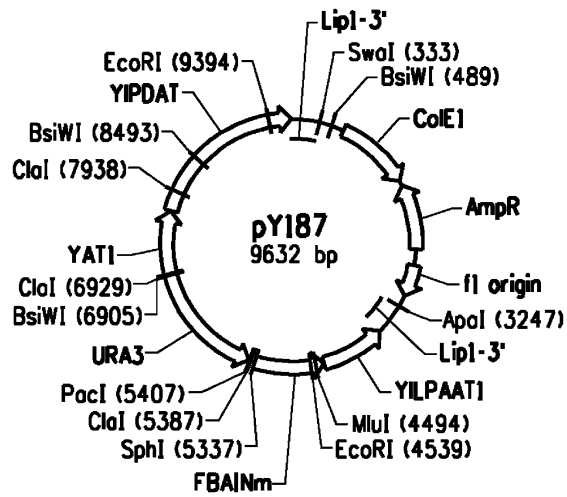
도면4a



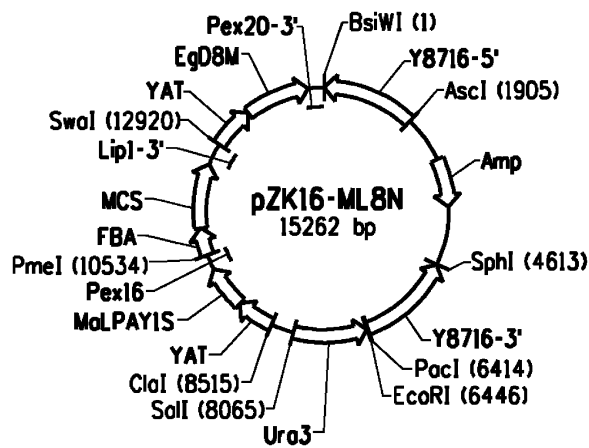
도면4b



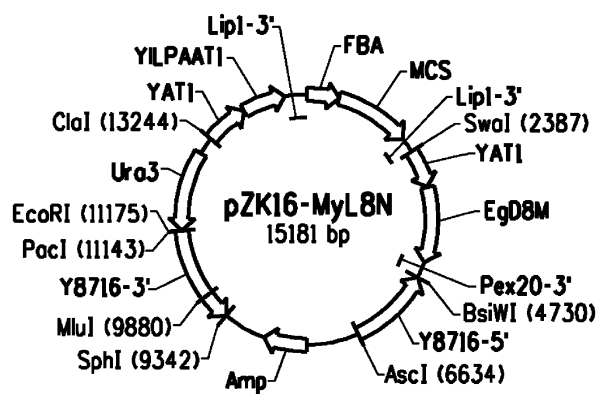
도면5a



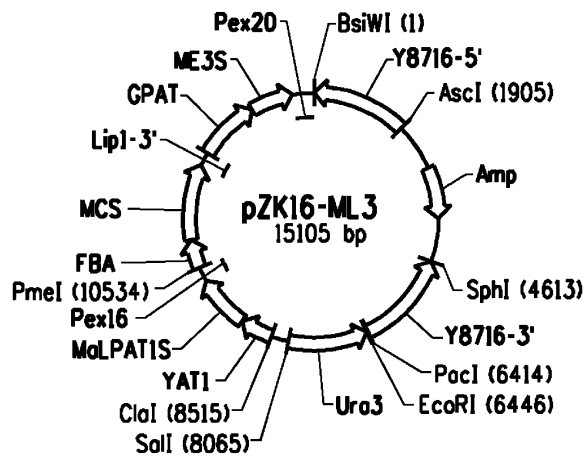
도면5b



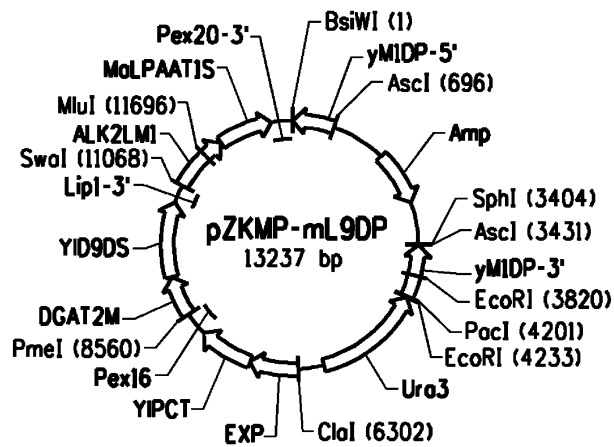
도면6a



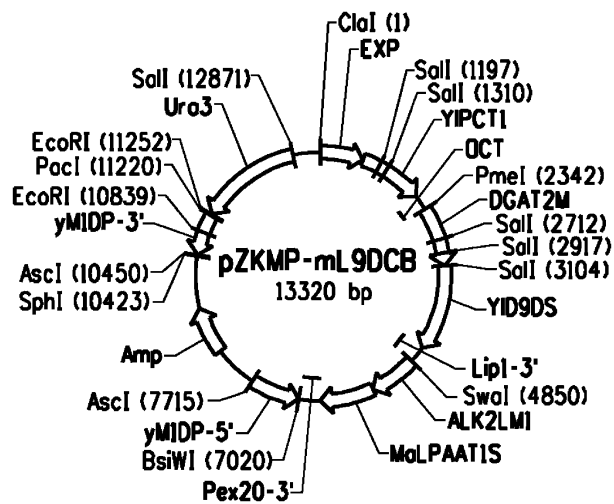
도면6b



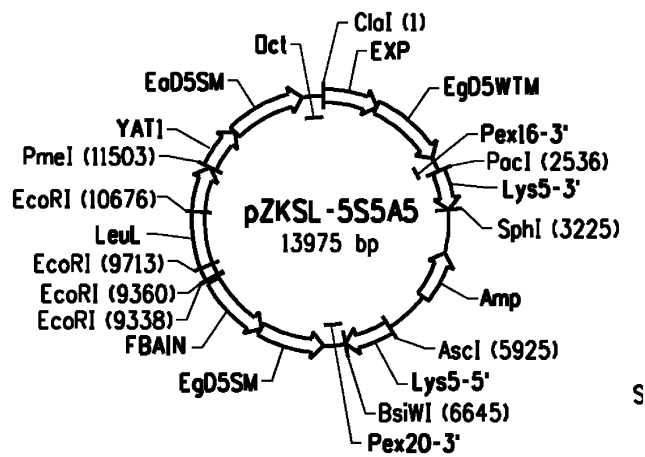
도면7a



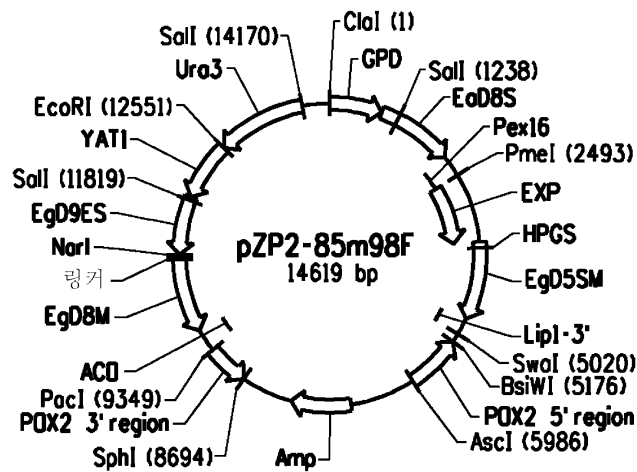
도면7b



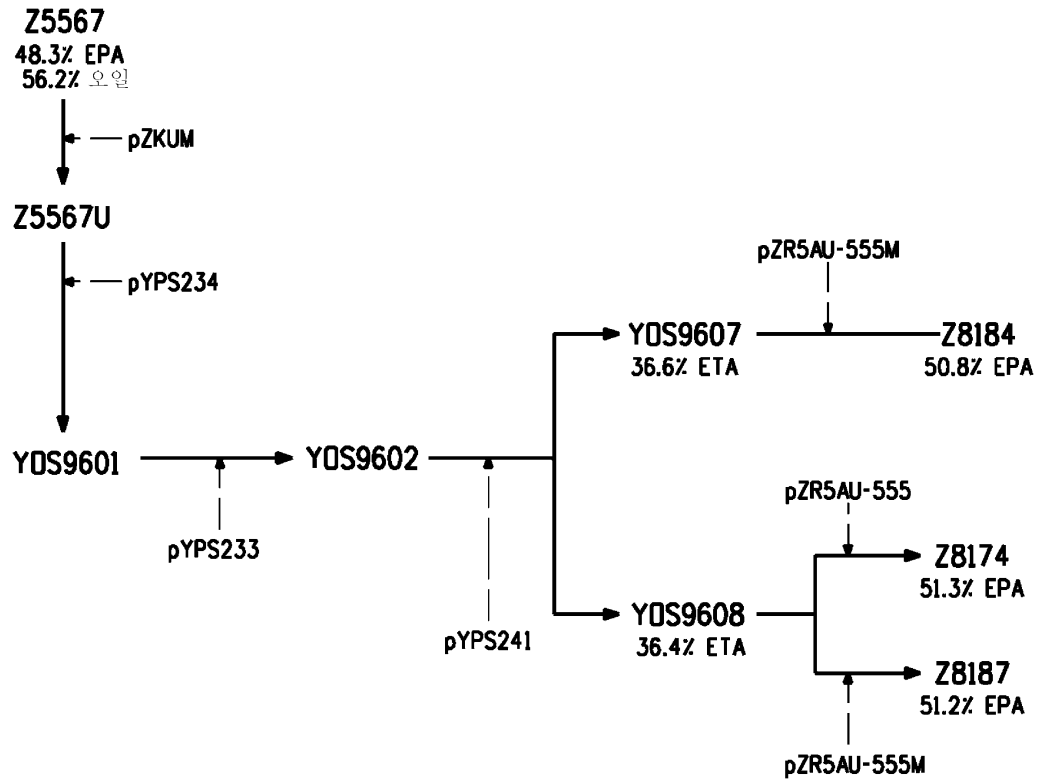
도면8a



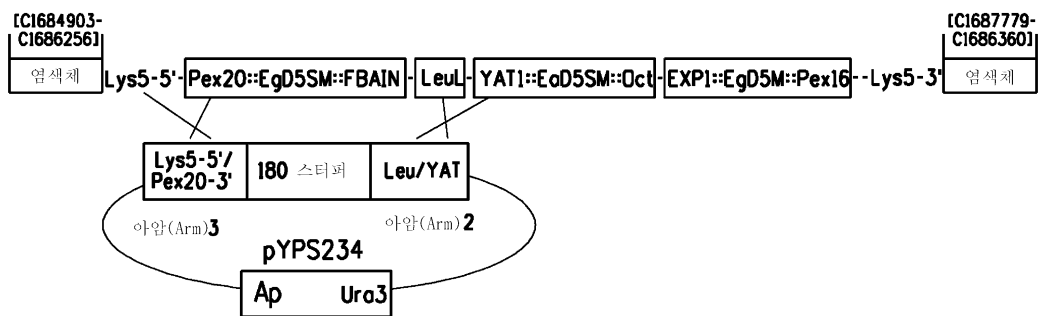
도면8b



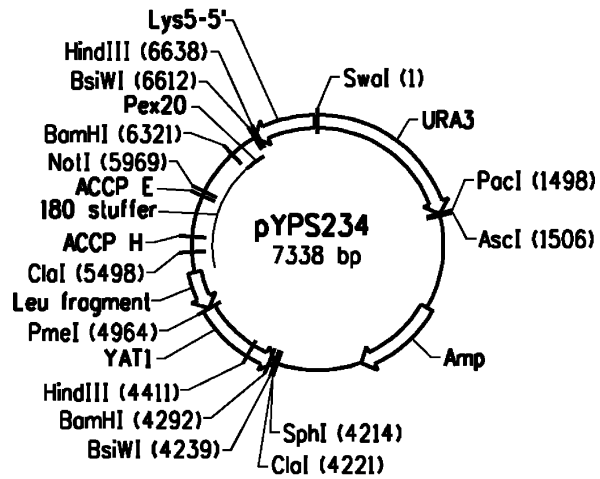
도면9



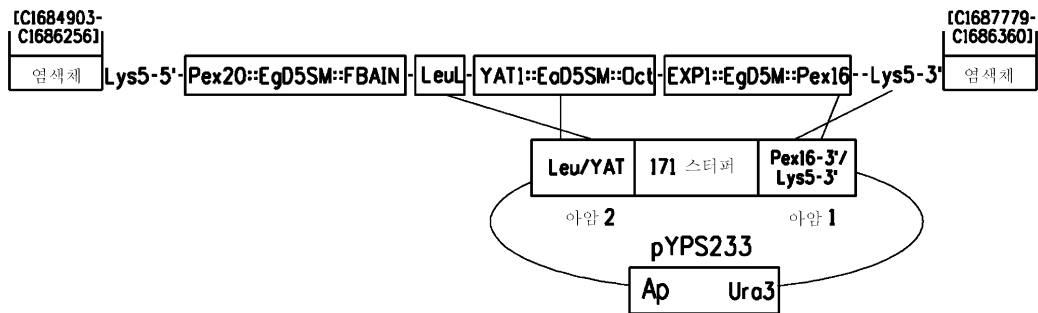
도면10a



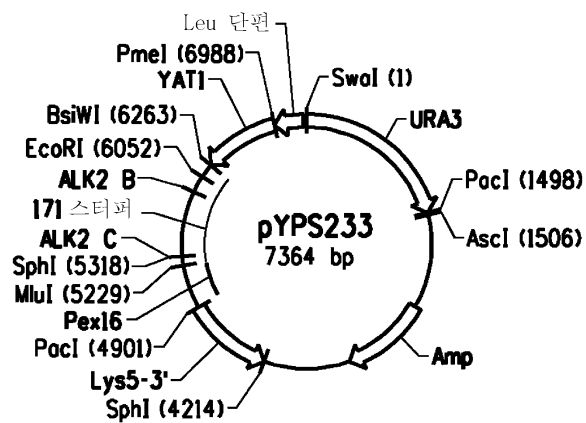
도면10b



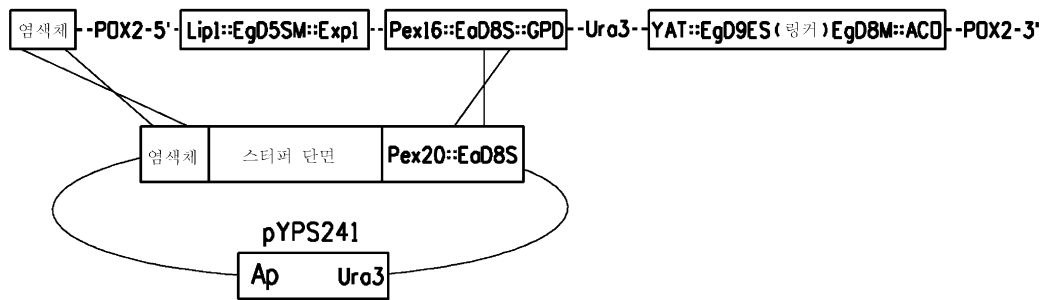
도면11a



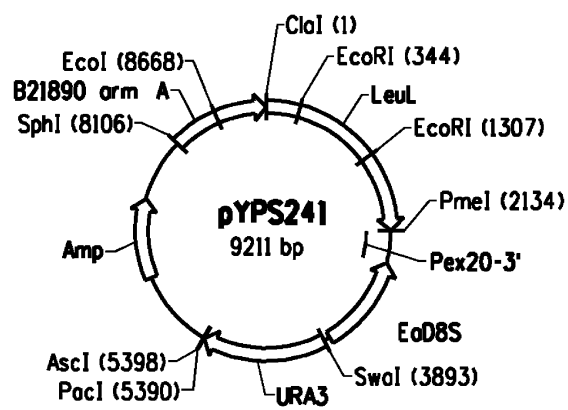
도면11b



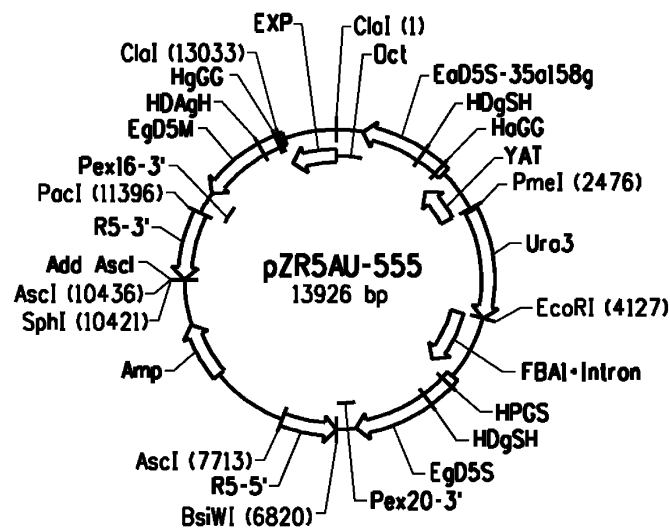
도면12a



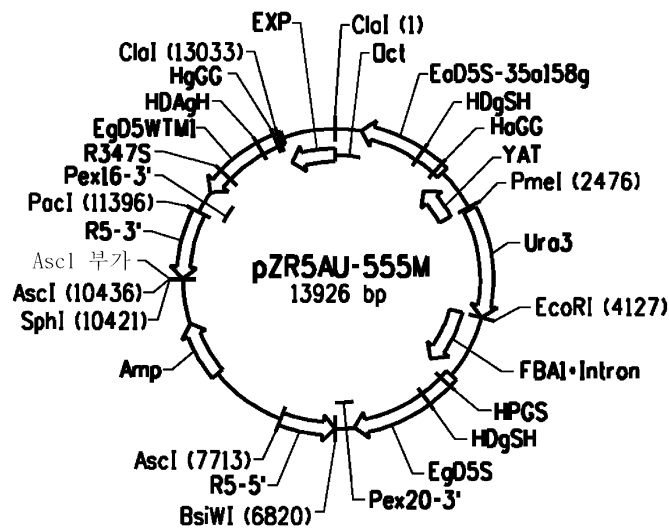
도면12b



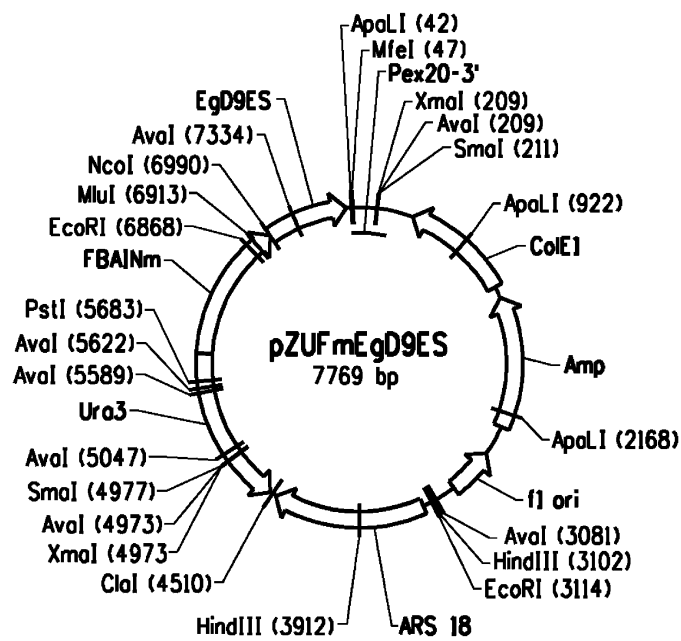
도면13a



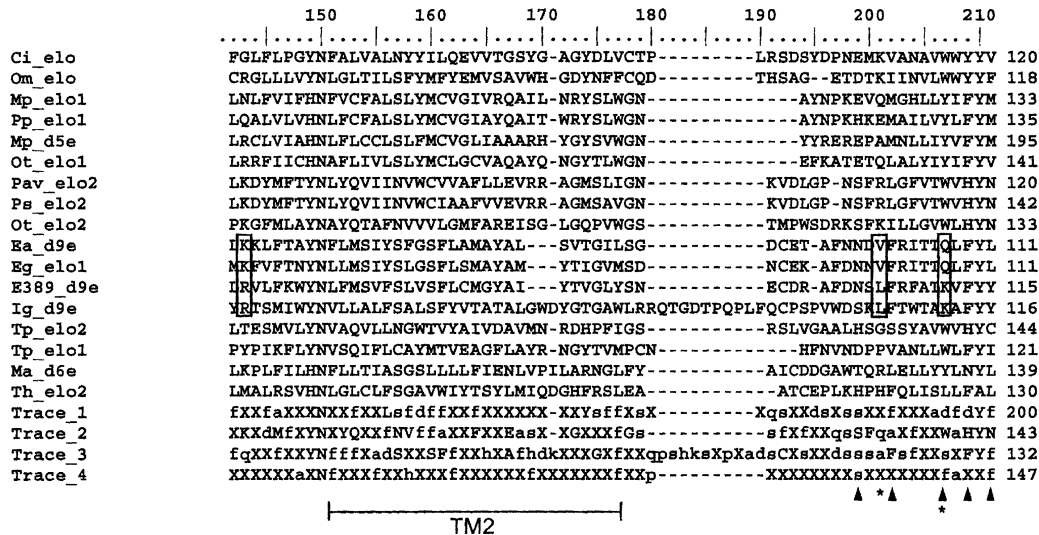
도면13b



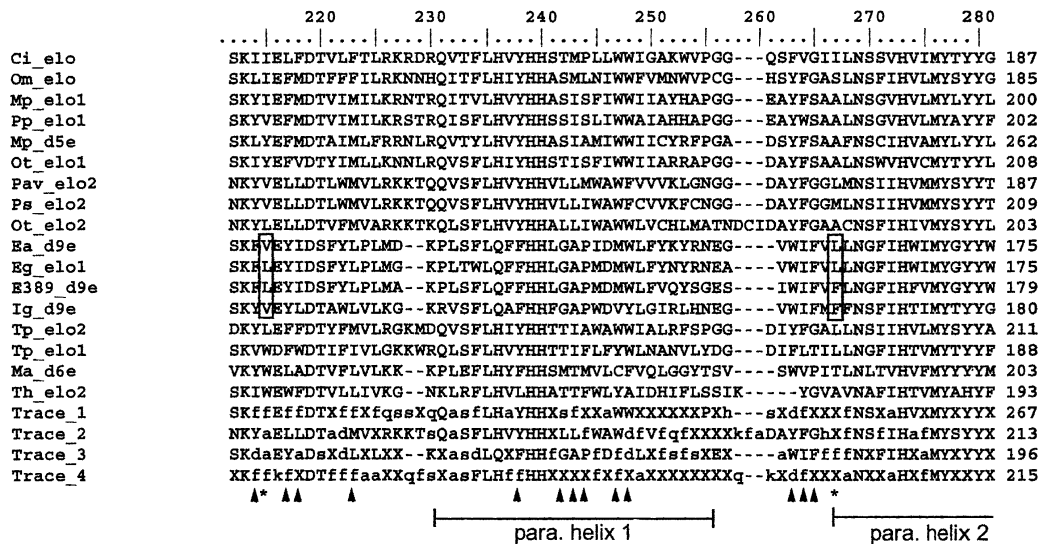
도면14



도면15c



도면15d



도면15g

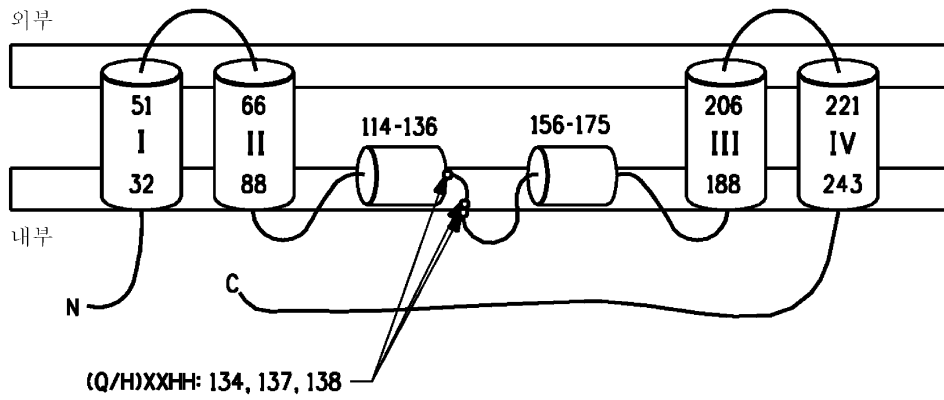
	430	
		
Ci_elo	-----	289
Om_elo	-----	295
Mp_elo1	-----	290
Pp_elo1	-----	290
Mp_d5e	-----	348
Ot_elo1	-----	292
Pav_elo2	-----	277
Ps_elo2	-----	302
Ot_elo2	-----	300
Ea_d9e	-----	258
Eg_elo1	-----	258
E389_d9e	-----	263
Ig_d9e	-----	263
Tp_elo2	SKAVKRRKGTRVTGAM	358
Tp_elo1	-----	272
Ma_d6e	-----	317
Th_elo2	-----	271
Trace_1	-----	382
Trace_2	-----	310
Trace_3	-----	280
Trace_4	sqhaqqqhsqashhf	370

도면15h

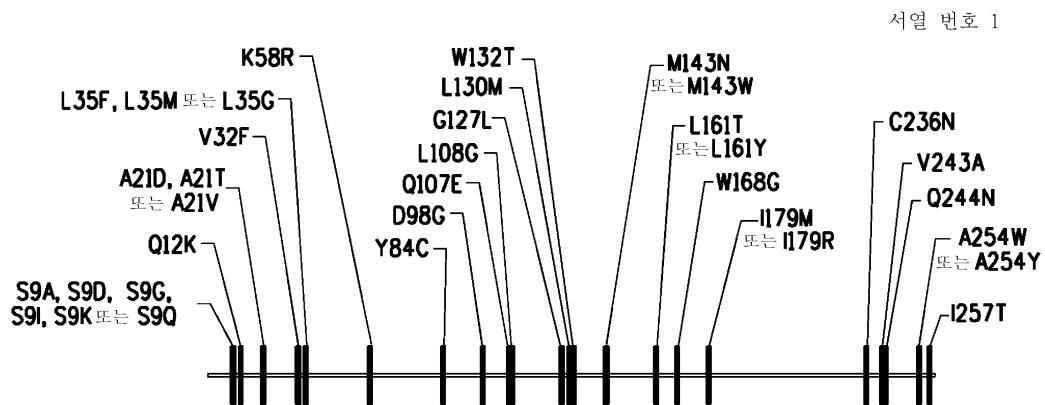
약어

Ci_elo = 서열번호:133
 Mp_elo1 = 서열번호:135
 Mp_d5e = 서열번호:137
 Pav_elo2 = 서열번호:139
 Ot_elo2 = 서열번호:141
 Eg_d9e = 서열번호:32
 Ig_d9e = 서열번호:42
 Tp_elo1 = 서열번호:143
 Th_elo2 = 서열번호:145
 Om_elo = 서열번호:134
 Pp_elo1 = 서열번호:136
 Ot_elo1 = 서열번호:138
 Ps_elo2 = 서열번호:140
 Ea_d9e = 서열번호:34
 E398_d9e = 서열번호:38
 Tp_elo2 = 서열번호:142
 Ma_d6e = 서열번호:144

도면16a



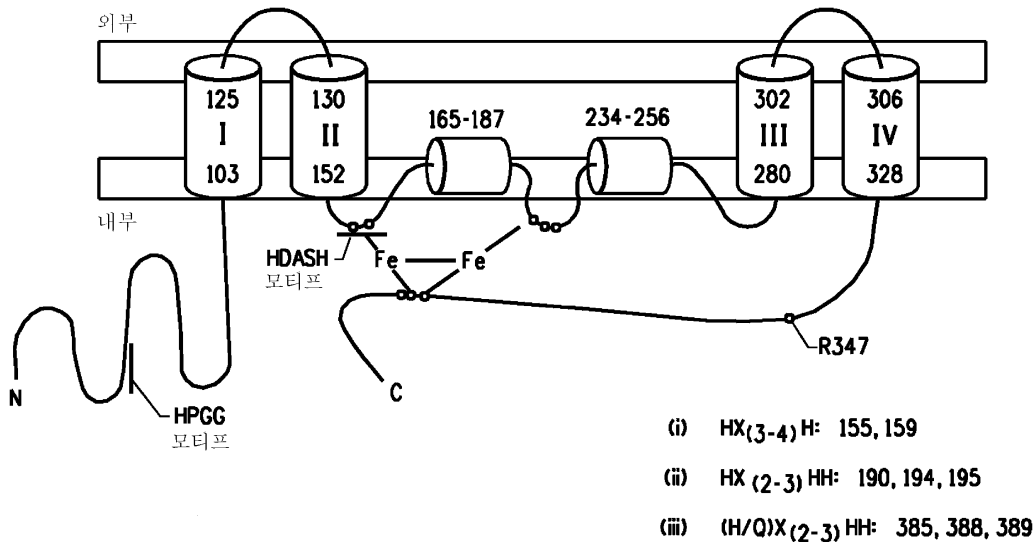
도면16b



도면17

		1		*	*		*	*	*		60
E389D9e (서열번호 : 38)	(1)	MAAVIEVANEFVAITAETLPKVDYQRLWRDIYSCELLYFSIAFVILKFTLGELSDSGKK-									
EaD9e (서열번호 : 34)	(1)	---MEAAKELVSI VQEEELPKVDYAQLWODASSCEVLYLSVAFVAIKFMLRPLDLKRQA-									
EgD9e (서열번호 : 32)	(1)	---MEVVNEIVSIGQEVLPKVDYAQLWSDASHCEVLYLSIAFVILKFTLGPLGPKGQS-									
IgD9e (서열번호 : 42)	(1)	-----MALANDAGERIWAAVTDPEILIGTFSYLLKPLLRNSGLVDEKK									
공통	(1)	MEVANEIVSI QE LPKVDYARLWADASSCEVLYLSIAFVILKFTLRPLGLKGQK									
		61	*							*	120
E389D9e (서열번호 : 38)	(60)	-ILRVLFKWNLFMSVPSLVSFLCMGYAIYTVG-----LYSNECDRAFDN									
EaD9e (서열번호 : 34)	(56)	-TLKKLFTAYNFMLSIYSGSFLAMAYALSVTG-----ILSGDCETAFNN									
EgD9e (서열번호 : 32)	(56)	-RMKFVFTNYNLLMSIYSLGSFLSMAYAMYTIG-----VMSDNCEKAFDN									
IgD9e (서열번호 : 42)	(45)	GAYRTSMIWNVLLALFSAISFYVTATLGDWDYGTGAWLRRTGDTQPPLFQCPSPVWDS									
공통	(61)	LK LFTWYNLLMSISFLGSFLAMAYALYTIG ILSNDCEKAFDN									
		121	**			*	*	*	*		180
E389D9e (서열번호 : 38)	(104)	SLERFATKVFFYYSKFLEYIDSFYFLPMKPLSFLQFFHHLGAPMDMWLFVQYSGESIWIIF									
EaD9e (서열번호 : 34)	(100)	DVERITTQLFYLSKFVEYIDSFYFLPMKPLSFLQFFHHLGAPIDMWLFYKYRNEGVIWIF									
EgD9e (서열번호 : 32)	(100)	NVERITTQLFYLSKFLEYIDSFYFLPMCKPLTWLQFFHHLGAPMDMWLFYNYRNEAVWIF									
IgD9e (서열번호 : 42)	(105)	KLFTWTAKAFYYSKYVEYLDTAWLVLKCKRVSFLQAFHHFGAPWDVYLGIRLHNEGVIWIF									
공통	(121)	LERITTQLFYYSKFLEYIDSFYFLPMCKPLSFLQFFHHLGAPMDMWLFYQYRNEGVIWIF									
		181	*			*					240
E389D9e (서열번호 : 38)	(164)	VLNNGFIHFWMYGYWTRLMKFNFPMPKQLITAMQIQFNVGFIYVWYKDIPIFYRKDPM									
EaD9e (서열번호 : 34)	(160)	VLNNGFIHFWMYGYWTRLIKLNFPMPKQLITSMQIQFNVGFIYVWKYRNVPICYRQDGM									
EgD9e (서열번호 : 32)	(160)	VLNNGFIHFWMYGYWTRRLIKLKFPMPKSLITSMQIQFNVGFIYVWKYRNIPCYRQDGM									
IgD9e (서열번호 : 42)	(165)	MFNNSFIHTIMYTYVGLTAAGYKFA-KPLITAMQIQFVGGFLLVWDYINVPFCNSDKG									
공통	(181)	VLNNGFIHFWMYGYWTRLIKLNFPMPKQLITAMQIQFNVGFIYVWKYRNIPCYRQDGM									
		241		*	*	*	*	*	*		280
E389D9e (서열번호 : 38)	(224)	RMLAWIFENYWFVGTVLLLFNFFVKSIVFPKPKTADKKVQ									
EaD9e (서열번호 : 34)	(220)	RMFAMIFENYWFVGTVLLLFNFFVQTYIRKPRKNRGKKE-									
EgD9e (서열번호 : 32)	(220)	RMFGMIFENYWFVGTVLLCLFLNFFVQTYIVRKHKGAKKIQ-									
IgD9e (서열번호 : 42)	(224)	KLFSWAFENYAVVGSVFLLFCHFFVQDNLATKKSAGKQQL									
공통	(241)	RMFAMIFENYWFVGTVLLLFNFFVQTYI KKKKAAAKKQ									

도면18



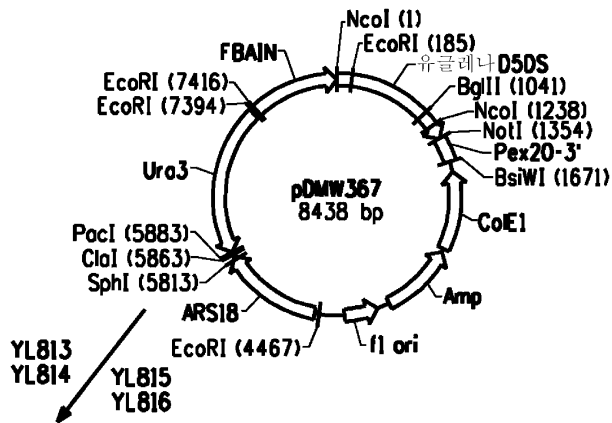
도면19a

EgD5 (서열번호 :184)	10	20	30	40	50	60	70
EgD5R (서열번호 :192)	ATGGCTCTCAGTCTTACCACAGAACAGCTGTTAGAACGCCCTGATTTGGTTGCGATTGATGGCATCCTCT						
EgD5 (서열번호 :184)	80	90	100	110	120	130	140
EgD5R (서열번호 :192)	ACGACCTTGAAGGGCTTGCCAAAGTTTCATCCAGGAGGAGATTTGATTCTCGCTTCTGGTGCCTCTGATGC						
EgD5 (서열번호 :184)	150	160	170	180	190	200	210
EgD5R (서열번호 :192)	CTCCCTCTCTTTTATTCAATGCATCCATACGTCAAACCGGAGAATTCCAAATTGCTTCAACAGTTCTGTC						
EgD5 (서열번호 :184)	220	230	240	250	260	270	280
EgD5R (서열번호 :192)	CGAGGGAAGCATGACCGCACCTCGAAGGACATTGTCTACACGTATGATTCTCCCTTCGCACAAGACGTTA						
EgD5 (서열번호 :184)	290	300	310	320	330	340	350
EgD5R (서열번호 :192)	CGAGGGAAGCATGACCGCACCTCGAAGGACATTGTCTACACGTATGATTCTCCCTTCGCACAAGACGTTA						
EgD5 (서열번호 :184)	360	370	380	390	400	410	420
EgD5R (서열번호 :192)	AGCGGACAATGCGCGAGGTGATGAAAGGGAAGAACTGGTACGCAACCCCTGGCTTCTGGCTGCGCACCGT						
EgD5 (서열번호 :184)	430	440	450	460	470	480	490
EgD5R (서열번호 :192)	AGCGGACAATGCGCGAGGTGATGAAAGGGAAGAACTGGTACGCAACCCCTGGCTTCTGGCTGCGCACCGT						
EgD5 (서열번호 :184)	500	510	520	530	540	550	560
EgD5R (서열번호 :192)	TGGGATCATCGCCGTGACGGCCTTTTGCAGTGGCACTGGGCTACCAACGGGGATGGTGCTGTGGGGCTG						
EgD5 (서열번호 :184)	570	580	590	600	610	620	630
EgD5R (서열번호 :192)	TGGGATCATCGCCGTGACGGCCTTTTGCAGTGGCACTGGGCTACCAACGGGGATGGTGCTGTGGGGCTG						
EgD5 (서열번호 :184)	640	650	660	670	680	690	700
EgD5R (서열번호 :192)	TTGACTGGATTTCATGCACATGCAGATCGGCTTATCCATCCAGCATGATGCGTCCCACGGGGCCATCAGCA						
EgD5 (서열번호 :184)	710	720	730	740	750	760	770
EgD5R (서열번호 :192)	TTGACTGGATTTCATGCACATGCAGATCGGCTTATCCATCCAGCATGATGCGTCCCACGGGGCCATCAGCA						
EgD5 (서열번호 :184)	780	790	800	810	820	830	840
EgD5R (서열번호 :192)	AGAAGCCTTGGGTCAACGCCCTCTTCGCCTACGGCATTGACGTATCGGATCGTCCCGGTGGATTGGCT						
EgD5 (서열번호 :184)	850	860	870	880	890	900	910
EgD5R (서열번호 :192)	AGAAGCCTTGGGTCAACGCCCTCTTCGCCTACGGCATTGACGTATCGGATCGTCCCGGTGGATTGGCT						
EgD5 (서열번호 :184)	920	930	940	950	960	970	980
EgD5R (서열번호 :192)	GCAGTCGCACATCATGCGGCACCAACCTACACCAACCAAGCACGGCCTCGACCTGGATGCGGAGTCGGCA						
EgD5 (서열번호 :184)	990	1000	1010	1020	1030	1040	1050
EgD5R (서열번호 :192)	GCAGTCGCACATCATGCGGCACCAACCTACACCAACCAAGCACGGCCTCGACCTGGATGCGGAGTCGGCA						
EgD5 (서열번호 :184)	1060	1070	1080	1090	1100	1110	1120
EgD5R (서열번호 :192)	GAGCGGTCCTGGTGTCCACAACCTACCCCGCCGCAACACCGCCGAAAGTGGTTCACCGCTTCCAAG						
EgD5 (서열번호 :184)	1130	1140	1150	1160	1170	1180	1190
EgD5R (서열번호 :192)	GAGCGGTCCTGGTGTCCACAACCTACCCCGCCGCAACACCGCCGAAAGTGGTTCACCGCTTCCAAG						
EgD5 (서열번호 :184)	1200	1210	1220	1230	1240	1250	1260
EgD5R (서열번호 :192)	GAGCGGTCCTGGTGTCCACAACCTACCCCGCCGCAACACCGCCGAAAGTGGTTCACCGCTTCCAAG						
EgD5 (서열번호 :184)	1270	1280	1290	1300	1310	1320	1330
EgD5R (서열번호 :192)	TCGCCATGTTCTTCACTGGCGGTCTCAATTTCCAGATCGAGCACCACCTCTTCCCGGATGTCGTCTTG						
EgD5 (서열번호 :184)	1340	1350					
EgD5R (서열번호 :192)	TCGCCATGTTCTTCACTGGCGGTCTCAATTTCCAGATCGAGCACCACCTCTTCCCGGATGTCGTCTTG						
EgD5 (서열번호 :184)							
EgD5R (서열번호 :192)	GCACTACCCCTTCGTCCAGCAGGCGGTCCGGAGTGTTGCGAACGCCATGGAGTGCATATGTTTCTAC						
EgD5 (서열번호 :184)							
EgD5R (서열번호 :192)	GCACTACCCCTTCGTCCAGCAGGCGGTCCGGAGTGTTGCGAACGCCATGGAGTGCATATGTTTCTAC						
EgD5 (서열번호 :184)							
EgD5R (서열번호 :192)	CCTACCATCGTCGGCAACATCATCTCCACCTGAAGTACATGCATAAGGTGGGTGTCGTCCACTGCGTGA						
EgD5 (서열번호 :184)							
EgD5R (서열번호 :192)	CCTACCATCGTCGGCAACATCATCTCCACCTGAAGTACATGCATAAGGTGGGTGTCGTCCACTGCGTGA						
EgD5 (서열번호 :184)							
EgD5R (서열번호 :192)	AGGACGCACAGGATTCTTGA						
EgD5 (서열번호 :184)							
EgD5R (서열번호 :192)	AGGACGCACAGGATTCTTGA						

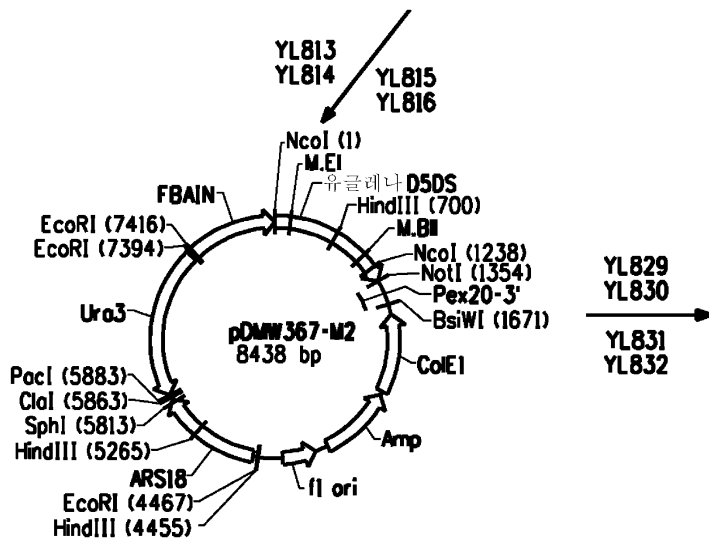
도면19b

EgD5 (서열번호 :184)	710	720	730	740	750	760	770
EgD5R (서열번호 :192)	CTTGGTACATGTACCTTGTGCTGGGGGCATACGGGGTATCGCTGGTGTACAACCCGCTCTACATTTTCCG						
EgD5 (서열번호 :184)	780	790	800	810	820	830	840
EgD5R (서열번호 :192)	CTTGGTACATGTACCTTGTGCTGGGGGCATACGGGGTATCGCTGGTGTACAACCCGCTCTACATTTTCCG						
EgD5 (서열번호 :184)	850	860	870	880	890	900	910
EgD5R (서열번호 :192)	GATGCAGACAATGACACCATCCAGAGTCTGTACGGCCATGCGGGAATGGCTTCTGCGGCGCTAC						
EgD5 (서열번호 :184)	920	930	940	950	960	970	980
EgD5R (서열번호 :192)	GATGCAGACAATGACACCATCCAGAGTCTGTACGGCCATGCGGGAATGGCTTCTGCGGCGCTAC						
EgD5 (서열번호 :184)	990	1000	1010	1020	1030	1040	1050
EgD5R (서열번호 :192)	CGCACACTTGCATTTCGTGATGCGAGCTTTCTTTCATCTTCCGGACCGCATTCCTGCCCTGGTACCTCACTG						
EgD5 (서열번호 :184)	1060	1070	1080	1090	1100	1110	1120
EgD5R (서열번호 :192)	CGCACACTTGCATTTCGTGATGCGAGCTTTCTTTCATCTTCCGGACCGCATTCCTGCCCTGGTACCTCACTG						
EgD5 (서열번호 :184)	1130	1140	1150	1160	1170	1180	1190
EgD5R (서열번호 :192)	GGACCTCATTGCTGATCACCAATTCCTCTGGTGCCCACTGCAACTGGTGCCTTCTTGACGTTCTTCTCAT						
EgD5 (서열번호 :184)	1200	1210	1220	1230	1240	1250	1260
EgD5R (서열번호 :192)	GGACCTCATTGCTGATCACCAATTCCTCTGGTGCCCACTGCAACTGGTGCCTTCTTGACGTTCTTCTCAT						
EgD5 (서열번호 :184)	1270	1280	1290	1300	1310	1320	1330
EgD5R (서열번호 :192)	TTTGTCCCACAATTTTGATGGCTCCGAACGGATCCCGACAAGAACTGCAAGGTTAAGAGCTCTGAGAAG						
EgD5 (서열번호 :184)	1340	1350					
EgD5R (서열번호 :192)	TTTGTCCCACAATTTTGATGGCTCCGAACGGATCCCGACAAGAACTGCAAGGTTAAGAGCTCTGAGAAG						
EgD5 (서열번호 :184)							
EgD5R (서열번호 :192)	GACGTTGAGGCTGACCAAATTTGACTGGTATCGGGCGCAGGTGGAGACGTCCTCCACATACGGTGCCCCA						
EgD5 (서열번호 :184)							
EgD5R (서열번호 :192)	GACGTTGAGGCTGACCAAATTTGACTGGTATCGGGCGCAGGTGGAGACGTCCTCCACATACGGTGCCCCA						
EgD5 (서열번호 :184)							
EgD5R (서열번호 :192)	TCGCCATGTTCTTCACTGGCGGTCTCAATTTCCAGATCGAGCACCACCTCTTCCCGGATGTCGTCTTG						
EgD5 (서열번호 :184)							
EgD5R (서열번호 :192)	TCGCCATGTTCTTCACTGGCGGTCTCAATTTCCAGATCGAGCACCACCTCTTCCCGGATGTCGTCTTG						
EgD5 (서열번호 :184)							
EgD5R (서열번호 :192)	GCACTACCCCTTCGTCCAGCAGGCGGTCCGGAGTGTTGCGAACGCCATGGAGTGCATATGTTTCTAC						
EgD5 (서열번호 :184)							
EgD5R (서열번호 :192)	GCACTACCCCTTCGTCCAGCAGGCGGTCCGGAGTGTTGCGAACGCCATGGAGTGCATATGTTTCTAC						
EgD5 (서열번호 :184)							
EgD5R (서열번호 :192)	CCTACCATCGTCGGCAACATCATCTCCACCTGAAGTACATGCATAAGGTGGGTGTCGTCCACTGCGTGA						
EgD5 (서열번호 :184)							
EgD5R (서열번호 :192)	CCTACCATCGTCGGCAACATCATCTCCACCTGAAGTACATGCATAAGGTGGGTGTCGTCCACTGCGTGA						
EgD5 (서열번호 :184)							
EgD5R (서열번호 :192)	AGGACGCACAGGATTCTTGA						
EgD5 (서열번호 :184)							
EgD5R (서열번호 :192)	AGGACGCACAGGATTCTTGA						

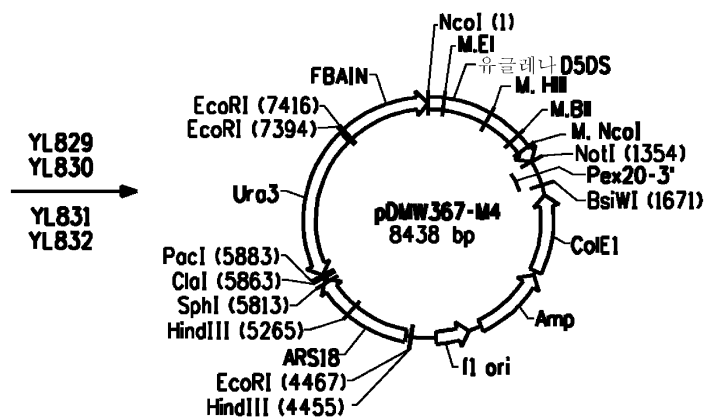
도면20a



도면20b



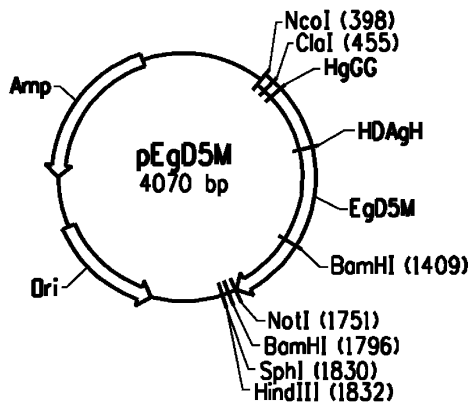
도면20c



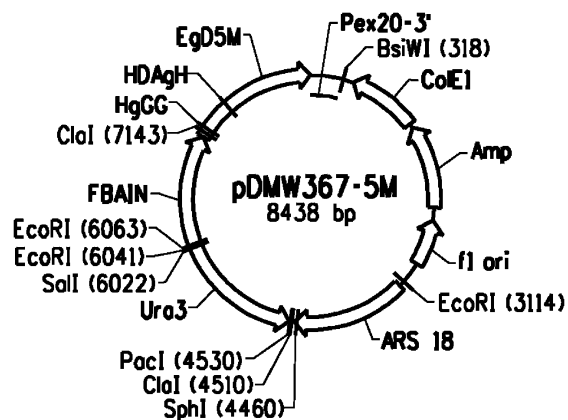
도면21

EgD5R (서열번호 :192)	ATGGCTCTCAGTCTTACCAACAGACAGCTGTTAGAACGCCCTGATTGGTTGCGATTGAT
EgD5M (서열번호 :105)	ATGGCCTGTCTCTCCTACCGAACAGCTCCTGGAGCGACCTGATCTCGTGGCTATCGAT
EgD5R (서열번호 :192)	GGCATCCTCTACGACCTTGAAGGGCTTGCCAAAGTTTCATCCAGGAGGAGATTTGATTCTC
EgD5M (서열번호 :105)	GGTATCCTGTACGACCTCGAAGGGCTTGCCAAAGTGCATGGTGGTGGAGACCTCAATTCTG
EgD5R (서열번호 :192)	GCTTCTGGTGCCTCTGATGCCTCCCTCTCTTTATTCATGCATCCATACGTCAAACCG
EgD5M (서열번호 :105)	GCCTCTGGAGCCTCCGACGCCTCTCCCTCTTCTACTCTATGCATCCCTACGTCAAAGCCC
EgD5R (서열번호 :192)	GAGAAATCCAAATTTGCTTCAACAGTTCGTCCGAGGGAAGCATGACCGCACCTCGAAGGAC
EgD5M (서열번호 :105)	GAGAACTCCAAGCTCCTGCAGCAATTTCGTCCGAGGGAAGCATGACCGCACCTCGAAGGAC

도면22a



도면22b



서열 목록

SEQUENCE LISTING

<110> E.I. duPont de Nemours & Co., Inc.

Zhu, Quinn

<120> FURTHER IMPROVED OPTIMIZED STRAINS OF YARROWIA LIPOLYTICA FOR
HIGH EICOSAPENTAENOIC ACID PRODUCTION

<130> CL5411

<150> US61/479921

<151> 2011-04-28

<160> 435

<170> PatentIn version 3.5

<210> 1

<211> 258

<212> PRT

<213> *Euglena gracilis*

<220><221> MISC_FEATURE

<222> (1)..(258)

<223> mutant delta-9 elongase; EgD9eS-mutant consensus

<220><221> MISC_FEATURE

<222> (9)..(9)

<223> Xaa = Ser [S] (synthetic codon-optimized) or (Ala [A] or Asp [D]
or Gly [G] or Ile [I] or Lys [K] or Gln [Q]) (mutant)

<220><221> MISC_FEATURE

<222> (12)..(12)

<223> Xaa = Gln [Q] (synthetic codon-optimized) or Lys [K] (mutant)

<220><221> MISC_FEATURE

<222> (21)..(21)

<223> Xaa = Ala [A] (synthetic codon-optimized) or (Asp [D] or Thr [T]
or Val [V]) (mutant)

<220><221> MISC_FEATURE

<222> (32)..(32)

<223> Xaa = Val [V] (synthetic codon-optimized) or Phe [F] (mutant)

<220><221> MISC_FEATURE

<222> (35)..(35)

<223> Xaa = Leu [L] (synthetic codon-optimized) or (Phe [F] or Gly [G]
or Met [M]) (mutant)

<220><221> MISC_FEATURE

<222> (58)..(58)

<223> Xaa = Lys [K] (synthetic codon-optimized) or Arg [R] (mutant)

<220><221> MISC_FEATURE

<222> (84)..(84)

<223> Xaa = Tyr [Y] (synthetic codon-optimized) or Cys [C] (mutant)

<220><221> MISC_FEATURE

<222> (98)..(98)

<223> Xaa = Asp [D] (synthetic codon-optimized) or Gly [G] (mutant)

<220>

><221> MISC_FEATURE

<222> (107)..(107)

<223> Xaa = Gln [Q] (synthetic codon-optimized) or Glu [E] (mutant)

<220><221> MISC_FEATURE

<222> (108)..(108)

<223> Xaa = Leu [L] (synthetic codon-optimized) or Gly [G] (mutant)

<220><221> MISC_FEATURE

<222> (127)..(127)

<223> Xaa = Gly [G] (synthetic codon-optimized) or Leu [L] (mutant)

<220><221> MISC_FEATURE

<222> (130)..(130)

<223> Xaa = Leu [L] (synthetic codon-optimized) or Met [M] (mutant)

<220><221> MISC_FEATURE

<222> (132)..(132)

<223> Xaa = Trp [W] (synthetic codon-optimized) or Thr [T] (mutant)

<220><221> MISC_FEATURE

<222> (143)..(143)

<223> Xaa = Met [M] (synthetic codon-optimized) or (Asn [N] or Trp [W]) (mutant)

<220><221> MISC_FEATURE

<222> (161)..(161)

<223> Xaa = Leu [L] (synthetic codon-optimized) or (Thr [T] or Tyr [Y]) (mutant)

<220><221> MISC_FEATURE

<222> (168)..(168)

<223> Xaa = Trp [W] (synthetic codon-optimized) or Gly [G] (mutant)

<220><221> MISC_FEATURE

<222> (179)..(179)

<223> Xaa = Ile [I] (synthetic codon-optimized) or (Met [M] or Arg [R]) (mutant)

<220><221> MISC_FEATURE

<222> (236)..(236)

<223> Xaa = Cys [C] (synthetic codon-optimized) or Asn [N] (mutant)

<220><221> MISC_FEATURE

<222> (243)..(243)

<223> Xaa = Val [V] (synthetic codon-optimized) or Ala [A] (mutant)

<220><221> MISC_FEATURE

<222> (244)..(244)

<223> Xaa = Gln [Q] (synthetic codon-optimized) or Asn [N] (mutant)

<220><221> MISC_FEATURE

<222> (254)..(254)

<223> Xaa = Ala [A] (synthetic codon-optimized) or (Trp [W] or Tyr [Y]) (mutant)

<220><221> MISC_FEATURE

<222> (257)..(257)

<223> Xaa = Ile [I] (synthetic codon-optimized) or Thr [T] (mutant)

<400> 1

Met Glu Val Val Asn Glu Ile Val Xaa Ile Gly Xaa Glu Val Leu Pro

1 5 10 15

Lys Val Asp Tyr Xaa Gln Leu Trp Ser Asp Ala Ser His Cys Glu Xaa

 20 25 30

Leu Tyr Xaa Ser Ile Ala Phe Val Ile Leu Lys Phe Thr Leu Gly Pro

 35 40 45

Leu Gly Pro Lys Gly Gln Ser Arg Met Xaa Phe Val Phe Thr Asn Tyr

 50 55 60

Asn Leu Leu Met Ser Ile Tyr Ser Leu Gly Ser Phe Leu Ser Met Ala

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

 $\langle 210 \rangle$ 2

<211> 777

<212> DNA

<213> Euglena gracilis

<220><221> CDS

<222> (1)..(777)

<223> synthetic delta-9 elongase (codon-optimized for Yarrowia

lipolytica)

<300><302> DELTA-9 ELONGASES AND THEIR USE IN MAKING POLYUNSATURATED FATTY
ACIDS

<310> WO 2007/061742

<311> 2006-11-16

<312> 2007-05-31

<313> (1)..(777)

<300><302> DELTA-9 ELONGASES AND THEIR USE IN MAKING POLYUNSATURATED FATTY
ACIDS

<310> US 7,645,604

<311> 2006-11-16

<312> 2010-01-12

<313> (1)..(777)

<400> 2

atg gag gtc gtg aac gaa atc gtc tcc att ggc cag gag gtt ctt ccc 48

Met Glu Val Val Asn Glu Ile Val Ser Ile Gly Gln Glu Val Leu Pro

1 5 10 15

aag gtc gac tat gct cag ctc tgg tct gat gcc tcg cac tgc gag gtg 96

Lys Val Asp Tyr Ala Gln Leu Trp Ser Asp Ala Ser His Cys Glu Val

20 25 30

ctg tac ctc tcc atc gcc ttc gtc atc ctg aag ttc acc ctt ggt cct 144

Leu Tyr Leu Ser Ile Ala Phe Val Ile Leu Lys Phe Thr Leu Gly Pro

35 40 45

ctc gga ccc aag ggt cag tct cga atg aag ttt gtg ttc acc aac tac 192

Leu Gly Pro Lys Gly Gln Ser Arg Met Lys Phe Val Phe Thr Asn Tyr

50 55 60

aac ctg ctc atg tcc atc tac tcg ctg ggc tcc ttc ctc tct atg gcc 240

Asn Leu Leu Met Ser Ile Tyr Ser Leu Gly Ser Phe Leu Ser Met Ala

65 70 75 80

tac gcc atg tac acc att ggt gtc atg tcc gac aac tgc gag aag gct 288

Tyr Ala Met Tyr Thr Ile Gly Val Met Ser Asp Asn Cys Glu Lys Ala

85 90 95

ttc gac aac aat gtc ttc cga atc acc act cag ctg ttc tac ctc agc 336

 Phe Asp Asn Asn Val Phe Arg Ile Thr Thr Gln Leu Phe Tyr Leu Ser
 100 105 110
 aag ttc ctc gag tac att gac tcc ttc tat ctg ccc ctc atg ggc aag 384
 Lys Phe Leu Glu Tyr Ile Asp Ser Phe Tyr Leu Pro Leu Met Gly Lys
 115 120 125
 cct ctg acc tgg ttg cag ttc ttt cac cat ctc gga gct cct atg gac 432
 Pro Leu Thr Trp Leu Gln Phe Phe His His Leu Gly Ala Pro Met Asp

 130 135 140
 atg tgg ctg ttc tac aac tac cga aac gaa gcc gtt tgg atc ttt gtg 480
 Met Trp Leu Phe Tyr Asn Tyr Arg Asn Glu Ala Val Trp Ile Phe Val
 145 150 155 160
 ctg ctc aac ggc ttc att cac tgg atc atg tac ggc tac tat tgg acc 528
 Leu Leu Asn Gly Phe Ile His Trp Ile Met Tyr Gly Tyr Tyr Trp Thr
 165 170 175

 cga ctg atc aag ctc aag ttc cct atg ccc aag tcc ctg att act tct 576
 Arg Leu Ile Lys Leu Lys Phe Pro Met Pro Lys Ser Leu Ile Thr Ser
 180 185 190
 atg cag atc att cag ttc aac gtt ggc ttc tac atc gtc tgg aag tac 624
 Met Gln Ile Ile Gln Phe Asn Val Gly Phe Tyr Ile Val Trp Lys Tyr
 195 200 205
 cgg aac att ccc tgc tac cga caa gat gga atg aga atg ttt ggc tgg 672

 Arg Asn Ile Pro Cys Tyr Arg Gln Asp Gly Met Arg Met Phe Gly Trp
 210 215 220
 ttt ttc aac tac ttc tac gtt ggt act gtc ctg tgt ctg ttc ctc aac 720
 Phe Phe Asn Tyr Phe Tyr Val Gly Thr Val Leu Cys Leu Phe Leu Asn
 225 230 235 240
 ttc tac gtg cag acc tac atc gtc cga aag cac aag gga gcc aaa aag 768
 Phe Tyr Val Gln Thr Tyr Ile Val Arg Lys His Lys Gly Ala Lys Lys

245 250 255

att cag tga 777

Ile Gln

<210> 3

<211> 258

<212> PRT

<213> Euglena gracilis

<400> 3

Met Glu Val Val Asn Glu Ile Val Ser Ile Gly Gln Glu Val Leu Pro

1 5 10 15

Lys Val Asp Tyr Ala Gln Leu Trp Ser Asp Ala Ser His Cys Glu Val

20 25 30

Leu Tyr Leu Ser Ile Ala Phe Val Ile Leu Lys Phe Thr Leu Gly Pro

35 40 45

Leu Gly Pro Lys Gly Gln Ser Arg Met Lys Phe Val Phe Thr Asn Tyr

50 55 60

Asn Leu Leu Met Ser Ile Tyr Ser Leu Gly Ser Phe Leu Ser Met Ala

65 70 75 80

Tyr Ala Met Tyr Thr Ile Gly Val Met Ser Asp Asn Cys Glu Lys Ala

85 90 95

Phe Asp Asn Asn Val Phe Arg Ile Thr Thr Gln Leu Phe Tyr Leu Ser

100 105 110

Lys Phe Leu Glu Tyr Ile Asp Ser Phe Tyr Leu Pro Leu Met Gly Lys

115 120 125

Pro Leu Thr Trp Leu Gln Phe Phe His His Leu Gly Ala Pro Met Asp

130 135 140

Met Trp Leu Phe Tyr Asn Tyr Arg Asn Glu Ala Val Trp Ile Phe Val

145 150 155 160

Leu Leu Asn Gly Phe Ile His Trp Ile Met Tyr Gly Tyr Tyr Trp Thr

165 170 175

Arg Leu Ile Lys Leu Lys Phe Pro Met Pro Lys Ser Leu Ile Thr Ser

180 185 190
Met Gln Ile Ile Gln Phe Asn Val Gly Phe Tyr Ile Val Trp Lys Tyr
195 200 205

Arg Asn Ile Pro Cys Tyr Arg Gln Asp Gly Met Arg Met Phe Gly Trp
210 215 220
Phe Phe Asn Tyr Phe Tyr Val Gly Thr Val Leu Cys Leu Phe Leu Asn
225 230 235 240
Phe Tyr Val Gln Thr Tyr Ile Val Arg Lys His Lys Gly Ala Lys Lys
245 250 255
Ile Gln

<210> 4

<211> 24

<212> PRT

<213> Artificial Sequence

<220><223> Multizyme linker

<300><302> MULTIZYMES AND THEIR USE IN MAKING POLYUNSATURATED FATTY ACIDS

<310> WO 2008/124048

<311> 2008-04-03

<312> 2008-10-16

<313> (1)..(24)

<300><302> MULTIZYMES AND THEIR USE IN MAKING POLYUNSATURATED FATTY ACIDS

<310> US-2008-0254191-A1

<311> 2008-04-03

<312> 2008-10-16

<313> (1)..(24)

<400> 4

Gly Ala Gly Pro Ala Arg Pro Ala Gly Leu Pro Pro Ala Thr Tyr Tyr
1 5 10 15
Asp Ser Leu Ala Val Met Gly Ser
20

<210> 5

<211> 19
 <212> PRT
 <213> Artificial Sequence
 <220><223> Multizyme linker
 <300><302> MULTIZYMES AND THEIR USE IN MAKING POLYUNSATURATED FATTY ACIDS
 <310> US-2008-0254191-A1
 <311> 2008-04-03
 <312> 2008-10-16
 <313> (1)..(19)
 <300><302> MULTIZYMES AND THEIR USE IN MAKING POLYUNSATURATED FATTY ACIDS
 <310> WO 2008/124048
 <311> 2008-04-03
 <312> 2008-10-16
 <313> (1)..(19)
 <400> 5

Gly Pro Ala Arg Pro Ala Gly Leu Pro Pro Ala Thr Tyr Tyr Asp Ser

1 5 10 15

Leu Ala Val

<210> 6
 <211> 18
 <212> PRT
 <213> Artificial Sequence
 <220><223> Multizyme linker
 <300><302> MULTIZYMES AND THEIR USE IN MAKING POLYUNSATURATED FATTY ACIDS
 <310> WO 2008/124048
 <311> 2008-04-03
 <312> 2008-10-16
 <313> (1)..(18)
 <300><302> MULTIZYMES AND THEIR USE IN MAKING POLYUNSATURATED FATTY ACIDS
 <310> US-2008-0254191-A1
 <311> 2008-04-03
 <312> 2008-10-16
 <313> (1)..(18)

<400> 6

Pro Ala Arg Pro Ala Gly Leu Pro Pro Ala Thr Tyr Tyr Asp Ser Leu

1 5 10 15

Ala Val

<210> 7

<211> 18

<212> PRT

<213> Artificial Sequence

<220><223> Multizyme linker

<300><302> MULTIZYMES AND THEIR USE IN MAKING POLYUNSATURATED FATTY ACIDS

<310> WO 2008/124048

<311> 2008-04-03

<312> 2008-10-16

<313> (1)..(18)

<300><302> MULTIZYMES AND THEIR USE IN MAKING POLYUNSATURATED FATTY ACIDS

<310> US-2008-0254191-A1

<311> 2008-04-03

<312> 2008-10-16

<313> (1)..(18)

<400>

7

Pro Thr Arg Pro Ala Gly Pro Pro Pro Ala Thr Tyr Tyr Asp Ser Leu

1 5 10 15

Ala Val

<210> 8

<211> 33

<212> PRT

<213> Artificial Sequence

<220><223> Multizyme linker

<300><302> MULTIZYMES AND THEIR USE IN MAKING POLYUNSATURATED FATTY ACIDS

<310> WO 2008/124048

<311> 2008-04-03

<312> 2008-10-16

<313> (1)..(33)

<300><302> MULTIZYMES AND THEIR USE IN MAKING POLYUNSATURATED FATTY ACIDS

<310> US-2008-0254191-A1

<311> 2008-04-03

<312> 2008-10-16

<313> (1)..(33)

<400> 8

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

1 5 10 15

Ile Arg Pro Val Gly Asn Pro Pro Ala Ala Tyr Tyr Asp Ala Leu Ala

20 25 30

Thr

<210> 9

<211> 22

<212> PRT

<213> Artificial Sequence

<220><223> Multizyme linker

<300><302> MULTIZYMES AND THEIR USE IN MAKING POLYUNSATURATED FATTY ACIDS

<310> WO 2008/124048

<311> 2008-04-03

<312> 2008-10-16

<313> (1)..(22)

<300><302> MULTIZYMES AND THEIR USE IN MAKING POLYUNSATURATED FATTY ACIDS

<310> US-2008-0254191-A1

<311> 2008-04-03

<312> 2008-10-16

<313> (1)..(22)

<400> 9

Pro Ala Arg Pro Ala Gly Leu Pro Pro Ala Thr Tyr Tyr Asp Ser Leu

1 5 10 15

Ala Val Ser Gly Arg Thr

20

<210> 10

<211> 36

<212> PRT

<213> Artificial Sequence

<220

><223> Multizyme linker

<300><302> MULTIZYMES AND THEIR USE IN MAKING POLYUNSATURATED FATTY ACIDS

<310> WO 2008/124048

<311> 2008-04-03

<312> 2008-10-16

<313> (1)..(36)

<300><302> MULTIZYMES AND THEIR USE IN MAKING POLYUNSATURATED FATTY ACIDS

<310> US-2008-0254191-A1

<311> 2008-04-03

<312> 2008-10-16

<313> (1)..(36)

<400> 10

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

1 5 10 15

Ile Arg Pro Val Gly Asn Pro Pro Ala Ala Tyr Tyr Asp Ala Leu Ala

20

25

30

Thr Gly Arg Thr

35

<210> 11

<211> 2112

<212> DNA

<213> Artificial Sequence

<220><223> Multizyme: EgD9ES/EgD8M gene fusion

<220><221> CDS

<222> (1)..(2112)

<223> Multizyme: EgD9ES/EgD8M gene fusion

<300><302> MULTIZYMES AND THEIR USE IN MAKING POLYUNSATURATED FATTY ACIDS

<310> US-2008-0254191-A1

<311> 2008-04-03

<312> 2008-10-16

<313> (1)..(2112)

<300><302> MULTIZYMES AND THEIR USE IN MAKING POLYUNSATURATED FATTY ACIDS

<310> WO 2008/124048

<311> 2008-04-03

<312> 2008-10-16

<313> (1)..(2112)

<400> 11

atg gag gtc gtg aac gaa atc gtc tcc att ggc cag gag gtt ctt ccc 48

Met Glu Val Val Asn Glu Ile Val Ser Ile Gly Gln Glu Val Leu Pro

1 5 10 15

aag gtc gac tat gct cag ctc tgg tct gat gcc tcg cac tgc gag gtg 96

Lys Val Asp Tyr Ala Gln Leu Trp Ser Asp Ala Ser His Cys Glu Val

20 25 30

ctg tac ctc tcc atc gcc ttc gtc atc ctg aag ttc acc ctt ggt cct 144

Leu Tyr Leu Ser Ile Ala Phe Val Ile Leu Lys Phe Thr Leu Gly Pro

35 40 45

ctc gga ccc aag ggt cag tct cga atg aag ttt gtg ttc acc aac tac 192

Leu Gly Pro Lys Gly Gln Ser Arg Met Lys Phe Val Phe Thr Asn Tyr

50 55 60

aac ctg ctc atg tcc atc tac tcg ctg ggc tcc ttc ctc tct atg gcc 240

Asn Leu Leu Met Ser Ile Tyr Ser Leu Gly Ser Phe Leu Ser Met Ala

65 70 75 80

tac gcc atg tac acc att ggt gtc atg tcc gac aac tgc gag aag gct 288

Tyr Ala Met Tyr Thr Ile Gly Val Met Ser Asp Asn Cys Glu Lys Ala

85 90 95

ttc gac aac aat gtc ttc cga atc acc act cag ctg ttc tac ctc agc 336

Phe Asp Asn Asn Val Phe Arg Ile Thr Thr Gln Leu Phe Tyr Leu Ser

100 105 110

aag ttc ctc gag tac att gac tcc ttc tat ctg ccc ctc atg ggc aag 384

Lys Phe Leu Glu Tyr Ile Asp Ser Phe Tyr Leu Pro Leu Met Gly Lys
115 120 125
cct ctg acc tgg ttg cag ttc ttt cac cat ctc gga gct cct atg gac 432
Pro Leu Thr Trp Leu Gln Phe Phe His His Leu Gly Ala Pro Met Asp
130 135 140
atg tgg ctg ttc tac aac tac cga aac gaa gcc gtt tgg atc ttt gtg 480
Met Trp Leu Phe Tyr Asn Tyr Arg Asn Glu Ala Val Trp Ile Phe Val
145 150 155 160
ctg ctc aac ggc ttc att cac tgg atc atg tac ggc tac tat tgg acc 528
Leu Leu Asn Gly Phe Ile His Trp Ile Met Tyr Gly Tyr Tyr Trp Thr
165 170 175
cga ctg atc aag ctc aag ttc cct atg ccc aag tcc ctg att act tct 576
Arg Leu Ile Lys Leu Lys Phe Pro Met Pro Lys Ser Leu Ile Thr Ser
180 185 190
atg cag atc att cag ttc aac gtt ggc ttc tac atc gtc tgg aag tac 624
Met Gln Ile Ile Gln Phe Asn Val Gly Phe Tyr Ile Val Trp Lys Tyr
195 200 205
cgg aac att ccc tgc tac cga caa gat gga atg aga atg ttt ggc tgg 672
Arg Asn Ile Pro Cys Tyr Arg Gln Asp Gly Met Arg Met Phe Gly Trp
210 215 220
ttt ttc aac tac ttc tac gtt ggt act gtc ctg tgt ctg ttc ctc aac 720
Phe Phe Asn Tyr Phe Tyr Val Gly Thr Val Leu Cys Leu Phe Leu Asn
225 230 235 240
ttc tac gtg cag acc tac atc gtc cga aag cac aag gga gcc aaa aag 768
Phe Tyr Val Gln Thr Tyr Ile Val Arg Lys His Lys Gly Ala Lys Lys
245 250 255
att cag ggc gcc ggt ccc gct cga cct gcc gga ctt cct ccc gct acc 816
Ile Gln Gly Ala Gly Pro Ala Arg Pro Ala Gly Leu Pro Pro Ala Thr
260 265 270
tac tac gac tct ctg gcc gtc atg gga tcc gtg aag gct tct cga cag 864
Tyr Tyr Asp Ser Leu Ala Val Met Gly Ser Val Lys Ala Ser Arg Gln

275	280	285	
gct ctg ccc ctc gtc atc gac gga aag gtg tac gac gtc tcc gct tgg			912
Ala Leu Pro Leu Val Ile Asp Gly Lys Val Tyr Asp Val Ser Ala Trp			
290	295	300	
gtg aac ttc cac cct ggt gga gct gaa atc att gag aac tac cag gga			960
Val Asn Phe His Pro Gly Gly Ala Glu Ile Ile Glu Asn Tyr Gln Gly			
305	310	315	320
cga gat gct act gac gcc ttc atg gtt atg cac tct cag gaa gcc ttc			1008
Arg Asp Ala Thr Asp Ala Phe Met Val Met His Ser Gln Glu Ala Phe			
325	330	335	
gac aag ctc aag cga atg ccc aag atc aac cag gct tcc gag ctg cct			1056
Asp Lys Leu Lys Arg Met Pro Lys Ile Asn Gln Ala Ser Glu Leu Pro			
340	345	350	
ccc cag gct gcc gtc aac gaa gct cag gag gat ttc cga aag ctc cga			1104
Pro Gln Ala Ala Val Asn Glu Ala Gln Glu Asp Phe Arg Lys Leu Arg			
355	360	365	
gaa gag ctg atc gcc act ggc atg ttt gac gcc tct ccc ctc tgg tac			1152
Glu Glu Leu Ile Ala Thr Gly Met Phe Asp Ala Ser Pro Leu Trp Tyr			
370	375	380	
tcg tac aag atc ttg acc acc ctg ggt ctt ggc gtg ctt gcc ttc ttc			1200
Ser Tyr Lys Ile Leu Thr Thr Leu Gly Leu Gly Val Leu Ala Phe Phe			
385	390	395	400
atg ctg gtc cag tac cac ctg tac ttc att ggt gct ctc gtg ctc ggt			1248
Met Leu Val Gln Tyr His Leu Tyr Phe Ile Gly Ala Leu Val Leu Gly			
405	410	415	
atg cac tac cag caa atg gga tgg ctg tct cat gac atc tgc cac cac			1296
Met His Tyr Gln Gln Met Gly Trp Leu Ser His Asp Ile Cys His His			
420	425	430	
cag acc ttc aag aac cga aac tgg aat aac gtc ctg ggt ctg gtc ttt			1344

Gln Thr Phe Lys Asn Arg Asn Trp Asn Asn Val Leu Gly Leu Val Phe

435 440 445

ggc aac gga ctc cag ggc ttc tcc gtg acc tgg tgg aag gac aga cac 1392

Gly Asn Gly Leu Gln Gly Phe Ser Val Thr Trp Trp Lys Asp Arg His

450 455 460

aac gcc cat cat tct gct acc aac gtt cag ggt cac gat ccc gac att 1440

Asn Ala His His Ser Ala Thr Asn Val Gln Gly His Asp Pro Asp Ile

465 470 475 480

gat aac ctg cct ctg ctc gcc tgg tcc gag gac gat gtc act cga gct 1488

Asp Asn Leu Pro Leu Leu Ala Trp Ser Glu Asp Asp Val Thr Arg Ala

485 490 495

tct ccc atc tcc cga aag ctc att cag ttc caa cag tac tat ttc ctg 1536

Ser Pro Ile Ser Arg Lys Leu Ile Gln Phe Gln Gln Tyr Tyr Phe Leu

500 505 510

gtc atc tgt att ctc ctg cga ttc atc tgg tgt ttc cag tct gtg ctg 1584

Val Ile Cys Ile Leu Leu Arg Phe Ile Trp Cys Phe Gln Ser Val Leu

515 520 525

acc gtt cga tcc ctc aag gac cga gac aac cag ttc tac cga tct cag 1632

Thr Val Arg Ser Leu Lys Asp Arg Asp Asn Gln Phe Tyr Arg Ser Gln

530 535 540

tac aag aaa gag gcc att gga ctc gct ctg cac tgg act ctc aag acc 1680

Tyr Lys Lys Glu Ala Ile Gly Leu Ala Leu His Trp Thr Leu Lys Thr

545 550 555 560

ctg ttc cac ctc ttc ttt atg ccc tcc atc ctg acc tcg atg ctg gtg 1728

Leu Phe His Leu Phe Phe Met Pro Ser Ile Leu Thr Ser Met Leu Val

565 570 575

ttc ttt gtt tcc gag ctc gtc ggt ggc ttc gga att gcc atc gtg gtc 1776

Phe Phe Val Ser Glu Leu Val Gly Gly Phe Gly Ile Ala Ile Val Val

580 585 590

ttc atg aac cac tac cct ctg gag aag atc ggt gat tcc gtc tgg gac 1824
 Phe Met Asn His Tyr Pro Leu Glu Lys Ile Gly Asp Ser Val Trp Asp
 595 600 605
 gga cat ggc ttc tct gtg ggt cag atc cat gag acc atg aac att cga 1872
 Gly His Gly Phe Ser Val Gly Gln Ile His Glu Thr Met Asn Ile Arg
 610 615 620
 cga ggc atc att act gac tgg ttc ttt gga ggc ctg aac tac cag atc 1920

 Arg Gly Ile Ile Thr Asp Trp Phe Phe Gly Gly Leu Asn Tyr Gln Ile
 625 630 635 640
 gag cac cat ctc tgg ccc acc ctg cct cga cac aac ctc act gcc gtt 1968
 Glu His His Leu Trp Pro Thr Leu Pro Arg His Asn Leu Thr Ala Val
 645 650 655
 tcc tac cag gtg gaa cag ctg tgc cag aag cac aac ctc ccc tac cga 2016
 Ser Tyr Gln Val Glu Gln Leu Cys Gln Lys His Asn Leu Pro Tyr Arg
 660 665 670
 aac cct ctg ccc cat gaa ggt ctc gtc atc ctg ctc cga tac ctg tcc 2064
 Asn Pro Leu Pro His Glu Gly Leu Val Ile Leu Leu Arg Tyr Leu Ser
 675 680 685
 cag ttc gct cga atg gcc gag aag cag ccc ggt gcc aag gct cag taa 2112
 Gln Phe Ala Arg Met Ala Glu Lys Gln Pro Gly Ala Lys Ala Gln
 690 695 700

 <210> 12
 <211> 703
 <212> PRT
 <213> Artificial Sequence
 <220><223> Synthetic Construct
 <400> 12
 Met Glu Val Val Asn Glu Ile Val Ser Ile Gly Gln Glu Val Leu Pro
 1 5 10 15
 Lys Val Asp Tyr Ala Gln Leu Trp Ser Asp Ala Ser His Cys Glu Val
 20 25 30
 Leu Tyr Leu Ser Ile Ala Phe Val Ile Leu Lys Phe Thr Leu Gly Pro

35 40 45
 Leu Gly Pro Lys Gly Gln Ser Arg Met Lys Phe Val Phe Thr Asn Tyr

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

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

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

 245 250 255
 Ile Gln Gly Ala Gly Pro Ala Arg Pro Ala Gly Leu Pro Pro Ala Thr
 260 265 270
 Tyr Tyr Asp Ser Leu Ala Val Met Gly Ser Val Lys Ala Ser Arg Gln
 275 280 285

Ala Leu Pro Leu Val Ile Asp Gly Lys Val Tyr Asp Val Ser Ala Trp
290 295 300
Val Asn Phe His Pro Gly Gly Ala Glu Ile Ile Glu Asn Tyr Gln Gly
305 310 315 320
Arg Asp Ala Thr Asp Ala Phe Met Val Met His Ser Gln Glu Ala Phe
325 330 335
Asp Lys Leu Lys Arg Met Pro Lys Ile Asn Gln Ala Ser Glu Leu Pro
340 345 350
Pro Gln Ala Ala Val Asn Glu Ala Gln Glu Asp Phe Arg Lys Leu Arg
355 360 365
Glu Glu Leu Ile Ala Thr Gly Met Phe Asp Ala Ser Pro Leu Trp Tyr
370 375 380
Ser Tyr Lys Ile Leu Thr Thr Leu Gly Leu Gly Val Leu Ala Phe Phe
385 390 395 400
Met Leu Val Gln Tyr His Leu Tyr Phe Ile Gly Ala Leu Val Leu Gly
405 410 415
Met His Tyr Gln Gln Met Gly Trp Leu Ser His Asp Ile Cys His His
420 425 430
Gln Thr Phe Lys Asn Arg Asn Trp Asn Asn Val Leu Gly Leu Val Phe
435 440 445
Gly Asn Gly Leu Gln Gly Phe Ser Val Thr Trp Trp Lys Asp Arg His
450 455 460
Asn Ala His His Ser Ala Thr Asn Val Gln Gly His Asp Pro Asp Ile
465 470 475 480
Asp Asn Leu Pro Leu Leu Ala Trp Ser Glu Asp Asp Val Thr Arg Ala
485 490 495
Ser Pro Ile Ser Arg Lys Leu Ile Gln Phe Gln Gln Tyr Tyr Phe Leu
500 505 510
Val Ile Cys Ile Leu Leu Arg Phe Ile Trp Cys Phe Gln Ser Val Leu
515 520 525
Thr Val Arg Ser Leu Lys Asp Arg Asp Asn Gln Phe Tyr Arg Ser Gln

530 535 540
Tyr Lys Lys Glu Ala Ile Gly Leu Ala Leu His Trp Thr Leu Lys Thr
545 550 555 560
Leu Phe His Leu Phe Phe Met Pro Ser Ile Leu Thr Ser Met Leu Val

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

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

690 695 700
<210> 13
<211> 2109
<212> DNA
<213> Artificial Sequence
<220><223> Multizyme: EaD9ES/EaD8S gene fusion
<220><221> CDS
<222> (1)..(2109)
<223> Multizyme: EaD9ES/EaD8S gene fusion
<300><302> MULTIZYMES AND THEIR USE IN MAKING POLYUNSATURATED FATTY ACIDS
<310> WO 2008/124048
<311> 2008-04-03

<312> 2008-10-16

<313> (1)..(2109)

<300><302> MULTIZYMES AND THEIR USE IN MAKING POLYUNSATURATED FATTY ACIDS

<310> US-2008-0254191-A1

<311> 2008-04-03

<312> 2008-10-16

<313> (1)..(2109)

<400> 13

atg gag gct gcc aag gag ctg gtc tcc atc gtc cag gag gaa ctt ccc 48

Met Glu Ala Ala Lys Glu Leu Val Ser Ile Val Gln Glu Glu Leu Pro

1 5 10 15

aag gtg gac tac gcc cag ctc tgg cag gac gcc tcc tct tgc gag gtt 96

Lys Val Asp Tyr Ala Gln Leu Trp Gln Asp Ala Ser Ser Cys Glu Val

20 25 30

ctg tac ctc tcg gtc gct ttc gtg gcc atc aag ttc atg ctt cga cct 144

Leu Tyr Leu Ser Val Ala Phe Val Ala Ile Lys Phe Met Leu Arg Pro

35 40 45

ctg gac ctc aag cga caa gcc acc ctc aaa aag ctg ttc acc gca tac 192

Leu Asp Leu Lys Arg Gln Ala Thr Leu Lys Lys Leu Phe Thr Ala Tyr

50 55 60

aac ttt ctc atg tcc atc tac tcg ttc ggc tcc ttc ctg gcg atg gcc 240

Asn Phe Leu Met Ser Ile Tyr Ser Phe Gly Ser Phe Leu Ala Met Ala

65 70 75 80

tac gct ctc tct gtc act ggt att ctt tcc ggc gat tgt gag act gcc 288

Tyr Ala Leu Ser Val Thr Gly Ile Leu Ser Gly Asp Cys Glu Thr Ala

85 90 95

ttc aac aat gac gtg ttc cga atc acc act cag ctg ttc tac ctc agc 336

Phe Asn Asn Asp Val Phe Arg Ile Thr Thr Gln Leu Phe Tyr Leu Ser

100 105 110

aag ttc gtc gag tac atc gac tcc ttc tac ctt ccc ctc atg gac aag 384

Lys Phe Val Glu Tyr Ile Asp Ser Phe Tyr Leu Pro Leu Met Asp Lys

115	120	125	
ccc ttg tgc ttt ctg cag ttc ttt cac cat ctc gga gct ccc atc gac			432
Pro Leu Ser Phe Leu Gln Phe Phe His His Leu Gly Ala Pro Ile Asp			
130	135	140	
atg tgg ctg ttc tac aag tat cga aac gaa ggc gtc tgg atc ttt gtt			480
Met Trp Leu Phe Tyr Lys Tyr Arg Asn Glu Gly Val Trp Ile Phe Val			
145	150	155	160
ctg ctc aac ggc ttc att cac tgg atc atg tac ggt tac tat tgg acg			528
Leu Leu Asn Gly Phe Ile His Trp Ile Met Tyr Gly Tyr Tyr Trp Thr			
165	170	175	
cga ctc atc aag ctg aac ttc cct atg ccc aag aac ctc att acc tcc			576
Arg Leu Ile Lys Leu Asn Phe Pro Met Pro Lys Asn Leu Ile Thr Ser			
180	185	190	
atg caa att atc cag ttc aac gtc gga ttc tac atc gtc tgg aag tac			624
Met Gln Ile Ile Gln Phe Asn Val Gly Phe Tyr Ile Val Trp Lys Tyr			
195	200	205	
cga aac gtg ccc tgc tac cgg cag gac ggt atg cga atg ttt gcc tgg			672
Arg Asn Val Pro Cys Tyr Arg Gln Asp Gly Met Arg Met Phe Ala Trp			
210	215	220	
atc ttc aac tac tgg tat gtc ggc acg gtg ctg ctt ctg ttc ctc aac			720
Ile Phe Asn Tyr Trp Tyr Val Gly Thr Val Leu Leu Leu Phe Leu Asn			
225	230	235	240
ttc tac gtc cag acc tac att cgg aag cct cga aag aac cga ggc aaa			768
Phe Tyr Val Gln Thr Tyr Ile Arg Lys Pro Arg Lys Asn Arg Gly Lys			
245	250	255	
aag gag ggc gcc ggt ccc gct cga cct gcc gga ctt cct ccc gct acc			816
Lys Glu Gly Ala Gly Pro Ala Arg Pro Ala Gly Leu Pro Pro Ala Thr			
260	265	270	
tac tac gac tct ctg gcc gtc atg gga tcc atg gtc aag cga ccc gct			864
Tyr Tyr Asp Ser Leu Ala Val Met Gly Ser Met Val Lys Arg Pro Ala			
275	280	285	

ctg cct ctc acc gtg gac ggt gtc acc tac gac gtt tct gcc tgg ctc	912
Leu Pro Leu Thr Val Asp Gly Val Thr Tyr Asp Val Ser Ala Trp Leu	
290 295 300	
aac cac cat ccc gga ggt gcc gac att atc gag aac tac cga ggt cgg	960
Asn His His Pro Gly Gly Ala Asp Ile Ile Glu Asn Tyr Arg Gly Arg	
305 310 315 320	
gat gct acc gac gtc ttc atg gtt atg cac tcc gag aac gcc gtg tcc	1008
Asp Ala Thr Asp Val Phe Met Val Met His Ser Glu Asn Ala Val Ser	
325 330 335	
aaa ctc aga cga atg ccc atc atg gaa cct tcc tct ccc ctg act cca	1056
Lys Leu Arg Arg Met Pro Ile Met Glu Pro Ser Ser Pro Leu Thr Pro	
340 345 350	
aca cct ccc aag cca aac tcc gac gaa cct cag gag gat ttc cga aag	1104
Thr Pro Pro Lys Pro Asn Ser Asp Glu Pro Gln Glu Asp Phe Arg Lys	
355 360 365	
ctg cga gac gag ctc att gct gca ggc atg ttc gat gcc tct ccc atg	1152
Leu Arg Asp Glu Leu Ile Ala Ala Gly Met Phe Asp Ala Ser Pro Met	
370 375 380	
tgg tac gct tac aag acc ctg tgc act ctc gga ctg ggt gtc ctt gcc	1200
Trp Tyr Ala Tyr Lys Thr Leu Ser Thr Leu Gly Leu Gly Val Leu Ala	
385 390 395 400	
gtg ctg ttg atg acc cag tgg cac tgg tac ctg gtt ggt gct atc gtc	1248
Val Leu Leu Met Thr Gln Trp His Trp Tyr Leu Val Gly Ala Ile Val	
405 410 415	
ctc ggc att cac ttt caa cag atg gga tgg ctc tgc cac gac att tgc	1296
Leu Gly Ile His Phe Gln Gln Met Gly Trp Leu Ser His Asp Ile Cys	
420 425 430	
cat cac cag ctg ttc aag gac cga tcc atc aac aat gcc att ggc ctg	1344
His His Gln Leu Phe Lys Asp Arg Ser Ile Asn Asn Ala Ile Gly Leu	

435	440	445	
ctc ttc gga aac gtg ctt cag ggc ttt tct gtc act tgg tgg aag gac			1392
Leu Phe Gly Asn Val Leu Gln Gly Phe Ser Val Thr Trp Trp Lys Asp			
450	455	460	
cga cac aac gct cat cac tcc gcc acc aac gtg cag ggt cac gat ccc			1440
Arg His Asn Ala His His Ser Ala Thr Asn Val Gln Gly His Asp Pro			
465	470	475	480
gac atc gac aac ctg cct ctc ctg gcg tgg tcc aag gag gac gtc gag			1488
Asp Ile Asp Asn Leu Pro Leu Leu Ala Trp Ser Lys Glu Asp Val Glu			
485	490	495	
cga gct ggc ccg ttt tct cga cgg atg atc aag tac caa cag tat tac			1536
Arg Ala Gly Pro Phe Ser Arg Arg Met Ile Lys Tyr Gln Gln Tyr Tyr			
500	505	510	
ttc ttt ttc atc tgt gcc ctt ctg cga ttc atc tgg tgc ttt cag tcc			1584
Phe Phe Phe Ile Cys Ala Leu Leu Arg Phe Ile Trp Cys Phe Gln Ser			
515	520	525	
att cat act gcc acg ggt ctc aag gat cga agc aat cag tac tat cga			1632
Ile His Thr Ala Thr Gly Leu Lys Asp Arg Ser Asn Gln Tyr Tyr Arg			
530	535	540	
aga cag tac gag aag gag tcc gtc ggt ctg gca ctc cac tgg ggt etc			1680
Arg Gln Tyr Glu Lys Glu Ser Val Gly Leu Ala Leu His Trp Gly Leu			
545	550	555	560
aag gcc ttg ttc tac tat ttc tac atg ccc tcg ttt ctc acc gga ctc			1728
Lys Ala Leu Phe Tyr Tyr Phe Tyr Met Pro Ser Phe Leu Thr Gly Leu			
565	570	575	
atg gtg ttc ttt gtc tcc gag ctg ctt ggt ggc ttc gga att gcc atc			1776
Met Val Phe Phe Val Ser Glu Leu Leu Gly Gly Phe Gly Ile Ala Ile			
580	585	590	
gtt gtc ttc atg aac cac tac cct ctg gag aag att cag gac tcc gtg			1824
Val Val Phe Met Asn His Tyr Pro Leu Glu Lys Ile Gln Asp Ser Val			
595	600	605	

tgg gat ggt cat ggc ttc tgt gct gga cag att cac gag acc atg aac 1872
 Trp Asp Gly His Gly Phe Cys Ala Gly Gln Ile His Glu Thr Met Asn
 610 615 620
 gtt cag cga ggc ctc gtc aca gac tgg ttt ttc ggt ggc ctc aac tac 1920

Val Gln Arg Gly Leu Val Thr Asp Trp Phe Phe Gly Gly Leu Asn Tyr
 625 630 635 640
 cag atc gaa cat cac ctg tgg cct act ctt ccc aga cac aac ctc acc 1968
 Gln Ile Glu His His Leu Trp Pro Thr Leu Pro Arg His Asn Leu Thr
 645 650 655
 gct gcc tcc atc aaa gtg gag cag ctg tgc aag aag cac aac ctg ccc 2016
 Ala Ala Ser Ile Lys Val Glu Gln Leu Cys Lys Lys His Asn Leu Pro

660 665 670
 tac cga tct cct ccc atg ctc gaa ggt gtc ggc att ctt atc tcc tac 2064
 Tyr Arg Ser Pro Pro Met Leu Glu Gly Val Gly Ile Leu Ile Ser Tyr
 675 680 685
 ctg ggc acc ttc gct cga atg gtt gcc aag gca gac aag gcc taa 2109
 Leu Gly Thr Phe Ala Arg Met Val Ala Lys Ala Asp Lys Ala
 690 695 700

<210> 14
 <211> 702
 <212> PRT
 <213> Artificial Sequence
 <220><223> Synthetic Construct
 <400> 14

Met Glu Ala Ala Lys Glu Leu Val Ser Ile Val Gln Glu Glu Leu Pro
 1 5 10 15
 Lys Val Asp Tyr Ala Gln Leu Trp Gln Asp Ala Ser Ser Cys Glu Val
 20 25 30
 Leu Tyr Leu Ser Val Ala Phe Val Ala Ile Lys Phe Met Leu Arg Pro
 35 40 45
 Leu Asp Leu Lys Arg Gln Ala Thr Leu Lys Lys Leu Phe Thr Ala Tyr

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

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

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

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

Asn His His Pro Gly Gly Ala Asp Ile Ile Glu Asn Tyr Arg Gly Arg

305 310 315 320

Asp Ala Thr Asp Val Phe Met Val Met His Ser Glu Asn Ala Val Ser

325 330 335

Lys Leu Arg Arg Met Pro Ile Met Glu Pro Ser Ser Pro Leu Thr Pro

340 345 350

Thr Pro Pro Lys Pro Asn Ser Asp Glu Pro Gln Glu Asp Phe Arg Lys

355 360 365

Leu Arg Asp Glu Leu Ile Ala Ala Gly Met Phe Asp Ala Ser Pro Met

370 375 380

Trp Tyr Ala Tyr Lys Thr Leu Ser Thr Leu Gly Leu Gly Val Leu Ala

385 390 395 400

Val Leu Leu Met Thr Gln Trp His Trp Tyr Leu Val Gly Ala Ile Val

405 410 415

Leu Gly Ile His Phe Gln Gln Met Gly Trp Leu Ser His Asp Ile Cys

420 425 430

His His Gln Leu Phe Lys Asp Arg Ser Ile Asn Asn Ala Ile Gly Leu

435 440 445

Leu Phe Gly Asn Val Leu Gln Gly Phe Ser Val Thr Trp Trp Lys Asp

450 455 460

Arg His Asn Ala His His Ser Ala Thr Asn Val Gln Gly His Asp Pro

465 470 475 480

Asp Ile Asp Asn Leu Pro Leu Leu Ala Trp Ser Lys Glu Asp Val Glu

485 490 495

Arg Ala Gly Pro Phe Ser Arg Arg Met Ile Lys Tyr Gln Gln Tyr Tyr

500 505 510

Phe Phe Phe Ile Cys Ala Leu Leu Arg Phe Ile Trp Cys Phe Gln Ser

515 520 525

Ile His Thr Ala Thr Gly Leu Lys Asp Arg Ser Asn Gln Tyr Tyr Arg

530 535 540

Arg Gln Tyr Glu Lys Glu Ser Val Gly Leu Ala Leu His Trp Gly Leu

545 550 555 560
Lys Ala Leu Phe Tyr Tyr Phe Tyr Met Pro Ser Phe Leu Thr Gly Leu

565 570 575
Met Val Phe Phe Val Ser Glu Leu Leu Gly Gly Phe Gly Ile Ala Ile

580 585 590
Val Val Phe Met Asn His Tyr Pro Leu Glu Lys Ile Gln Asp Ser Val

595 600 605
Trp Asp Gly His Gly Phe Cys Ala Gly Gln Ile His Glu Thr Met Asn

610 615 620
Val Gln Arg Gly Leu Val Thr Asp Trp Phe Phe Gly Gly Leu Asn Tyr

625 630 635 640
Gln Ile Glu His His Leu Trp Pro Thr Leu Pro Arg His Asn Leu Thr

645 650 655
Ala Ala Ser Ile Lys Val Glu Gln Leu Cys Lys Lys His Asn Leu Pro

660 665 670
Tyr Arg Ser Pro Pro Met Leu Glu Gly Val Gly Ile Leu Ile Ser Tyr

675 680 685
Leu Gly Thr Phe Ala Arg Met Val Ala Lys Ala Asp Lys Ala

690 695 700

<210> 15

<211> 2127

<212> DNA

<213> Artificial Sequence

<220><223> Multizyme: E389D9eS/EgD8M gene fusion

<220><221> CDS

<222> (1)..(2127)

<223> Multizyme: E389D9eS/EgD8M gene fusion

<300><302> MULTIZYMES AND THEIR USE IN MAKING POLYUNSATURATED FATTY ACIDS

<310> US-2008-0254191-A1

<311> 2008-04-03

<312> 2008-10-16

<313> (1)..(2127)

<300><302> MULTIZYMES AND THEIR USE IN MAKING POLYUNSATURATED FATTY ACIDS

<310>

WO 2008/124048

<311> 2008-04-03

<312> 2008-10-16

<313> (1)..(2127)

<400> 15

atg gct gcc gtc atc gag gtg gcc aac gag ttc gtc gct atc act gcc 48

Met Ala Ala Val Ile Glu Val Ala Asn Glu Phe Val Ala Ile Thr Ala

1 5 10 15

gag acc ctt ccc aag gtg gac tat cag cga ctc tgg cga gac atc tac 96

Glu Thr Leu Pro Lys Val Asp Tyr Gln Arg Leu Trp Arg Asp Ile Tyr

20 25 30

tcc tgc gag ctc ctg tac ttc tcc att gct ttc gtc atc ctc aag ttt 144

Ser Cys Glu Leu Leu Tyr Phe Ser Ile Ala Phe Val Ile Leu Lys Phe

35 40 45

acc ctt ggc gag ctc tgc gat tct ggc aaa aag att ctg cga gtg ctg 192

Thr Leu Gly Glu Leu Ser Asp Ser Gly Lys Lys Ile Leu Arg Val Leu

50 55 60

ttc aag tgg tac aac ctc ttc atg tcc gtc ttt tgc ctg gtg tcc ttc 240

Phe Lys Trp Tyr Asn Leu Phe Met Ser Val Phe Ser Leu Val Ser Phe

65 70 75 80

ctc tgt atg ggt tac gcc atc tac acc gtt gga ctg tac tcc aac gaa 288

Leu Cys Met Gly Tyr Ala Ile Tyr Thr Val Gly Leu Tyr Ser Asn Glu

85 90 95

tgc gac aga gct ttc gac aac agc ttg ttc cga ttt gcc acc aag gtc 336

Cys Asp Arg Ala Phe Asp Asn Ser Leu Phe Arg Phe Ala Thr Lys Val

100 105 110

ttc tac tat tcc aag ttt ctg gag tac atc gac tct ttc tac ctt ccc 384

Phe Tyr Tyr Ser Lys Phe Leu Glu Tyr Ile Asp Ser Phe Tyr Leu Pro

115 120 125

ctc atg gcc aag cct ctg tcc ttt ctg cag ttc ttt cat cac ttg gga 432

Leu Met Ala Lys Pro Leu Ser Phe Leu Gln Phe Phe His His Leu Gly	
130	135
140	
gct cct atg gac atg tgg ctc ttc gtg cag tac tct ggc gaa tcc att	480
Ala Pro Met Asp Met Trp Leu Phe Val Gln Tyr Ser Gly Glu Ser Ile	
145	150
155	160
tgg atc ttt gtg ttc ctg aac gga ttc att cac ttt gtc atg tac ggc	528
Trp Ile Phe Val Phe Leu Asn Gly Phe Ile His Phe Val Met Tyr Gly	
165	170
175	
tac tat tgg aca cgg ctg atg aag ttc aac ttt ccc atg ccc aag cag	576
Tyr Tyr Trp Thr Arg Leu Met Lys Phe Asn Phe Pro Met Pro Lys Gln	
180	185
190	
ctc att acc gca atg cag atc acc cag ttc aac gtt ggc ttc tac ctc	624
Leu Ile Thr Ala Met Gln Ile Thr Gln Phe Asn Val Gly Phe Tyr Leu	
195	200
205	
gtg tgg tgg tac aag gac att ccc tgt tac cga aag gat ccc atg cga	672
Val Trp Trp Tyr Lys Asp Ile Pro Cys Tyr Arg Lys Asp Pro Met Arg	
210	215
220	
atg ctg gcc tgg atc ttc aac tac tgg tac gtc ggt acc gtt ctt ctg	720
Met Leu Ala Trp Ile Phe Asn Tyr Trp Tyr Val Gly Thr Val Leu Leu	
225	230
235	240
ctc ttc atc aac ttc ttt gtc aag tcc tac gtg ttt ccc aag cct aag	768
Leu Phe Ile Asn Phe Phe Val Lys Ser Tyr Val Phe Pro Lys Pro Lys	
245	250
255	
act gcc gac aaa aag gtc cag ggc gcc ggt ccc gct cga cct gcc gga	816
Thr Ala Asp Lys Lys Val Gln Gly Ala Gly Pro Ala Arg Pro Ala Gly	
260	265
270	
ctt cct ccc get acc tac tac gac tct ctg gcc gtc atg gga tcc gtg	864
Leu Pro Pro Ala Thr Tyr Tyr Asp Ser Leu Ala Val Met Gly Ser Val	
275	280
285	

aag gct tct cga cag gct ctg ccc ctc gtc atc gac gga aag gtg tac 912

Lys Ala Ser Arg Gln Ala Leu Pro Leu Val Ile Asp Gly Lys Val Tyr
290 295 300

gac gtc tcc gct tgg gtg aac ttc cac cct ggt gga gct gaa atc att 960
Asp Val Ser Ala Trp Val Asn Phe His Pro Gly Gly Ala Glu Ile Ile
305 310 315 320

gag aac tac cag gga cga gat gct act gac gcc ttc atg gtt atg cac 1008
Glu Asn Tyr Gln Gly Arg Asp Ala Thr Asp Ala Phe Met Val Met His
325 330 335

tct cag gaa gcc ttc gac aag ctc aag cga atg ccc aag atc aac cag 1056
Ser Gln Glu Ala Phe Asp Lys Leu Lys Arg Met Pro Lys Ile Asn Gln
340 345 350

gct tcc gag ctg cct ccc cag gct gcc gtc aac gaa gct cag gag gat 1104
Ala Ser Glu Leu Pro Pro Gln Ala Ala Val Asn Glu Ala Gln Glu Asp
355 360 365

ttc cga aag ctc cga gaa gag ctg atc gcc act ggc atg ttt gac gcc 1152
Phe Arg Lys Leu Arg Glu Glu Leu Ile Ala Thr Gly Met Phe Asp Ala
370 375 380

tct ccc ctc tgg tac tcg tac aag atc ttg acc acc ctg ggt ctt ggc 1200
Ser Pro Leu Trp Tyr Ser Tyr Lys Ile Leu Thr Thr Leu Gly Leu Gly
385 390 395 400

gtg ctt gcc ttc ttc atg ctg gtc cag tac cac ctg tac ttc att ggt 1248

Val Leu Ala Phe Phe Met Leu Val Gln Tyr His Leu Tyr Phe Ile Gly
405 410 415

gct ctc gtg ctc ggt atg cac tac cag caa atg gga tgg ctg tct cat 1296
Ala Leu Val Leu Gly Met His Tyr Gln Gln Met Gly Trp Leu Ser His
420 425 430

gac atc tgc cac cac cag acc ttc aag aac cga aac tgg aat aac gtc 1344
Asp Ile Cys His His Gln Thr Phe Lys Asn Arg Asn Trp Asn Asn Val

435	440	445	
ctg ggt ctg gtc ttt ggc aac gga ctc cag ggc ttc tcc gtg acc tgg			1392
Leu Gly Leu Val Phe Gly Asn Gly Leu Gln Gly Phe Ser Val Thr Trp			
450	455	460	
tgg aag gac aga cac aac gcc cat cat tct gct acc aac gtt cag ggt			1440
Trp Lys Asp Arg His Asn Ala His His Ser Ala Thr Asn Val Gln Gly			
465	470	475	480
cac gat ccc gac att gat aac ctg cct ctg ctc gcc tgg tcc gag gac			1488
His Asp Pro Asp Ile Asp Asn Leu Pro Leu Leu Ala Trp Ser Glu Asp			
485	490	495	
gat gtc act cga gct tct ccc atc tcc cga aag ctc att cag ttc caa			1536
Asp Val Thr Arg Ala Ser Pro Ile Ser Arg Lys Leu Ile Gln Phe Gln			
500	505	510	
cag tac tat ttc ctg gtc atc tgt att ctc ctg cga ttc atc tgg tgt			1584
Gln Tyr Tyr Phe Leu Val Ile Cys Ile Leu Leu Arg Phe Ile Trp Cys			
515	520	525	
ttc cag tct gtg ctg acc gtt cga tcc ctc aag gac cga gac aac cag			1632
Phe Gln Ser Val Leu Thr Val Arg Ser Leu Lys Asp Arg Asp Asn Gln			
530	535	540	
ttc tac cga tct cag tac aag aaa gag gcc att gga ctc gct ctg cac			1680
Phe Tyr Arg Ser Gln Tyr Lys Lys Glu Ala Ile Gly Leu Ala Leu His			
545	550	555	560
tgg act ctc aag acc ctg ttc cac ctc ttc ttt atg ccc tcc atc ctg			1728
Trp Thr Leu Lys Thr Leu Phe His Leu Phe Phe Met Pro Ser Ile Leu			
565	570	575	
acc tcg atg ctg gtg ttc ttt gtt tcc gag ctc gtc ggt ggc ttc gga			1776
Thr Ser Met Leu Val Phe Phe Val Ser Glu Leu Val Gly Gly Phe Gly			
580	585	590	
att gcc atc gtg gtc ttc atg aac cac tac cct ctg gag aag atc ggt			1824
Ile Ala Ile Val Val Phe Met Asn His Tyr Pro Leu Glu Lys Ile Gly			
595	600	605	

gat tcc gtc tgg gac gga cat ggc ttc tct gtg ggt cag atc cat gag 1872

Asp Ser Val Trp Asp Gly His Gly Phe Ser Val Gly Gln Ile His Glu

610 615 620

acc atg aac att cga cga ggc atc att act gac tgg ttc ttt gga ggc 1920

Thr Met Asn Ile Arg Arg Gly Ile Ile Thr Asp Trp Phe Phe Gly Gly

625 630 635 640

ctg aac tac cag atc gag cac cat ctc tgg ccc acc ctg cct cga cac 1968

Leu Asn Tyr Gln Ile Glu His His Leu Trp Pro Thr Leu Pro Arg His

645 650 655

aac ctc act gcc gtt tcc tac cag gtg gaa cag ctg tgc cag aag cac 2016

Asn Leu Thr Ala Val Ser Tyr Gln Val Glu Gln Leu Cys Gln Lys His

660 665 670

aac ctc ccc tac cga aac cct ctg ccc cat gaa ggt ctc gtc atc ctg 2064

Asn Leu Pro Tyr Arg Asn Pro Leu Pro His Glu Gly Leu Val Ile Leu

675 680 685

ctc cga tac ctg tcc cag ttc gct cga atg gcc gag aag cag ccc ggt 2112

Leu Arg Tyr Leu Ser Gln Phe Ala Arg Met Ala Glu Lys Gln Pro Gly

690 695 700

gcc aag gct cag taa 2127

Ala Lys Ala Gln

705

<210> 16

<211> 708

<212> PRT

<213> Artificial Sequence

<220><223> Synthetic Construct

<400> 16

Met Ala Ala Val Ile Glu Val Ala Asn Glu Phe Val Ala Ile Thr Ala

1 5 10 15

Glu Thr Leu Pro Lys Val Asp Tyr Gln Arg Leu Trp Arg Asp Ile Tyr

20 25 30

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

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

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

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

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

Phe Gln Ser Val Leu Thr Val Arg Ser Leu Lys Asp Arg Asp Asn Gln

530

535

540

Phe Tyr Arg Ser Gln Tyr Lys Lys Glu Ala Ile Gly Leu Ala Leu His

545

550

555

560

Trp Thr Leu Lys Thr Leu Phe His Leu Phe Phe Met Pro Ser Ile Leu

565

570

575

Thr Ser Met Leu Val Phe Phe Val Ser Glu Leu Val Gly Gly Phe Gly

580

585

590

Ile Ala Ile Val Val Phe Met Asn His Tyr Pro Leu Glu Lys Ile Gly

595

600

605

Asp Ser Val Trp Asp Gly His Gly Phe Ser Val Gly Gln Ile His Glu

610

615

620

Thr Met Asn Ile Arg Arg Gly Ile Ile Thr Asp Trp Phe Phe Gly Gly

625

630

635

640

Leu Asn Tyr Gln Ile Glu His His Leu Trp Pro Thr Leu Pro Arg His

645

650

655

Asn Leu Thr Ala Val Ser Tyr Gln Val Glu Gln Leu Cys Gln Lys His

660

665

670

Asn Leu Pro Tyr Arg Asn Pro Leu Pro His Glu Gly Leu Val Ile Leu

675

680

685

Leu Arg Tyr Leu Ser Gln Phe Ala Arg Met Ala Glu Lys Gln Pro Gly

690

695

700

Ala Lys Ala Gln

705

<210> 17

<211> 945

<212> DNA

<213> Mortierella alpina

<220><221> CDS

<222> (1)..(945)

<223> LPAAT1

<300><302> A MORTIERELLA ALPINA LYSOPHOSPHATIDIC ACID ACYLTRANSFERASE FOR

IMPROVEMENT OF LONG CHAIN OMEGA-3 AND OMEGA-6 POLYUNSATURATED
FATTY ACID BIOSYNTHESIS IN OLEAGINOUS ORGANISMS

<310> US 7,879,591

<311> 2009-06-16

<312> 2011-02-01

<313> (1)..(945)

<400> 17

atg tcc ata ggt tct tcc aat cct gtc ctg ctg gca gcg atc ccc ttc 48

Met Ser Ile Gly Ser Ser Asn Pro Val Leu Leu Ala Ala Ile Pro Phe

1 5 10 15

gtc tac ctc ttc gtc ctc cct cgt gtc ctc gcc ttc ctc cct caa aag 96

Val Tyr Leu Phe Val Leu Pro Arg Val Leu Ala Phe Leu Pro Gln Lys

20 25 30

gcc cag ttc ctc gca aaa tgc atc gtg gtc ttg atc gcc acc ctt atc 144

Ala Gln Phe Leu Ala Lys Cys Ile Val Val Leu Ile Ala Thr Leu Ile

35 40 45

atg tcc gtc gca ggc tgc ttc att tcc atc gtc tgt gcg ctc ctc gat 192

Met Ser Val Ala Gly Cys Phe Ile Ser Ile Val Cys Ala Leu Leu Asp

50 55 60

aaa cgc tat gtg atc aac tac gtc gtc tca aga ctc ttc tca ttc ctc 240

Lys Arg Tyr Val Ile Asn Tyr Val Val Ser Arg Leu Phe Ser Phe Leu

65 70 75 80

gct gca aga ccc tgc ggt gtc acc tac aag atc gtc ggc gag gaa cat 288

Ala Ala Arg Pro Cys Gly Val Thr Tyr Lys Ile Val Gly Glu Glu His

85 90 95

ctg gac aag tac ccc gcc att gtc gtc tgc aac cac cag agc tcc atg 336

Leu Asp Lys Tyr Pro Ala Ile Val Val Cys Asn His Gln Ser Ser Met

100 105 110

gac atg atg gtc ctg gga cgc gtc ttc cca aag cac tgt gtc gtc atg 384

Asp Met Met Val Leu Gly Arg Val Phe Pro Lys His Cys Val Val Met

115	120	125	
gca aag aag gaa ctt ctt tac ttt ccg ttc ctg ggc atg ttt atg aag			432
Ala Lys Lys Glu Leu Leu Tyr Phe Pro Phe Leu Gly Met Phe Met Lys			
130	135	140	
ctg agt aac gcc atc ttc att gac cgc aag aac cac aag aag gcg atc			480
Leu Ser Asn Ala Ile Phe Ile Asp Arg Lys Asn His Lys Lys Ala Ile			
145	150	155	160
gag tcc acc acc caa gct gtc gcc gac atg aag aag cac aac tct gga			528
Glu Ser Thr Thr Gln Ala Val Ala Asp Met Lys Lys His Asn Ser Gly			
165	170	175	
atc tgg att ttc ccc gaa gga aca cgt tcc cgc ttg gac aag gcc gat			576
Ile Trp Ile Phe Pro Glu Gly Thr Arg Ser Arg Leu Asp Lys Ala Asp			
180	185	190	
ctc ttg ccc ttc aag aag gga gcc ttc cac ctc gcc att caa gcc caa			624
Leu Leu Pro Phe Lys Lys Gly Ala Phe His Leu Ala Ile Gln Ala Gln			
195	200	205	
ctc ccg atc ctc ccc atc atc tcg caa gga tac tca cac atc tac gat			672
Leu Pro Ile Leu Pro Ile Ile Ser Gln Gly Tyr Ser His Ile Tyr Asp			
210	215	220	
tcg tca aaa cgc tac ttc ccc ggt gga gag ctc gag atc aga gtc ctg			720
Ser Ser Lys Arg Tyr Phe Pro Gly Gly Glu Leu Glu Ile Arg Val Leu			
225	230	235	240
gaa cct atc ccc acc acg gga ttg acc aca gac gat gtg aac gac ctg			768
Glu Pro Ile Pro Thr Thr Gly Leu Thr Thr Asp Asp Val Asn Asp Leu			
245	250	255	
atg gac aag act cgc aac ctg atg ctg aag cac ctc aag gag atg gat			816
Met Asp Lys Thr Arg Asn Leu Met Leu Lys His Leu Lys Glu Met Asp			
260	265	270	
tct caa tac tcc tcc tcc acc gct gaa aac gga tcg acc cat att gac			864

Ser Gln Tyr Ser Ser Ser Thr Ala Glu Asn Gly Ser Thr His Ile Asp

275	280	285	
gcc gat atc gca aag tca act gcc aca tcg atc gga aac acg gac gat			912
Ala Asp Ile Ala Lys Ser Thr Ala Thr Ser Ile Gly Asn Thr Asp Asp			
290	295	300	
gct atc aca aag agg agg aca cca aaa gag tag			945
Ala Ile Thr Lys Arg Arg Thr Pro Lys Glu			
305	310		

<210> 18

<211> 314

<212> PRT

<213> Mortierella alpina

<400> 18

Met Ser Ile Gly Ser Ser Asn Pro Val Leu Leu Ala Ala Ile Pro Phe
1 5 10 15
Val Tyr Leu Phe Val Leu Pro Arg Val Leu Ala Phe Leu Pro Gln Lys
20 25 30
Ala Gln Phe Leu Ala Lys Cys Ile Val Val Leu Ile Ala Thr Leu Ile
35 40 45
Met Ser Val Ala Gly Cys Phe Ile Ser Ile Val Cys Ala Leu Leu Asp

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

115	120	125
Ala Lys Lys Glu Leu Leu Tyr Phe Pro Phe Leu Gly Met Phe Met Lys		
130	135	140

Leu Ser Asn Ala Ile Phe Ile Asp Arg Lys Asn His Lys Lys Ala Ile
 145 150 155 160
 Glu Ser Thr Thr Gln Ala Val Ala Asp Met Lys Lys His Asn Ser Gly
 165 170 175
 Ile Trp Ile Phe Pro Glu Gly Thr Arg Ser Arg Leu Asp Lys Ala Asp
 180 185 190
 Leu Leu Pro Phe Lys Lys Gly Ala Phe His Leu Ala Ile Gln Ala Gln
 195 200 205
 Leu Pro Ile Leu Pro Ile Ile Ser Gln Gly Tyr Ser His Ile Tyr Asp
 210 215 220
 Ser Ser Lys Arg Tyr Phe Pro Gly Gly Glu Leu Glu Ile Arg Val Leu
 225 230 235 240
 Glu Pro Ile Pro Thr Thr Gly Leu Thr Thr Asp Asp Val Asn Asp Leu
 245 250 255
 Met Asp Lys Thr Arg Asn Leu Met Leu Lys His Leu Lys Glu Met Asp
 260 265 270
 Ser Gln Tyr Ser Ser Ser Thr Ala Glu Asn Gly Ser Thr His Ile Asp
 275 280 285
 Ala Asp Ile Ala Lys Ser Thr Ala Thr Ser Ile Gly Asn Thr Asp Asp
 290 295 300
 Ala Ile Thr Lys Arg Arg Thr Pro Lys Glu
 305 310

<210> 19

<211> 955

<212> DNA

<213> *Mortierella alpina*

<220><221> CDS

<222> (3)..(947)

<223> synthetic LPAAT1 (codon-optimized for *Yarrowia lipolytica*)

<300><302> A MORTIERELLA ALPINA LYSOPHOSPHATIDIC ACID ACYLTRANSFERASE FOR
 IMPROVEMENT OF LONG CHAIN OMEGA-3 AND OMEGA-6 POLYUNSATURATED
 FATTY ACID BIOSYNTHESIS IN OLEAGINOUS ORGANISMS

<310> US 7,879,591

<311> 2009-06-16

<312> 2011-02-01

<313> (1)..(955)

<400> 19

ac atg tct att ggt tgc tcc aac ccc gtg ctc ttg gct gcg att ccc 47

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

1 5 10 15

ttc gtc tac ctg ttt gtc ctc cca cga gtc ctg gct ttc ctg cct cag 95

Phe Val Tyr Leu Phe Val Leu Pro Arg Val Leu Ala Phe Leu Pro Gln

20 25 30

aag gct cag ttc ctg gcc aaa tgt att gtg gtc ctg att gcc acg ctt 143

Lys Ala Gln Phe Leu Ala Lys Cys Ile Val Val Leu Ile Ala Thr Leu

35 40 45

atc atg tcc gtt gca ggc tgc ttc atc tgc atc gtg tgc gct ctt ctg 191

Ile Met Ser Val Ala Gly Cys Phe Ile Ser Ile Val Cys Ala Leu Leu

50 55 60

gac aag aga tac gtc atc aat tac gtt gtg tgc cga ttg ttc tcc ttc 239

Asp Lys Arg Tyr Val Ile Asn Tyr Val Val Ser Arg Leu Phe Ser Phe

65 70 75

ctt gcc gct cga ccg tgt ggt gtg acc tat aag att gtt ggt gag gaa 287

Leu Ala Ala Arg Pro Cys Gly Val Thr Tyr Lys Ile Val Gly Glu Glu

80 85 90 95

cac ctc gat aag tac cct gct atc gtg gtc tgt aac cat caa tcc tct 335

His Leu Asp Lys Tyr Pro Ala Ile Val Val Cys Asn His Gln Ser Ser

100 105 110

atg gat atg atg gtt ttg gga cga gtt ttt cca aag cac tgc gtt gtc 383

Met Asp Met Met Val Leu Gly Arg Val Phe Pro Lys His Cys Val Val

115 120 125

atg gcg aag aag gaa ctc ctg tac ttt ccc ttt ttg gga atg ttt atg 431

Met Ala Lys Lys Glu Leu Leu Tyr Phe Pro Phe Leu Gly Met Phe Met

130	135	140	
aaa ctg agc aac gct atc ttc atc gac cgg aag aac cac aag aaa gcc			479
Lys Leu Ser Asn Ala Ile Phe Ile Asp Arg Lys Asn His Lys Lys Ala			
145	150	155	
atc gag tct acc acc caa gcc gtg gcg gac atg aag aag cac aac tct			527
Ile Glu Ser Thr Thr Gln Ala Val Ala Asp Met Lys Lys His Asn Ser			
160	165	170	175
gga atc tgg att ttc cca gag ggc acc cgg tct aga ctg gac aag gca			575
Gly Ile Trp Ile Phe Pro Glu Gly Thr Arg Ser Arg Leu Asp Lys Ala			
180	185	190	
gac ctg ctg ccc ttc aag aaa ggt gcc ttt cat ctt gca att cag gcc			623
Asp Leu Leu Pro Phe Lys Lys Gly Ala Phe His Leu Ala Ile Gln Ala			
195	200	205	
cag ctc cct att ctc ccc att atc tgc cag ggc tat tcc cat atc tac			671
Gln Leu Pro Ile Leu Pro Ile Ile Ser Gln Gly Tyr Ser His Ile Tyr			
210	215	220	
gac tct tgc aag cgg tac ttc ccc ggt gga gag ctc gag atc aga gtc			719
Asp Ser Ser Lys Arg Tyr Phe Pro Gly Gly Glu Leu Glu Ile Arg Val			
225	230	235	
ctg gag ccc att cct aca act ggc ctc act act gat gat gtg aac gac			767
Leu Glu Pro Ile Pro Thr Thr Gly Leu Thr Thr Asp Asp Val Asn Asp			
240	245	250	255
ctg atg gac aag aca cga aac ctt atg ctc aag cac ttg aag gag atg			815
Leu Met Asp Lys Thr Arg Asn Leu Met Leu Lys His Leu Lys Glu Met			
260	265	270	
gat tcc cag tat tgc tgc agc act gct gaa aat gga tcc acg cac atc			863
Asp Ser Gln Tyr Ser Ser Ser Thr Ala Glu Asn Gly Ser Thr His Ile			
275	280	285	
gac gcc gat att gcc aag tct aca gcc acc agc att ggc aac act gac			911
Asp Ala Asp Ile Ala Lys Ser Thr Ala Thr Ser Ile Gly Asn Thr Asp			

290 295 300 955

gac gca att aca aaa cgt cgt acc cct aag gaa taa gcggccgc

Asp Ala Ile Thr Lys Arg Arg Thr Pro Lys Glu

305 310

<210> 20

<211> 314

<212> PRT

<213> Mortierella alpina

<400> 20

Met Ser Ile Gly Ser Ser Asn Pro Val Leu Leu Ala Ala Ile Pro Phe

1 5 10 15

Val Tyr Leu Phe Val Leu Pro Arg Val Leu Ala Phe Leu Pro Gln Lys

20 25 30

Ala Gln Phe Leu Ala Lys Cys Ile Val Val Leu Ile Ala Thr Leu Ile

35 40 45

Met Ser Val Ala Gly Cys Phe Ile Ser Ile Val Cys Ala Leu Leu Asp

50 55 60

Lys Arg Tyr Val Ile Asn Tyr Val Val Ser Arg Leu Phe Ser Phe Leu

65 70 75 80

Ala Ala Arg Pro Cys Gly Val Thr Tyr Lys Ile Val Gly Glu Glu His

85 90 95

Leu Asp Lys Tyr Pro Ala Ile Val Val Cys Asn His Gln Ser Ser Met

100 105 110

Asp Met Met Val Leu Gly Arg Val Phe Pro Lys His Cys Val Val Met

115 120 125

Ala Lys Lys Glu Leu Leu Tyr Phe Pro Phe Leu Gly Met Phe Met Lys

130 135 140

Leu Ser Asn Ala Ile Phe Ile Asp Arg Lys Asn His Lys Lys Ala Ile

145 150 155 160

Glu Ser Thr Thr Gln Ala Val Ala Asp Met Lys Lys His Asn Ser Gly

165 170 175

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

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

Ala Asp Ile Ala Lys Ser Thr Ala Thr Ser Ile Gly Asn Thr Asp Asp
290 295 300
Ala Ile Thr Lys Arg Arg Thr Pro Lys Glu
305 310

<210> 21

<211> 849

<212> DNA

<213> *Yarrowia lipolytica*

<220><221> CDS

<222> (1)..(849)

<300><302> A MORTIERELLA ALPINA LYSOPHOSPHATIDIC ACID ACYLTRANSFERASE
HOMOLOG FOR ALTERATION OF POLYUNSATURATED FATTY ACIDS AND OIL
CONTENT IN OLEAGINOUS ORGANISMS

<310>

US7189559

<311> 2005-10-14

<312> 2007-03-13

<313> (1)..(849)

<400> 21

atg tcc gtt gca tcc aag ctc gtc ttc tac gtc cgc gcc gcc atc gcc	48
Met Ser Val Ala Ser Lys Leu Val Phe Tyr Val Arg Ala Ala Ile Ala	
1 5 10 15	
gtg gtc atc ttt gcc gcc tgt gcc acc tac ggc gtg ctg gcg tcc acc	96
Val Val Ile Phe Ala Ala Cys Ala Thr Tyr Gly Val Leu Ala Ser Thr	
20 25 30	
att ctc acc gcc atc ggc aag cag ggc ctg gcc caa tgg acc gtt gcc	144
Ile Leu Thr Ala Ile Gly Lys Gln Gly Leu Ala Gln Trp Thr Val Ala	
35 40 45	
aga gcc ttc tac tac tcg gtg cgc atc ttc ctg ggt atc agc atc aag	192
Arg Ala Phe Tyr Tyr Ser Val Arg Ile Phe Leu Gly Ile Ser Ile Lys	
50 55 60	
ctg cgt agc cgg cag gtg acc gga acc gcc ggt ctg gat gcc tcc aag	240
Leu Arg Ser Arg Gln Val Thr Gly Thr Ala Gly Leu Asp Ala Ser Lys	
65 70 75 80	
atc cag gtc gcc aac acc acc aag ccc att gac gac atc acc aaa cac	288
Ile Gln Val Ala Asn Thr Thr Lys Pro Ile Asp Asp Ile Thr Lys His	
85 90 95	
ctg ccc cga cca tgc att ctg att tcc aac cac cag aac gaa atg gac	336
Leu Pro Arg Pro Cys Ile Leu Ile Ser Asn His Gln Asn Glu Met Asp	
100 105 110	
att ctg gtg ctc ggt cgc atc ttc ccc cag tac tgc tcc gtc acc gcc	384
Ile Leu Val Leu Gly Arg Ile Phe Pro Gln Tyr Cys Ser Val Thr Ala	
115 120 125	
aaa aag gcc ctc aag tgg tac cct ctg ctg ggc cag ttc atg gcg ctg	432
Lys Lys Ala Leu Lys Trp Tyr Pro Leu Leu Gly Gln Phe Met Ala Leu	
130 135 140	
tcc ggc acc atc ttc ctg gac cga aag gac cga acc aag tcc gtg cag	480
Ser Gly Thr Ile Phe Leu Asp Arg Lys Asp Arg Thr Lys Ser Val Gln	
145 150 155 160	
acc ctc ggc ggc gcc gtc aag acc atc cag agc ggc aac gga ggc aag	528

Thr Leu Gly Gly Ala Val Lys Thr Ile Gln Ser Gly Asn Gly Gly Lys
 165 170 175
 ggc cag agc gtc ttc atg ttc ccc gag gga acc cga tcc tac tcc aag 576

Gly Gln Ser Val Phe Met Phe Pro Glu Gly Thr Arg Ser Tyr Ser Lys
 180 185 190
 gac gtc ggc atc atg ccc ttc aag aag ggc tgt ttc cac ctg gcg gtc 624

Asp Val Gly Ile Met Pro Phe Lys Lys Gly Cys Phe His Leu Ala Val
 195 200 205
 cag tcg ggc gct ccc att gtc ccc gtg gtg gtc cag aac acc tcc cga 672

Gln Ser Gly Ala Pro Ile Val Pro Val Val Val Gln Asn Thr Ser Arg

210 215 220
 atg ttt tctttc ggc cga ggc aag ctg gac gcc gga gag atc ctt gtc 720

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

gac gtc ctg agc ccc att gag acc aag ggt ctg gac gcc agc aac gtc 768
 Asp Val Leu Ser Pro Ile Glu Thr Lys Gly Leu Asp Ala Ser Asn Val

245 250 255

gac gct ctc atg gcc acc act tat aag gcc atg tgc gag act gcc gac 816
 Asp Ala Leu Met Ala Thr Thr Tyr Lys Ala Met Cys Glu Thr Ala Asp

260 265 270

cag att ggc tac gct ggc cag aag act cag tag 849

Gln Ile Gly Tyr Ala Gly Gln Lys Thr Gln

275 280

<210> 22

<211> 282

<212> PRT

<213> Yarrowia lipolytica

<400> 22

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

1 5 10 15

Val Val Ile Phe Ala Ala Cys Ala Thr Tyr Gly Val Leu Ala Ser Thr

20 25 30
 Ile Leu Thr Ala Ile Gly Lys Gln Gly Leu Ala Gln Trp Thr Val Ala
 35 40 45
 Arg Ala Phe Tyr Tyr Ser Val Arg Ile Phe Leu Gly Ile Ser Ile Lys
 50 55 60

 Leu Arg Ser Arg Gln Val Thr Gly Thr Ala Gly Leu Asp Ala Ser Lys
 65 70 75 80
 Ile Gln Val Ala Asn Thr Thr Lys Pro Ile Asp Asp Ile Thr Lys His
 85 90 95
 Leu Pro Arg Pro Cys Ile Leu Ile Ser Asn His Gln Asn Glu Met Asp
 100 105 110
 Ile Leu Val Leu Gly Arg Ile Phe Pro Gln Tyr Cys Ser Val Thr Ala
 115 120 125

 Lys Lys Ala Leu Lys Trp Tyr Pro Leu Leu Gly Gln Phe Met Ala Leu
 130 135 140
 Ser Gly Thr Ile Phe Leu Asp Arg Lys Asp Arg Thr Lys Ser Val Gln
 145 150 155 160
 Thr Leu Gly Gly Ala Val Lys Thr Ile Gln Ser Gly Asn Gly Gly Lys
 165 170 175
 Gly Gln Ser Val Phe Met Phe Pro Glu Gly Thr Arg Ser Tyr Ser Lys
 180 185 190

 Asp Val Gly Ile Met Pro Phe Lys Lys Gly Cys Phe His Leu Ala Val
 195 200 205
 Gln Ser Gly Ala Pro Ile Val Pro Val Val Val Gln Asn Thr Ser Arg
 210 215 220
 Met Phe Ser Phe Gly Arg Gly Lys Leu Asp Ala Gly Glu Ile Leu Val
 225 230 235 240
 Asp Val Leu Ser Pro Ile Glu Thr Lys Gly Leu Asp Ala Ser Asn Val
 245 250 255

 Asp Ala Leu Met Ala Thr Thr Tyr Lys Ala Met Cys Glu Thr Ala Asp
 260 265 270

Gln Ile Gly Tyr Ala Gly Gln Lys Thr Gln

275 280

<210> 23

<211> 303

<212> PRT

<213> *Saccharomyces cerevisiae*

<220><221> MISC_FEATURE

<222> (1)..(303)

<223> Slc1p; GenBank Accession No. NP_010231

<400> 23

Met Ser Val Ile Gly Arg Phe Leu Tyr Tyr Leu Arg Ser Val Leu Val

1 5 10 15

Val Leu Ala Leu Ala Gly Cys Gly Phe Tyr Gly Val Ile Ala Ser Ile

20 25 30

Leu Cys Thr Leu Ile Gly Lys Gln His Leu Ala Gln Trp Ile Thr Ala

35 40 45

Arg Cys Phe Tyr His Val Met Lys Leu Met Leu Gly Leu Asp Val Lys

50 55 60

Val Val Gly Glu Glu Asn Leu Ala Lys Lys Pro Tyr Ile Met Ile Ala

65 70 75 80

Asn His Gln Ser Thr Leu Asp Ile Phe Met Leu Gly Arg Ile Phe Pro

85 90 95

Pro Gly Cys Thr Val Thr Ala Lys Lys Ser Leu Lys Tyr Val Pro Phe

100 105 110

Leu Gly Trp Phe Met Ala Leu Ser Gly Thr Tyr Phe Leu Asp Arg Ser

115 120 125

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

130 135 140

Lys Lys Asn Lys Arg Ala Leu Trp Val Phe Pro Glu Gly Thr Arg Ser

145 150 155 160

Tyr Thr Ser Glu Leu Thr Met Leu Pro Phe Lys Lys Gly Ala Phe His

165 170 175

Leu Ala Gln Gln Gly Lys Ile Pro Ile Val Pro Val Val Val Ser Asn
180 185 190
Thr Ser Thr Leu Val Ser Pro Lys Tyr Gly Val Phe Asn Arg Gly Cys
195 200 205
Met Ile Val Arg Ile Leu Lys Pro Ile Ser Thr Glu Asn Leu Thr Lys
210 215 220
Asp Lys Ile Gly Glu Phe Ala Glu Lys Val Arg Asp Gln Met Val Asp
225 230 235 240
Thr Leu Lys Glu Ile Gly Tyr Ser Pro Ala Ile Asn Asp Thr Thr Leu
245 250 255
Pro Pro Gln Ala Ile Glu Tyr Ala Ala Leu Gln His Asp Lys Lys Val
260 265 270
Asn Lys Lys Ile Lys Asn Glu Pro Val Pro Ser Val Ser Ile Ser Asn
275 280 285
Asp Val Asn Thr His Asn Glu Gly Ser Ser Val Lys Lys Met His
290 295 300
<210> 24
<211> 926
<212> DNA
<213> *Saccharomyces cerevisiae*
<220><221> CDS
<222> (4)..(915)
<400> 24
aac atg tct gtt att gga cga ttt ctt tac tac ctg aga tcg gtg ctc 48
Met Ser Val Ile Gly Arg Phe Leu Tyr Tyr Leu Arg Ser Val Leu
1 5 10 15
gtc gtt ttg gcc ctc gct gga tgt ggc ttc tat ggc gtg att gcc tct 96
Val Val Leu Ala Leu Ala Gly Cys Gly Phe Tyr Gly Val Ile Ala Ser
20 25 30
atc ctg tgt act ctc atc ggc aag cag cat ctc gcg caa tgg att acc 144
Ile Leu Cys Thr Leu Ile Gly Lys Gln His Leu Ala Gln Trp Ile Thr

35	40	45	
gcc cga tgc ttt tac cac gtg atg aaa ctg atg ctg gga ttg gac gtc			192
Ala Arg Cys Phe Tyr His Val Met Lys Leu Met Leu Gly Leu Asp Val			
50	55	60	
aaa gtc gtg ggt gaa gag aac ctg gct aag aag ccc tat atc atg atc			240
Lys Val Val Gly Glu Glu Asn Leu Ala Lys Lys Pro Tyr Ile Met Ile			
65	70	75	
gct aac cac cag tcc acc ctc gat atc ttt atg ctc ggc aga atc ttc			288
Ala Asn His Gln Ser Thr Leu Asp Ile Phe Met Leu Gly Arg Ile Phe			
80	85	90	95
cct ccc gga tgc acc gtt acc gca aag aag agc ctt aag tac gtc ccc			336
Pro Pro Gly Cys Thr Val Thr Ala Lys Lys Ser Leu Lys Tyr Val Pro			
100	105	110	
ttc ctg ggc tgg ttt atg gcg ctt tcc ggt aca tac ttc ctc gac cgt			384
Phe Leu Gly Trp Phe Met Ala Leu Ser Gly Thr Tyr Phe Leu Asp Arg			
115	120	125	
tcc aag cgg caa gag gct att gat acc ctg aac aaa gga ctg gag aac			432
Ser Lys Arg Gln Glu Ala Ile Asp Thr Leu Asn Lys Gly Leu Glu Asn			
130	135	140	
gtc aaa aag aat aag cga gcc ctc tgg gtt ttt ccc gaa ggt acc cgg			480
Val Lys Lys Asn Lys Arg Ala Leu Trp Val Phe Pro Glu Gly Thr Arg			
145	150	155	
tct tac acg tcg gag ctt acc atg ctg ccg ttc aag aag ggc gcc ttt			528
Ser Tyr Thr Ser Glu Leu Thr Met Leu Pro Phe Lys Lys Gly Ala Phe			
160	165	170	175
cat ttg gcc cag cag gga aag atc cct atc gtc cca gtg gtt gtg tct			576
His Leu Ala Gln Gln Gly Lys Ile Pro Ile Val Pro Val Val Val Ser			
180	185	190	
aac act agc acg ctc gtc agc cct aag tac ggt gtg ttc aac cga ggc			624

Asn Thr Ser Thr Leu Val Ser Pro Lys Tyr Gly Val Phe Asn Arg Gly
195 200 205
tgt atg att gtc cga att ctg aag ccc atc tcg act gag aac ctg aca 672
Cys Met Ile Val Arg Ile Leu Lys Pro Ile Ser Thr Glu Asn Leu Thr
210 215 220
aaa gat aag att ggc gag ttc gct gag aaa gtg cga gat cag atg gtt 720
Lys Asp Lys Ile Gly Glu Phe Ala Glu Lys Val Arg Asp Gln Met Val
225 230 235
gac aca ctg aag gaa atc ggt tat tcc ccc gca atc aac gac acc acc 768
Asp Thr Leu Lys Glu Ile Gly Tyr Ser Pro Ala Ile Asn Asp Thr Thr
240 245 250 255
ctt cct ccg cag gca att gag tac gcc gca ttg cag cat gac aag aag 816
Leu Pro Pro Gln Ala Ile Glu Tyr Ala Ala Leu Gln His Asp Lys Lys
260 265 270
gtg aac aaa aag att aag aac gaa cct gtt cca tcg gtg tcc att tct 864
Val Asn Lys Lys Ile Lys Asn Glu Pro Val Pro Ser Val Ser Ile Ser
275 280 285
aat gat gtg aat act cac aac gaa gga tcg tct gtc aag aag atg cac 912
Asn Asp Val Asn Thr His Asn Glu Gly Ser Ser Val Lys Lys Met His
290 295 300
tag gcggccgcgat g 926

<210> 25

<211> 303

<212> PRT

<213> *Saccharomyces cerevisiae*

<400> 25

Met Ser Val Ile Gly Arg Phe Leu Tyr Tyr Leu Arg Ser Val Leu Val
1 5 10 15
Val Leu Ala Leu Ala Gly Cys Gly Phe Tyr Gly Val Ile Ala Ser Ile
20 25 30
Leu Cys Thr Leu Ile Gly Lys Gln His Leu Ala Gln Trp Ile Thr Ala
35 40 45

Arg Cys Phe Tyr His Val Met Lys Leu Met Leu Gly Leu Asp Val Lys

50 55 60

Val Val Gly Glu Glu Asn Leu Ala Lys Lys Pro Tyr Ile Met Ile Ala

65 70 75 80

Asn His Gln Ser Thr Leu Asp Ile Phe Met Leu Gly Arg Ile Phe Pro

85 90 95

Pro Gly Cys Thr Val Thr Ala Lys Lys Ser Leu Lys Tyr Val Pro Phe

100 105 110

Leu Gly Trp Phe Met Ala Leu Ser Gly Thr Tyr Phe Leu Asp Arg Ser

115 120 125

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

130 135 140

Lys Lys Asn Lys Arg Ala Leu Trp Val Phe Pro Glu Gly Thr Arg Ser

145 150 155 160

Tyr Thr Ser Glu Leu Thr Met Leu Pro Phe Lys Lys Gly Ala Phe His

165 170 175

Leu Ala Gln Gln Gly Lys Ile Pro Ile Val Pro Val Val Val Ser Asn

180 185 190

Thr Ser Thr Leu Val Ser Pro Lys Tyr Gly Val Phe Asn Arg Gly Cys

195 200 205

Met Ile Val Arg Ile Leu Lys Pro Ile Ser Thr Glu Asn Leu Thr Lys

210 215 220

Asp Lys Ile Gly Glu Phe Ala Glu Lys Val Arg Asp Gln Met Val Asp

225 230 235 240

Thr Leu Lys Glu Ile Gly Tyr Ser Pro Ala Ile Asn Asp Thr Thr Leu

245 250 255

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

260 265 270

Asn Lys Lys Ile Lys Asn Glu Pro Val Pro Ser Val Ser Ile Ser Asn

275 280 285

Asp Val Asn Thr His Asn Glu Gly Ser Ser Val Lys Lys Met His

290 295 300

<210> 26

<211> 7

<212> PRT

<213> Artificial Sequence

<220><223> 1-acyl-sn-glycerol-3-phosphate acyltransferase motif

<220><221> misc_feature

<222> (3)..(6)

<223> Xaa can be any naturally occurring amino acid

<400> 26

Asn His Xaa Xaa Xaa Xaa Asp

1 5

<210> 27

<211> 4

<212> PRT

<213> Artificial Sequence

<220><223> 1-acyl-sn-glycerol-3-phosphate acyltransferase motif

<400> 27

Glu Gly Thr Arg

1

<210> 28

<211> 1947

<212> DNA

<213> Yarrowia lipolytica

<220><221> CDS

<222> (1)..(1947)

<223> phospholipid:diacylglycerol acyltransferase

<300><302> ACYLTRANSFERASES FOR ALTERATION OF POLYUNSATURATED FATTY ACIDS
AND OIL CONTENT IN OLEAGINOUS YEASTS

<310> US 7,901,928

<311> 2009-03-11

<312> 2011-03-08

<313> (1)..(1947)

<400> 28

atg aca caa cct gtg aat cgg aag gcg act gtc gag cgg gtc gag cca 48

Met Thr Gln Pro Val Asn Arg Lys Ala Thr Val Glu Arg Val Glu Pro

1 5 10 15

gca gtg gag gtg gct gac tcc gag tcc gag gcc aag acc gac gtc cac 96

Ala Val Glu Val Ala Asp Ser Glu Ser Glu Ala Lys Thr Asp Val His

20 25 30

gtt cac cac cat cat cac cac cac aag cga aaa tcc gtc aag ggc aag 144

Val His His His His His His His Lys Arg Lys Ser Val Lys Gly Lys

35 40 45

att ctc aac ttc ttc acc cga agt cga cgt atc acc ttc gtc ctc ggc 192

Ile Leu Asn Phe Phe Thr Arg Ser Arg Arg Ile Thr Phe Val Leu Gly

50 55 60

gcc gtg gtc ggt gtg ata gcc gcg gga tac tac gct gcg cca ccg gag 240

Ala Val Val Gly Val Ile Ala Ala Gly Tyr Tyr Ala Ala Pro Pro Glu

65 70 75 80

ctc agc att gat atc gat gct ctt ctc ggc gac ttg ccc tcg ttc gac 288

Leu Ser Ile Asp Ile Asp Ala Leu Leu Gly Asp Leu Pro Ser Phe Asp

85 90 95

ttt gac gct cta tct ctc gac aac ttg tcc atg gac agt gtg tcg gac 336

Phe Asp Ala Leu Ser Leu Asp Asn Leu Ser Met Asp Ser Val Ser Asp

100 105 110

ttt gta caa gac atg aaa tcg cgg ttt ccg acc aag att ctg cag gag 384

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

115 120 125

gcg gcc aag atc gag aag cac cag aaa agc gaa cag aag gct gcc cct 432

Ala Ala Lys Ile Glu Lys His Gln Lys Ser Glu Gln Lys Ala Ala Pro

130 135 140

ttt gct gtg ggc aag gct atg aag agc gag gga ctc aac gcc aag tac 480

Phe Ala Val Gly Lys Ala Met Lys Ser Glu Gly Leu Asn Ala Lys Tyr

145 150 155 160

ccg gtg gtg ctg gtg ccc ggc gtc atc tcc acg gga ctg gag agc tgg 528
Pro Val Val Leu Val Pro Gly Val Ile Ser Thr Gly Leu Glu Ser Trp

165 170 175
tcc ctg gag gga acc gag gag tgt ccc acc gag tcg cac ttc aga aag 576
Ser Leu Glu Gly Thr Glu Glu Cys Pro Thr Glu Ser His Phe Arg Lys

180 185 190
cga atg tgg ggc tcc tgg tac atg atc cga gtc atg ctg ctg gac aag 624
Arg Met Trp Gly Ser Trp Tyr Met Ile Arg Val Met Leu Leu Asp Lys

195 200 205

tac tgc tgg ctg cag aac ctg atg ctg gac aca gag acc ggt cta gac 672
Tyr Cys Trp Leu Gln Asn Leu Met Leu Asp Thr Glu Thr Gly Leu Asp

210 215 220
cct ccc cat ttc aag ctg cga gcc gcc cag gga ttt gcc tcc gcc gac 720
Pro Pro His Phe Lys Leu Arg Ala Ala Gln Gly Phe Ala Ser Ala Asp

225 230 235 240
ttc ttt atg gca ggc tac tgg ctg tgg aac aag ctg ctc gag aac ctg 768

Phe Phe Met Ala Gly Tyr Trp Leu Trp Asn Lys Leu Leu Glu Asn Leu

245 250 255
gct gtt att gga tac gat acg gat aca atg tct gct gcg gcg tac gac 816
Ala Val Ile Gly Tyr Asp Thr Asp Thr Met Ser Ala Ala Ala Tyr Asp

260 265 270
tgg aga ctg tcc tac cct gat ttg gag cac cga gac gga tac ttc tcc 864
Trp Arg Leu Ser Tyr Pro Asp Leu Glu His Arg Asp Gly Tyr Phe Ser

275 280 285
aag ctc aaa gct tca atc gaa gag act aag cgt atg aca ggt gag aag 912
Lys Leu Lys Ala Ser Ile Glu Glu Thr Lys Arg Met Thr Gly Glu Lys

290 295 300
aca gtt ctg acg ggc cat tcc atg ggc tcc cag gtc atc ttc tac ttc 960
Thr Val Leu Thr Gly His Ser Met Gly Ser Gln Val Ile Phe Tyr Phe

305	310	315	320	
atg aag tgg gct gag gcc gag gga tat gga gga gga ggt ccc aac tgg				1008
Met Lys Trp Ala Glu Ala Glu Gly Tyr Gly Gly Gly Gly Pro Asn Trp				
325	330	335		
gtc aat gac cat att gaa tcc ttt gtc gac att tcc ggc tcc atg ctg				1056
Val Asn Asp His Ile Glu Ser Phe Val Asp Ile Ser Gly Ser Met Leu				
340	345	350		
ggg act ccc aag acc ctg gtt gct ctt ctg tct gga gaa atg aag gat				1104
Gly Thr Pro Lys Thr Leu Val Ala Leu Leu Ser Gly Glu Met Lys Asp				
355	360	365		
acc gtg cag ctg aac gcg atg gct gtg tat gga ctg gag cag ttc ttc				1152
Thr Val Gln Leu Asn Ala Met Ala Val Tyr Gly Leu Glu Gln Phe Phe				
370	375	380		
tct cga cga gag cga gcc gat ctg ctg cga aca tgg gga gga att gct				1200
Ser Arg Arg Glu Arg Ala Asp Leu Leu Arg Thr Trp Gly Gly Ile Ala				
385	390	395	400	
tcc atg att ccc aag ggt ggt aag gct atc tgg ggt gat cat tct gga				1248
Ser Met Ile Pro Lys Gly Gly Lys Ala Ile Trp Gly Asp His Ser Gly				
405	410	415		
gcc cct gat gac gag ccc ggc cag aat gtc acc ttt ggc aac ttc atc				1296
Ala Pro Asp Asp Glu Pro Gly Gln Asn Val Thr Phe Gly Asn Phe Ile				
420	425	430		
aag ttc aag gag tcc ttg acc gag tac tct gct aag aac ctc acc atg				1344
Lys Phe Lys Glu Ser Leu Thr Glu Tyr Ser Ala Lys Asn Leu Thr Met				
435	440	445		
gat gaa acc gtt gac ttc ctg tat tct cag tct ccc gag tgg ttt gtg				1392
Asp Glu Thr Val Asp Phe Leu Tyr Ser Gln Ser Pro Glu Trp Phe Val				
450	455	460		
aac cga acc gag ggt gct tac tcc ttt gga att gcc aag act cga aag				1440

Asn	Arg	Thr	Glu	Gly	Ala	Tyr	Ser	Phe	Gly	Ile	Ala	Lys	Thr	Arg	Lys	
465			470					475					480			
cag	gtt	gag	cag	aat	gag	aag	cga	cct	tct	acc	tgg	agc	aac	cct	ctg	1488
Gln	Val	Glu	Gln	Asn	Glu	Lys	Arg	Pro	Ser	Thr	Trp	Ser	Asn	Pro	Leu	
485			490					495								
gaa	gct	gct	ctc	ccc	aat	gcc	ccc	gat	ctc	aag	atc	tac	tgc	ttc	tat	1536
Glu	Ala	Ala	Leu	Pro	Asn	Ala	Pro	Asp	Leu	Lys	Ile	Tyr	Cys	Phe	Tyr	
500			505					510								
gga	gtc	ggt	aag	gat	acc	gag	cga	gcc	tac	tac	tac	cag	gat	gag	ccc	1584
Gly	Val	Gly	Lys	Asp	Thr	Glu	Arg	Ala	Tyr	Tyr	Tyr	Gln	Asp	Glu	Pro	
515			520					525								
aat	ccc	gag	cag	acc	aac	ttg	aac	gtc	agt	atc	gct	gga	aac	gac	cct	1632
Asn	Pro	Glu	Gln	Thr	Asn	Leu	Asn	Val	Ser	Ile	Ala	Gly	Asn	Asp	Pro	
530			535					540								
gat	ggt	gtg	ctt	atg	ggt	cag	ggc	gat	gga	acc	gtc	tcc	ctt	gtg	acc	1680
Asp	Gly	Val	Leu	Met	Gly	Gln	Gly	Asp	Gly	Thr	Val	Ser	Leu	Val	Thr	
545			550					555					560			
cat	acc	atg	tgt	cac	cga	tgg	aag	gac	gag	aat	tcc	aag	ttc	aac	cct	1728
His	Thr	Met	Cys	His	Arg	Trp	Lys	Asp	Glu	Asn	Ser	Lys	Phe	Asn	Pro	
565			570					575								
ggt	aac	gcc	cag	gtc	aag	gtt	gtg	gag	atg	ttg	cac	cag	cct	gat	cga	1776
Gly	Asn	Ala	Gln	Val	Lys	Val	Val	Glu	Met	Leu	His	Gln	Pro	Asp	Arg	
580			585					590								
ctt	gat	att	cga	ggc	ggt	gct	cag	act	gcc	gag	cat	gtg	gac	att	ctg	1824
Leu	Asp	Ile	Arg	Gly	Gly	Ala	Gln	Thr	Ala	Glu	His	Val	Asp	Ile	Leu	
595			600					605								
ggg	cgt	tct	gag	ttg	aac	gag	atg	gtt	ctg	aag	gtg	gct	agt	gga	aag	1872
Gly	Arg	Ser	Glu	Leu	Asn	Glu	Met	Val	Leu	Lys	Val	Ala	Ser	Gly	Lys	
610			615					620								
gga	aat	gag	att	gaa	gag	aga	gtc	atc	tcc	aac	att	gat	gag	tgg	gtg	1920
Gly	Asn	Glu	Ile	Glu	Glu	Arg	Val	Ile	Ser	Asn	Ile	Asp	Glu	Trp	Val	

625 630 635 640

tgg aag att gat ctc ggc agc aat tag 1947

Trp Lys Ile Asp Leu Gly Ser Asn

645

<210> 29

<211> 648

<212> PRT

<213> Yarrowia lipolytica

<400> 29

Met Thr Gln Pro Val Asn Arg Lys Ala Thr Val Glu Arg Val Glu Pro

1 5 10 15

Ala Val Glu Val Ala Asp Ser Glu Ser Glu Ala Lys Thr Asp Val His

20 25 30

Val His His His His His His His Lys Arg Lys Ser Val Lys Gly Lys

35 40 45

Ile Leu Asn Phe Phe Thr Arg Ser Arg Arg Ile Thr Phe Val Leu Gly

50 55 60

Ala Val Val Gly Val Ile Ala Ala Gly Tyr Tyr Ala Ala Pro Pro Glu

65 70 75 80

Leu Ser Ile Asp Ile Asp Ala Leu Leu Gly Asp Leu Pro Ser Phe Asp

85 90 95

Phe Asp Ala Leu Ser Leu Asp Asn Leu Ser Met Asp Ser Val Ser Asp

100 105 110

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

115 120 125

Ala Ala Lys Ile Glu Lys His Gln Lys Ser Glu Gln Lys Ala Ala Pro

130 135 140

Phe Ala Val Gly Lys Ala Met Lys Ser Glu Gly Leu Asn Ala Lys Tyr

145 150 155 160

Pro Val Val Leu Val Pro Gly Val Ile Ser Thr Gly Leu Glu Ser Trp

165 170 175

Ser Leu Glu Gly Thr Glu Glu Cys Pro Thr Glu Ser His Phe Arg Lys

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

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

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

370 375 380
Ser Arg Arg Glu Arg Ala Asp Leu Leu Arg Thr Trp Gly Gly Ile Ala
385 390 395 400
Ser Met Ile Pro Lys Gly Gly Lys Ala Ile Trp Gly Asp His Ser Gly
405 410 415
Ala Pro Asp Asp Glu Pro Gly Gln Asn Val Thr Phe Gly Asn Phe Ile

420 425 430
Lys Phe Lys Glu Ser Leu Thr Glu Tyr Ser Ala Lys Asn Leu Thr Met

435 440 445
Asp Glu Thr Val Asp Phe Leu Tyr Ser Gln Ser Pro Glu Trp Phe Val

450 455 460

Asn Arg Thr Glu Gly Ala Tyr Ser Phe Gly Ile Ala Lys Thr Arg Lys

465 470 475 480

Gln Val Glu Gln Asn Glu Lys Arg Pro Ser Thr Trp Ser Asn Pro Leu

485 490 495

Glu Ala Ala Leu Pro Asn Ala Pro Asp Leu Lys Ile Tyr Cys Phe Tyr

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

515 520 525

Asn Pro Glu Gln Thr Asn Leu Asn Val Ser Ile Ala Gly Asn Asp Pro

530 535 540

Asp Gly Val Leu Met Gly Gln Gly Asp Gly Thr Val Ser Leu Val Thr

545 550 555 560

His Thr Met Cys His Arg Trp Lys Asp Glu Asn Ser Lys Phe Asn Pro

565 570 575
Gly Asn Ala Gln Val Lys Val Val Glu Met Leu His Gln Pro Asp Arg

580 585 590

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

595 600 605

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

610 615 620

Gly Asn Glu Ile Glu Glu Arg Val Ile Ser Asn Ile Asp Glu Trp Val

625 630 635 640

Trp Lys Ile Asp Leu Gly Ser Asn

645

<210> 30

<211> 661

<212> PRT

<213> *Saccharomyces cerevisiae*

<220><221> MISC_FEATURE

<222> (1)..(661)

<223> GenBank Accession No. P40345

<400> 30

Met Gly Thr Leu Phe Arg Arg Asn Val Gln Asn Gln Lys Ser Asp Ser

1 5 10 15

Asp Glu Asn Asn Lys Gly Gly Ser Val His Asn Lys Arg Glu Ser Arg

20 25 30

Asn His Ile His His Gln Gln Gly Leu Gly His Lys Arg Arg Arg Gly

35 40 45

Ile Ser Gly Ser Ala Lys Arg Asn Glu Arg Gly Lys Asp Phe Asp Arg

50 55 60

Lys Arg Asp Gly Asn Gly Arg Lys Arg Trp Arg Asp Ser Arg Arg Leu

65 70 75 80

Ile Phe Ile Leu Gly Ala Phe Leu Gly Val Leu Leu Pro Phe Ser Phe

85 90 95

Gly Ala Tyr His Val His Asn Ser Asp Ser Asp Leu Phe Asp Asn Phe

100 105 110

Val Asn Phe Asp Ser Leu Lys Val Tyr Leu Asp Asp Trp Lys Asp Val

115 120 125

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

130 135 140

Tyr Ser Thr Ser Ser Leu Asp Asp Leu Ser Glu Asn Phe Ala Val Gly

145 150 155 160

Lys Gln Leu Leu Arg Asp Tyr Asn Ile Glu Ala Lys His Pro Val Val

165 170 175

Met Val Pro Gly Val Ile Ser Thr Gly Ile Glu Ser Trp Gly Val Ile

180 185 190

Gly Asp Asp Glu Cys Asp Ser Ser Ala His Phe Arg Lys Arg Leu Trp

195 200 205

Gly Ser Phe Tyr Met Leu Arg Thr Met Val Met Asp Lys Val Cys Trp
210 215 220

Leu Lys His Val Met Leu Asp Pro Glu Thr Gly Leu Asp Pro Pro Asn
225 230 235 240

Phe Thr Leu Arg Ala Ala Gln Gly Phe Glu Ser Thr Asp Tyr Phe Ile
245 250 255

Ala Gly Tyr Trp Ile Trp Asn Lys Val Phe Gln Asn Leu Gly Val Ile
260 265 270

Gly Tyr Glu Pro Asn Lys Met Thr Ser Ala Ala Tyr Asp Trp Arg Leu
275 280 285

Ala Tyr Leu Asp Leu Glu Arg Arg Asp Arg Tyr Phe Thr Lys Leu Lys
290 295 300

Glu Gln Ile Glu Leu Phe His Gln Leu Ser Gly Glu Lys Val Cys Leu
305 310 315 320

Ile Gly His Ser Met Gly Ser Gln Ile Ile Phe Tyr Phe Met Lys Trp
325 330 335

Val Glu Ala Glu Gly Pro Leu Tyr Gly Asn Gly Gly Arg Gly Trp Val
340 345 350

Asn Glu His Ile Asp Ser Phe Ile Asn Ala Ala Gly Thr Leu Leu Gly
355 360 365

Ala Pro Lys Ala Val Pro Ala Leu Ile Ser Gly Glu Met Lys Asp Thr
370 375 380

Ile Gln Leu Asn Thr Leu Ala Met Tyr Gly Leu Glu Lys Phe Phe Ser
385 390 395 400

Arg Ile Glu Arg Val Lys Met Leu Gln Thr Trp Gly Gly Ile Pro Ser
405 410 415

Met Leu Pro Lys Gly Glu Glu Val Ile Trp Gly Asp Met Lys Ser Ser
420 425 430

Ser Glu Asp Ala Leu Asn Asn Asn Thr Asp Thr Tyr Gly Asn Phe Ile
435 440 445

Arg Phe Glu Arg Asn Thr Ser Asp Ala Phe Asn Lys Asn Leu Thr Met

450 455 460
 Lys Asp Ala Ile Asn Met Thr Leu Ser Ile Ser Pro Glu Trp Leu Gln
 465 470 475 480

 Arg Arg Val His Glu Gln Tyr Ser Phe Gly Tyr Ser Lys Asn Glu Glu
 485 490 495
 Glu Leu Arg Lys Asn Glu Leu His His Lys His Trp Ser Asn Pro Met
 500 505 510
 Glu Val Pro Leu Pro Glu Ala Pro His Met Lys Ile Tyr Cys Ile Tyr
 515 520 525
 Gly Val Asn Asn Pro Thr Glu Arg Ala Tyr Val Tyr Lys Glu Glu Asp
 530 535 540

 Asp Ser Ser Ala Leu Asn Leu Thr Ile Asp Tyr Glu Ser Lys Gln Pro
 545 550 555 560
 Val Phe Leu Thr Glu Gly Asp Gly Thr Val Pro Leu Val Ala His Ser
 565 570 575
 Met Cys His Lys Trp Ala Gln Gly Ala Ser Pro Tyr Asn Pro Ala Gly
 580 585 590
 Ile Asn Val Thr Ile Val Glu Met Lys His Gln Pro Asp Arg Phe Asp
 595 600 605

 Ile Arg Gly Gly Ala Lys Ser Ala Glu His Val Asp Ile Leu Gly Ser
 610 615 620
 Ala Glu Leu Asn Asp Tyr Ile Leu Lys Ile Ala Ser Gly Asn Gly Asp
 625 630 635 640
 Leu Val Glu Pro Arg Gln Leu Ser Asn Leu Ser Gln Trp Val Ser Gln
 645 650 655
 Met Pro Phe Pro Met
 660

 <210> 31
 <211> 777
 <212> DNA
 <213> Euglena gracilis
 <220><

221> CDS

<222> (1)..(777)

<223> delta-9 elongase

<300><302> DELTA-9 ELONGASES AND THEIR USE IN MAKING POLYUNSATURATED FATTY
ACIDS

<310> WO 2007/061742

<311> 2006-11-16

<312> 2007-05-31

<313> (1)..(777)

<300><302> DELTA-9 ELONGASES AND THEIR USE IN MAKING POLYUNSATURATED FATTY
ACIDS

<310> US 7,645,604

<311> 2006-11-16

<312> 2010-01-12

<313> (1)..(777)

<400> 31

atg gag gtg gtg aat gaa ata gtc tca att ggg cag gaa gtt tta ccc 48
Met Glu Val Val Asn Glu Ile Val Ser Ile Gly Gln Glu Val Leu Pro

1 5 10 15
aaa gtt gat tat gcc caa ctc tgg agt gat gcc agt cac tgt gag gtg 96
Lys Val Asp Tyr Ala Gln Leu Trp Ser Asp Ala Ser His Cys Glu Val

20 25 30
ctt tac ttg tcc atc gca ttt gtc atc ttg aag ttc act ctt ggc ccc 144
Leu Tyr Leu Ser Ile Ala Phe Val Ile Leu Lys Phe Thr Leu Gly Pro

35 40 45
ctt ggt cca aaa ggt cag tct cgt atg aag ttt gtt ttc acc aat tac 192
Leu Gly Pro Lys Gly Gln Ser Arg Met Lys Phe Val Phe Thr Asn Tyr

50 55 60
aac ctt ctc atg tcc att tat tcg ttg gga tca ttc ctc tca atg gca 240
Asn Leu Leu Met Ser Ile Tyr Ser Leu Gly Ser Phe Leu Ser Met Ala

65 70 75 80
tat gcc atg tac acc atc ggt gtt atg tct gac aac tgc gag aag gct 288

Tyr Ala Met Tyr Thr Ile Gly Val Met Ser Asp Asn Cys Glu Lys Ala	
85 90 95	
ttt gac aac aac gtc ttc agg atc acc acg cag ttg ttc tat ttg agc	336
Phe Asp Asn Asn Val Phe Arg Ile Thr Thr Gln Leu Phe Tyr Leu Ser	
100 105 110	
aag ttc ctg gag tat att gac tcc ttc tat ttg cca ctg atg ggc aag	384
Lys Phe Leu Glu Tyr Ile Asp Ser Phe Tyr Leu Pro Leu Met Gly Lys	
115 120 125	
cct ctg acc tgg ttg caa ttc ttc cat cat ttg ggg gca ccg atg gat	432
Pro Leu Thr Trp Leu Gln Phe Phe His His Leu Gly Ala Pro Met Asp	
130 135 140	
atg tgg ctg ttc tat aat tac cga aat gaa gct gtt tgg att ttt gtg	480
Met Trp Leu Phe Tyr Asn Tyr Arg Asn Glu Ala Val Trp Ile Phe Val	
145 150 155 160	
ctg ttg aat ggt ttc atc cac tgg atc atg tac ggt tat tat tgg acc	528
Leu Leu Asn Gly Phe Ile His Trp Ile Met Tyr Gly Tyr Tyr Trp Thr	
165 170 175	
aga ttg atc aag ctg aag ttc ccc atg cca aaa tcc ctg att aca tca	576
Arg Leu Ile Lys Leu Lys Phe Pro Met Pro Lys Ser Leu Ile Thr Ser	
180 185 190	
atg cag atc att caa ttc aat gtt ggt ttc tac att gtc tgg aag tac	624
Met Gln Ile Ile Gln Phe Asn Val Gly Phe Tyr Ile Val Trp Lys Tyr	
195 200 205	
agg aac att ccc tgt tat cgc caa gat ggg atg agg atg ttt ggc tgg	672
Arg Asn Ile Pro Cys Tyr Arg Gln Asp Gly Met Arg Met Phe Gly Trp	
210 215 220	
ttc ttc aat tac ttt tat gtt ggc aca gtc ttg tgt ttg ttc ttg aat	720
Phe Phe Asn Tyr Phe Tyr Val Gly Thr Val Leu Cys Leu Phe Leu Asn	
225 230 235 240	
ttc tat gtg caa acg tat atc gtc agg aag cac aag gga gcc aaa aag	768
Phe Tyr Val Gln Thr Tyr Ile Val Arg Lys His Lys Gly Ala Lys Lys	

245 250 255

att cag tga 777

Ile Gln

<210> 32

<211> 258

<212> PRT

<213> Euglena gracilis

<400> 32

Met Glu Val Val Asn Glu Ile Val Ser Ile Gly Gln Glu Val Leu Pro

1 5 10 15

Lys Val Asp Tyr Ala Gln Leu Trp Ser Asp Ala Ser His Cys Glu Val

20 25 30

Leu Tyr Leu Ser Ile Ala Phe Val Ile Leu Lys Phe Thr Leu Gly Pro

35 40 45

Leu Gly Pro Lys Gly Gln Ser Arg Met Lys Phe Val Phe Thr Asn Tyr

50 55 60

Asn Leu Leu Met Ser Ile Tyr Ser Leu Gly Ser Phe Leu Ser Met Ala

65 70 75 80

Tyr Ala Met Tyr Thr Ile Gly Val Met Ser Asp Asn Cys Glu Lys Ala

85 90 95

Phe Asp Asn Asn Val Phe Arg Ile Thr Thr Gln Leu Phe Tyr Leu Ser

100 105 110

Lys Phe Leu Glu Tyr Ile Asp Ser Phe Tyr Leu Pro Leu Met Gly Lys

115 120 125

Pro Leu Thr Trp Leu Gln Phe Phe His His Leu Gly Ala Pro Met Asp

130 135 140

Met Trp Leu Phe Tyr Asn Tyr Arg Asn Glu Ala Val Trp Ile Phe Val

145 150 155 160

Leu Leu Asn Gly Phe Ile His Trp Ile Met Tyr Gly Tyr Tyr Trp Thr

165 170 175

Arg Leu Ile Lys Leu Lys Phe Pro Met Pro Lys Ser Leu Ile Thr Ser

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

245 250 255
Ile Gln

<210> 33

<211> 774

<212> DNA

<213> Euglena anabaena UTEX 373

<220><221> CDS

<222> (1)..(774)

<223> delta-9 elongase

<300><302> DELTA-9 ELONGASES AND THEIR USE IN MAKING POLYUNSATURATED FATTY
ACIDS

<310> U.S. 7,794,701

<311> 2008-04-15

<312> 2010-09-14

<313> (1)..(774)

<300><302> DELTA-9 ELONGASES AND THEIR USE IN MAKING POLYUNSATURATED FATTY
ACIDS

<310> WO 2008/128241

<311> 2008-04-16

<312> 2008-10-23

<313> (1)..(774)

<400> 33

atg gaa gca gcc aaa gaa ttg gtt tcc atc gtc caa gag gag ctc ccc	48
Met Glu Ala Ala Lys Glu Leu Val Ser Ile Val Gln Glu Glu Leu Pro	
1 5 10 15	
aag gtg gac tat gcc cag ctt tgg cag gat gcc agc agc tgt gag gtc	96
Lys Val Asp Tyr Ala Gln Leu Trp Gln Asp Ala Ser Ser Cys Glu Val	
20 25 30	
ctt tac ctc tcg gtg gca ttc gtg gcg atc aag ttc atg ctg cgc cca	144
Leu Tyr Leu Ser Val Ala Phe Val Ala Ile Lys Phe Met Leu Arg Pro	
35 40 45	
ctg gac ctg aag cgc cag gcc acc ttg aag aag ctg ttc aca gca tac	192
Leu Asp Leu Lys Arg Gln Ala Thr Leu Lys Lys Leu Phe Thr Ala Tyr	
50 55 60	
aac ttc ctc atg tcg atc tat tcc ttt ggc tcc ttc ctg gcc atg gcc	240
Asn Phe Leu Met Ser Ile Tyr Ser Phe Gly Ser Phe Leu Ala Met Ala	
65 70 75 80	
tat gcc cta tca gta act ggc atc ctc tcc ggc gac tgt gag acg gcg	288
Tyr Ala Leu Ser Val Thr Gly Ile Leu Ser Gly Asp Cys Glu Thr Ala	
85 90 95	
ttc aac aac gat gtg ttc agg atc aca act cag ctg ttc tac ctc agc	336
Phe Asn Asn Asp Val Phe Arg Ile Thr Thr Gln Leu Phe Tyr Leu Ser	
100 105 110	
aag ttc gta gag tac atc gac tcc ttc tac ctt ccc ctt atg gac aag	384
Lys Phe Val Glu Tyr Ile Asp Ser Phe Tyr Leu Pro Leu Met Asp Lys	
115 120 125	
cca ctg tcg ttc ctt cag ttc ttc cat cat ttg ggg gcc ccc att gac	432
Pro Leu Ser Phe Leu Gln Phe Phe His His Leu Gly Ala Pro Ile Asp	
130 135 140	
atg tgg cta ttc tac aaa tac cgc aac gaa gga gtc tgg atc ttt gtc	480
Met Trp Leu Phe Tyr Lys Tyr Arg Asn Glu Gly Val Trp Ile Phe Val	
145 150 155 160	
ctg ttg aat ggg ttc att cac tgg atc atg tac ggt tac tat tgg acg	528

Leu Leu Asn Gly Phe Ile His Trp Ile Met Tyr Gly Tyr Tyr Trp Thr
 165 170 175
 cgg ctc atc aag ctg aac ttc ccc atg ccc aag aac ctg atc acc tcc 576

Arg Leu Ile Lys Leu Asn Phe Pro Met Pro Lys Asn Leu Ile Thr Ser
 180 185 190
 atg cag atc atc cag ttc aat gtc ggg ttc tac atc gtc tgg aag tac 624
 Met Gln Ile Ile Gln Phe Asn Val Gly Phe Tyr Ile Val Trp Lys Tyr
 195 200 205
 cgc aat gtg cca tgc tac cgc cag gat ggg atg cgc atg ttt gcc tgg 672
 Arg Asn Val Pro Cys Tyr Arg Gln Asp Gly Met Arg Met Phe Ala Trp

210 215 220
 atc ttc aac tac tgg tat gtc ggg acg gtc ttg ctg ctg ttc ctc aac 720
 Ile Phe Asn Tyr Trp Tyr Val Gly Thr Val Leu Leu Leu Phe Leu Asn
 225 230 235 240
 ttt tac gtg cag acg tac atc cgg aag ccg agg aag aac cga ggg aag 768
 Phe Tyr Val Gln Thr Tyr Ile Arg Lys Pro Arg Lys Asn Arg Gly Lys
 245 250 255

aag gag 774
 Lys Glu

<210> 34

<211> 258

<212> PRT

<213> Euglena anabaena UTEX 373

<400> 34

Met Glu Ala Ala Lys Glu Leu Val Ser Ile Val Gln Glu Glu Leu Pro
 1 5 10 15
 Lys Val Asp Tyr Ala Gln Leu Trp Gln Asp Ala Ser Ser Cys Glu Val

20 25 30
 Leu Tyr Leu Ser Val Ala Phe Val Ala Ile Lys Phe Met Leu Arg Pro
 35 40 45

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

<210> 35

<211> 774

<212> DNA

<213> Euglena anabaena UTEX 373

<220><221> CDS

<222> (1)..(774)

<223> synthetic delta-9 elongase (codon-optimized for Yarrowia

lipolytica)

<300><302> DELTA-9 ELONGASES AND THEIR USE IN MAKING POLYUNSATURATED FATTY ACIDS

<310> U.S. 7,794,701

<311> 2008-04-15

<312> 2010-09-14

<313> (1)..(774)

<300><302> DELTA-9 ELONGASES AND THEIR USE IN MAKING POLYUNSATURATED FATTY ACIDS

<310> WO 2008/128241

<311> 2008-04-16

<312> 2008-10-23

<313> (1)..(774)

<400> 35

atg gag gct gcc aag gag ctg gtc tcc atc gtc cag gag gaa ctt ccc 48
Met Glu Ala Ala Lys Glu Leu Val Ser Ile Val Gln Glu Glu Leu Pro

1 5 10 15
aag gtg gac tac gcc cag ctc tgg cag gac gcc tcc tct tgc gag gtt 96
Lys Val Asp Tyr Ala Gln Leu Trp Gln Asp Ala Ser Ser Cys Glu Val

20 25 30
ctg tac ctc tcg gtc gct ttc gtg gcc atc aag ttc atg ctt cga cct 144
Leu Tyr Leu Ser Val Ala Phe Val Ala Ile Lys Phe Met Leu Arg Pro

35 40 45
ctg gac ctc aag cga caa gcc acc ctc aaa aag ctg ttc acc gca tac 192
Leu Asp Leu Lys Arg Gln Ala Thr Leu Lys Lys Leu Phe Thr Ala Tyr

50 55 60
aac ttt ctc atg tcc atc tac tcg ttc ggc tcc ttc ctg gcg atg gcc 240
Asn Phe Leu Met Ser Ile Tyr Ser Phe Gly Ser Phe Leu Ala Met Ala

65 70 75 80

tac gct ctc tct gtc act ggt att ctt tcc ggc gat tgt gag act gcc	288
Tyr Ala Leu Ser Val Thr Gly Ile Leu Ser Gly Asp Cys Glu Thr Ala	
85 90 95	
ttc aac aat gac gtg ttc cga atc acc act cag ctg ttc tac ctc agc	336
Phe Asn Asn Asp Val Phe Arg Ile Thr Thr Gln Leu Phe Tyr Leu Ser	
100 105 110	
aag ttc gtc gag tac atc gac tcc ttc tac ctt ccc ctc atg gac aag	384
Lys Phe Val Glu Tyr Ile Asp Ser Phe Tyr Leu Pro Leu Met Asp Lys	
115 120 125	
ccc ttg tcg ttt ctg cag ttc ttt cac cat ctc gga gct ccc atc gac	432
Pro Leu Ser Phe Leu Gln Phe Phe His His Leu Gly Ala Pro Ile Asp	
130 135 140	
atg tgg ctg ttc tac aag tat cga aac gaa ggc gtc tgg atc ttt gtt	480
Met Trp Leu Phe Tyr Lys Tyr Arg Asn Glu Gly Val Trp Ile Phe Val	
145 150 155 160	
ctg ctc aac ggc ttc att cac tgg atc atg tac ggt tac tat tgg acg	528
Leu Leu Asn Gly Phe Ile His Trp Ile Met Tyr Gly Tyr Tyr Trp Thr	
165 170 175	
cga ctc atc aag ctg aac ttc cct atg ccc aag aac ctc att acc tcc	576
Arg Leu Ile Lys Leu Asn Phe Pro Met Pro Lys Asn Leu Ile Thr Ser	
180 185 190	
atg caa att atc cag ttc aac gtc gga ttc tac atc gtc tgg aag tac	624
Met Gln Ile Ile Gln Phe Asn Val Gly Phe Tyr Ile Val Trp Lys Tyr	
195 200 205	
cga aac gtg ccc tgc tac cgg cag gac ggt atg cga atg ttt gcc tgg	672
Arg Asn Val Pro Cys Tyr Arg Gln Asp Gly Met Arg Met Phe Ala Trp	
210 215 220	
atc ttc aac tac tgg tat gtc ggc acg gtg ctg ctt ctg ttc ctc aac	720
Ile Phe Asn Tyr Trp Tyr Val Gly Thr Val Leu Leu Leu Phe Leu Asn	

225	230	235	240	
ttc tac gtc cag acc tac att cgg aag cct cga aag aac cga ggc aaa				768
Phe Tyr Val Gln Thr Tyr Ile Arg Lys Pro Arg Lys Asn Arg Gly Lys				
245	250	255		
aag gag				774
Lys Glu				

<210> 36
<211> 258
<212> PRT
<213> Euglena anabaena UTEX 373
<400> 36

Met Glu Ala Ala Lys Glu Leu Val Ser Ile Val Gln Glu Glu Leu Pro
1 5 10 15
Lys Val Asp Tyr Ala Gln Leu Trp Gln Asp Ala Ser Ser Cys Glu Val
20 25 30
Leu Tyr Leu Ser Val Ala Phe Val Ala Ile Lys Phe Met Leu Arg Pro
35 40 45
Leu Asp Leu Lys Arg Gln Ala Thr Leu Lys Lys Leu Phe Thr Ala Tyr
50 55 60
Asn Phe Leu Met Ser Ile Tyr Ser Phe Gly Ser Phe Leu Ala Met Ala
65 70 75 80
Tyr Ala Leu Ser Val Thr Gly Ile Leu Ser Gly Asp Cys Glu Thr Ala
85 90 95
Phe Asn Asn Asp Val Phe Arg Ile Thr Thr Gln Leu Phe Tyr Leu Ser
100 105 110
Lys Phe Val Glu Tyr Ile Asp Ser Phe Tyr Leu Pro Leu Met Asp Lys
115 120 125
Pro Leu Ser Phe Leu Gln Phe Phe His His Leu Gly Ala Pro Ile Asp
130 135 140
Met Trp Leu Phe Tyr Lys Tyr Arg Asn Glu Gly Val Trp Ile Phe Val
145 150 155 160

Leu Leu Asn Gly Phe Ile His Trp Ile Met Tyr Gly Tyr Tyr Trp Thr

165 170 175

Arg Leu Ile Lys Leu Asn Phe Pro Met Pro Lys Asn Leu Ile Thr Ser

180 185 190

Met Gln Ile Ile Gln Phe Asn Val Gly Phe Tyr Ile Val Trp Lys Tyr

195 200 205

Arg Asn Val Pro Cys Tyr Arg Gln Asp Gly Met Arg Met Phe Ala Trp

210 215 220

Ile Phe Asn Tyr Trp Tyr Val Gly Thr Val Leu Leu Leu Phe Leu Asn

225 230 235 240

Phe Tyr Val Gln Thr Tyr Ile Arg Lys Pro Arg Lys Asn Arg Gly Lys

245 250 255

Lys Glu

<210> 37

<211> 792

<212> DNA

<213> Eutreptiella sp. CCMP389

<220><221> CDS

<222> (1)..(792)

<223> delta-9 elongase

<300><302> DELTA-9 ELONGASES AND THEIR USE IN MAKING POLYUNSATURATED FATTY

ACIDS

<310> WO 2007/061742

<311> 2006-11-16

<312> 2007-05-31

<313> (1)..(792)

<300><302> DELTA-9 ELONGASES AND THEIR USE IN MAKING POLYUNSATURATED FATTY

ACIDS

<310> US 7,645,604

<311> 2006-11-16

<312> 2010-01-12

<313> (1)..(792)

<400> 37

atg gct gcg gtg ata gag gtc gcc aac gag ttt gta gcc atc acg gca 48

Met Ala Ala Val Ile Glu Val Ala Asn Glu Phe Val Ala Ile Thr Ala

1 5 10 15

gaa acg ctc ccc aaa gtt gac tat caa cga cta tgg cga gac att tac 96

Glu Thr Leu Pro Lys Val Asp Tyr Gln Arg Leu Trp Arg Asp Ile Tyr

20 25 30

agt tgt gag cta ctg tat ttc tcc att gcc ttc gtg atc ttg aag ttt 144

Ser Cys Glu Leu Leu Tyr Phe Ser Ile Ala Phe Val Ile Leu Lys Phe

35 40 45

acg ttg ggc gag ttg agc gac agc gga aaa aag att ttg aga gtg ttg 192

Thr Leu Gly Glu Leu Ser Asp Ser Gly Lys Lys Ile Leu Arg Val Leu

50 55 60

ttc aag tgg tac aat ctc ttc atg tcc gtg ttc tcc ttg gtg tct ttc 240

Phe Lys Trp Tyr Asn Leu Phe Met Ser Val Phe Ser Leu Val Ser Phe

65 70 75 80

ctt tgc atg ggc tat gcc att tat acc gtg ggc cta tac tct aac gaa 288

Leu Cys Met Gly Tyr Ala Ile Tyr Thr Val Gly Leu Tyr Ser Asn Glu

85 90 95

tgc gac agg gct ttc gac aac tgc ttg ttc cgc ttt gca aca aag gtg 336

Cys Asp Arg Ala Phe Asp Asn Ser Leu Phe Arg Phe Ala Thr Lys Val

100 105 110

ttc tac tac agt aag ttt ttg gag tac atc gac tct ttt tat ctt ccg 384

Phe Tyr Tyr Ser Lys Phe Leu Glu Tyr Ile Asp Ser Phe Tyr Leu Pro

115 120 125

ctc atg gcc aag ccg ctg tct ttc ctg caa ttc ttc cat cac ttg gga 432

Leu Met Ala Lys Pro Leu Ser Phe Leu Gln Phe Phe His His Leu Gly

130 135 140

gcc ccc atg gac atg tgg ctc ttt gtc caa tat tct ggg gaa tct att 480

Ala Pro Met Asp Met Trp Leu Phe Val Gln Tyr Ser Gly Glu Ser Ile

145	150	155	160	
tgg atc ttt gtg ttt ttg aat ggg ttc att cac ttt gtt atg tac ggg				528
Trp Ile Phe Val Phe Leu Asn Gly Phe Ile His Phe Val Met Tyr Gly				
	165	170	175	
tac tac tgg act cgg ctg atg aag ttc aat ttc cca atg ccc aag cag				576
Tyr Tyr Trp Thr Arg Leu Met Lys Phe Asn Phe Pro Met Pro Lys Gln				
	180	185	190	
ttg att acc gcg atg cag atc acg cag ttc aac gtt ggt ttc tac ctc				624
Leu Ile Thr Ala Met Gln Ile Thr Gln Phe Asn Val Gly Phe Tyr Leu				
	195	200	205	
gtg tgg tgg tac aaa gat att ccc tgc tac cga aag gat ccc atg cga				672
Val Trp Trp Tyr Lys Asp Ile Pro Cys Tyr Arg Lys Asp Pro Met Arg				
	210	215	220	
atg ttg gcc tgg atc ttc aat tac tgg tat gtt ggg act gtc ttg ctg				720
Met Leu Ala Trp Ile Phe Asn Tyr Trp Tyr Val Gly Thr Val Leu Leu				
	225	230	235	240
ctg ttc att aat ttc ttc gtc aaa tcc tat gtg ttc cca aag ccg aag				768
Leu Phe Ile Asn Phe Phe Val Lys Ser Tyr Val Phe Pro Lys Pro Lys				
	245	250	255	
act gca gat aaa aag gtc caa tag				792
Thr Ala Asp Lys Lys Val Gln				
	260			
<210> 38				
<211> 263				
<212> PRT				
<213> Eutreptiella sp. CCMP389				
<400> 38				
Met Ala Ala Val Ile Glu Val Ala Asn Glu Phe Val Ala Ile Thr Ala				
1	5	10	15	
Glu Thr Leu Pro Lys Val Asp Tyr Gln Arg Leu Trp Arg Asp Ile Tyr				
	20	25	30	

Ser Cys Glu Leu Leu Tyr Phe Ser Ile Ala Phe Val Ile Leu Lys Phe
 35 40 45
 Thr Leu Gly Glu Leu Ser Asp Ser Gly Lys Lys Ile Leu Arg Val Leu
 50 55 60
 Phe Lys Trp Tyr Asn Leu Phe Met Ser Val Phe Ser Leu Val Ser Phe
 65 70 75 80
 Leu Cys Met Gly Tyr Ala Ile Tyr Thr Val Gly Leu Tyr Ser Asn Glu
 85 90 95
 Cys Asp Arg Ala Phe Asp Asn Ser Leu Phe Arg Phe Ala Thr Lys Val
 100 105 110
 Phe Tyr Tyr Ser Lys Phe Leu Glu Tyr Ile Asp Ser Phe Tyr Leu Pro
 115 120 125
 Leu Met Ala Lys Pro Leu Ser Phe Leu Gln Phe Phe His His Leu Gly
 130 135 140
 Ala Pro Met Asp Met Trp Leu Phe Val Gln Tyr Ser Gly Glu Ser Ile
 145 150 155 160
 Trp Ile Phe Val Phe Leu Asn Gly Phe Ile His Phe Val Met Tyr Gly
 165 170 175
 Tyr Tyr Trp Thr Arg Leu Met Lys Phe Asn Phe Pro Met Pro Lys Gln
 180 185 190
 Leu Ile Thr Ala Met Gln Ile Thr Gln Phe Asn Val Gly Phe Tyr Leu
 195 200 205
 Val Trp Trp Tyr Lys Asp Ile Pro Cys Tyr Arg Lys Asp Pro Met Arg
 210 215 220
 Met Leu Ala Trp Ile Phe Asn Tyr Trp Tyr Val Gly Thr Val Leu Leu
 225 230 235 240
 Leu Phe Ile Asn Phe Phe Val Lys Ser Tyr Val Phe Pro Lys Pro Lys
 245 250 255
 Thr Ala Asp Lys Lys Val Gln
 260
 <210> 39
 <211> 792

<212> DNA
 <213> Eutreptiella sp. CCMP389
 <220><221> CDS
 <222> (1)..(792)
 <223> synthetic delta-9 elongase (codon-optimized for Yarrowia

lipolytica)

<300><302> DELTA-9 ELONGASES AND THEIR USE IN MAKING POLYUNSATURATED FATTY
 ACIDS

<310> WO 2007/061742

<311> 2006-11-16

<312> 2007-05-31

<313> (1)..(792)

<300><302> DELTA-9 ELONGASES AND THEIR USE IN MAKING POLYUNSATURATED FATTY
 ACIDS

<310> US 7,645,604

<311> 2006-11-16

<312> 2010-01-12

<313> (1)..(792)

<400> 39

atg get gcc gtc atc gag gtg gcc aac gag ttc gtc gct atc act gcc 48
 Met Ala Ala Val Ile Glu Val Ala Asn Glu Phe Val Ala Ile Thr Ala

1 5 10 15
 gag acc ctt ccc aag gtg gac tat cag cga ctc tgg cga gac atc tac 96
 Glu Thr Leu Pro Lys Val Asp Tyr Gln Arg Leu Trp Arg Asp Ile Tyr

20 25 30
 tcc tgc gag ctc ctg tac ttc tcc att gct ttc gtc atc ctc aag ttt 144
 Ser Cys Glu Leu Leu Tyr Phe Ser Ile Ala Phe Val Ile Leu Lys Phe

35 40 45
 acc ctt ggc gag ctc tgc gat tct ggc aaa aag att ctg cga gtg ctg 192
 Thr Leu Gly Glu Leu Ser Asp Ser Gly Lys Lys Ile Leu Arg Val Leu

50 55 60
 ttc aag tgg tac aac ctc ttc atg tcc gtc ttt tgc ctg gtg tcc ttc 240

Phe	Lys	Trp	Tyr	Asn	Leu	Phe	Met	Ser	Val	Phe	Ser	Leu	Val	Ser	Phe		
				70				75				80					
ctc	tgt	atg	ggt	tac	gcc	atc	tac	acc	gtt	gga	ctg	tac	tcc	aac	gaa	288	
Leu	Cys	Met	Gly	Tyr	Ala	Ile	Tyr	Thr	Val	Gly	Leu	Tyr	Ser	Asn	Glu		
				85				90				95					
tgc	gac	aga	gct	ttc	gac	aac	agc	ttg	ttc	cga	ttt	gcc	acc	aag	gtc	336	
Cys	Asp	Arg	Ala	Phe	Asp	Asn	Ser	Leu	Phe	Arg	Phe	Ala	Thr	Lys	Val		
				100				105				110					
ttc	tac	tat	tcc	aag	ttt	ctg	gag	tac	atc	gac	tct	ttc	tac	ctt	ccc	384	
Phe	Tyr	Tyr	Ser	Lys	Phe	Leu	Glu	Tyr	Ile	Asp	Ser	Phe	Tyr	Leu	Pro		
				115				120				125					
ctc	atg	gcc	aag	cct	ctg	tcc	ttt	ctg	cag	ttc	ttt	cat	cac	ttg	gga	432	
Leu	Met	Ala	Lys	Pro	Leu	Ser	Phe	Leu	Gln	Phe	Phe	His	His	Leu	Gly		
				130				135				140					
gct	cct	atg	gac	atg	tgg	ctc	ttc	gtg	cag	tac	tct	ggc	gaa	tcc	att	480	
Ala	Pro	Met	Asp	Met	Trp	Leu	Phe	Val	Gln	Tyr	Ser	Gly	Glu	Ser	Ile		
				145				150				155				160	
tgg	atc	ttt	gtg	ttc	ctg	aac	gga	ttc	att	cac	ttt	gtc	atg	tac	ggc	528	
Trp	Ile	Phe	Val	Phe	Leu	Asn	Gly	Phe	Ile	His	Phe	Val	Met	Tyr	Gly		
				165				170				175					
tac	tat	tgg	aca	cgg	ctg	atg	aag	ttc	aac	ttt	ccc	atg	ccc	aag	cag	576	
Tyr	Tyr	Trp	Thr	Arg	Leu	Met	Lys	Phe	Asn	Phe	Pro	Met	Pro	Lys	Gln		
				180				185				190					
ctc	att	acc	gca	atg	cag	atc	acc	cag	ttc	aac	gtt	ggc	ttc	tac	ctc	624	
Leu	Ile	Thr	Ala	Met	Gln	Ile	Thr	Gln	Phe	Asn	Val	Gly	Phe	Tyr	Leu		
				195				200				205					
gtg	tgg	tgg	tac	aag	gac	att	ccc	tgt	tac	cga	aag	gat	ccc	atg	cga	672	
Val	Trp	Trp	Tyr	Lys	Asp	Ile	Pro	Cys	Tyr	Arg	Lys	Asp	Pro	Met	Arg		
				210				215				220					
atg	ctg	gcc	tgg	atc	ttc	aac	tac	tgg	tac	gtc	ggc	acc	gtt	ctt	ctg	720	

Met Leu Ala Trp Ile Phe Asn Tyr Trp Tyr Val Gly Thr Val Leu Leu

225 230 235 240
 ctc ttc atc aac ttc ttt gtc aag tcc tac gtg ttt ccc aag cct aag 768
 Leu Phe Ile Asn Phe Phe Val Lys Ser Tyr Val Phe Pro Lys Pro Lys
 245 250 255
 act gcc gac aaa aag gtc cag tag 792
 Thr Ala Asp Lys Lys Val Gln
 260

<210> 40
 <211> 263
 <212> PRT
 <213> Eutreptiella sp. CCMP389
 <400> 40

Met Ala Ala Val Ile Glu Val Ala Asn Glu Phe Val Ala Ile Thr Ala
 1 5 10 15
 Glu Thr Leu Pro Lys Val Asp Tyr Gln Arg Leu Trp Arg Asp Ile Tyr
 20 25 30
 Ser Cys Glu Leu Leu Tyr Phe Ser Ile Ala Phe Val Ile Leu Lys Phe
 35 40 45
 Thr Leu Gly Glu Leu Ser Asp Ser Gly Lys Lys Ile Leu Arg Val Leu

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

115 120 125
 Leu Met Ala Lys Pro Leu Ser Phe Leu Gln Phe Phe His His Leu Gly
 130 135 140

Ala Pro Met Asp Met Trp Leu Phe Val Gln Tyr Ser Gly Glu Ser Ile
 145 150 155 160
 Trp Ile Phe Val Phe Leu Asn Gly Phe Ile His Phe Val Met Tyr Gly
 165 170 175
 Tyr Tyr Trp Thr Arg Leu Met Lys Phe Asn Phe Pro Met Pro Lys Gln
 180 185 190
 Leu Ile Thr Ala Met Gln Ile Thr Gln Phe Asn Val Gly Phe Tyr Leu
 195 200 205
 Val Trp Trp Tyr Lys Asp Ile Pro Cys Tyr Arg Lys Asp Pro Met Arg
 210 215 220
 Met Leu Ala Trp Ile Phe Asn Tyr Trp Tyr Val Gly Thr Val Leu Leu
 225 230 235 240
 Leu Phe Ile Asn Phe Phe Val Lys Ser Tyr Val Phe Pro Lys Pro Lys

245 250 255
 Thr Ala Asp Lys Lys Val Gln
 260

<210> 41

<211> 1064

<212> DNA

<213> Isochrysis galbana (GenBank Accession No. AF390174)

<220><221> CDS

<222> (2)..(793)

<223> delta-9 elongase (IgD9e)

<400> 41

g atg gcc ctc gca aac gac gcg gga gag cgc atc tgg gcg gct gtg acc 49

Met Ala Leu Ala Asn Asp Ala Gly Glu Arg Ile Trp Ala Ala Val Thr

1 5 10 15

gac ccg gaa atc ctc att ggc acc ttc tcg tac ttg cta ctc aaa ccg 97

Asp Pro Glu Ile Leu Ile Gly Thr Phe Ser Tyr Leu Leu Leu Lys Pro

20 25 30

ctg ctc cgc aat tcc ggg ctg gtg gat gag aag aag ggc gca tac agg 145

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

35	40	45	
acg tcc atg atc tgg tac aac gtt ctg ctg gcg ctc ttc tct gcg ctg			193
Thr Ser Met Ile Trp Tyr Asn Val Leu Leu Ala Leu Phe Ser Ala Leu			
50	55	60	
agc ttc tac gtg acg gcg acc gcc ctc ggc tgg gac tat ggt acg ggc			241
Ser Phe Tyr Val Thr Ala Thr Ala Leu Gly Trp Asp Tyr Gly Thr Gly			
65	70	75	80
gcg tgg ctg cgc agg caa acc ggc gac aca ccg cag ccg ctc ttc cag			289
Ala Trp Leu Arg Arg Gln Thr Gly Asp Thr Pro Gln Pro Leu Phe Gln			
85	90	95	
tgc ccg tcc ccg gtt tgg gac tgc aag ctc ttc aca tgg acc gcc aag			337
Cys Pro Ser Pro Val Trp Asp Ser Lys Leu Phe Thr Trp Thr Ala Lys			
100	105	110	
gca ttc tat tac tcc aag tac gtg gag tac ctc gac acg gcc tgg ctg			385
Ala Phe Tyr Tyr Ser Lys Tyr Val Glu Tyr Leu Asp Thr Ala Trp Leu			
115	120	125	
gtg ctc aag ggc aag agg gtc tcc ttt ctc cag gcc ttc cac cac ttt			433
Val Leu Lys Gly Lys Arg Val Ser Phe Leu Gln Ala Phe His His Phe			
130	135	140	
ggc gcg ccg tgg gat gtg tac ctc ggc att cgg ctg cac aac gag ggc			481
Gly Ala Pro Trp Asp Val Tyr Leu Gly Ile Arg Leu His Asn Glu Gly			
145	150	155	160
gta tgg atc ttc atg ttt ttc aac tcg ttc att cac acc atc atg tac			529
Val Trp Ile Phe Met Phe Phe Asn Ser Phe Ile His Thr Ile Met Tyr			
165	170	175	
acc tac tac ggc ctc acc gcc gcc ggg tat aag ttc aag gcc aag ccg			577
Thr Tyr Tyr Gly Leu Thr Ala Ala Gly Tyr Lys Phe Lys Ala Lys Pro			
180	185	190	
ctc atc acc gcg atg cag atc tgc cag ttc gtg ggc ggc ttc ctg ttg			625

Leu Ile Thr Ala Met Gln Ile Cys Gln Phe Val Gly Gly Phe Leu Leu

195 200 205
gtc tgg gac tac atc aac gtc ccc tgc ttc aac tgc gac aaa ggg aag 673

Val Trp Asp Tyr Ile Asn Val Pro Cys Phe Asn Ser Asp Lys Gly Lys

210 215 220
ttg ttc agc tgg gct ttc aac tat gca tac gtc ggc tgc gtc ttc ttg 721

Leu Phe Ser Trp Ala Phe Asn Tyr Ala Tyr Val Gly Ser Val Phe Leu

225 230 235 240

ctc ttc tgc cac ttt ttc tac cag gac aac ttg gca acg aag aaa tgc 769

Leu Phe Cys His Phe Phe Tyr Gln Asp Asn Leu Ala Thr Lys Lys Ser

245 250 255
gcc aag gcg ggc aag cag ctc tag gcctcgagcc ggctcgcggg ttcaaggagg 823

Ala Lys Ala Gly Lys Gln Leu

260
gcgacacggg ggtgggacgt ttgcatggag atggattgtg gatgtcctta cgccttactc 883

atcaatgtcc tcccatctct cccctctaga ccttctacta gccatctaga agggcagctc 943

agagacggat accgttcccc ctccccttcc ttttctcttt tgctttgcca ttgtttgttt 1003

gtctctatatt tttaaactat tgacgctaac gcgttacgct cgcaaaaaaa aaaaaaaaaa 1063

a 1064

<210> 42

<211> 263

<212> PRT

<213> Isochrysis galbana (GenBank Accession No. AF390174)

<400> 42

Met Ala Leu Ala Asn Asp Ala Gly Glu Arg Ile Trp Ala Ala Val Thr

1 5 10 15

Asp Pro Glu Ile Leu Ile Gly Thr Phe Ser Tyr Leu Leu Leu Lys Pro

20 25 30

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

35 40 45

Thr Ser Met Ile Trp Tyr Asn Val Leu Leu Ala Leu Phe Ser Ala Leu

50 55 60
 Ser Phe Tyr Val Thr Ala Thr Ala Leu Gly Trp Asp Tyr Gly Thr Gly
 65 70 75 80

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

 Gly Ala Pro Trp Asp Val Tyr Leu Gly Ile Arg Leu His Asn Glu Gly
 145 150 155 160
 Val Trp Ile Phe Met Phe Phe Asn Ser Phe Ile His Thr Ile Met Tyr
 165 170 175
 Thr Tyr Tyr Gly Leu Thr Ala Ala Gly Tyr Lys Phe Lys Ala Lys Pro
 180 185 190
 Leu Ile Thr Ala Met Gln Ile Cys Gln Phe Val Gly Gly Phe Leu Leu
 195 200 205

 Val Trp Asp Tyr Ile Asn Val Pro Cys Phe Asn Ser Asp Lys Gly Lys
 210 215 220
 Leu Phe Ser Trp Ala Phe Asn Tyr Ala Tyr Val Gly Ser Val Phe Leu
 225 230 235 240
 Leu Phe Cys His Phe Phe Tyr Gln Asp Asn Leu Ala Thr Lys Lys Ser
 245 250 255
 Ala Lys Ala Gly Lys Gln Leu
 260

 <210> 43
 <211> 777
 <212> DNA
 <213> Euglena gracilis

<220><221> CDS

<222> (1)..(777)

<223> mutant delta-9 elongase "EgD9eS-L35G"

<400> 43

atg gag gtc gtg aac gaa atc gtc tcc att ggc cag gag gtt ctt ccc 48

Met Glu Val Val Asn Glu Ile Val Ser Ile Gly Gln Glu Val Leu Pro

1 5 10 15

aag gtc gac tat gct cag ctc tgg tct gat gcc tcg cac tgc gag gtg 96

Lys Val Asp Tyr Ala Gln Leu Trp Ser Asp Ala Ser His Cys Glu Val

20 25 30

ctg tac ggg tcc atc gcc ttc gtc atc ctg aag ttc acc ctt ggt cct 144

Leu Tyr Gly Ser Ile Ala Phe Val Ile Leu Lys Phe Thr Leu Gly Pro

35 40 45

ctc gga ccc aag ggt cag tct cga atg aag ttt gtg ttc acc aac tac 192

Leu Gly Pro Lys Gly Gln Ser Arg Met Lys Phe Val Phe Thr Asn Tyr

50 55 60

aac ctg ctc atg tcc atc tac tcg ctg ggc tcc ttc ctc tct atg gcc 240

Asn Leu Leu Met Ser Ile Tyr Ser Leu Gly Ser Phe Leu Ser Met Ala

65 70 75 80

tac gcc atg tac acc att ggt gtc atg tcc gac aac tgc gag aag gct 288

Tyr Ala Met Tyr Thr Ile Gly Val Met Ser Asp Asn Cys Glu Lys Ala

85 90 95

ttc gac aac aat gtc ttc cga atc acc act cag ctg ttc tac ctc agc 336

Phe Asp Asn Asn Val Phe Arg Ile Thr Thr Gln Leu Phe Tyr Leu Ser

100 105 110

aag ttc ctc gag tac att gac tcc ttc tat ctg ccc ctc atg ggc aag 384

Lys Phe Leu Glu Tyr Ile Asp Ser Phe Tyr Leu Pro Leu Met Gly Lys

115 120 125

cct ctg acc tgg ttg cag ttc ttt cac cat ctc gga gct cct atg gac 432

Pro Leu Thr Trp Leu Gln Phe Phe His His Leu Gly Ala Pro Met Asp

130 135 140

atg tgg ctg ttc tac aac tac cga aac gaa gcc gtt tgg atc ttt gtg 480
Met Trp Leu Phe Tyr Asn Tyr Arg Asn Glu Ala Val Trp Ile Phe Val
145 150 155 160
ctg ctc aac ggc ttc att cac tgg atc atg tac ggc tac tat tgg acc 528
Leu Leu Asn Gly Phe Ile His Trp Ile Met Tyr Gly Tyr Tyr Trp Thr
165 170 175
cga ctg atc aag ctc aag ttc cct atg ccc aag tcc ctg att act tct 576

Arg Leu Ile Lys Leu Lys Phe Pro Met Pro Lys Ser Leu Ile Thr Ser
180 185 190
atg cag atc att cag ttc aac gtt ggc ttc tac atc gtc tgg aag tac 624
Met Gln Ile Ile Gln Phe Asn Val Gly Phe Tyr Ile Val Trp Lys Tyr
195 200 205
cgg aac att ccc tgc tac cga caa gat gga atg aga atg ttt ggc tgg 672
Arg Asn Ile Pro Cys Tyr Arg Gln Asp Gly Met Arg Met Phe Gly Trp

210 215 220
ttt ttc aac tac ttc tac gtt ggt act gtc ctg tgt ctg ttc ctc aac 720
Phe Phe Asn Tyr Phe Tyr Val Gly Thr Val Leu Cys Leu Phe Leu Asn
225 230 235 240
ttc tac gtg cag acc tac atc gtc cga aag cac aag gga gcc aaa aag 768
Phe Tyr Val Gln Thr Tyr Ile Val Arg Lys His Lys Gly Ala Lys Lys
245 250 255

att cag tga 777
Ile Gln

<210> 44
<211> 258
<212> PRT
<213> Euglena gracilis
<400> 44
Met Glu Val Val Asn Glu Ile Val Ser Ile Gly Gln Glu Val Leu Pro
1 5 10 15
Lys Val Asp Tyr Ala Gln Leu Trp Ser Asp Ala Ser His Cys Glu Val

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

<210> 45
 <211> 1101
 <212> DNA
 <213> Yarrowia lipolytica
 <220><221> CDS
 <222> (1)..(1101)
 <223> cholinephosphate cytidyltransferase; GenBank Accession No.
 XM_502978
 <400> 45
 atg gcc aaa agc aaa cga cgg tcg gag gct gtg gaa gag cac gtg acc 48

 Met Ala Lys Ser Lys Arg Arg Ser Glu Ala Val Glu Glu His Val Thr
 1 5 10 15
 ggc tcg gac gag ggc ttg acc gat act tcg ggt cac gtg agc cct gcc 96
 Gly Ser Asp Glu Gly Leu Thr Asp Thr Ser Gly His Val Ser Pro Ala
 20 25 30
 gcc aag aag cag aag aac tcg gag att cat ttc acc acc cag gct gcc 144
 Ala Lys Lys Gln Lys Asn Ser Glu Ile His Phe Thr Thr Gln Ala Ala
 35 40 45
 cag cag ttg gat cgg gag cgc aag gag gag tat ctg gac tcg ctg atc 192
 Gln Gln Leu Asp Arg Glu Arg Lys Glu Glu Tyr Leu Asp Ser Leu Ile
 50 55 60
 gac aac aag gac tat ctc aag tac cgt cct cga ggc tgg aag ctc aac 240
 Asp Asn Lys Asp Tyr Leu Lys Tyr Arg Pro Arg Gly Trp Lys Leu Asn
 65 70 75 80
 aac ccg cct acc gac cga cct gtg cga atc tac gcc gat gga gtg ttt 288
 Asn Pro Pro Thr Asp Arg Pro Val Arg Ile Tyr Ala Asp Gly Val Phe
 85 90 95
 gat ttg ttc cat ctg gga cac atg cgt cag ctg gag cag tcc aag aag 336
 Asp Leu Phe His Leu Gly His Met Arg Gln Leu Glu Gln Ser Lys Lys
 100 105 110
 gcc ttc ccc aac gca gtg ttg att gtg ggc att ccc agc gac aag gag 384

Ala Phe Pro Asn Ala Val Leu Ile Val Gly Ile Pro Ser Asp Lys Glu
115 120 125
acc cac aag cgg aag gga ttg acc gtg ctg agt gac gtc cag cgg tac 432
Thr His Lys Arg Lys Gly Leu Thr Val Leu Ser Asp Val Gln Arg Tyr
130 135 140
gag acg gtg cga cac tgc aag tgg gtg gac gag gtg gtg gag gat gct 480
Glu Thr Val Arg His Cys Lys Trp Val Asp Glu Val Val Glu Asp Ala
145 150 155 160
ccc tgg tgt gtc acc atg gac ttt ctg gaa aaa cac aaa atc gac tac 528
Pro Trp Cys Val Thr Met Asp Phe Leu Glu Lys His Lys Ile Asp Tyr
165 170 175
gtg gcc cat gac gat ctg ccc tac gct tcc ggc aac gac gat gat atc 576
Val Ala His Asp Asp Leu Pro Tyr Ala Ser Gly Asn Asp Asp Asp Ile
180 185 190
tac aag ccc atc aag gag aag ggc atg ttt ctg gcc acc cag cga acc 624
Tyr Lys Pro Ile Lys Glu Lys Gly Met Phe Leu Ala Thr Gln Arg Thr
195 200 205
gag ggc att tcc acc tcg gac atc atc acc aag att atc cga gac tac 672
Glu Gly Ile Ser Thr Ser Asp Ile Ile Thr Lys Ile Ile Arg Asp Tyr
210 215 220
gac aag tat tta atg cga aac ttt gcc cgg ggt gct aac cga aag gat 720
Asp Lys Tyr Leu Met Arg Asn Phe Ala Arg Gly Ala Asn Arg Lys Asp
225 230 235 240
ctc aac gtc tcg tgg ctc aag aag aac gag ctg gac ttc aag cgt cat 768
Leu Asn Val Ser Trp Leu Lys Lys Asn Glu Leu Asp Phe Lys Arg His
245 250 255
gtg gcc gag ttc cga aac tcg ttc aag cga aag aag gtc ggt aag gat 816
Val Ala Glu Phe Arg Asn Ser Phe Lys Arg Lys Lys Val Gly Lys Asp
260 265 270
ctc tac ggc gag att cgc ggt ctg ctg cag aat gtg ctc att tgg aac 864
Leu Tyr Gly Glu Ile Arg Gly Leu Leu Gln Asn Val Leu Ile Trp Asn

275	280	285	
ggc gac aac tcc ggc act tcc act ccc cag cga aag acg ctg cag acc			912
Gly Asp Asn Ser Gly Thr Ser Thr Pro Gln Arg Lys Thr Leu Gln Thr			
290	295	300	
aac gcc aag aag atg tac atg aac gtg ctc aag act ctg cag gct cct			960
Asn Ala Lys Lys Met Tyr Met Asn Val Leu Lys Thr Leu Gln Ala Pro			
305	310	315	320
gac gct gtt gac gtg gac tcc tcg gag aac gtg tct gag aac gtc act			1008
Asp Ala Val Asp Val Asp Ser Ser Glu Asn Val Ser Glu Asn Val Thr			
325	330	335	
gat gag gag gag gaa gac gac gac gag gtt gat gag gac gaa gaa gcc			1056
Asp Glu Glu Glu Glu Asp Asp Asp Glu Val Asp Glu Asp Glu Glu Ala			
340	345	350	
gac gac gac gac gaa gac gac gaa gac gag gaa gac gac gag tag			1101
Asp Asp Asp Asp Glu Asp Asp Glu Asp Glu Glu Asp Asp Glu			
355	360	365	
<210> 46			
<211> 366			
<212> PRT			
<213> Yarrowia lipolytica			
<400> 46			
Met Ala Lys Ser Lys Arg Arg Ser Glu Ala Val Glu Glu His Val Thr			
1	5	10	15
Gly Ser Asp Glu Gly Leu Thr Asp Thr Ser Gly His Val Ser Pro Ala			
20	25	30	
Ala Lys Lys Gln Lys Asn Ser Glu Ile His Phe Thr Thr Gln Ala Ala			
35	40	45	
Gln Gln Leu Asp Arg Glu Arg Lys Glu Glu Tyr Leu Asp Ser Leu Ile			
50	55	60	
Asp Asn Lys Asp Tyr Leu Lys Tyr Arg Pro Arg Gly Trp Lys Leu Asn			
65	70	75	80

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

325 330 335
Asp Glu Glu Glu Glu Asp Asp Asp Glu Val Asp Glu Asp Glu Glu Ala
340 345 350
Asp Asp Asp Asp Glu Asp Asp Glu Asp Glu Glu Asp Asp Glu
355 360 365
<210> 47
<211> 1185
<212> DNA
<213> Yarrowia lipolytica
<220><221> CDS
<222> (1)..(1185)
<223> diacylglycerol cholinephosphotransferase (YICPT1)
<300><302> HIGH EICOSAPENTAENOIC ACID PRODUCING STRAINS OF
<310> WO 2006/052870
<311> 2005-11-03
<312> 2006-05-18
<313> (1)..(1185)
<300><302> HIGH EICOSAPENTAENOIC ACID PRODUCING STRAINS OF
<310> US 2006-0115881-A1
<311> 2005-11-02
<312> 2006-06-01
<313> (1)..(1185)
<400> 47
atg ggc gta ttc att aaa cag gag cag ctt ccg gct ctc aag aag tac 48
Met Gly Val Phe Ile Lys Gln Glu Gln Leu Pro Ala Leu Lys Lys Tyr
1 5 10 15
aag tac tcc gcc gag gat cac tcg ttc atc tcc aac aac att ctg cgc 96
Lys Tyr Ser Ala Glu Asp His Ser Phe Ile Ser Asn Asn Ile Leu Arg
20 25 30
ccc ttc tgg cga cag ttt gtc aaa atc ttc cct ctg tgg atg gcc ccc 144
Pro Phe Trp Arg Gln Phe Val Lys Ile Phe Pro Leu Trp Met Ala Pro
35 40 45
aac atg gtg act ctg ttg ggc ttc ttc ttt gtc att gtg aac ttc atc 192

Asn Met Val Thr Leu Leu Gly Phe Phe Phe Val Ile Val Asn Phe Ile

50	55	60	
acc atg ctc att gtt gat ccc acc cac gac cgc gag cct ccc aga tgg			240
Thr Met Leu Ile Val Asp Pro Thr His Asp Arg Glu Pro Pro Arg Trp			
65	70	75	80
gtc tac ctc acc tac gct ctg ggt ctg ttc ctt tac cag aca ttt gat			288
Val Tyr Leu Thr Tyr Ala Leu Gly Leu Phe Leu Tyr Gln Thr Phe Asp			
85	90	95	
gcc tgt gac gga tcc cat gcc cga cga act ggc cag agt gga ccc ctt			336
Ala Cys Asp Gly Ser His Ala Arg Arg Thr Gly Gln Ser Gly Pro Leu			
100	105	110	
gga gag ctg ttt gac cac tgt gtc gac gcc atg aat acc tct ctg att			384
Gly Glu Leu Phe Asp His Cys Val Asp Ala Met Asn Thr Ser Leu Ile			
115	120	125	
ctc acg gtg gtg gtg tcc acc acc cat atg gga tat aac atg aag ctg			432
Leu Thr Val Val Val Ser Thr Thr His Met Gly Tyr Asn Met Lys Leu			
130	135	140	
ctg att gtg cag att gcc gct ctc gga aac ttc tac ctg tcg acc tgg			480
Leu Ile Val Gln Ile Ala Ala Leu Gly Asn Phe Tyr Leu Ser Thr Trp			
145	150	155	160
gag acc tac cat acc gga act ctg tac ctt tct ggc ttc tct ggt cct			528
Glu Thr Tyr His Thr Gly Thr Leu Tyr Leu Ser Gly Phe Ser Gly Pro			
165	170	175	
gtt gaa ggt atc ttg att ctg gtg gct ctt ttc gtc ctc acc ttc ttc			576
Val Glu Gly Ile Leu Ile Leu Val Ala Leu Phe Val Leu Thr Phe Phe			
180	185	190	
act ggt ccc aac gtg tac gct ctg acc gtc tac gag gct ctt ccc gaa			624
Thr Gly Pro Asn Val Tyr Ala Leu Thr Val Tyr Glu Ala Leu Pro Glu			
195	200	205	

tcc atc act tcg ctg ctg cct gcc agc ttc ctg gac gtc acc atc acc	672
Ser Ile Thr Ser Leu Leu Pro Ala Ser Phe Leu Asp Val Thr Ile Thr	
210 215 220	
cag atc tac att gga ttc gga gtg ctg ggc atg gtg ttc aac atc tac	720
Gln Ile Tyr Ile Gly Phe Gly Val Leu Gly Met Val Phe Asn Ile Tyr	
225 230 235 240	
ggc gcc tgc gga aac gtg atc aag tac tac aac aac aag ggc aag agc	768
Gly Ala Cys Gly Asn Val Ile Lys Tyr Tyr Asn Asn Lys Gly Lys Ser	
245 250 255	
gct ctc ccc gcc att ctc gga atc gcc ccc ttt ggc atc ttc tac gtc	816
Ala Leu Pro Ala Ile Leu Gly Ile Ala Pro Phe Gly Ile Phe Tyr Val	
260 265 270	
ggc gtc ttt gcc tgg gcc cat gtt gct cct ctg ctt ctc tcc aag tac	864
Gly Val Phe Ala Trp Ala His Val Ala Pro Leu Leu Leu Ser Lys Tyr	
275 280 285	
gcc atc gtc tat ctg ttt gcc att ggg gct gcc ttt gcc atg caa gtc	912
Ala Ile Val Tyr Leu Phe Ala Ile Gly Ala Ala Phe Ala Met Gln Val	
290 295 300	
ggc cag atg att ctt gcc cat ctc gtg ctt gct ccc ttc ccc cac tgg	960
Gly Gln Met Ile Leu Ala His Leu Val Leu Ala Pro Phe Pro His Trp	
305 310 315 320	
aac gtg ctg ctc ttc ttc ccc ttt gtg gga ctg gca gtg cac tac att	1008
Asn Val Leu Leu Phe Phe Pro Phe Val Gly Leu Ala Val His Tyr Ile	
325 330 335	
gca ccc gtg ttt ggc tgg gac gcc gat atc gtg tgc gtt aac act ctc	1056
Ala Pro Val Phe Gly Trp Asp Ala Asp Ile Val Ser Val Asn Thr Leu	
340 345 350	
ttc acc tgt ttt ggc gcc acc ctc tcc att tac gcc ttc ttt gtg ctt	1104
Phe Thr Cys Phe Gly Ala Thr Leu Ser Ile Tyr Ala Phe Phe Val Leu	
355 360 365	
gag atc atc gac gag atc acc aac tac ctc gat atc tgg tgt ctg cga	1152

Glu Ile Ile Asp Glu Ile Thr Asn Tyr Leu Asp Ile Trp Cys Leu Arg
 370 375 380

atc aag tac cct cag gag aag aag act gag taa 1185

Ile Lys Tyr Pro Gln Glu Lys Lys Thr Glu

385 390

<210> 48

<211> 394

<212> PRT

<213> Yarrowia lipolytica

<400> 48

Met Gly Val Phe Ile Lys Gln Glu Gln Leu Pro Ala Leu Lys Lys Tyr

1 5 10 15

Lys Tyr Ser Ala Glu Asp His Ser Phe Ile Ser Asn Asn Ile Leu Arg

20 25 30

Pro Phe Trp Arg Gln Phe Val Lys Ile Phe Pro Leu Trp Met Ala Pro

35 40 45

Asn Met Val Thr Leu Leu Gly Phe Phe Phe Val Ile Val Asn Phe Ile

50 55 60

Thr Met Leu Ile Val Asp Pro Thr His Asp Arg Glu Pro Pro Arg Trp

65 70 75 80

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

85 90 95

Ala Cys Asp Gly Ser His Ala Arg Arg Thr Gly Gln Ser Gly Pro Leu

100 105 110

Gly Glu Leu Phe Asp His Cys Val Asp Ala Met Asn Thr Ser Leu Ile

115 120 125

Leu Thr Val Val Val Ser Thr Thr His Met Gly Tyr Asn Met Lys Leu

130 135 140

Leu Ile Val Gln Ile Ala Ala Leu Gly Asn Phe Tyr Leu Ser Thr Trp

145 150 155 160

Glu Thr Tyr His Thr Gly Thr Leu Tyr Leu Ser Gly Phe Ser Gly Pro

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

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

 Gly Gln Met Ile Leu Ala His Leu Val Leu Ala Pro Phe Pro His Trp
 305 310 315 320
 Asn Val Leu Leu Phe Phe Pro Phe Val Gly Leu Ala Val His Tyr Ile
 325 330 335
 Ala Pro Val Phe Gly Trp Asp Ala Asp Ile Val Ser Val Asn Thr Leu
 340 345 350
 Phe Thr Cys Phe Gly Ala Thr Leu Ser Ile Tyr Ala Phe Phe Val Leu
 355 360 365

 Glu Ile Ile Asp Glu Ile Thr Asn Tyr Leu Asp Ile Trp Cys Leu Arg
 370 375 380
 Ile Lys Tyr Pro Gln Glu Lys Lys Thr Glu
 385 390
 <210> 49
 <211> 1518

<212> DNA

<213> Rhizobium leguminosarum bv. viciae 3841

<220><221> CDS

<222> (1)..(1518)

<223> synthetic malonyl-CoA synthetase (codon-optimized for Yarrowia
lipolytica)

<400> 49

atg gtc tcc aac cac ctg ttc gac gcc atg cga gct gcc gct ccc gga 48

Met Val Ser Asn His Leu Phe Asp Ala Met Arg Ala Ala Ala Pro Gly

1 5 10 15

gac gca cct ttc att cga atc gac aac gct cgg acc tgg act tac gat 96

Asp Ala Pro Phe Ile Arg Ile Asp Asn Ala Arg Thr Trp Thr Tyr Asp

20 25 30

gac gcc att gct ctt tcc ggt cga atc gct gga gct atg gac gca ctc 144

Asp Ala Ile Ala Leu Ser Gly Arg Ile Ala Gly Ala Met Asp Ala Leu

35 40 45

ggc att cga ccc gga gac aga gtt gcc gtg cag gtc gag aag tct gcc 192

Gly Ile Arg Pro Gly Asp Arg Val Ala Val Gln Val Glu Lys Ser Ala

50 55 60

gag gcg ttg att ctc tac ctg gcc tgt ctt cga acc gga gct gtc tac 240

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

65 70 75 80

ctg cct ctc aac act gcc tac acc ctg gcc gag ctc gac tac ttc atc 288

Leu Pro Leu Asn Thr Ala Tyr Thr Leu Ala Glu Leu Asp Tyr Phe Ile

85 90 95

ggc gat gcc gaa ccg cgt ctg gtg gtc gtt gct ccc gca gct cga ggt 336

Gly Asp Ala Glu Pro Arg Leu Val Val Val Ala Pro Ala Ala Arg Gly

100 105 110

ggc gtg gag aca att gcc aag cga cac ggt gct atc gtc gaa acc ctc 384

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

115 120 125

gac gcc gat gga cga ggc tcc ttg ctg gac ctt gct aga gat gag cct 432
 Asp Ala Asp Gly Arg Gly Ser Leu Leu Asp Leu Ala Arg Asp Glu Pro
 130 135 140

gcc gac ttt gtc gat gct tcg cga tct gcc gac gat ctg gct gct att 480
 Ala Asp Phe Val Asp Ala Ser Arg Ser Ala Asp Asp Leu Ala Ala Ile

145 150 155 160
 ctc tac act tcc ggt aca acc gga cga tcg aag ggt gcc atg ctt act 528
 Leu Tyr Thr Ser Gly Thr Thr Gly Arg Ser Lys Gly Ala Met Leu Thr
 165 170 175

cat ggc aat ctg ctc tcc aac gct ctc acc ttg cga gac tat tgg aga 576
 His Gly Asn Leu Leu Ser Asn Ala Leu Thr Leu Arg Asp Tyr Trp Arg
 180 185 190

gtt acc gca gac gat cga ctc atc cat gcc ttg cca atc ttt cac act 624
 Val Thr Ala Asp Asp Arg Leu Ile His Ala Leu Pro Ile Phe His Thr
 195 200 205

cat ggt ctg ttc gtt gct acg aac gtc aca ctg ctt gca gga gcc tcg 672
 His Gly Leu Phe Val Ala Thr Asn Val Thr Leu Leu Ala Gly Ala Ser
 210 215 220

atg ttt ctg ctc tcc aag ttc gat gcc gac gag gtc gtt tct ctc atg 720

Met Phe Leu Leu Ser Lys Phe Asp Ala Asp Glu Val Val Ser Leu Met
 225 230 235 240

cca cag gcc acc atg ctt atg ggc gtg ccc aca ttc tac gtt cga ttg 768
 Pro Gln Ala Thr Met Leu Met Gly Val Pro Thr Phe Tyr Val Arg Leu
 245 250 255

ctg cag agt cct cga ctc gag aag ggt gct gtg gcc agc atc aga ctg 816
 Leu Gln Ser Pro Arg Leu Glu Lys Gly Ala Val Ala Ser Ile Arg Leu

260 265 270
 ttc att tct gga tca gct ccc ttg ctt gcc gaa acc cac gcc gag ttt 864
 Phe Ile Ser Gly Ser Ala Pro Leu Leu Ala Glu Thr His Ala Glu Phe
 275 280 285

cat gct cgt act ggt cac gcc att ctc gag cga tac ggc atg acg gaa 912

His Ala Arg Thr Gly His Ala Ile Leu Glu Arg Tyr Gly Met Thr Glu	
290	295
300	
acc aac atg aat act tcc aac ccc tac gag ggc aag cgt att gcc gga	960
Thr Asn Met Asn Thr Ser Asn Pro Tyr Glu Gly Lys Arg Ile Ala Gly	
305	310
315	320
acc gtt ggt ttt cct ctg ccc gac gtc act gtg cga gtc acc gat ccc	1008
Thr Val Gly Phe Pro Leu Pro Asp Val Thr Val Arg Val Thr Asp Pro	
325	330
335	
gcc acc ggt ctc gtt ctt cca cct gaa gag act ggc atg atc gag atc	1056
Ala Thr Gly Leu Val Leu Pro Pro Glu Glu Thr Gly Met Ile Glu Ile	
340	345
350	
aag gga ccc aac gtc ttc aag ggc tat tgg cga atg ccc gaa aag acc	1104
Lys Gly Pro Asn Val Phe Lys Gly Tyr Trp Arg Met Pro Glu Lys Thr	
355	360
365	
gct gcc gag ttt acc gca gac ggt ttc ttt atc tct gga gat ctc ggc	1152
Ala Ala Glu Phe Thr Ala Asp Gly Phe Phe Ile Ser Gly Asp Leu Gly	
370	375
380	
aag atc gac cga gaa ggt tac gtt cac att gtg gga cga ggc aag gac	1200
Lys Ile Asp Arg Glu Gly Tyr Val His Ile Val Gly Arg Gly Lys Asp	
385	390
395	400
ctg gtc att tcc ggt ggc tac aac atc tat ccc aaa gag gtc gaa ggc	1248
Leu Val Ile Ser Gly Gly Tyr Asn Ile Tyr Pro Lys Glu Val Glu Gly	
405	410
415	
gag atc gac cag atc gag ggt gtg gtc gag tct gct gtc att ggt gtt	1296
Glu Ile Asp Gln Ile Glu Gly Val Val Glu Ser Ala Val Ile Gly Val	
420	425
430	
cct cat ccc gat ttc gga gaa ggt gtc acc gct gtt gtc gtg tgc aaa	1344
Pro His Pro Asp Phe Gly Glu Gly Val Thr Ala Val Val Val Cys Lys	
435	440
445	

cct ggt gcc gtt ctc gac gaa aag acc atc gtg tct gct ctg cag gac 1392

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

450 455 460

cgt ctt gcc cga tac aag caa ccc aag cgg att atc ttt gcc gac gat 1440

Arg Leu Ala Arg Tyr Lys Gln Pro Lys Arg Ile Ile Phe Ala Asp Asp

465 470 475 480

ctg cct cga aac act atg gga aag gtt cag aag aac att ctt cga cag 1488

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

485 490 495

caa tac gcc gat ctc tac acc aga cga taa 1518

Gln Tyr Ala Asp Leu Tyr Thr Arg Arg

500 505

<210> 50

<211> 505

<212> PRT

<213> Rhizobium leguminosarum bv. viciae 3841

<400> 50

Met Val Ser Asn His Leu Phe Asp Ala Met Arg Ala Ala Ala Pro Gly

1 5 10 15

Asp Ala Pro Phe Ile Arg Ile Asp Asn Ala Arg Thr Trp Thr Tyr Asp

20 25 30

Asp Ala Ile Ala Leu Ser Gly Arg Ile Ala Gly Ala Met Asp Ala Leu

35 40 45

Gly Ile Arg Pro Gly Asp Arg Val Ala Val Gln Val Glu Lys Ser Ala

50 55 60

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

65 70 75 80

Leu Pro Leu Asn Thr Ala Tyr Thr Leu Ala Glu Leu Asp Tyr Phe Ile

85 90 95

Gly Asp Ala Glu Pro Arg Leu Val Val Val Ala Pro Ala Ala Arg Gly

100 105 110

Gly Val Glu Thr Ile Ala Lys Arg His Gly Ala Ile Val Glu Thr Leu
115 120 125

Asp Ala Asp Gly Arg Gly Ser Leu Leu Asp Leu Ala Arg Asp Glu Pro
130 135 140

Ala Asp Phe Val Asp Ala Ser Arg Ser Ala Asp Asp Leu Ala Ala Ile
145 150 155 160

Leu Tyr Thr Ser Gly Thr Thr Gly Arg Ser Lys Gly Ala Met Leu Thr
165 170 175

His Gly Asn Leu Leu Ser Asn Ala Leu Thr Leu Arg Asp Tyr Trp Arg
180 185 190

Val Thr Ala Asp Asp Arg Leu Ile His Ala Leu Pro Ile Phe His Thr
195 200 205

His Gly Leu Phe Val Ala Thr Asn Val Thr Leu Leu Ala Gly Ala Ser
210 215 220

Met Phe Leu Leu Ser Lys Phe Asp Ala Asp Glu Val Val Ser Leu Met
225 230 235 240

Pro Gln Ala Thr Met Leu Met Gly Val Pro Thr Phe Tyr Val Arg Leu
245 250 255

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

Phe Ile Ser Gly Ser Ala Pro Leu Leu Ala Glu Thr His Ala Glu Phe
275 280 285

His Ala Arg Thr Gly His Ala Ile Leu Glu Arg Tyr Gly Met Thr Glu
290 295 300

Thr Asn Met Asn Thr Ser Asn Pro Tyr Glu Gly Lys Arg Ile Ala Gly
305 310 315 320

Thr Val Gly Phe Pro Leu Pro Asp Val Thr Val Arg Val Thr Asp Pro
325 330 335

Ala Thr Gly Leu Val Leu Pro Pro Glu Glu Thr Gly Met Ile Glu Ile
340 345 350

Lys Gly Pro Asn Val Phe Lys Gly Tyr Trp Arg Met Pro Glu Lys Thr

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

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

Arg Leu Ala Arg Tyr Lys Gln Pro Lys Arg Ile Ile Phe Ala Asp Asp
465 470 475 480
Leu Pro Arg Asn Thr Met Gly Lys Val Gln Lys Asn Ile Leu Arg Gln
485 490 495
Gln Tyr Ala Asp Leu Tyr Thr Arg Arg
500 505

<210> 51

<211> 1024

<212> PRT

<213> Yarrowia lipolytica

<220><221> MISC_FEATURE

<222> (1)..(1024)

<223> YIPex1p; GenBank Accession No. CAG82178

<400> 51

Met Thr Ser Lys Ser Asp Tyr Ser Gly Lys Asp Lys Ile Glu Leu Asp
1 5 10 15
Pro Val Phe Ala Lys Ser Ile Asp Leu Leu Pro Asn Thr Gln Val Val
20 25 30
Ile Asp Ile Gln Leu Asn Pro Lys Ile Ala His Thr Ile His Leu Glu

35 40 45
 Pro Val Thr Val Ala Asp Trp Glu Ile Val Glu Leu His Ala Ala Tyr
 50 55 60

 Leu Glu Ser Arg Met Ile Asn Gln Val Arg Ala Val Ser Pro Asn Gln
 65 70 75 80
 Pro Val Thr Val Tyr Pro Ser Ser Thr Thr Ser Ala Thr Leu Lys Val
 85 90 95
 Ile Arg Ile Glu Pro Asp Leu Gly Ala Ala Gly Phe Ala Lys Leu Ser
 100 105 110
 Pro Asp Ser Glu Val Val Val Ala Pro Lys Gln Arg Lys Lys Glu Glu
 115 120 125

 Lys Gln Val Lys Lys Arg Ser Gly Ser Ala Arg Ser Thr Gly Ser Gln
 130 135 140
 Lys Arg Lys Gly Gly Arg Gly Pro His Ala Leu Arg Arg Ala Ile Ser
 145 150 155 160
 Glu Asp Phe Asp Gly His Leu Arg Leu Glu Val Ser Leu Asp Val Ser
 165 170 175
 Gln Leu Pro Pro Glu Phe His Gln Leu Lys Asn Val Ser Ile Lys Val
 180 185 190

 Ile Thr Pro Pro Asn Leu Ala Ser Pro Gln Gln Ala Ala Ser Ile Ala
 195 200 205
 Val Glu Glu Lys Ser Glu Glu Ser Leu Ser Gln Asn Lys Pro Pro Ser
 210 215 220
 Ser Glu Pro Lys Val Glu Val Pro Pro Asp Ile Ile Asn Pro Ala Ser
 225 230 235 240
 Glu Ile Val Ala Thr Leu Val Asn Asp Thr Thr Ser Pro Thr Gly His
 245 250 255

 Ala Lys Leu Ser Tyr Ala Leu Ala Asp Ala Leu Gly Ile Pro Ser Ser
 260 265 270
 Val Gly His Val Ile Arg Phe Glu Ser Ala Ser Lys Pro Leu Ser Gln
 275 280 285

Lys Pro Gly Ala Leu Val Ile His Arg Phe Ile Thr Lys Thr Val Gly
 290 295 300
 Ala Ala Glu Gln Lys Ser Leu Arg Leu Lys Gly Glu Lys Asn Ala Asp
 305 310 315 320

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

 Phe Val Ser Gly Ser Glu Ile Leu Arg Ser Glu Ser Ser Leu Ser Pro
 385 390 395 400
 Ser Asn Ile Glu Ser Glu Asp Lys Arg Val Val Gly Leu Asp Asn Met
 405 410 415
 Leu Asn Lys Ile Asn Glu Val Leu Ser Arg Asp Ser Ile Gly Cys Leu
 420 425 430
 Val Tyr Gly Ser Arg Gly Ser Gly Lys Ser Ala Val Leu Asn His Ile
 435 440 445

 Lys Lys Glu Cys Lys Val Ser His Thr His Thr Val Ser Ile Ala Cys
 450 455 460
 Gly Leu Ile Ala Gln Asp Arg Val Gln Ala Val Arg Glu Ile Leu Thr
 465 470 475 480
 Lys Ala Phe Leu Glu Ala Ser Trp Phe Ser Pro Ser Val Leu Phe Leu
 485 490 495
 Asp Asp Ile Asp Ala Leu Met Pro Ala Glu Val Glu His Ala Asp Ser
 500 505 510

 Ser Arg Thr Arg Gln Leu Thr Gln Leu Phe Leu Glu Leu Ala Leu Pro
 515 520 525
 Ile Met Lys Ser Arg His Val Ser Val Val Ala Ser Ala Gln Ala Lys

530 535 540
 Glu Ser Leu His Met Asn Leu Val Thr Gly His Val Phe Glu Glu Leu
 545 550 555 560
 Phe His Leu Lys Ser Pro Asp Lys Glu Ala Arg Leu Ala Ile Leu Ser
 565 570 575

 Glu Ala Val Lys Leu Met Asp Gln Asn Val Ser Phe Ser Gln Asn Asp
 580 585 590
 Val Leu Glu Ile Ala Ser Gln Val Asp Gly Tyr Leu Pro Gly Asp Leu
 595 600 605
 Trp Thr Leu Ser Glu Arg Ala Gln His Glu Met Ala Leu Arg Gln Ile
 610 615 620
 Glu Ile Gly Leu Glu Asn Pro Ser Ile Gln Leu Ala Asp Phe Met Lys
 625 630 635 640

 Ala Leu Glu Asp Phe Val Pro Ser Ser Leu Arg Gly Val Lys Leu Gln
 645 650 655
 Lys Ser Asn Val Lys Trp Asn Asp Ile Gly Gly Leu Lys Glu Thr Lys
 660 665 670
 Ala Val Leu Leu Glu Thr Leu Glu Trp Pro Thr Lys Tyr Ala Pro Ile
 675 680 685
 Phe Ala Ser Cys Pro Leu Arg Leu Arg Ser Gly Leu Leu Leu Tyr Gly
 690 695 700

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

 Ala Pro Lys Arg Gly His Asp Ser Thr Gly Val Thr Asp Arg Val Val
 770 775 780

Asn Gln Met Leu Thr Gln Met Asp Gly Ala Glu Gly Leu Asp Gly Val

785 790 795 800

Tyr Val Leu Ala Ala Thr Ser Arg Pro Asp Leu Ile Asp Pro Ala Leu

805 810 815

Leu Arg Pro Gly Arg Leu Asp Lys Met Leu Ile Cys Asp Leu Pro Ser

820 825 830

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

835 840 845

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

850 855 860

Gly Phe Ser Gly Ala Asp Leu Gln Ala Val Met Phe Asn Ala Tyr Leu

865 870 875 880

Glu Ala Ile His Glu Val Val Asp Val Ala Asp Asp Thr Ala Ala Asp

885 890 895

Thr Pro Ala Leu Glu Asp Lys Arg Leu Glu Phe Phe Gln Thr Thr Leu

900 905 910

Gly Asp Ala Lys Lys Asp Pro Ala Ala Val Gln Asn Glu Val Met Asn

915 920 925

Ala Arg Ala Ala Val Ala Glu Lys Ala Arg Val Thr Ala Lys Leu Glu

930 935 940

Ala Leu Phe Lys Gly Met Ser Val Gly Val Asp Asn Asp Asp Asp Lys

945 950 955 960

Pro Arg Lys Lys Ala Val Val Val Ile Lys Pro Gln His Met Asn Lys

965 970 975

Ser Leu Asp Glu Thr Ser Pro Ser Ile Ser Lys Lys Glu Leu Leu Lys

980 985 990

Leu Lys Gly Ile Tyr Ser Gln Phe Val Ser Gly Arg Ser Gly Asp Met

995 1000 1005

Pro Pro Gly Thr Ala Ser Thr Asp Val Gly Gly Arg Ala Thr Leu

1010 1015 1020

Ala

<210> 52

<211> 381

<212> PRT

<213> *Yarrowia lipolytica*

<220><221> MISC_FEATURE

<222> (1)..(381)

<223> YlPex2p; GenBank Accession No. CAG77647

<400> 52

Met Ser Ser Val Leu Arg Leu Phe Lys Ile Gly Ala Pro Val Pro Asn

1 5 10 15

Val Arg Val His Gln Leu Asp Ala Ser Leu Leu Asp Ala Glu Leu Val

20 25 30

Asp Leu Leu Lys Asn Gln Leu Phe Lys Gly Phe Thr Asn Phe His Pro

35 40 45

Glu Phe Arg Asp Lys Tyr Glu Ser Glu Leu Val Leu Ala Leu Lys Leu

50 55 60

Ile Leu Phe Lys Leu Thr Val Trp Asp His Ala Ile Thr Tyr Gly Gly

65 70 75 80

Lys Leu Gln Asn Leu Lys Phe Ile Asp Ser Arg His Ser Ser Lys Leu

85 90 95

Gln Ile Gln Pro Ser Val Ile Gln Lys Leu Gly Tyr Gly Ile Leu Val

100 105 110

Val Gly Gly Gly Tyr Leu Trp Ser Lys Ile Glu Gly Tyr Leu Leu Ala

115 120 125

Arg Ser Glu Asp Asp Val Ala Thr Asp Gly Thr Ser Val Arg Gly Ala

130 135 140

Ser Ala Ala Arg Gly Ala Leu Lys Val Ala Asn Phe Ala Ser Leu Leu

145 150 155 160

Tyr Ser Ala Ala Thr Leu Gly Asn Phe Val Ala Phe Leu Tyr Thr Gly

165 170 175

Arg Tyr Ala Thr Val Ile Met Arg Leu Leu Arg Ile Arg Leu Val Pro

180 185 190
 Ser Gln Arg Thr Ser Ser Arg Gln Val Ser Tyr Glu Phe Gln Asn Arg
 195 200 205
 Gln Leu Val Trp Asn Ala Phe Thr Glu Phe Leu Ile Phe Ile Leu Pro
 210 215 220
 Leu Leu Gln Leu Pro Lys Leu Lys Arg Arg Ile Glu Arg Lys Leu Gln

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

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

 355 360 365
 Asp Asp Asp Glu Asp Ser Asn Phe Gln Leu Met Lys Asp
 370 375 380
 <210> 53
 <211> 431
 <212> PRT
 <213> Yarrowia lipolytica
 <220><221> MISC_FEATURE
 <222> (1)..(431)
 <223> YIPex3p; GenBank Accession No. CAG78565

<400> 53

Met Asp Phe Phe Arg Arg His Gln Lys Lys Val Leu Ala Leu Val Gly

1 5 10 15

Val Ala Leu Ser Ser Tyr Leu Phe Ile Asp Tyr Val Lys Lys Lys Phe

20 25 30

Phe Glu Ile Gln Gly Arg Leu Ser Ser Glu Arg Thr Ala Lys Gln Asn

35 40 45

Leu Arg Arg Arg Phe Glu Gln Asn Gln Gln Asp Ala Asp Phe Thr Ile

50 55 60

Met Ala Leu Leu Ser Ser Leu Thr Thr Pro Val Met Glu Arg Tyr Pro

65 70 75 80

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

85 90 95

Arg Val Leu Ala Leu Glu Ser Ser Thr Ser Ser Ser Ala Thr Ala Gln

100 105 110

Thr Val Pro Thr Met Thr Ser Gly Ala Thr Glu Glu Gly Glu Lys Ser

115 120 125

Lys Thr Gln Leu Trp Gln Asp Leu Lys Arg Thr Thr Ile Ser Arg Ala

130 135 140

Phe Ser Leu Val Tyr Ala Asp Ala Leu Leu Ile Phe Phe Thr Arg Leu

145 150 155 160

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

165 170 175

Leu Ala Gln Gln Gly Arg Glu Gly Asn Ala Glu Gly Arg Val Ala Pro

180 185 190

Ser Phe Gly Asp Leu Ala Asp Met Gly Tyr Phe Gly Asp Leu Ser Gly

195 200 205

Ser Ser Ser Phe Gly Glu Thr Ile Val Asp Pro Asp Leu Asp Glu Gln

210 215 220

Tyr Leu Thr Phe Ser Trp Trp Leu Leu Asn Glu Gly Trp Val Ser Leu

225 230 235 240

Ser Glu Arg Val Glu Glu Ala Val Arg Arg Val Trp Asp Pro Val Ser
245 250 255
Pro Lys Ala Glu Leu Gly Phe Asp Glu Leu Ser Glu Leu Ile Gly Arg
260 265 270
Thr Gln Met Leu Ile Asp Arg Pro Leu Asn Pro Ser Ser Pro Leu Asn
275 280 285
Phe Leu Ser Gln Leu Leu Pro Pro Arg Glu Gln Glu Glu Tyr Val Leu
290 295 300
Ala Gln Asn Pro Ser Asp Thr Ala Ala Pro Ile Val Gly Pro Thr Leu
305 310 315 320
Arg Arg Leu Leu Asp Glu Thr Ala Asp Phe Ile Glu Ser Pro Asn Ala
325 330 335
Ala Glu Val Ile Glu Arg Leu Val His Ser Gly Leu Ser Val Phe Met
340 345 350
Asp Lys Leu Ala Val Thr Phe Gly Ala Thr Pro Ala Asp Ser Gly Ser
355 360 365
Pro Tyr Pro Val Val Leu Pro Thr Ala Lys Val Lys Leu Pro Ser Ile
370 375 380
Leu Ala Asn Met Ala Arg Gln Ala Gly Gly Met Ala Gln Gly Ser Pro
385 390 395 400
Gly Val Glu Asn Glu Tyr Ile Asp Val Met Asn Gln Val Gln Glu Leu

405 410 415
Thr Ser Phe Ser Ala Val Val Tyr Ser Ser Phe Asp Trp Ala Leu
420 425 430

<210> 54

<211> 395

<212> PRT

<213> Yarrowia lipolytica

<220><221> MISC_FEATURE

<222> (1)..(395)

<223> YIPex3Bp; GenBank Accession No. CAG83356

<400> 54

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

245 250 255
Glu Leu Glu Gln Leu Arg Leu Glu Lys Glu Lys Gln Ser Leu Ser Pro

260 265 270
Asn Tyr Thr Tyr Gly Ser Pro Leu Lys Asp Leu Val Phe Glu Ser Ala

275 280 285
Gln His Ile Gln Ser Pro Gln Gly Met Ser Ser Phe Arg Ala Ile Ile

290 295 300
Asp Gln Ser Phe Lys Val Phe Leu Glu Lys Val Asn Glu Ser Gln Tyr

305 310 315 320
Val Asn Pro Pro Ser Thr Gly Gly Lys Arg Ile Ala Val Gly Ala Leu

325 330 335
Gln Pro Pro Ile Ile Ser Gly Gly Pro Lys Lys Val Lys Leu Ala Ser

340 345 350
Leu Leu Ser Val Ala Thr Arg Gln Ser Ser Val Ile Ser His Ala Gln

355 360 365
Pro Asn Pro Tyr Val Asp Ala Ile Asn Ser Val Ala Glu Tyr Asn Gly

370 375 380
Leu Cys Ala Val Ile Tyr Ser Ser Phe Glu Gln

385 390 395

<210> 55

<211> 153

<212> PRT

<213> *Yarrowia lipolytica*

<220><221> MISC_FEATURE

<222> (1)..(153)

<223> YIPex4p; GenBank Accession No. CAG79130

<400> 55

Met Ala Ser Gln Lys Arg Leu Ile Lys Glu Leu Ala Ala Tyr Lys Lys
1 5 10 15

Asp Pro Asn Pro Cys Leu Ala Ser Leu Thr Ala Asp Gly Asp Ser Leu
20 25 30

Tyr Lys Trp Thr Ala Val Met Arg Gly Thr Glu Gly Thr Ala Tyr Glu

35 40 45
Asn Gly Leu Trp Gln Val Glu Ile Asn Ile Pro Glu Asn Tyr Pro Leu

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

115 120 125
Leu Val Arg Cys Gly Asp Glu Ser Ala Met Glu Gly Leu Val Arg Tyr
130 135 140
Tyr Val Asn Lys Tyr Ala Ser Gly Asn

145 150

<210> 56

<211> 598

<212> PRT

<213> *Yarrowia lipolytica*

<220><221> MISC_FEATURE

<222> (1)..(598)

<223> YIPex5p; GenBank Accession No. CAG78803

<400> 56

Met Ser Phe Met Arg Gly Gly Ser Glu Cys Ser Thr Gly Arg Asn Pro

1 5 10 15
Leu Ser Gln Phe Thr Lys His Thr Ala Glu Asp Arg Ser Leu Gln His
20 25 30
Asp Arg Val Ala Gly Pro Ser Gly Gly Arg Val Gly Gly Met Arg Ser
35 40 45
Asn Thr Gly Glu Met Ser Gln Gln Asp Arg Glu Met Met Ala Arg Phe
50 55 60

Gly Ala Ala Gly Pro Glu Gln Ser Ser Phe Asn Tyr Glu Gln Met Arg

65 70 75 80

His Glu Leu His Asn Met Gly Ala Gln Gly Gly Gln Ile Pro Gln Val

85 90 95

Pro Ser Gln Gln Gly Ala Ala Asn Gly Gly Gln Trp Ala Arg Asp Phe

100 105 110

Gly Gly Gln Gln Thr Ala Pro Gly Ala Ala Pro Gln Asp Ala Lys Asn

115 120 125

Trp Asn Ala Glu Phe Gln Arg Gly Gly Ser Pro Ala Glu Ala Met Gln

130 135 140

Gln Gln Gly Pro Gly Pro Met Gln Gly Gly Met Gly Met Gly Gly Met

145 150 155 160

Pro Met Tyr Gly Met Ala Arg Pro Met Tyr Ser Gly Met Ser Ala Asn

165 170 175

Met Ala Pro Gln Phe Gln Pro Gln Gln Ala Asn Ala Arg Val Val Glu

180 185 190

Leu Asp Glu Gln Asn Trp Glu Glu Gln Phe Lys Gln Met Asp Ser Ala

195 200 205

Val Gly Lys Gly Lys Glu Val Glu Glu Gln Thr Ala Glu Thr Ala Thr

210 215 220

Ala Thr Glu Thr Val Thr Glu Thr Glu Thr Thr Thr Glu Asp Lys Pro

225 230 235 240

Met Asp Ile Lys Asn Met Asp Phe Glu Asn Ile Trp Lys Asn Leu Gln

245 250 255

Val Asn Val Leu Asp Asn Met Asp Glu Trp Leu Glu Glu Thr Asn Ser

260 265 270

Pro Ala Trp Glu Arg Asp Phe His Glu Tyr Thr His Asn Arg Pro Glu

275 280 285

Phe Ala Asp Tyr Gln Phe Glu Glu Asn Asn Gln Phe Met Glu His Pro

290 295 300

Asp Pro Phe Lys Ile Gly Val Glu Leu Met Glu Thr Gly Gly Arg Leu

305 310 315 320
 Ser Glu Ala Ala Leu Ala Phe Glu Ala Ala Val Gln Lys Asn Thr Glu

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

 385 390 395 400
 Thr Lys Tyr Pro Glu Val Val Asp Gln Ala Arg Asn Gln Glu Pro Arg
 405 410 415
 Leu Gly Asn Glu Asp Lys Phe Gln Leu His Ser Arg Val Thr Glu Leu
 420 425 430
 Phe Ile Arg Ala Ala Gln Leu Ser Pro Asp Gly Ala Asn Ile Asp Ala
 435 440 445
 Asp Val Gln Val Gly Leu Gly Val Leu Phe Tyr Gly Asn Glu Glu Tyr

 450 455 460
 Asp Lys Ala Ile Asp Cys Phe Asn Ala Ala Ile Ala Val Arg Pro Asp
 465 470 475 480
 Asp Ala Leu Leu Trp Asn Arg Leu Gly Ala Thr Leu Ala Asn Ser His
 485 490 495
 Arg Ser Glu Glu Ala Ile Asp Ala Tyr Tyr Lys Ala Leu Glu Leu Arg
 500 505 510
 Pro Ser Phe Val Arg Ala Arg Tyr Asn Leu Gly Val Ser Cys Ile Asn

 515 520 525
 Ile Gly Cys Tyr Lys Glu Ala Ala Gln Tyr Leu Leu Gly Ala Leu Ser
 530 535 540
 Met His Lys Val Glu Gly Val Gln Asp Asp Val Leu Ala Asn Gln Ser
 545 550 555 560

Thr Asn Leu Tyr Asp Thr Leu Lys Arg Val Phe Leu Gly Met Asp Arg
565 570 575
Arg Asp Leu Val Ala Lys Val Gly Asn Gly Met Asp Val Asn Gln Phe
580 585 590
Arg Asn Glu Phe Glu Phe
595
<210> 57
<211> 1024
<212> PRT
<213> Yarrowia lipolytica
<220><221> MISC_FEATURE
<222> (1)..(1024)
<223> YIPex6p; GenBank Accession No. CAG82306
<400> 57
Met Pro Ser Ile Ser His Lys Pro Ile Thr Ala Lys Leu Val Ala Ala
1 5 10 15
Pro Asp Ala Thr Lys Leu Glu Leu Ser Ser Tyr Leu Tyr Gln Gln Leu
20 25 30
Phe Ser Asp Lys Pro Ala Glu Pro Tyr Val Ala Phe Glu Ala Pro Gly
35 40 45
Ile Lys Trp Ala Leu Tyr Pro Ala Ser Glu Asp Arg Ser Leu Pro Gln
50 55 60
Tyr Thr Cys Lys Ala Asp Ile Arg His Val Ala Gly Ser Leu Lys Lys
65 70 75 80
Phe Met Pro Val Val Leu Lys Arg Val Asn Pro Val Thr Ile Glu His
85 90 95
Ala Ile Val Thr Val Pro Ala Ser Gln Tyr Glu Thr Leu Asn Thr Pro
100 105 110
Glu Gln Val Leu Lys Ala Leu Glu Pro Gln Leu Asp Lys Asp Arg Pro
115 120 125
Val Ile Arg Gln Gly Asp Val Leu Leu Asn Gly Cys Arg Val Arg Leu
130 135 140

Cys Glu Pro Val Asn Gln Gly Lys Val Val Lys Gly Thr Thr Lys Leu
145 150 155 160

Thr Val Ala Lys Glu Gln Glu Thr Ile Gln Pro Ala Asp Glu Ala Ala
165 170 175

Asp Val Ala Phe Asp Ile Ala Glu Phe Leu Asp Phe Asp Thr Ser Val
180 185 190

Ala Lys Thr Arg Glu Ser Thr Asn Leu Gln Val Ala Pro Leu Glu Gly
195 200 205

Ala Ile Pro Thr Pro Leu Ser Asp Arg Phe Asp Asp Cys Glu Ser Arg
210 215 220

Gly Phe Val Lys Ser Glu Thr Met Ser Lys Leu Gly Val Phe Ser Gly
225 230 235 240

Asp Ile Val Ser Ile Lys Thr Lys Asn Gly Ala Glu Arg Val Leu Arg
245 250 255

Leu Phe Ala Tyr Pro Glu Pro Asn Thr Val Lys Tyr Asp Val Val Tyr
260 265 270

Val Ser Pro Ile Leu Tyr His Asn Ile Gly Asp Lys Glu Ile Glu Val
275 280 285

Thr Pro Asn Gly Glu Thr His Lys Ser Val Gly Glu Ala Leu Asp Ser
290 295 300

Val Leu Glu Ala Ala Glu Glu Val Lys Leu Ala Arg Val Leu Gly Pro
305 310 315 320

Thr Thr Thr Asp Arg Thr Phe Gln Thr Ala Tyr His Ala Gly Leu Gln
325 330 335

Ala Tyr Phe Lys Pro Val Lys Arg Ala Val Arg Val Gly Asp Leu Ile
340 345 350

Pro Ile Pro Phe Asp Ser Ile Leu Ala Arg Thr Ile Gly Glu Asp Pro
355 360 365

Glu Met Ser His Ile Pro Leu Glu Ala Leu Ala Val Lys Pro Asp Ser
370 375 380

Val Ala Trp Phe Gln Val Thr Ser Leu Asn Gly Ser Glu Asp Pro Ala

385 390 395 400
 Ser Lys Gln Tyr Leu Val Asp Ser Ser Gln Thr Lys Leu Ile Glu Gly
 405 410 415

 Gly Thr Thr Ser Ser Ala Val Ile Pro Thr Ser Val Pro Trp Arg Glu
 420 425 430
 Tyr Leu Gly Leu Asp Thr Leu Pro Lys Phe Gly Ser Glu Phe Ala Tyr
 435 440 445
 Ala Asp Lys Ile Arg Asn Leu Val Gln Ile Ser Thr Ser Ala Leu Ser
 450 455 460
 His Ala Lys Leu Asn Thr Ser Val Leu Leu His Ser Ala Lys Arg Gly
 465 470 475 480

 Val Gly Lys Ser Thr Val Leu Arg Ser Val Ala Ala Gln Cys Gly Ile
 485 490 495
 Ser Val Phe Glu Ile Ser Cys Phe Gly Leu Ile Gly Asp Asn Glu Ala
 500 505 510
 Gln Thr Leu Gly Thr Leu Arg Ala Lys Leu Asp Arg Ala Tyr Gly Cys
 515 520 525
 Ser Pro Cys Val Val Val Leu Gln His Leu Glu Ser Ile Ala Lys Lys
 530 535 540

 Ser Asp Gln Asp Gly Lys Asp Glu Gly Ile Val Ser Lys Leu Val Asp
 545 550 555 560
 Val Leu Ala Asp Tyr Ser Gly His Gly Val Leu Leu Ala Ala Thr Ser
 565 570 575
 Asn Asp Pro Asp Lys Ile Ser Glu Ala Ile Arg Ser Arg Phe Gln Phe
 580 585 590
 Glu Ile Glu Ile Gly Val Pro Ser Glu Pro Gln Arg Arg Gln Ile Phe
 595 600 605

 Ser His Leu Thr Lys Ser Gly Pro Gly Gly Asp Ser Ile Arg Asn Ala
 610 615 620
 Pro Ile Ser Leu Arg Ser Asp Val Ser Val Glu Asn Leu Ala Leu Gln
 625 630 635 640

Ser Ala Gly Leu Thr Pro Pro Asp Leu Thr Ala Ile Val Gln Thr Thr
645 650 655

Arg Leu Arg Ala Ile Asp Arg Leu Asn Lys Leu Thr Lys Asp Ser Asp
660 665 670

Thr Thr Leu Asp Asp Leu Leu Thr Leu Ser His Gly Thr Leu Gln Leu
675 680 685

Thr Pro Ser Asp Phe Asp Asp Ala Ile Ala Asp Ala Arg Gln Lys Tyr
690 695 700

Ser Asp Ser Ile Gly Ala Pro Arg Ile Pro Asn Val Gly Trp Asp Asp
705 710 715 720

Val Gly Gly Met Glu Gly Val Lys Lys Asp Ile Leu Asp Thr Ile Glu
725 730 735

Thr Pro Leu Lys Tyr Pro His Trp Phe Ser Asp Gly Val Lys Lys Arg
740 745 750

Ser Gly Ile Leu Phe Tyr Gly Pro Pro Gly Thr Gly Lys Thr Leu Leu
755 760 765

Ala Lys Ala Ile Ala Thr Thr Phe Ser Leu Asn Phe Phe Ser Val Lys
770 775 780

Gly Pro Glu Leu Leu Asn Met Tyr Ile Gly Glu Ser Glu Ala Asn Val
785 790 795 800

Arg Arg Val Phe Gln Lys Ala Arg Asp Ala Lys Pro Cys Val Val Phe
805 810 815

Phe Asp Glu Leu Asp Ser Val Ala Pro Gln Arg Gly Asn Gln Gly Asp
820 825 830

Ser Gly Gly Val Met Asp Arg Ile Val Ser Gln Leu Leu Ala Glu Leu
835 840 845

Asp Gly Met Ser Thr Ala Gly Gly Glu Gly Val Phe Val Val Gly Ala
850 855 860

Thr Asn Arg Pro Asp Leu Leu Asp Glu Ala Leu Leu Arg Pro Gly Arg
865 870 875 880

Phe Asp Lys Met Leu Tyr Leu Gly Ile Ser Asp Thr His Glu Lys Gln

885 890 895
 Gln Thr Ile Met Glu Ala Leu Thr Arg Lys Phe Arg Leu Ala Ala Asp
 900 905 910
 Val Ser Leu Glu Ala Ile Ser Lys Arg Cys Pro Phe Thr Phe Thr Gly
 915 920 925

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

 Phe Glu Lys Ala Lys Asp Glu Leu Ser Pro Ser Val Ser Ala Glu Glu
 995 1000 1005
 Leu Gln His Tyr Leu Lys Leu Arg Gln Gln Phe Glu Gly Gly Lys
 1010 1015 1020
 Lys

<210> 58

<211> 356

<212> PRT

<213> *Yarrowia lipolytica*

<220><221> MISC_FEATURE

<222> (1)..(356)

<223> YIPex7p; GenBank Accession No. CAG78389

<400> 58

Met Leu Gly Phe Lys Thr Gln Gly Phe Asn Gly Tyr Ala Ala Asn Tyr

1 5 10 15
 Ser Pro Phe Phe Asn Asp Lys Ile Ala Val Gly Thr Ala Ala Asn Tyr
 20 25 30
 Gly Leu Val Gly Asn Gly Lys Leu Phe Ile Leu Gly Ile Ser Pro Glu

35 40 45
 Gly Arg Met Val Cys Glu Gly Gln Phe Asp Thr Gln Asp Gly Ile Phe
 50 55 60
 Asp Val Ala Trp Ser Glu Gln His Glu Asn His Val Ala Thr Ala Cys

 65 70 75 80
 Gly Asp Gly Ser Val Lys Leu Phe Asp Ile Lys Ala Gly Ala Phe Pro
 85 90 95
 Leu Val Ser Phe Lys Glu His Thr Arg Glu Val Phe Ser Val Asn Trp
 100 105 110
 Asn Met Ala Asn Lys Ala Leu Phe Cys Thr Ser Ser Trp Asp Ser Thr
 115 120 125
 Ile Lys Ile Trp Thr Pro Glu Arg Thr Asn Ser Ile Met Thr Leu Gly

 130 135 140
 Gln Pro Ala Pro Ala Gln Gly Thr Asn Ala Ser Ala His Ile Gly Arg
 145 150 155 160
 Gln Thr Ala Pro Asn Gln Ala Ala Ala Gln Glu Cys Ile Tyr Ser Ala
 165 170 175
 Lys Phe Ser Pro His Thr Asp Ser Ile Ile Ala Ser Ala His Ser Thr
 180 185 190
 Gly Met Val Lys Val Trp Asp Thr Arg Ala Pro Gln Pro Leu Gln Gln

 195 200 205
 Gln Phe Ser Thr Gln Gln Thr Glu Ser Gly Gly Pro Pro Glu Val Leu
 210 215 220
 Ser Leu Asp Trp Asn Lys Tyr Arg Pro Thr Val Ile Ala Thr Gly Gly
 225 230 235 240
 Val Asp Arg Ser Val Gln Val Tyr Asp Ile Arg Met Thr Gln Pro Ala
 245 250 255
 Ala Asn Gln Pro Val Gln Pro Leu Ser Leu Ile Leu Gly His Arg Leu

 260 265 270
 Pro Val Arg Gly Val Ser Trp Ser Pro His His Ala Asp Leu Leu Leu
 275 280 285

Ser Cys Ser Tyr Asp Met Thr Ala Arg Val Trp Arg Asp Ala Ser Thr
290 295 300

Gly Gly Asn Tyr Leu Ala Arg Gln Arg Gly Gly Thr Glu Val Lys Cys
305 310 315 320

Met Asp Arg His Thr Glu Phe Val Ile Gly Gly Asp Trp Ser Leu Trp

325 330 335
Gly Asp Pro Gly Trp Ile Thr Thr Val Gly Trp Asp Gln Met Val Tyr
340 345 350

Val Trp His Ala
355

<210> 59

<211> 671

<212> PRT

<213> Yarrowia lipolytica

<220><221> MISC_FEATURE

<222> (1)..(671)

<223> YIPex8p; GenBank Accession No. CAG80447

<400> 59

Met Asn Lys Tyr Leu Val Pro Pro Pro Gln Ala Asn Arg Thr Val Thr
1 5 10 15

Asn Leu Asp Leu Leu Ile Asn Asn Leu Arg Gly Ser Ser Thr Pro Gly
20 25 30

Ala Ala Glu Val Asp Thr Arg Asp Ile Leu Gln Arg Ile Val Phe Ile
35 40 45

Leu Pro Thr Ile Lys Asn Pro Leu Asn Leu Asp Leu Val Ile Lys Glu
50 55 60

Ile Ile Asn Ser Pro Arg Leu Leu Pro Pro Leu Ile Asp Leu His Asp
65 70 75 80

Tyr Gln Gln Leu Thr Asp Ala Phe Arg Ala Thr Ile Lys Arg Lys Ala
85 90 95

Leu Val Thr Asp Pro Thr Ile Ser Phe Glu Ala Trp Leu Glu Thr Cys
100 105 110

Phe Gln Val Ile Thr Arg Phe Ala Gly Pro Gly Trp Lys Lys Leu Pro
 115 120 125
 Leu Leu Ala Gly Leu Ile Leu Ala Asp Tyr Asp Ile Ser Ala Asp Gly
 130 135 140

 Pro Thr Leu Glu Arg Lys Pro Gly Phe Pro Ser Lys Leu Lys His Leu
 145 150 155 160
 Leu Lys Arg Glu Phe Val Thr Thr Phe Asp Gln Cys Leu Ser Ile Asp
 165 170 175
 Thr Arg Asn Arg Ser Asp Ala Thr Lys Trp Val Pro Val Leu Ala Cys
 180 185 190
 Ile Ser Ile Ala Gln Val Tyr Ser Leu Leu Gly Asp Val Ala Ile Asn
 195 200 205

 Tyr Arg Arg Phe Leu Gln Val Gly Leu Asp Leu Ile Phe Ser Asn Tyr
 210 215 220
 Gly Leu Glu Met Gly Thr Ala Leu Ala Arg Leu His Ala Glu Ser Gly
 225 230 235 240
 Gly Asp Ala Thr Thr Ala Gly Gly Leu Ile Gly Lys Lys Leu Lys Glu
 245 250 255
 Pro Val Val Ala Leu Leu Asn Thr Phe Ala His Ile Ala Ser Ser Cys
 260 265 270

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

 Leu Leu Thr Gln Lys Met Asn Gln Phe Met Tyr Leu Gln Leu Ala Tyr
 340 345 350
 Lys Gln Leu Gln Ala Leu His Ser Ile Tyr Phe Ile Val Asp Gln Met

355	360	365	
Gly Ser Gln Phe Ala Ala Tyr Asp Tyr Val Phe Phe Ser Ala Ile Asp			
370	375	380	
Val Leu Leu Ser Glu Tyr Ala Pro Tyr Ile Lys Asn Arg Gly Thr Ile			
385	390	395	400
Pro Pro Asn Lys Glu Phe Val Ala Glu Arg Leu Ala Ala Asn Leu Ala			
405	410	415	
Gly Thr Ser Asn Val Gly Ser His Leu Pro Ile Asp Arg Ser Arg Val			
420	425	430	
Leu Phe Ala Leu Asn Tyr Tyr Glu Gln Leu Val Thr Val Cys His Asp			
435	440	445	
Ser Cys Val Glu Thr Ile Ile Tyr Pro Met Ala Arg Ser Phe Leu Tyr			
450	455	460	
Pro Thr Ser Asp Ile Gln Gln Leu Lys Pro Leu Val Glu Ala Ala His			
465	470	475	480
Ser Val Ile Leu Ala Gly Leu Ala Val Pro Thr Asn Ala Val Val Asn			
485	490	495	
Ala Lys Leu Ile Pro Glu Tyr Met Gly Gly Val Leu Pro Leu Phe Pro			
500	505	510	
Gly Val Phe Ser Trp Asn Gln Phe Val Leu Ala Ile Gln Ser Ile Val			
515	520	525	
Asn Thr Val Ser Pro Pro Ser Glu Val Phe Lys Thr Asn Gln Lys Leu			
530	535	540	
Phe Arg Leu Val Leu Asp Ser Leu Met Lys Lys Cys Arg Asp Thr Pro			
545	550	555	560
Val Gly Ile Pro Val Pro His Ser Val Thr Val Ser Gln Glu Gln Glu			
565	570	575	
Asp Ile Pro Pro Thr Gln Arg Ala Val Val Met Leu Ala Leu Ile Asn			
580	585	590	
Ser Leu Pro Tyr Val Asp Ile Arg Ser Phe Glu Leu Trp Leu Gln Glu			
595	600	605	

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

Ile Arg Trp Trp Tyr Lys Lys Asn Ala Arg Val His Gly Thr Leu
660 665 670

<210> 60

<211> 377

<212> PRT

<213> Yarrowia lipolytica

<220><221> MISC_FEATURE

<222> (1)..(377)

<223> YIPex10p; GenBank Accession No. CAG81606

<400> 60

Met Trp Gly Ser Ser His Ala Phe Ala Gly Glu Ser Asp Leu Thr Leu
1 5 10 15
Gln Leu His Thr Arg Ser Asn Met Ser Asp Asn Thr Thr Ile Lys Lys
20 25 30

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

Ala Tyr Leu Gly Leu Cys Thr Leu Val Gly Ser Lys Thr Leu Gly Glu
100 105 110
Glu Tyr Thr Asn Leu Met Tyr Thr Ile Arg Asp Arg Thr Ala Leu Pro
115 120 125

Gly Val Val Arg Arg Phe Gly Tyr Val Leu Ser Asn Thr Leu Phe Pro
130 135 140

Tyr Leu Phe Val Arg Tyr Met Gly Lys Leu Arg Ala Lys Leu Met Arg
145 150 155 160

Glu Tyr Pro His Leu Val Glu Tyr Asp Glu Asp Glu Pro Val Pro Ser
165 170 175

Pro Glu Thr Trp Lys Glu Arg Val Ile Lys Thr Phe Val Asn Lys Phe
180 185 190

Asp Lys Phe Thr Ala Leu Glu Gly Phe Thr Ala Ile His Leu Ala Ile
195 200 205

Phe Tyr Val Tyr Gly Ser Tyr Tyr Gln Leu Ser Lys Arg Ile Trp Gly
210 215 220

Met Arg Tyr Val Phe Gly His Arg Leu Asp Lys Asn Glu Pro Arg Ile
225 230 235 240

Gly Tyr Glu Met Leu Gly Leu Leu Ile Phe Ala Arg Phe Ala Thr Ser
245 250 255

Phe Val Gln Thr Gly Arg Glu Tyr Leu Gly Ala Leu Leu Glu Lys Ser
260 265 270

Val Glu Lys Glu Ala Gly Glu Lys Glu Asp Glu Lys Glu Ala Val Val
275 280 285

Pro Lys Lys Lys Ser Ser Ile Pro Phe Ile Glu Asp Thr Glu Gly Glu
290 295 300

Thr Glu Asp Lys Ile Asp Leu Glu Asp Pro Arg Gln Leu Lys Phe Ile
305 310 315 320

Pro Glu Ala Ser Arg Ala Cys Thr Leu Cys Leu Ser Tyr Ile Ser Ala
325 330 335

Pro Ala Cys Thr Pro Cys Gly His Phe Phe Cys Trp Asp Cys Ile Ser
340 345 350

Glu Trp Val Arg Glu Lys Pro Glu Cys Pro Leu Cys Arg Gln Gly Val
355 360 365

Arg Glu Gln Asn Leu Leu Pro Ile Arg

370 375

<210> 61

<211> 408

<212> PRT

<213> *Yarrowia lipolytica*

<220><221> MISC_FEATURE

<222> (1)..(408)

<223> YIPex12p; GenBank Accession No. CAG81532

<400> 61

Met Asp Tyr Phe Ser Ser Leu Asn Ala Ser Gln Leu Asp Pro Asp Val

1 5 10 15

Pro Thr Leu Phe Glu Leu Leu Ser Ala Lys Gln Leu Glu Gly Leu Ile

20 25 30

Ala Pro Ser Val Arg Tyr Ile Leu Ala Phe Tyr Ala Gln Arg His Pro

35 40 45

Arg Tyr Leu Leu Arg Ile Val Asn Arg Tyr Asp Glu Leu Tyr Ala Leu

50 55 60

Phe Met Gly Leu Val Glu Tyr Tyr Asn Leu Lys Thr Trp Asn Ala Ser

65 70 75 80

Phe Thr Glu Lys Phe Tyr Gly Leu Lys Arg Thr Gln Ile Leu Thr Asn

85 90 95

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

100 105 110

Lys Arg Leu Ser Lys Lys Lys Ile Trp Gly Ser Leu Phe Phe Leu Ile

115 120 125

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

130 135 140

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

145 150 155 160

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

165 170 175

Trp Leu Leu Lys Leu Tyr Pro Ile Val Thr Met Gly Cys Thr Thr Ala

180 185 190
 Thr Leu Ala Phe His Met Leu Phe Leu Phe Ser Val Thr Arg Ala Tyr
 195 200 205

Ser Ile Asp Asp Phe Leu Leu Asn Ile Gln Phe Ser Arg Met Thr Arg
 210 215 220
 Tyr Asp Tyr Gln Met Glu Thr Gln Arg Asp Ser Arg Asn Ala Ala Asn
 225 230 235 240
 Val Ala His Thr Met Lys Ser Ile Ser Glu Tyr Pro Val Ala Glu Arg
 245 250 255
 Val Met Leu Leu Leu Thr Thr Lys Ala Gly Ala Asn Ala Met Arg Ser
 260 265 270

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

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

Thr Gly Leu Arg Arg Leu Met Val
 405

<210> 62

<211> 412

<212> PRT

<213> Yarrowia lipolytica

<220><221> MISC_FEATURE

<222> (1)..(412)

<223> YIPex13p; GenBank Accession No. CAG81789

<400> 62

Met Ser Val Pro Arg Pro Lys Pro Trp Glu Gly Ala Ser Gly Ser Ser

1 5 10 15

Ala Ala Thr Ala Thr Pro Ala Ala Thr Ala Thr Pro Ala Ser Thr Asp

20 25 30

Ala Val Ser Ser Ser Ala Gly Ser Ala Thr Gly Ala Pro Glu Leu Pro

35 40 45

Ser Arg Pro Ser Ala Met Gly Ser Thr Ser Asn Ala Leu Ser Ser Pro

50 55 60

Met Gly Ser Ser Met Asn Ser Gly Tyr Gly Gly Met Asn Ser Gly Tyr

65 70 75 80

Gly Gly Met Gly Ser Ser Tyr Gly Ser Gly Tyr Gly Ser Ser Tyr Gly

85 90 95

Met Gly Ser Ser Tyr Gly Ser Gly Tyr Gly Ser Gly Leu Gly Gly Tyr

100 105 110

Gly Ser Tyr Gly Gly Met Gly Gly Met Gly Gly Met Tyr Gly Ser Arg

115 120 125

Tyr Gly Gly Tyr Gly Ser Tyr Gly Gly Met Gly Gly Tyr Gly Gly Tyr

130 135 140

Gly Gly Met Gly Gly Gly Pro Met Gly Gln Asn Gly Leu Ala Gly Gly

145 150 155 160

Thr Gln Ala Thr Phe Gln Leu Ile Glu Ser Ile Val Gly Ala Val Gly

165 170 175

Gly Phe Ala Gln Met Leu Glu Ser Thr Tyr Met Ala Thr Gln Ser Ser

180 185 190

Phe Phe Ala Met Val Ser Val Ala Glu Gln Phe Gly Asn Leu Lys Asn

195 200 205

Thr Leu Gly Ser Leu Leu Gly Ile Tyr Ala Ile Met Arg Trp Ala Arg
 210 215 220
 Arg Leu Val Ala Lys Leu Ser Gly Gln Pro Val Thr Gly Ala Asn Gly

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

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

 355 360 365
 Trp Trp Arg Val Arg Thr Arg Asp Gly Arg Ser Gly Tyr Val Pro Ala
 370 375 380
 Asn Tyr Leu Glu Val Ile Pro Arg Pro Ala Val Glu Ala Pro Lys Lys
 385 390 395 400
 Val Glu Glu Ile Gly Ala Ser Ala Val Pro Val Asn
 405 410

 <210> 63
 <211> 380
 <212> PRT
 <213> Yarrowia lipolytica
 <220><221> MISC_FEATURE

<222> (1)..(380)

<223> YIPex14p; GenBank Accession No. CAG79323

<400> 63

Met Ile Pro Ser Cys Leu Ser Thr Gln His Met Ala Pro Arg Glu Asp

1 5 10 15

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

20 25 30

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

35 40 45

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

50 55 60

Gln Ser His Pro Gly Ser Val Val Ser His Gly Gly Ala Ala Pro Thr

65 70 75 80

Val Pro Ala Ser Tyr Ala Phe Gln Ser Ala Pro Pro Leu Pro Glu Arg

85 90 95

Asp Trp Lys Asp Val Phe Ile Met Ala Thr Val Thr Val Gly Val Gly

100 105 110

Phe Gly Leu Tyr Thr Val Ala Lys Arg Tyr Leu Met Pro Leu Ile Leu

115 120 125

Pro Pro Thr Pro Pro Ser Leu Glu Ala Asp Lys Glu Ala Leu Glu Ala

130 135 140

Glu Phe Ala Arg Val Gln Gly Leu Leu Asp Gln Val Gln Gln Asp Thr

145 150 155 160

Glu Glu Val Lys Asn Ser Gln Val Glu Val Ala Lys Arg Val Thr Asp

165 170 175

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

180 185 190

Lys Lys Arg Asp Asp Glu Met Lys Leu Val Thr Ala Glu Val Glu Arg

195 200 205

Ile Arg Asp Arg Leu Pro Lys Asn Ile Asp Lys Leu Lys Asp Ser Gln

210 215 220

Glu Gln Gly Leu Ala Asp Ile Gln Ser Glu Leu Lys Ser Leu Lys Gln

225 230 235 240
Leu Leu Ser Thr Arg Thr Ala Ala Ser Ser Gly Pro Lys Leu Pro Pro

245 250 255
Ile Pro Pro Pro Ser Ser Tyr Leu Thr Arg Lys Ala Ser Pro Ala Val

260 265 270
Pro Ala Ala Ala Pro Ala Pro Val Thr Pro Gly Ser Pro Val His Asn

275 280 285
Val Ser Ser Ser Ser Thr Val Pro Ala Asp Arg Asp Asp Phe Ile Pro

290 295 300
Thr Pro Ala Gly Ala Val Pro Met Ile Pro Gln Pro Ala Ser Met Ser

305 310 315 320
Ser Ser Ser Thr Ser Thr Val Pro Asn Ser Ala Ile Ser Ser Ala Pro

325 330 335
Ser Pro Ile Gln Glu Pro Glu Pro Phe Val Pro Glu Pro Gly Asn Ser

340 345 350
Ala Val Lys Lys Pro Ala Pro Lys Ala Ser Ile Pro Ala Trp Gln Leu

355 360 365
Ala Ala Leu Glu Lys Glu Lys Glu Lys Glu Lys Glu

370 375 380

<210> 64

<211> 391

<212> PRT

<213> Yarrowia lipolytica

<220><221> MISC_FEATURE

<222> (1)..(391)

<223> YlPex16p; GenBank Accession No. CAG79622

<400> 64

Met Thr Asp Lys Leu Val Lys Val Met Gln Lys Lys Lys Ser Ala Pro

1 5 10 15

Gln Thr Trp Leu Asp Ser Tyr Asp Lys Phe Leu Val Arg Asn Ala Ala

20 25 30

Ser Ile Gly Ser Ile Glu Ser Thr Leu Arg Thr Val Ser Tyr Val Leu

35 40 45

Pro Gly Arg Phe Asn Asp Val Glu Ile Ala Thr Glu Thr Leu Tyr Ala

50 55 60

Val Leu Asn Val Leu Gly Leu Tyr His Asp Thr Ile Ile Ala Arg Ala

65 70 75 80

Val Ala Ala Ser Pro Asn Ala Ala Ala Val Tyr Arg Pro Ser Pro His

85 90 95

Asn Arg Tyr Thr Asp Trp Phe Ile Lys Asn Arg Lys Gly Tyr Lys Tyr

100 105 110

Ala Ser Arg Ala Val Thr Phe Val Lys Phe Gly Glu Leu Val Ala Glu

115 120 125

Met Val Ala Lys Lys Asn Gly Gly Glu Met Ala Arg Trp Lys Cys Ile

130 135 140

Ile Gly Ile Glu Gly Ile Lys Ala Gly Leu Arg Ile Tyr Met Leu Gly

145 150 155 160

Ser Thr Leu Tyr Gln Pro Leu Cys Thr Thr Pro Tyr Pro Asp Arg Glu

165 170 175

Val Thr Gly Glu Leu Leu Glu Thr Ile Cys Arg Asp Glu Gly Glu Leu

180 185 190

Asp Ile Glu Lys Gly Leu Met Asp Pro Gln Trp Lys Met Pro Arg Thr

195 200 205

Gly Arg Thr Ile Pro Glu Ile Ala Pro Thr Asn Val Glu Gly Tyr Leu

210 215 220

Leu Thr Lys Val Leu Arg Ser Glu Asp Val Asp Arg Pro Tyr Asn Leu

225 230 235 240

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

245 250 255

Leu Arg Pro Leu Ile Tyr Ala Cys Leu Leu Phe Arg Gln His Val Asn

260 265 270

Lys Thr Val Pro Ala Ser Thr Lys Ser Lys Phe Pro Phe Leu Asn Ser

275 280 285
Pro Trp Ala Pro Trp Ile Ile Gly Leu Val Ile Glu Ala Leu Ser Arg

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

355 360 365
Leu Phe Gly Ala Leu Ile Ser Asp Tyr Leu Tyr Leu Phe Asp Arg Tyr

370 375 380

Tyr Phe Thr Ala Ser Thr Leu

385 390

<210> 65

<211> 225

<212> PRT

<213> Yarrowia lipolytica

<220><221> MISC_FEATURE

<222> (1)..(225)

<223> YIPex17p; GenBank Accession No. CAG84025

<400> 65

Met Ser Ala Phe Pro Glu Pro Ser Ser Phe Glu Ile Glu Phe Ala Lys

1 5 10 15

Gln Met Asn Arg Pro Arg Thr Val Gln Phe Lys Gln Leu Val Ala Val

20 25 30

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

35 40 45

Thr Ile Leu Asn Pro Leu Phe Glu Glu Leu Thr Phe Ala Arg Ser Glu

50 55 60

Tyr Ala Ile His Ala Arg Arg Leu Met Glu Gln Leu Asn Ala Lys Leu

65					70					75					80
Ser	Ser	Met	Ala	Ser	Tyr	Ile	Pro	Pro	Val	Arg	Ala	Leu	Gln	Gly	Gln
				85				90				95			
Arg	Phe	Val	Asp	Ala	Gln	Thr	Gln	Thr	Glu	Asp	Glu	Glu	Gly	Glu	Asp
				100				105				110			
Ile	Pro	Asn	Pro	Ser	Leu	Gly	Lys	Ser	Ser	His	Val	Ser	Phe	Gly	Glu
				115				120				125			
Ser	Pro	Met	Gln	Leu	Lys	Leu	Ala	Glu	Lys	Glu	Lys	Gln	Gln	Lys	Leu
130				135				140							
Ile	Asp	Asp	Ser	Val	Asp	Asn	Leu	Glu	Arg	Leu	Ala	Asp	Ser	Leu	Lys
145				150				155				160			
His	Ala	Gly	Glu	Val	Ser	Asp	Leu	Ser	Ala	Leu	Ser	Gly	Phe	Lys	Tyr
				165				170				175			
Gln	Val	Glu	Glu	Leu	Thr	Asn	Tyr	Ser	Asp	Gln	Leu	Ala	Met	Ser	Gly
				180				185				190			
Tyr	Ser	Met	Met	Lys	Ser	Gly	Leu	Pro	Gly	His	Glu	Thr	Ala	Met	Ser
195				200				205							
Glu	Thr	Lys	Lys	Glu	Ile	Arg	Ser	Leu	Lys	Gly	Ser	Val	Leu	Ser	Val
210				215				220							
Arg															
225															
<210> 66															
<211> 324															
<212> PRT															
<213> Yarrowia lipolytica															
<220><221> MISC_FEATURE															
<222> (1)..(324)															
<223> YIPex19p; GenBank Accession No. AAK84827															
<400> 66															
Met	Ser	His	Glu	Glu	Asp	Leu	Asp	Asp	Leu	Asp	Asp	Phe	Leu	Asp	Glu
1				5				10				15			

Phe Asp Glu Gln Val Leu Ser Lys Pro Pro Gly Ala Gln Lys Asp Ala

20 25 30

Thr Pro Thr Thr Ser Thr Ala Pro Thr Thr Ala Glu Ala Lys Pro Asp

35 40 45

Ala Thr Lys Lys Ser Thr Glu Thr Ser Gly Thr Asp Ser Lys Thr Glu

50 55 60

Gly Ala Asp Thr Ala Asp Lys Asn Ala Ala Thr Asp Ser Ala Glu Ala

65 70 75 80

Gly Ala Glu Lys Val Ser Leu Pro Asn Leu Glu Asp Gln Leu Ala Gly

85 90 95

Leu Lys Met Asp Asp Phe Leu Lys Asp Ile Glu Ala Asp Pro Glu Ser

100 105 110

Lys Ala Gln Phe Glu Ser Leu Leu Lys Glu Ile Asn Asn Val Thr Ser

115 120 125

Ala Thr Ala Ser Glu Lys Ala Gln Gln Pro Lys Ser Phe Lys Glu Thr

130 135 140

Ile Ser Ala Thr Ala Asp Arg Leu Asn Gln Ser Asn Gln Glu Met Gly

145 150 155 160

Asp Met Pro Leu Gly Asp Asp Met Leu Ala Gly Leu Met Glu Gln Leu

165 170 175

Ser Gly Ala Gly Gly Phe Gly Glu Gly Gly Glu Gly Asp Phe Gly Asp

180 185 190

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

195 200 205

Gln Pro Leu Lys Glu Met His Asp Asn Tyr Pro Lys Trp Trp Asp Glu

210 215 220

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

225 230 235 240

Gln Gln Asp Ile Val Gly Lys Ile Cys Ala Lys Phe Glu Asp Pro Ser

245 250 255

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

260 265 270
Asp Glu Met Gln Glu Thr Gly Ala Pro Pro Asp Glu Ile Met Ser Asn

275 280 285
Val Ala Asp Gly Ser Ile Pro Gly Gly Leu Asp Gly Leu Gly Leu Gly
290 295 300
Gly Leu Gly Gly Gly Lys Met Pro Glu Met Pro Glu Asn Met Pro Glu
305 310 315 320
Cys Asn Gln Gln

<210> 67

<211> 417

<212> PRT

<213> Yarrowia lipolytica

<220><221> MISC_FEATURE

<222> (1)..(417)

<223> YIPex20p; GenBank Accession No. CAG79226

<400> 67

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

Asn Ser Phe Glu Gln Ser Gln Met Gly Pro His Phe Gly Gln Gln His
65 70 75 80
Phe Gly Gln Pro His Gln Pro Gln Met Gly Gln His Ala Pro Met Ala
85 90 95
His Gly Gln Gln Ser Asp Trp Ala Gln Ser Phe Ser Gln Leu Asn Leu
100 105 110
Gly Pro Gln Thr Gly Pro Gln His Thr Gln Gln Ser Asn Trp Gly Gln

115	120	125	
Asp Phe Met Arg Gln Ser Pro Gln Ser His Gln Val Gln Pro Gln Met			
130	135	140	
Ala Asn Gly Val Met Gly Ser Met Ser Gly Met Ser Ser Phe Gly Pro			
145	150	155	160
Met Tyr Ser Asn Ser Gln Leu Met Asn Ser Thr Tyr Gly Leu Gln Thr			
165	170	175	
Glu His Gln Gln Thr His Lys Thr Glu Thr Lys Ser Ser Gln Asp Ala			
180	185	190	
Ala Phe Glu Ala Ala Phe Gly Ala Val Glu Glu Ser Ile Thr Lys Thr			
195	200	205	
Ser Asp Lys Gly Lys Glu Val Glu Lys Asp Pro Met Glu Gln Thr Tyr			
210	215	220	
Arg Tyr Asp Gln Ala Asp Ala Leu Asn Arg Gln Ala Glu His Ile Ser			
225	230	235	240
Asp Asn Ile Ser Arg Glu Glu Val Asp Ile Lys Thr Asp Glu Asn Gly			
245	250	255	
Glu Phe Ala Ser Ile Ala Arg Gln Ile Ala Ser Ser Leu Glu Glu Ala			
260	265	270	
Asp Lys Ser Lys Phe Glu Lys Ser Thr Phe Met Asn Leu Met Arg Arg			
275	280	285	
Ile Gly Asn His Glu Val Thr Leu Asp Gly Asp Lys Leu Val Asn Lys			
290	295	300	
Glu Gly Glu Asp Ile Arg Glu Glu Val Arg Asp Glu Leu Leu Arg Glu			
305	310	315	320
Gly Ala Ser Gln Glu Asn Gly Phe Gln Ser Glu Ala Gln Gln Thr Ala			
325	330	335	
Pro Leu Pro Val His His Glu Ala Pro Pro Pro Glu Gln Ile His Pro			
340	345	350	
His Thr Glu Thr Gly Asp Lys Gln Leu Glu Asp Pro Met Val Tyr Ile			
355	360	365	

Glu Gln Glu Ala Ala Arg Arg Ala Ala Glu Ser Gly Arg Thr Val Glu
370 375 380

Glu Glu Lys Leu Asn Phe Tyr Ser Pro Phe Glu Tyr Ala Gln Lys Leu
385 390 395 400
Gly Pro Gln Gly Val Ala Lys Gln Ser Asn Trp Glu Glu Asp Tyr Asp
405 410 415

Phe

<210> 68

<211> 195

<212> PRT

<213> Yarrowia lipolytica

<220><221> MISC_FEATURE

<222> (1)..(195)

<223> YIPex22p; GenBank Accession No. CAG77876

<400> 68

Val Pro Arg Cys Thr Ser His Pro Cys Asn Leu Thr Leu His Leu Pro

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

65 70 75 80
Lys Ile Thr Val Val Val Ser Gln Glu Leu Val Gln Ser Gln Leu Val
85 90 95
Asp Phe Lys His Leu Met Ser Val His Pro Asn Leu Val Val Ile Val
100 105 110
Pro Pro Met Val Ala Asn Lys Phe His Arg Ala Leu Lys Ser Ser Val
115 120 125

Gly His Asp His Gly Val Lys Val Ile Arg Cys Asp Thr Asp Val Gly

130 135 140

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

145 150 155 160

Asp Gly Val Gly Asp Asn Ile Gln Gly Glu Ile Lys Arg Phe Val Gly

165 170 175

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

180 185 190

Thr Gly Leu

195

<210> 69

<211> 386

<212> PRT

<213> *Yarrowia lipolytica*

<220><221> MISC_FEATURE

<222> (1)..(386)

<223> YIPex26p; GenBank Accession No. NC_006072, antisense translation
of nucleotides 117230-118387

<400> 69

Met Pro Pro Ala Met Pro Gln Met Thr Thr Ser Thr Leu Leu Thr Asp

1 5 10 15

Ser Val Thr Ser Ala Val Asn Gln Ala Ala Thr Pro Lys Val Asp Gln

20 25 30

Met Tyr Gln Thr Phe Gly Glu Ser Ala Arg Glu Phe Val Asn Lys Asn

35 40 45

Phe Tyr Asn Ser Tyr Glu Leu Ile Arg Pro Phe Phe Asp Glu Ile Thr

50 55 60

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

65 70 75 80

Pro His Asn Ile Pro Leu Ser Leu Trp Ile Lys Val Trp Ser Leu Tyr

85 90 95

Leu Ala Ile Leu Asp Ala Ser Cys Lys Gln Ala Gly Glu Ala Leu Leu

100 105 110

Asn Ser Thr Gly Asp Leu Ser Gly Ser Asp Ser Gly Glu Trp Asn Gln

115 120 125

Thr Arg Lys Leu Leu Ala Arg Lys Leu Thr Ser Gly Ser Val Trp Asp

130 135 140

Glu Leu Val Thr Ala Ser Gly Gly Thr Gly Asn Ile His Pro Thr Ile

145 150 155 160

Leu Ala Leu Leu Ala Ser Leu Ser Ile Arg His Asp Thr Asp Ala Lys

165 170 175

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

180 185 190

Gly Ser Asp Asp Val Lys Thr Lys Thr Ala Phe Tyr Lys Val Leu Asp

195 200 205

Leu Tyr Leu Leu Arg Val Leu Pro Asp Leu Gly Gln Trp Asp Val Ala

210 215 220

His Ser Phe Val Asn Asn Thr Asn Leu Phe Ser His Glu Gln Lys Lys

225 230 235 240

Glu Met Thr His Lys Leu Asp Gln Ser Gln Lys His Ala Glu Gln Glu

245 250 255

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

260 265 270

Lys Glu Lys Glu Arg Glu Glu Arg Val Ser Arg Asp Thr Gln Ser Arg

275 280 285

Glu Ile Lys Ser Pro Ile Val Asp Ser Ser Thr Ser Ser Arg Asp Val

290 295 300

Thr Arg Asp Thr Thr Arg Glu Leu Ser Lys Ser Ser Arg Gln Pro Arg

305 310 315 320

Thr Leu Ser Gln Ile Ile Ser Thr Ser Leu Lys Ser Gln Phe Asp Gly

325 330 335

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

340 345 350
Ala Ala Asn Pro Leu Ile Arg Lys Arg Val Val Asp Thr Leu Lys Met

355 360 365
Leu Trp Ile Lys Ile Leu Gln Thr Leu Ser Met Gly Phe Lys Val Ser

370 375 380
Tyr Leu

385

<210> 70

<211> 10

<212> DNA

<213> Yarrowia lipolytica

<220><221> misc_feature

<222> (8)..(8)

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

<300><302> CODON-OPTIMIZED GENES FOR THE PRODUCTION OF POLYUNSATURATED FATTY
ACIDS IN OLEAGINOUS YEAST

<310> US 7,838,651

<311> 2008-01-23

<312> 2010-11-23

<313> (1)..(10)

<400> 70

mammatgnhs

10

<210> 71

<211> 5

<212> PRT

<213> Artificial Sequence

<220><223> Variant His box

<220><221> misc_feature

<222> (2)..(3)

<223> Xaa can be any naturally occurring amino acid

<400> 71

Gln Xaa Xaa His His

1

5

<210> 72

<211> 5

<212> PRT

<213> Artificial Sequence

<220><223> His box

<220><221> misc_feature

<222> (2)..(3)

<

<223> Xaa can be any naturally occurring amino acid

<400> 72

His Xaa Xaa His His

1 5

<210> 73

<211> 13

<212> PRT

<213> Artificial Sequence

<220><223> delta-9 elongase motif

<220><221> misc_feature

<222> (3)..(3)

<223> Xaa can be any naturally occurring amino acid

<220><221> MISC_FEATURE

<222> (4)..(4)

<223> X = L or F

<220><221> misc_feature

<222> (5)..(8)

<223> Xaa can be any naturally occurring amino acid

<220><221> misc_feature

<222> (10)..(11)

<223> Xaa can be any naturally occurring amino acid

<300><302> DELTA-9 ELONGASES AND THEIR USE IN MAKING POLYUNSATURATED FATTY

ACIDS

<310> US 7,645,604

<311> 2006-11-16

<312> 2010-01-12

<313> (1)..(13)

<400> 73

Tyr Asn Xaa Xaa Xaa Xaa Xaa Xaa Ser Xaa Xaa Ser Phe

1 5 10

<210> 74

<211> 15

<212> PRT

<213> Artificial Sequence

<220><223> delta-9 elongase motif

<220><221> misc_feature

<222> (3)..(3)

<223> Xaa can be any naturally occurring amino acid

<220><221> misc_feature

<222> (6)..(7)

<223> Xaa can be any naturally occurring amino acid

<220><221> MISC_FEATURE

<222> (8)..(8)

<223> X = E or D

<220><221> misc_feature

<222> (10)..(10)

<223> Xaa can be any naturally occurring amino acid

<220><221> MISC_FEATURE

<222> (12)..(12)

<223> X = T or S

<220><221> misc_feature

<222> (13)..(14)

<223> Xaa can be any naturally occurring amino acid

<300><302> DELTA-9 ELONGASES AND THEIR USE IN MAKING POLYUNSATURATED FATTY

ACIDS

<310> U.S. 7,645,604

<311> 2006-11-16

<312> 2010-01-12

<313> (1)..(15)

<400> 74

Phe Tyr Xaa Ser Lys Xaa Xaa Xaa Tyr Xaa Asp Xaa Xaa Xaa Leu

1 5 10 15

<210> 75

<211> 9

<212> PRT

<213> Artificial Sequence

<220><223> delta-9 elongase motif

<220><221> MISC_FEATURE

<222> (2)..(2)

<223> X = Q or H

<220><221> misc_feature

<222> (3)..(3)

<223> Xaa can be any naturally occurring amino acid

<220><221> misc_feature

<222> (7)..(7)

<223> Xaa can be any naturally occurring amino acid

<300><302> DELTA-9 ELONGASES AND THEIR USE IN MAKING POLYUNSATURATED FATTY
ACIDS

<310> U.S. 7,645,604

<311> 2006-11-16

<312> 2010-01-12

<313> (1)..(9)

<400> 75

Leu Xaa Xaa Phe His His Xaa Gly Ala

1 5

<210> 76

<211> 14

<212> PRT

<213> Artificial Sequence

<220><223> delta-9 elongase motif

<220><221> misc_feature

<222> (3)..(3)

<223> Xaa can be any naturally occurring amino acid

<220><221> misc_feature

<222> (6)..(12)

<223> Xaa can be any naturally occurring amino acid

<220><221> MISC_FEATURE

<222> (13)..(13)

<223> X = K or R or N

<300><302> DELTA-9 ELONGASES AND THEIR USE IN MAKING POLYUNSATURATED FATTY
ACIDS

<310> U.S. 7,645,604

<311> 2006-11-16

<312> 2010-01-12

<313> (1)..(14)

<400> 76

Met Tyr Xaa Tyr Tyr Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Phe

1 5 10

<210> 77

<211> 8

<212> PRT

<213> Artificial Sequence

<220><223> delta-9 elongase motif

<220><221> misc_feature

<222> (2)..(2)

<223> Xaa can be any naturally occurring amino acid

<220><221> MISC_FEATURE

<222> (4)..(4)

<223> X = I or L or M

<220><221> misc_feature

<222> (6)..(7)

<223> Xaa can be any naturally occurring amino acid

<300><302> DELTA-9 ELONGASES AND THEIR USE IN MAKING POLYUNSATURATED FATTY
ACIDS

<310> U.S. 7,645,604

<311> 2006-11-16

<312> 2010-01-12

<313> (1)..(8)

<400

> 77

Lys Xaa Leu Xaa Thr Xaa Xaa Gln

1 5

<210> 78

<211> 7

<212> PRT

<213> Artificial Sequence

<220><223> delta-9 elongase motif

<220><221> misc_feature

<222> (2)..(2)

<223> Xaa can be any naturally occurring amino acid

<220><221> misc_feature

<222> (6)..(6)

<223> Xaa can be any naturally occurring amino acid

<300><302> DELTA-9 ELONGASES AND THEIR USE IN MAKING POLYUNSATURATED FATTY ACIDS

<310> U.S. 7,645,604

<311> 2006-11-16

<312> 2010-01-12

<313> (1)..(7)

<400> 78

Trp Xaa Phe Asn Tyr Xaa Tyr

1 5

<210> 79

<211> 9

<212> PRT

<213> Artificial Sequence

<220><223> delta-9 elongase motif

<220><221> misc_feature

<222> (2)..(2)

<223> Xaa can be any naturally occurring amino acid

<220><221> misc_feature

<222> (4)..(4)

<223> Xaa can be any naturally occurring amino acid

<220><221> misc_feature

<222> (6)..(7)

<223> Xaa can be any naturally occurring amino acid

<300><302> DELTA-9 ELONGASES AND THEIR USE IN MAKING POLYUNSATURATED FATTY
ACIDS

<310> U.S. 7,645,604

<311> 2006-11-16

<312> 2010-01-12

<313> (1)..(9)

<400> 79

Tyr Xaa Gly Xaa Val Xaa Xaa Leu Phe

1 5

<210> 80

<211> 1449

<212> DNA

<213> Yarrowia lipolytica

<220><221> CDS

<222> (1)..(1449)

<223> delta-9 desaturase; GenBank Accession No. XM_501496

<400> 80

atg gtg aaa aac gtg gac caa gtg gat ctc tcg cag gtc gac acc att 48

Met Val Lys Asn Val Asp Gln Val Asp Leu Ser Gln Val Asp Thr Ile

1 5 10 15

gcc tcc ggc cga gat gtc aac tac aag gtc aag tac acc tcc ggc gtt 96

Ala Ser Gly Arg Asp Val Asn Tyr Lys Val Lys Tyr Thr Ser Gly Val

20 25 30

aag atg agc cag ggc gcc tac gac gac aag ggc cgc cac att tcc gag 144

Lys Met Ser Gln Gly Ala Tyr Asp Asp Lys Gly Arg His Ile Ser Glu

35	40	45	
cag ccc ttc acc tgg gcc aac tgg cac cag cac atc aac tgg ctc aac			192
Gln Pro Phe Thr Trp Ala Asn Trp His Gln His Ile Asn Trp Leu Asn			
50	55	60	
ttc att ctg gtg att gcg ctg cct ctg tcg tcc ttt gct gcc gct ccc			240
Phe Ile Leu Val Ile Ala Leu Pro Leu Ser Ser Phe Ala Ala Ala Pro			
65	70	75	80
ttc gtc tcc ttc aac tgg aag acc gcc gcg ttt gct gtc ggc tat tac			288
Phe Val Ser Phe Asn Trp Lys Thr Ala Ala Phe Ala Val Gly Tyr Tyr			
85	90	95	
atg tgc acc ggt ctc ggt atc acc gcc ggc tac cac cga atg tgg gcc			336
Met Cys Thr Gly Leu Gly Ile Thr Ala Gly Tyr His Arg Met Trp Ala			
100	105	110	
cat cga gcc tac aag gcc gct ctg ccc gtt cga atc atc ctt gct ctg			384
His Arg Ala Tyr Lys Ala Ala Leu Pro Val Arg Ile Ile Leu Ala Leu			
115	120	125	
ttt gga gga gga gct gtc gag ggc tcc atc cga tgg tgg gcc tcg tct			432
Phe Gly Gly Gly Ala Val Glu Gly Ser Ile Arg Trp Trp Ala Ser Ser			
130	135	140	
cac cga gtc cac cac cga tgg acc gac tcc aac aag gac cct tac gac			480
His Arg Val His His Arg Trp Thr Asp Ser Asn Lys Asp Pro Tyr Asp			
145	150	155	160
gcc cga aag gga ttc tgg ttc tcc cac ttt ggc tgg atg ctg ctt gtg			528
Ala Arg Lys Gly Phe Trp Phe Ser His Phe Gly Trp Met Leu Leu Val			
165	170	175	
ccc aac ccc aag aac aag ggc cga act gac att tct gac ctc aac aac			576
Pro Asn Pro Lys Asn Lys Gly Arg Thr Asp Ile Ser Asp Leu Asn Asn			
180	185	190	
gac tgg gtt gtc cga ctc cag cac aag tac tac gtt tac gtt ctc gtc			624
Asp Trp Val Val Arg Leu Gln His Lys Tyr Tyr Val Tyr Val Leu Val			
195	200	205	

ttc atg gcc att gtt ctg ccc acc ctc gtc tgt ggc ttt ggc tgg ggc	672
Phe Met Ala Ile Val Leu Pro Thr Leu Val Cys Gly Phe Gly Trp Gly	
210 215 220	
gac tgg aag gga ggt ctt gtc tac gcc ggt atc atg cga tac acc ttt	720
Asp Trp Lys Gly Gly Leu Val Tyr Ala Gly Ile Met Arg Tyr Thr Phe	
225 230 235 240	
gtg cag cag gtg act ttc tgt gtc aac tcc ctt gcc cac tgg att gga	768
Val Gln Gln Val Thr Phe Cys Val Asn Ser Leu Ala His Trp Ile Gly	
245 250 255	
gag cag ccc ttc gac gac cga cga act ccc cga gac cac gct ctt acc	816
Glu Gln Pro Phe Asp Asp Arg Arg Thr Pro Arg Asp His Ala Leu Thr	
260 265 270	
gcc ctg gtc acc ttt gga gag ggc tac cac aac ttc cac cac gag ttc	864
Ala Leu Val Thr Phe Gly Glu Gly Tyr His Asn Phe His His Glu Phe	
275 280 285	
ccc tcg gac tac cga aac gcc ctc atc tgg tac cag tac gac ccc acc	912
Pro Ser Asp Tyr Arg Asn Ala Leu Ile Trp Tyr Gln Tyr Asp Pro Thr	
290 295 300	
aag tgg ctc atc tgg acc ctc aag cag gtt ggt ctc gcc tgg gac ctc	960
Lys Trp Leu Ile Trp Thr Leu Lys Gln Val Gly Leu Ala Trp Asp Leu	
305 310 315 320	
cag acc ttc tcc cag aac gcc atc gag cag ggt ctc gtg cag cag cga	1008
Gln Thr Phe Ser Gln Asn Ala Ile Glu Gln Gly Leu Val Gln Gln Arg	
325 330 335	
cag aag aag ctg gac aag tgg cga aac aac ctc aac tgg ggt atc ccc	1056
Gln Lys Lys Leu Asp Lys Trp Arg Asn Asn Leu Asn Trp Gly Ile Pro	
340 345 350	
att gag cag ctg cct gtc att gag ttt gag gag ttc caa gag cag gcc	1104
Ile Glu Gln Leu Pro Val Ile Glu Phe Glu Glu Phe Gln Glu Gln Ala	
355 360 365	
aag acc cga gat ctg gtt ctc att tct ggc att gtc cac gac gtg tct	1152

Lys Thr Arg Asp Leu Val Leu Ile Ser Gly Ile Val His Asp Val Ser

370	375	380	
gcc ttt gtc gag cac cac cct ggt gga aag gcc ctc att atg agc gcc	1200		
Ala Phe Val Glu His His Pro Gly Gly Lys Ala Leu Ile Met Ser Ala			
385	390	395	400
gtc ggc aag gac ggt acc gct gtc ttc aac gga ggt gtc tac cga cac	1248		
Val Gly Lys Asp Gly Thr Ala Val Phe Asn Gly Gly Val Tyr Arg His			
405	410	415	
tcc aac gct ggc cac aac ctg ctt gcc acc atg cga gtt tcg gtc att	1296		
Ser Asn Ala Gly His Asn Leu Leu Ala Thr Met Arg Val Ser Val Ile			
420	425	430	
cga ggc ggc atg gag gtt gag gtg tgg aag act gcc cag aac gaa aag	1344		
Arg Gly Gly Met Glu Val Glu Val Trp Lys Thr Ala Gln Asn Glu Lys			
435	440	445	
aag gac cag aac att gtc tcc gat gag agt gga aac cga atc cac cga	1392		
Lys Asp Gln Asn Ile Val Ser Asp Glu Ser Gly Asn Arg Ile His Arg			
450	455	460	
gct ggt ctc cag gcc acc cgg gtc gag aac ccc ggt atg tct ggc atg	1440		
Ala Gly Leu Gln Ala Thr Arg Val Glu Asn Pro Gly Met Ser Gly Met			
465	470	475	480
gct gct tag			1449
Ala Ala			

<210> 81

<211> 482

<212> PRT

<213> Yarrowia lipolytica

<400> 81

Met Val Lys Asn Val Asp Gln Val Asp Leu Ser Gln Val Asp Thr Ile

1	5	10	15

Ala Ser Gly Arg Asp Val Asn Tyr Lys Val Lys Tyr Thr Ser Gly Val

				20				25				30			
Lys	Met	Ser	Gln	Gly	Ala	Tyr	Asp	Asp	Lys	Gly	Arg	His	Ile	Ser	Glu
35				40				45							
Gln	Pro	Phe	Thr	Trp	Ala	Asn	Trp	His	Gln	His	Ile	Asn	Trp	Leu	Asn
50				55				60							
Phe	Ile	Leu	Val	Ile	Ala	Leu	Pro	Leu	Ser	Ser	Phe	Ala	Ala	Ala	Pro
65				70				75				80			
Phe	Val	Ser	Phe	Asn	Trp	Lys	Thr	Ala	Ala	Phe	Ala	Val	Gly	Tyr	Tyr
				85				90				95			
Met	Cys	Thr	Gly	Leu	Gly	Ile	Thr	Ala	Gly	Tyr	His	Arg	Met	Trp	Ala
100				105				110							
His	Arg	Ala	Tyr	Lys	Ala	Ala	Leu	Pro	Val	Arg	Ile	Ile	Leu	Ala	Leu
115				120				125							
Phe	Gly	Gly	Gly	Ala	Val	Glu	Gly	Ser	Ile	Arg	Trp	Trp	Ala	Ser	Ser
130				135				140							
His	Arg	Val	His	His	Arg	Trp	Thr	Asp	Ser	Asn	Lys	Asp	Pro	Tyr	Asp
145				150				155				160			
Ala	Arg	Lys	Gly	Phe	Trp	Phe	Ser	His	Phe	Gly	Trp	Met	Leu	Leu	Val
				165				170				175			
Pro	Asn	Pro	Lys	Asn	Lys	Gly	Arg	Thr	Asp	Ile	Ser	Asp	Leu	Asn	Asn
180				185				190							
Asp	Trp	Val	Val	Arg	Leu	Gln	His	Lys	Tyr	Tyr	Val	Tyr	Val	Leu	Val
195				200				205							
Phe	Met	Ala	Ile	Val	Leu	Pro	Thr	Leu	Val	Cys	Gly	Phe	Gly	Trp	Gly
210				215				220							
Asp	Trp	Lys	Gly	Gly	Leu	Val	Tyr	Ala	Gly	Ile	Met	Arg	Tyr	Thr	Phe
225				230				235				240			
Val	Gln	Gln	Val	Thr	Phe	Cys	Val	Asn	Ser	Leu	Ala	His	Trp	Ile	Gly
				245				250				255			
Glu	Gln	Pro	Phe	Asp	Asp	Arg	Arg	Thr	Pro	Arg	Asp	His	Ala	Leu	Thr
260				265				270							

Ala Leu Val Thr Phe Gly Glu Gly Tyr His Asn Phe His His Glu Phe

275 280 285

Pro Ser Asp Tyr Arg Asn Ala Leu Ile Trp Tyr Gln Tyr Asp Pro Thr

290 295 300

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

305 310 315 320

Gln Thr Phe Ser Gln Asn Ala Ile Glu Gln Gly Leu Val Gln Gln Arg

325 330 335

Gln Lys Lys Leu Asp Lys Trp Arg Asn Asn Leu Asn Trp Gly Ile Pro

340 345 350

Ile Glu Gln Leu Pro Val Ile Glu Phe Glu Glu Phe Gln Glu Gln Ala

355 360 365

Lys Thr Arg Asp Leu Val Leu Ile Ser Gly Ile Val His Asp Val Ser

370 375 380

Ala Phe Val Glu His His Pro Gly Gly Lys Ala Leu Ile Met Ser Ala

385 390 395 400

Val Gly Lys Asp Gly Thr Ala Val Phe Asn Gly Gly Val Tyr Arg His

405 410 415

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

420 425 430

Arg Gly Gly Met Glu Val Glu Val Trp Lys Thr Ala Gln Asn Glu Lys

435 440 445

Lys Asp Gln Asn Ile Val Ser Asp Glu Ser Gly Asn Arg Ile His Arg

450 455 460

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

465 470 475 480

Ala Ala

<210> 82

<211> 4313

<212> DNA

<213> Artificial Sequence

<220><223> Plasmid pZKUM

<400> 82

taatcgagct tggcgtaatc atggcatag ctgtttcctg tgtgaaattg ttatccgctc	60
acaattccac acaacatacg agccggaagc ataaagtga aagcctgggg tgcctaatga	120
gtgagctaac tcacattaat tgcgttgccg tccactgccc ctttccagtc gggaaacctg	180
tcgtgccagc tgcattaatg aatcggccaa cgcgcgggga gaggcggttt gcgtattggg	240
cgctcttccg ctctctcgct cactgactcg ctgcgctcgg tcgttcggct gcggcgagcg	300
gtatcagctc actcaaagcg ggtaatacgg ttatccacag aatcagggga taacgcagga	360
aagaacatgt gagcaaaagg ccagcaaaag gccaggaacc gtaaaaaggc cgcgttgctg	420
gcgtttttcc ataggctccg ccccccgtac gagcatcaca aaaatcgacg ctcaagtcag	480
aggtggcgaa acccgacagg actataaaga taccaggcgt tccccctgg aagctccctc	540
gtgcgctctc ctgttccgac cctgcgcgtt accggatacc tgtccgcctt tctcccttcg	600
ggaagcgtgg cgcttttcca tagctcacgc ttaggtatc tcagttcggg ttaggtcggt	660
cgctccaagc tgggctgtgt gcacgaacc cccgttcagc ccgaccgtg cgccttatec	720
ggtaactatc gtcttgagtc caaccggta agacacgact tatcgccact ggacgcagcc	780
actggttaaca ggattagcag agcgaggtat gtaggcgggtg ctacagagtt ctggaagtgg	840
tgccctaact acggctacac tagaaggaca gtatttggta tctgcgctct gctgaagcca	900
gttaccttcg gaaaaagagt tggtagctct tgatccggca aacaaaccac cgctggtagc	960
ggtaggtttt ttgtttgcaa gcagcagatt acgcgcagaa aaaaaggatc tcaagaagat	1020
cctttgatct ttctacggg gtctgacgct cagtggaaacg aaaactcacg ttaagggtat	1080
ttggtcatga gattatcaa aaggatcttc acctagatcc ttttaatta aaaatgaagt	1140
tttaaatcaa tctaaagtat atatgagtaa acttggtctg acagttacca atgcttaatc	1200
agttaggcac ctatctcagc gatctgtcta tttcgttcat ccatagttgc ctgactcccc	1260
gtcgtgtaga taactacgat acgggagggc ttaccatctg gcccagtg tgcaatgata	1320
ccgcgagacc cagctcacc ggctccagat ttatcagcaa taaaccagcc agccggaagg	1380
gccgagcgca gaagtgggcc tgcaacttta tccgcctcca tccagtttat taattgttgc	1440
cgggaagcta gagtaagtag ttccgccagt aatagtttgc gcaacgttgt tgccattgct	1500
acaggcatcg tgggtgtcacg ctctcgtttt ggtaggtctt cattcagctc cggttcccaa	1560
cgatcaagcg gagttacatg atccccatg ttgtgcaaaa aagcggttag ctcttcgggt	1620

cctccgatcg ttgtcagaag taagttggcc gcagtggtat cactcatggt tatggcagca	1680
ctgcataatt ctcttactgt catgccatcc gtaagatgct tttctgtgac tggtagtac	1740
tcaaccaagt cattctgaga atagtgtatg cggcgaccga gttgctcttg cccggcgtea	1800
atacgggata ataccgcgcc acatagcaga actttaaaag tgctcatcat tggaaaacgt	1860
tcttcggggc gaaaactctc aaggatctta ccgctgttga gatccagttc gatgtaacce	1920
actcgtgcac ccaactgac ttcagcatct tttactttca ccagcgtttc tgggtgagca	1980
aaaacaggaa ggcaaaatgc cgcaaaaag ggaataaggg cgacacggaa atgttgaata	2040
ctcatactct tcctttttca atattattga agcatttate agggttattg tctcatgagc	2100
ggatacatat ttgaatgtat ttagaaaaat aaacaaatag gggttccgcg cacatttccc	2160
cgaaaagtgc cacctgacgc gccctgtagc ggcgcattaa gcgcggcggg tgtggtggtt	2220
acgcgcagcg tgaccgtac acttgccagc gccctagcgc ccgctccttt cgctttcttc	2280
ccttcctttc tcgccacgtt cgccggcttt ccccgtaag ctctaaatcg ggggctccct	2340
ttagggttcc gatttagtgc ttacggcac ctgcaccca aaaaacttga ttaggtgat	2400
ggttcacgta gtgggccatc gccctgatag acggtttttc gccctttgac gttggagtcc	2460
acgttcttta atagtggact cttgttccaa actggaacaa cactcaacce tatctcggtc	2520
tattcttttg atttataagg gattttgccg atttcggcct attggttaaa aaatgagctg	2580
atttaacaaa aatttaacgc gaattttaac aaaatatata cgcttacaat ttccattcgc	2640
cattcaggct gcgcaactgt tgggaagggc gatcggtgcg ggcctcttcg ctattacgcc	2700
agctggcgaa agggggatgt gctgcaaggc gattaagtig ggtaacgcca gggttttccc	2760
agtcacgacg ttgtaaaacg acggccagtg aattgtaata cgactcacta tagggcgaat	2820
tgggtaccgg gccccctc gaggtcgacg agtatctgtc tgactcgtca ttgccgcctt	2880
tggagtacga ctccaactat gagtgtgctt ggatcacttt gacgatacat tcttcgttgg	2940
aggctgtggg tctgacagct gcgttttcgg cgcggttggc cgacaacaat atcagctgca	3000
acgtcattgc tggttttcat catgatcaca tttttgtcgg caaaggcgac gccagagag	3060
ccattgacgt tctttctaatt ttggaccgat agccgtatag tccagtctat ctataagttc	3120
aactaactcg taactattac cataacatat acttactgc ccagataag gttccgataa	3180
aaagtctgc agactaaatt tatttcagtc tctcttcac caccaaaatg ccctcctacg	3240
aagctcgagt gctcaagctc gtggcagcca agaaaaccaa cctgtgtgct tctctggatg	3300
ttaccaccac caaggagctc attgagcttg ccgataaggt cggaccttat gtgtgcatga	3360

tcaaaaccca tatcgacatc attgacgact tcacctacgc cggcactgtg ctccccctca 3420
 aggaacttgc tcttaagcac ggtttcttcc tgttcgagga cagaaagttc gcagatattg 3480
 gcaacactgt caagcaccag taccggtgtc accgaatcgc cgagtgggtcc gatataacca 3540
 acgcccacgg tgiaccggga accggaatcg attgctggcc tgcgagctgg tgcgtacgag 3600
 gaaactgtct ctgaacagaa gaaggaggac gtctctgact acgagaactc ccagtacaag 3660
 gaggttcctag tccccctccc caacgagaag ctggccagag gtctgtcat gctggccgag 3720
 ctgtcttgca agggctctct ggccactggc gactactcca agcagacat tgagcttgcc 3780

cgatccgacc ccgagtttgt ggttggcttc attgccaga accgacctaa gggcgactct 3840
 gaggactggc ttattctgac ccccggggtg ggtcttgacg acaagggaga cgctctcgga 3900
 cagcagtacc gaactgttga ggatgtcatg tctaccgaa cggatatcat aattgtcggc 3960
 cgaggctgtg acggccagaa ccgagatcct attgaggagg ccaagcgata ccagaaggct 4020
 ggctgggagg cttaccagaa gattaactgt tagaggtag actatggata tgtaatttaa 4080
 ctgtgtatat agagagcgtg caagtatgga gcgcttggtc agcttgatg atggtcagac 4140
 gacctgtctg atcgagtatg tatgatactg cacaacctgt gtatccgcat gatctgtcca 4200

atggggcatg ttgttgtgt tctcgatacg gagatgctgg gtacagtgtc aatacgttga 4260
 actacttata cttatatgag gctcgaagaa agctgacttg tgtatgactt aat 4313

<210> 83

<211> 13565

<212> DNA

<213> Artificial Sequence

<220><223> Plasmid pZKL3-9DP9N

<400> 83

gtacggattg tgiatgtccc tgtacctgca tcttgatgga gagagctccg gaaagcggat 60
 caggagctgt ccaattttaa ttttataaca tggaaacgag tccttggagc tagaagacca 120
 ttttttcaac tgcctatcgc actatattta tctactccaa aaccgactgc ttccaagaa 180
 tcttcagcca aggtcttcaa agtaaccct cgcttccga cacttaattg aaaccttaga 240

tgcagtcact gcgagtgaag tggactctaa catctccaac atagcgacga tattgcgagg 300
 gtttgaatat aactaagatg catgatccat tacatttgta gaaatatcat aaacaacgaa 360
 gcacatagac agaattgtgt tggttgttac atctgaagcc gaggtaccga tgtcattttc 420
 agctgtcact gcagagacag gggatatgtca catttgaaga tcatacaacc gacgtttatg 480
 aaaaccagag atatagagaa tgtattgacg gttgtggcta tgtcataagt gcagtgaagt 540
 gcagtgatta taggtatagt acacttactg tagctacaag tacatactgc tacagtaata 600

ctcatgtatg caaaccgat tctgtgtcta cagaaggcga tacggaagag tcaatctctt 660

atgtagagcc atttctataa tcgaaggggc cttgtaattt ccaaacgagt aattgagtaa 720

ttgaagagca tcgtagacat tacttatcat gtattgtgag agggaggaga tgcagctgta 780

gctactgcac atactgtact cgcccatgca gggataatgc atagcgagac ttggcagtag 840

gtgacagttg ctagctgcta cttgtagtcg ggtgggtgat agcatggcgc gccagctgca 900

ttaatgaatc ggccaacgag cggggagagg cggtttgcgt attgggcgct cttccgcttc 960

ctcgctcact gactcgctgc gctcggtcgt tcggctgcgg cgagcggtat cagctcactc 1020

aaaggcggta atacggttat ccacagaatc aggggataac gcaggaaaga acatgtgagc 1080

aaaaggccag caaaaggcca ggaaccgtaa aaaggccgag ttgctggcgt tttccatag 1140

gctccgccc cctgacgagc atcacaaaaa tcgacgtca agtcagaggt ggcgaaaccc 1200

gacaggacta taaagatacc aggcgtttcc ccttggaagc tccctcgtgc gctctcctgt 1260

tccgaccctg ccgcttaccg gatactgtc cgcctttctc ccttcgggaa gcgtggcgt 1320

ttctcatagc tcacgtgta ggtatctcag ttcggtgtag gtcgttcgt ccaagctggg 1380

ctgtgtgca gaacccccg ttcagccccg ccgctgcgc ttatccggt actatcgtct 1440

tgagtccaac ccgtaagac acgacttat gccactggca gcagccactg gtaacaggat 1500

tagcagagcg aggtatgtag gcggtgtac agagttcttg aagtgggtgc ctaactacgg 1560

ctacactaga agaacagtat ttggtatctg cgtctgtctg aagccagtta ccttcgaaa 1620

aagagttggt agctcttgat ccgcaaaaca aaccaccgt ggtagcgtg gttttttgt 1680

ttgcaagcag cagattacgc gcagaaaaaa aggatctcaa gaagatcctt tgatctttc 1740

tacggggtct gacgtcagt ggaacgaaaa ctacgttaa gggattttgg tcatgagatt 1800

atcaaaaagg atcttcacct agatcctttt aaattaaaaa tgaagtttta aatcaatcta 1860

aagtatatat gagttaaactt ggtctgacag ttaccaatgc ttaatcagt aggcacctat 1920

ctcagcgatc tgtctatttc gttcatccat agttgcctga ctccccgtcg ttagataac 1980

tacgatacgg gagggcttac catctggccc cagtgtgca atgataccgc gagaccacg 2040

ctcacggct ccagatttat cagcaataaa ccagccagcc ggaagggccg agcgcagaag 2100

tggctctgca actttatccg cctccatcca gtctattaat tgttgccggg aagctagagt 2160

aagtagttcg ccagtttaata gtttcgcaa cgttgttgcc attgtctacag gcatcgttgt 2220

gtcacgtcg tcgttttgta tggcttcatt cagctccggt tcccaacgat caaggcgagt 2280

tacatgatcc ccatgttgt gcaaaaaagc ggttagctcc ttcggtctc cgatcgttgt 2340

cagaagtaag ttggccgcag tgttatcact catggttatg gcagcactgc ataattctct	2400
tactgtcatg ccatccgtaa gatgcttttc tgtgactggt gagtactcaa ccaagtcatt	2460
ctgagaatag tgtatgcggc gaccgagttg ctcttgcccg gcgtcaatac gggataatac	2520
cgcgccacat agcagaactt taaaagtgtc catcattgga aaacgttctt cggggcgaaa	2580
actctcaagg atcttaccgc tgttgagatc cagttcgatg taaccactc gtgcacccaa	2640
ctgatcttca gcatctttta ctttcaccag cgtttctggg tgagcaaaaa caggaaggca	2700
aaatgccgca aaaaagggaa taaggcgac acggaaatgt tgaatactca tactcttcct	2760
ttttcaatat tattgaagca tttatcaggg ttattgtctc atgagcggat acatatttga	2820
atgtatttag aaaaataaac aaataggggt tccgcgcaca tttccccgaa aagtgccacc	2880
tgatgcggtg tgaataaccg cacagatgcg taaggagaaa ataccgcac aggaaattgt	2940
aagcgttaat atttgttaa aattcgctt aaattttgt taaatcagct ctttttttaa	3000
ccaataggcc gaaatcgga aaatccctta taaatcaaaa gaatagaccg agatagggtt	3060
gagtgttgtt ccagtttga acaagagtcc actattaaag aacgtggact ccaacgtcaa	3120
agggcgaaaa accgtctatc agggcgatgg cccactacgt gaaccatcac cctaataag	3180
ttttttgggg tcgaggtgcc gtaaagcact aaatcggaac cctaaaggga gccccgatt	3240
tagagcttga cggggaaagc cggcgaacgt ggcgagaaag gaagggaaga aagcgaaagg	3300
agcggcgctt agggcgctgg caagtgtagc ggtcacgctg cgcgtaacca ccacacccgc	3360
cgcgcttaat gcgccgtac agggcgctc cattcgccat tcaggctgcg caactgttgg	3420
gaaggcgcat cgggtcgggc ctcttcgcta ttacgccagc tggcgaaaagg gggatgtgct	3480
gcaaggcgat taagtgggt aacgccaggg ttttccagt cacgacgttg taaaacgacg	3540
gccagtgaat tgtaatacga ctactatag ggcgaaattg gcccgacgtc gcatgcagga	3600
atagacatct tcaataggag cattaatacc tgtgggatca ctgatgtaaa ctctcccag	3660
agtatgtgaa taaccagcgg gccatccaac aaagaagtcg ttccagttag tgactcggta	3720
catccgtctt tcggggttga tggtaagtcc gtcgtctcct tgcttaaaga acagagcgtc	3780
cacgtagtct gcaaaagcct tgtttccaag tcgaggtgc ccatagtga ttagcgttgg	3840
atcatatcca agattcttca ggttgatgcc catgaataga gcagtgacag ctcttagaga	3900
gtggccagtt acgatcaatt ttagtcagt gttgtttcca aggaagtcga ccagacgatc	3960
ctgtacgttc accatagtct ctctgtatgc ctctgaaag ccatcatgaa ctgggcagcc	4020
aggacaattg atactggcag aagggtttgt ggagtttatg tcagttagtg taagaggagg	4080
gatactggtc atgtagggtt gttggatcgt ttggatgtca gtaatagcgt ctgcaatgga	4140
gaaagtgcct cggaatacaa tatacttttc ctttttggtg tgatcgtggg ccaaaaatcc	4200

agtaactgaa gtcgagaaga aatttcctcc aaactggtag tcaagagtca catcgggaaa 4260
 atgagcgcaa gagtttccac aggtaaaatc gctctgcagg gcaaatgggc caggggctct 4320
 gacacaatag gccacgttag atagccatcc gtacttgaga acaaagtcgt atgtctcctg 4380
 ggtgatagga gccgttaatt aagttgcgac acatgtcttg atagtatctt gaattctctc 4440

 tcttgagctt ttccataaca agttcttctg cctccaggaa gtccatgggt ggtttgatca 4500
 tggttttggt gtagtggtag tgcagtgggt gtattgtgac tggggatgta gttgagaata 4560
 agtcatacac aagtcagctt tcttcgagcc tcatataagt ataagtagtt caacgtatta 4620
 gcactgtacc cagcatctcc gtatcgagaa acacaacaac atgccccatt ggacagatca 4680
 tgcggataca caggttgtgc agtatcatac atactcgatc agacaggctg tctgaccatc 4740
 atacaagctg aacaagcgt ccatacttgc acgctctcta tatacacagt taaattacat 4800
 atccatagtc taacctctaa cagttaatct tctggtaagc ctcccagcca gccttctggt 4860

 atcgcttggc ctctcaata ggatctcgg tctggccgta cagacctcg cgcacaatta 4920
 tgatatccgt tccggtagac atgacatcct caacagttcg gtactgctgt ccgagagcgt 4980
 ctcccttgte gtcaagacce accccggggg tcagaataag ccagtcctca gagtcgccct 5040
 taggtcgggt ctgggcaatg aagccaacca caaactcggg gtcggatcgg gcaagctcaa 5100
 tggctctgct ggagtactcg ccagtggcca gagagccctt gcaagacagc tcggccagca 5160
 tgagcagacc tctggccagc ttctcgttgg gagaggggac taggaactcc ttgtactggg 5220
 agttctcgta gtcagagacg tctctcttct tctgttcaga gacagtttcc tcggcaccag 5280

 ctgcaggcc agcaatgatt ccggttccgg gtacacgtg ggcgttggtg atatcggacc 5340
 actcggcgat tcggtgacac cggtagtgggt gcttgacagt gttgccaata tctgcgaact 5400
 ttctgtctc gaacaggaag aaacctgtct taagagcaag ttcttgagg gggagcacag 5460
 tgccggcgta ggtgaagtcg tcaatgatgt cgatatgggt tttgatcatg cacacataag 5520
 gtccgacctt atcggcaagc tcaatgagct ccttgggtgg ggtaacatcc agagaagcac 5580
 acaggttgggt tttcttggt gccacgagct tgagcactcg agcggcaaag gcggacttgt 5640
 ggacgttagc tcgagcttcg taggagggca ttttgggtgg gaagaggaga ctgaaataaa 5700

 tttagtctgc agaacttttt atcggaacct tatctggggc agtgaagtat atgttatggt 5760
 aatagttagc agttagtga acttatagat agactggact atacggctat cggtcctaat 5820
 tagaaagaac gtcaatggct ctctgggcgt cgcctttgcc gacaaaaatg tgatcatgat 5880
 gaaagccagc aatgacgttg cagctgatat tgtgtcggc caaccgcgc gaaaacgcag 5940
 ctgtcagacc cacagctcc aacgaagaat gtatcgtaa agtgatcaa gcacactcat 6000
 agttggagtc gtactccaaa ggcggaatg acgagtcaga cagatctcg tcgacctttt 6060

ccttggaac caccaccgtc agcccttctg actcacgtat tgtagccacc gacacaggca 6120

 acagtccttg gatagcagaa tatgtcttgt cgggccatct ctcaccaact ttaggcgtca 6180
 agtgaatgtt gcagaagaag tatgtgcctt cattgagaat cgggtgttgc gatttcaata 6240
 aagtcttgag atcagtttgg ccagtcattg tgtggggggg aattggattg agttatcgcc 6300
 tacagtctgt acaggtatag tcgctgcca ctttatactt ttgtattccg ctgcacttga 6360
 agcaatgtcg ttaccacaaa gtgagaatgc tccacagaac acaccccagg gtatggttga 6420
 gcaaaaaata aacctccga tacggggaat cgaaccccg tctccacggt tctcaagaag 6480
 tattcttgat gagagcgtat cgatgggtta tctgctgtg tctgtgtgt gtgtgtgtt 6540

 tggcgctcat tgttgcgtta tgcagcgtac accacaatat tggaagctta ttagcctttc 6600
 tattttttcg ttgcaaggc ttaacaacat tctgttggag agggatgggg atatggaggc 6660
 cgctggaggg agtcggagag gcgttttga gcggcttggc ctggcgcca gctcgcaaaa 6720
 cgcacctagg accctttggc acgccgaaat gtgccacttt tcagtcctag aacgccttac 6780
 ctacgtcatt ccatgcgtgc atgtttgcgc cttttttccc ttgcccttga tcgccacaca 6840
 gtacagtga ctgtacagtg gaggttttgg gggggtctta gatgggagct aaaagcggcc 6900
 tagcgtaca ctatgggat tgtatggagt ggcatggagc ctatgggag cctgacagga 6960

 cgcacgaccg gctagcccg gacagacgat ggggtggctcc tgtgtccac cgcgtacaaa 7020
 tgtttgggcc aaagtcttgt cagccttgc tgcgaaccta attccaatt ttgtcacttc 7080
 gcaccccat tgatcgagcc ctaaccctg cccatcaggc aatccaatta agctcgcttc 7140
 gtctgccttg tttagtttgg ctctgcccg tttcggcgtc cacttgaca aacacaaaca 7200
 agcattatat ataaggctcg tctctccctc ccaaccacac tcactttttt gcccgcttc 7260
 ccttgctaac aaaaagtca agaacacaaa caaccacccc aacccctta cacacaagac 7320
 atatctacag caatggccat ggccaaaagc aaacgacggt cggaggctgt ggaagagcac 7380

 gtgaccggct cggacgaggg cttgaccgat acttcgggtc acgtgagccc tgccccaag 7440
 aagcagaaga actcgagat tcatctacc acccaggctg cccagcagtt ggatcgggag 7500
 cgcaaggagg agtatctgga ctgctgac gacaacaagg actatctcaa gtaccgtcct 7560
 cgaggttga agtcaacaa cccgcctacc gaccgacctg tgcgaatcta cgccgatgga 7620
 gtgtttgatt tgttccatc gggacacatg cgtcagctgg agcagtcaa gaaggccttc 7680
 cccaacgcag tgttgattgt gggcattccc agcgacaagg agaccacaa gcggaaggga 7740
 ttgaccgtgc tgagtacgt ccagcggtac gagacggtgc gacactgcaa gtgggtggac 7800

gaggtgggtgg aggatgctcc ctgggtgtgtc accatggact ttctggaaaa acacaaaatc 7860

gactacgtgg cccatgacga tctgccctac gcttccggca acgacgatga tatctacaag 7920

cccatcaagg agaagggcat gtttctggcc acccagcgaa ccgagggcat ttccacctcg 7980

gacatcatca ccaagattat ccgagactac gacaagtatt taatgcgaaa ctttgcccgg 8040

ggtgctaacc gaaaggatct caacgtctcg tggctcaaga agaacgagct ggacttcaag 8100

cgtcatgtgg ccgagttccg aaactcgttc aagcgaaaga aggtcggtaa ggatctctac 8160

ggcgagattc gcggtctgct gcagaatgtg ctcatTTTgga acggcgacaa ctccggcact 8220

tccactcccc agcgaaagac gctgcagacc aacgccaaga agatgtacat gaacgtgctc 8280

aagactctgc aggtctctga cgctgttgac gtggactcct cggagaacgt gtctgagaac 8340

gtcactgatg aggaggagga agacgacgac gaggttgatg aggacgaaga agccgacgac 8400

gacgacgaag acgacgaaga cgaggaagac gacgagtagg cggccgcatt gatgattgga 8460

aacacacaca tgggttatat ctaggtgaga gttagtTgga cagttatata ttaaTcagc 8520

tatgccaacg gtaactTcat tcatgtcaac gaggaaccag tgactgcaag taatatagaa 8580

tttgaccacc ttgccattct cttgcactcc ttTactatat ctcatTTTatt tcttatatac 8640

aaatcacttc ttcttcccag catcgagctc ggaaacctca tgagcaataa catcgtggat 8700

ctcgtcaata gagggttttt tggactcctt gctgttggcc acctTgtcct tgctgtTtaa 8760

acacgcagta ggatgtcctg cacgggtctt ttTgtgggt gtggagaaag gggTgctTgg 8820

agatggaagc cggtagaacc gggctgcttg tgcttggaga tggaaaccgg tagaaccggg 8880

ctgctTgggg ggattTgggg ccgctgggct ccaaagaggg gtaggcattt cgtTggggTt 8940

acgtaattgc ggcattTggg tcctgcgcgc atgtccatt ggtcagaatt agtccggata 9000

ggagacttat cagccaatca cagcggcgga tccacctgta ggtTgggttg ggtgggagca 9060

ccccccaca gagtagagtc aaacagcagc agcaacatga tagttggggg tgtgcgtgtt 9120

aaaggaaaaa aaagaagctt gggTtatatt cccgtctat ttagaggttg cgggatagac 9180

gccgacggag ggcaatggcg ctatggaacc ttgcggatat ccatacgccg cggcggactg 9240

cgtccgaacc agctccagca gcgtTttttc cgggccattg agccgactgc gaccccgcca 9300

acgtgtcttg gcccacgcac tcatgtcatg ttggtgttgg gagccactt tTaaTtagc 9360

acaaggcacc tagctcgag caaggtgtcc gaaccaaaga agcggtgca gtggtgcaaa 9420

cggggcgga acggcgggaa aaagccacgg gggcacgaat tgaggcacgc cctcgaattt 9480

gagacgagtc acggcccat tcgcccgcgc aatggctcgc caacggccgg tcttttgac 9540

cacatcaggt taccccaagc caaacctttg tgtTaaaaag cttaacatat tataccgaac 9600

gtaggttttg gcgggcttg tccgtctgtc caaggcaaca tttatataag ggtctgcac 9660

gccggctcaa ttgaatcttt tttcttcttc tcttctctat attcattctt gaattaaaca	9720
cacatcaaca tggccatcaa agtcggtatt aacggattcg ggcgaaatcg acgaattgtg	9780
agtaccatag aaggtgatgg aaacatgacc caacagaaac agatgacaag tgtcatcgac	9840
ccaccagagc ccaattgagc tcatactaac agtcgacaac ctgtcgaacc aattgatgac	9900
tccccgacaa tgtactaaca caggctcctgc ccatggtgaa aaacgtggac caagtggatc	9960
tctcgagggt cgacaccatt gcctccggcc gagatgtcaa ctacaaggtc aagtacacct	10020
ccggcggttaa gatgagccag ggcgctacg acgacaaggg ccgccacatt tccgagcagc	10080
ccttcacctg ggccaactgg caccagcaca tcaactggct caacttcatt ctggtgattg	10140
cgctgcctct gtcgtccttt gctgccgtc ccttcgtctc cttcaactgg aagaccgccg	10200
cgtttctgt cggtattac atgtgcaccg gtctcggtat caccgccggc taccaccgaa	10260
tgtgggcccc tcgagcctac aaggccgtc tgcctgttcg aatcatcctt gctctgtttg	10320
gaggaggagc tgtcgagggc tccatccgat ggtgggcctc gtctcaccga gtccaccacc	10380
gatggaccga ctccaacaag gaccttacg acgcccgaag gggattcttg ttctccact	10440
ttggctggat gctgcttggt cccaacccca agaacaaggg ccgaactgac atttctgacc	10500
tcaacaacga ctgggttgct cgactccagc acaagtacta cgtttacgtt ctgctcttca	10560
tggccattgt tctgcccacc ctgctctgtg gctttggctg gggcgactgg aaggagggtc	10620
ttgtctacgc cggtatcatg cgatacacct ttgtgcagca ggtgactttc tgtgtcaact	10680
cccttgcccc ctggattgga gacgagccct tcgacgaccg acgaactccc cgagaccacg	10740
ctcttacgc cctgggtcacc ttgagagagg gctaccacaa ctccaccac gaggctccct	10800
cggactaccg aaacgcctc atctggtacc agtacgacc caccaagtgg ctcatctgga	10860
ccctcaagca ggttggtctc gcctgggacc tccagacctt ctccagaac gccatcgagc	10920
agggtctcgt gcagcagcga cagaagaagc tggacaagtg gcgaacaac ctcaactggg	10980
glatcccat tgagcagctg cctgtcattg agtttgagga gttccaagag caggccaaga	11040
cccagatct ggttctcatt tctggcattg tccacgacgt gtctgccttt gtcgagcacc	11100
accttggtgg aaaggccctc attatgagcg ccgtcggcaa ggacgggtacc gctgtcttca	11160
acggagggtgt ctaccgacac tccaacgtg gccacaacct gcttgccacc atgcgagttt	11220
cggctattcg aggcggcatg gaggttgagg tgtggaagac tgcccagaac gaaaagaagg	11280
accagaacat tgtctccgat gagagtggaa accgaatcca ccgagctggt ctccaggcca	11340
cccgggtcga gaaccccggt atgtctggca tggtctgctta ggcgcccgca tgagaagata	11400
aatatataaa tacattgaga tattaatgc gctagattag agagcctcat actgctcgga	11460
gagaagccaa gacgagtact caaaggggat tacacatcc atatccacag acacaagctg	11520

gggaaagggtt ctatatacac ttccggaat accgtagttt ccgatgttat caatgggggc	11580
agccaggatt tcaggcactt cgggtgtctcg gggtgaaatg gcgttcttgg cctccatcaa	11640
gtcgtacat gtcttcattt gcctgtcaaa gtaaaacaga agcagatgaa gaatgaactt	11700
gaagtgaagg aatttaaata gttggagcaa gggagaaatg tagagtgtga aagactcact	11760
atggtccggg cttatctcga ccaatagcca aagtctggag tttctgagag aaaaaggcaa	11820
gatacgtatg taacaaagcg acgcatggta caataatacc ggaggcatgt atcatagaga	11880
gttagtgggt cgatgatggc actggtgcct ggtatgactt tatacggctg actacatatt	11940
tgctctcaga catacaatta cagtcaagca cttacccttg gacatctgta ggtaccccc	12000
ggccaagacg atctcagcgt gtcgtatgtc ggattggcgt agctccctcg ctctcaatt	12060
ggctcccatc tactttcttc tgcttgcta caccagcat gtctgctatg gctcgttttc	12120
gtgccttate taccctccca gtattaccaa ctctaaatga catgatgtga ttgggtctac	12180
actttcatat cagagataag gagtagcaca gttgcataaa aagcccaact ctaatcagct	12240
tcttcctttc ttgtaattag tacaagggtg attagcgaat tctggaagct tagttggccc	12300
taaaaaatc aaaaaagca aaaaacgaaa aacgaaaaac cacagttttg agaacaggga	12360
ggtaacgaag gatcgtatat atatatatat atatatatc ccacggatcc cgagaccggc	12420
ctttgattct tcctacaac caaccattct caccacccta attcacaacc atggaggctc	12480
tgaacgaaat cgctccatt ggccaggagg ttcttccaa ggtcgactat gtcagctct	12540
ggtctgatgc ctgcactgc gaggtgctgt acctctccat cgccttcgtc atcctgaagt	12600
tcacccttgg tctctcga cccaagggtc agtctcgaat gaagtttgtg ttcaccaact	12660
acaacctgct catgtccatc tactcgctgg gtccttctct ctctatggcc tacgccatgt	12720
acaccattgg tgtcatgtcc gacaactgcg agaaggcttt cgacaacaat gtcttccgaa	12780
tcaccactca gctgttctac ctacgaagt tctcgagta cattgactcc ttctatctgc	12840
ccctcatggg caagcctctg acctggttgc agttctttca ccatctcgga gtcctatgg	12900
acatgtggct gttctacaac taccgaaacg aagccgtttg gatctttgtg ctgctcaacg	12960
gcttcattca ctggatcatg tacggctact attggaccgg actgatcaag ctcaagttcc	13020
ctatgccaa gtccctgatt acttctatgc agatcattca gttcaacgtt ggcttctaca	13080
tcgtctggaa gtaccggaac attccctgct accgacaaga tggaaatgaga atgtttggct	13140
ggtttttcaa ctacttctac gttgttactg tctgtgtct gttctcaac ttctacgtgc	13200
agacctacat cgtccgaaag cacaaggag ccaaaaagat tcagtgcgc gccgcaagtg	13260

tggatgggga agtgagtgcc cggttctgtg tgcacaattg gcaatccaag atggatggat 13320
tcaacacagg gatatagcga gctacgtggt ggtgcgagga tatagcaacg gatatttatg 13380
tttgacactt gagaatgtac gatacaagca ctgtccaagt acaatactaa acatactgta 13440
catactcata ctctgacccg gcaacgggtt cacttgagtg cagtggctag tgctcttact 13500
cgtacagtgt gcaatactgc gtatcatagt ctttgatgta tatcgtattc attcatgtta 13560
gttgc 13565

<210> 84

<211> 9632

<212> DNA

<213> Artificial Sequence

<220><223> Plasmid pY187

<400> 84

ggccgcatga gaagataaat atataaatac attgagatat taaatgcgct agattagaga 60
gcctcatact gctcggagag aagccaagac gactactcaa aggggattac accatccata 120
tccacagaca caagctgggg aaaggttcta tatacacttt ccggaatacc gtagtttccg 180
atgttatcaa tgggggcagc caggatttca ggcaacttcg tgtctcgggg tgaaatggcg 240
ttcttggcct ccatcaagtc gtaccatgtc ttcatattgcc tgtcaaagta aaacagaagc 300
agatgaagaa tgaacttgaa gtgaaggaat ttaaattgaa cgaaactgaa atttgaccag 360
atattgtgtc cgcggtggag ctccagcttt tgttcccttt agtgagggtt aatttcgagc 420

ttggcgtaat catggtcata gctgtttcct gtgtgaaatt gttatccgct cacaagcttc 480
cacacaacgt acgagccgga agcataaagt gtaaagcctg ggggtcctaa tgagtgagct 540
aactcacatt aattgcgttg cgtcactgc ccgctttcca gtcgggaaac ctgtcgtgcc 600
agctgcatta atgaatcggc caacgcgcgg ggagaggcgg tttgcgtatt gggcgctctt 660
ccgcttcttc gctcactgac tcgtcgcgt cggtcgttcg gctgcggcga gcggtatcag 720
ctcactcaaa ggcggttaata cggttatcca cagaatcagg ggataacgca ggaaagaaca 780
tgtgagcaaa aggccagcaa aaggccagga accgtaaaaa ggccgcgttg ctggcgtttt 840

tccataggct ccgccccct gacgagcatc acaaaaatcg acgtcaagt cagaggtggc 900
gaaacccgac aggactataa agataccagg cgtttccccc tggaagctcc ctctgcgct 960
ctcctgttcc gaccctgccg cttaccgat acctgtccgc ctttctcct tcgggaagcg 1020
tggcgctttc tcatagctca cgctgtaggt atctcagttc ggtgtaggtc gttcgtccca 1080
agctgggctg tgtgcacgaa cccccgttc agcccagccg ctgcgcctta tccgtaact 1140
atcgtcttga gtccaacccg gtaagacacg acttatcgcc actggcagca gccactggta 1200

acaggattag cagagcgagg tatgtaggcg gtgctacaga gttcttgaag tggtagccta 1260

actacggcta cactagaagg acagtatttg gtatctgcgc tctgctgaag ccagttacct 1320

tcggaaaaag agttggtagc tcttgatccg gcaaaaaaac caccgctggt agcggtggtt 1380

tttttgtttg caagcagcag attacgcgca gaaaaaaagg atctcaagaa gatcctttga 1440

tcttttctac ggggtctgac gctcagtga acgaaaactc acgttaaggg attttggta 1500

tgagattatc aaaaaggatc ttcacctaga tccttttaaa ttaaaaatga agttttaaat 1560

caatctaaag tatatatgag taaacttggc ctgacagtta ccaatgctta atcagtgagg 1620

cacatatctc agcgatctgt ctatttcgtt catccatagt tgcctgactc cccgtcgtgt 1680

agataactac gatacgggag ggcttaccat ctggccccag tgctgcaatg ataccgcgag 1740

accacgctc accggctcca gatttatcag caataaacca gccagccgga agggccgagc 1800

gcagaagtgg tctgcaact ttatccgct ccatccagtc tattaattgt tgccggaag 1860

ctagagtaag tagttcgcca gttaatagtt tgcgcaacgt tgttgccatt gctacaggca 1920

tcgtggtgtc acgctcgtcg ttggtatgg cttcattcag ctccggttcc caacgatcaa 1980

ggcgagttac atgatecccc atgttgtagc aaaaagcggg tagctccttc ggtcctccga 2040

tcgttgcag aagtaagttg gccgcagtgt tatcactcat gggtatggca gcaactgcata 2100

attctcttac tgtcatgcca tccgtaagat gcttttctgt gactggtgag tactcaacca 2160

agtcattctg agaatagttg atgcggcgac cgagttgctc ttgcccggcg tcaatacggg 2220

ataataccgc gccacatagc agaactttaa aagtgtcat cattggaaaa cgttcttcgg 2280

ggcgaaaact ctcaaggatc ttaccgctgt tgagatccag ttcgatgtaa cccactcgtg 2340

cacccaactg atcttcagca tcttttactt tcaccagcgt ttcctgggtga gcaaaaacag 2400

gaaggcaaaa tgccgcaaaa aagggaataa gggcgacacg gaaatgttga atactcatac 2460

tcttcctttt tcaatattat tgaagcattt atcagggtta ttgtctcatg agcggataca 2520

tatttgaatg tatttagaaa aataaacaag taggggttcc gcgcacattt ccccgaaaag 2580

tgccacctga cgcgccctgt agcggcgcat taagcgcggc ggggtgtgtg gttacgcgca 2640

gcgtgaccgc tacacttgcc agcgccctag cgcccgtcc tttcgcttc ttccttcct 2700

ttctcgccac gttcgccggc tttccccgtc aagctctaaa tcgggggctc cctttagggt 2760

tccgatttag tgccttacgg cacctcgacc ccaaaaaact tgattagggt gatggttcac 2820

glagtgggcc atcgccctga tagacgggtt ttcgcccttt gacgttgag tccacgttct 2880

ttaatagtgg actcttgctc caaactggaa caacactcaa ccctatctcg gtctattctt 2940

ttgatttata agggattttg ccgatttcgg cctattgggt aaaaaatgag ctgatttaac	3000
aaaaatttaa cgcaatttt aacaaaatat taacgcttac aatttccatt cgccattcag	3060
gctgcgcaac tgttgggaag ggcgatcggg gggggcctct tcgctattac gccagctggc	3120
gaaaggggga tgtgctgcaa ggcgattaag ttgggtaacg ccagggtttt cccagtcacg	3180
acgttgtaaa acgacggcca gtgaattgta atacgactca ctataggcg aattgggtac	3240
cgggccctcc ttcaattcaa gttaattctt catctgcttc tgttttactt tgacaggcaa	3300
atgaagacat ggiacgactt gatggaggcc aagaacgcca tttcaccccg agacaccgaa	3360
gtgcctgaaa tcttggctgc cccattgat aacatcgga actacggtat tccgaaagt	3420
gtatatagaa ctttcccca gcttgtgtct gtggatatgg atggtgtaat cccctttgag	3480
tactgtcttt ggcttctctc cgagcagtat gaggtctctt aatctagcgc atttaatatc	3540
tcaatgtatt tatatattta tcttctcatg cggccgccta ctgagtcttc tggccagcgt	3600
agccaatctg gtcggcagtc tcgcacatgg ccttataagt ggtggccatg agagcgtcga	3660
cgttgctggc gtccagacct ttggctctca tggggctcag gacgtcgaca aggatctctc	3720
cggcgtccag cttgcctcgg ccgaaagaaa acattcggga ggtgttcttg accaccacgg	3780
ggacaatggg agcggccgac tggaccgcca ggtggaaaca gcccttcttg aagggcata	3840
tgccgacgtc cttggagtag gatcgggttc cctcggggaa catgaagacg ctctggccct	3900
tgcctccgtt gccgctctgg atggtcttga cggcgccgcc gaggtctgc acggacttgg	3960
ttcggtcctt tcggtccagg aagatgggtgc cggacagcgc catgaactgg cccagcagag	4020
ggtaccactt gagggccttt ttggcggtga cggagcagta ctgggggaag atgcaccga	4080
gcaccagaat gtccatttcg ttctgggtgt tggaaatcag aatgcatggt cggggcaggt	4140
gtttgggtgat gtcgtcaatg ggcttgggtg tgttggcgac ctggatcttg gaggcacca	4200
gaccggcgtt tccggtcacc tgccggctac gcagcttgat gctgataccc aggaagatgc	4260
gcaccgagta gtagaaggct ctggcaacgg tcatttgggc caggccctgc ttgccgatgg	4320
cggtgagaat ggtggacgcc agcacgccgt aggtggcaca ggcggaag atgaccacgg	4380
cgatggcggc gcggacgtag aagacgagct tggatgcaac ggacatgagc tgggttagtt	4440
tgtgtagaga gtgtgtgttg ctagcgactt tcggattgtg tcattacaca aaacgcgtcg	4500
tctcgacact gatcttctcg tggatactca cggctcggaa ttctgtgatg ttagtttag	4560
atttcgaatc tgtggggaaa gaaaggaaaa aagagactgg caaccgattg ggagagccac	4620
tgtttatata taccctagac aagcccccg cttgtaagat gttggtcaat gtaaaccagt	4680
attaaggttg gcaagtgcag gagaagcaag gtgtgggtac cgagcaatgg aatgtgcgg	4740
aaggcaaaaa aatgaggcca cggcctattg tcggggctat atccaggggg cgattgaagt	4800

acactaacat gacatgtgtc cacagaccct caatctggcc tgatgagcca aatccatacg	4860
cgctttcgca gctctaaagg ctataacaag tcacaccacc ctgctcgacc tcagcgccct	4920
cactttttgt taagacaaac tgtacacgct gttccagcgt tttctgcctg cacctgggtg	4980
gacatttggg gcaacctaaa gtgctcggaa cctctgtggg gtccagatca gcgcagcagt	5040
tccgaggtag ttttgaggcc cttagatgat gcaatgggtg cagtcgctgg atcacgagtc	5100
ttaatggcag tattcgttct tatttgtgcc attgagcccc gttatcctcg tatcttctac	5160
cccccatccc atccctttgt tgggtcaacc ctaccattt attgttgggt gcagcccaac	5220
cgacgtggag agcttggcct ggccatataa aaaggccccc ccctagtggc aatggcagaa	5280
agtcagctgt gagtgttga atttgtcatc taggcggcct ggccgtcttc tgcattctac	5340
ttgggtgtaa tattgggacg tgttcggaaa tcaacggatg ctcaatcgat ttcgacagta	5400
attaattaag tcatacacia gtcagctttc ttcgagcctc atataagtat aagtagttca	5460
acgtattagc actgtacca gcatctccgt atcgagaaac acaacaacat gccccattgg	5520
acagatcatg cggatacaca ggttgtgcag tatcatacat actcgatcag acaggtcgtc	5580
tgaccatcat acaagctgaa caagcgctcc atacttgca gctctctata tacacagtta	5640
aattacatat ccatagtcta acctctaaca gttaatcttc tggtaagcct ccagccagc	5700
cttctgggat cgcttggcct cctcaatagg atctcggttc tggccgtaca gacctcgcc	5760
gacaattatg ataccgttc cggtagacat gacatcctca acagttcggg actgctgtcc	5820
gagagcgtct ccctgtcgt caagaccac cccgggggtc agaataagcc agtcctcaga	5880
gtcgccctta ggtcggttct gggcaatgaa gccaaccaca aactcggggt cggatcgggc	5940
aagctcaatg gtctgcttgg agtactcgcc agtggccaga gagcccttgc aagacagctc	6000
ggccagcatg agcagacctc tggccagctt ctggttggga gaggggacta ggaactcctt	6060
gtactgggag ttctcgtagt cagagacgtc ctctttcttc tgttcagaga cagtttcttc	6120
ggcaccagct cgcaggccag caatgattcc ggttcggggt acaccgtggg cgttggatg	6180
atcgaccac tcggcgattc ggtgacaccg gtactggtgc ttgacagtgt tgccaatata	6240
tgccaacttt ctgtcctcga acaggaagaa accgtgctta agagcaagtt ccttgagggg	6300
gagcacagtg cggcgtagg tgaagtcgtc aatgatgtcg atatgggttt tgatcatgca	6360
cacataaggt ccgaccttat cggcaagctc aatgagctcc ttggtggtgg taacatccag	6420
agaagcacac aggttggttt tcttggctgc cagagcttg agcactcgag cggcaaaggc	6480
ggacttgtgg acgttagctc gagcttcgta ggagggcatt ttggtggtga agaggagact	6540
gaaataaatt tagtctgcag aactttttat cggaacctta tctggggcag tgaagtatat	6600
gttatggtaa tagttacgag ttagttgaac ttatagatag actggactat acggctatcg	6660

gtccaaatta gaaagaacgt caatggctct ctgggcgtcg cctttgccga caaaaatgtg	6720
atcatgatga aagccagcaa tgacgttgca gctgatattg ttgtcggcca accgcgccga	6780
aaacgcagct gtcagaccca cagcctccaa cgaagaatgt atcgtcaaag tgatccaagc	6840
acactcatag ttggagtcgt actccaaagg cggcaatgac gagtcagaca gatactcgtc	6900
gaccgtacga tagttagtag acaacaatcg atagtggag caaggagaa atgtagagt	6960
tgaaagactc actatgggcc gggcttatct cgaccaatag ccaaagtctg gagtttctga	7020
gagaaaaagg caagatacgt atgtaacaaa gcgacgcatg gtacaataat accggaggca	7080
tgtatcatag agagttagtg gttcgatgat ggcaactgtg cctggtatga ctttatcgg	7140
ctgactacat atttgtctc agacatacaa ttacagtcaa gcacttacc ttggacatct	7200
gtaggtaccc cccggccaag acgatctcag cgtgtcgtat gtcggattgg cgtagctccc	7260
tcgtctgca attggtccc atctactttc ttctgcttgg ctacaccag catgtctgt	7320
atggctcgtt ttctgtcctt atctatctc ccagtattac caactctaaa tgacatgat	7380
tgattgggtc tacactttca tatcagagat aaggagtagc acagtgcac aaaaagccca	7440
actctaata gcttcttctt ttcttgtaat tagtacaag gtgattagcg aaatctggaa	7500
gcttagttgg ccctaaaaaa atcaaaaaaa gcaaaaaacg aaaaacgaaa aaccacagtt	7560
ttgagaacag ggaggaacg aaggatcgta tatatatata tatatatata taccacgga	7620
tcccagagacc ggcctttgat tcttcctac aaccaacat tctcaccacc ctaattcaca	7680
accatgacac aacctgtgaa tcggaaggcg actgtcgagc gggtcgagcc agcagtggag	7740
gtggctgact ccgagtcga ggccaagacc gacgtccacg ttaccacca tcatcaccac	7800
cacaagcgaa aatccgtcaa gggcaagatt ctcaacttct tcaccgaag tcgacgtatc	7860
accttctgc tcggcgccgt ggtcgggtgt atagccgcgg gatactacgc tgcgccaccg	7920
gagctcagca ttgatatcga tgctcttctc ggcgacttgc cctcgttcga ctttgacgt	7980
ctatctctcg acaacttgc catggacagt gtgtcggact ttgtacaaga catgaaatcg	8040
cgttttccga ccaagattct gcaggaggcg gccaaagatc agaagcacca gaaaagcgaa	8100
cagaaggctg ccccttttgc tgtgggcaag gctatgaaga gcgagggact caacgccaag	8160
taccgggtgg tgctggtgcc cggcgtcatc tccacgggac tggagagctg gtccctggag	8220
ggaaccgagg agtgtccac cgagtgcac ttcagaaagc gaattgtggg ctccctggtac	8280
atgatccgag tcatgctgct ggacaagtac tgctggctgc agaacctgat gctggacaca	8340
gagaccggtc tagaccctcc ccatttcaag ctgcgagccg cccagggtt tgccctccgc	8400

gacttcttta tggcaggcta ctggctgtgg aacaagctgc tgcagaacct ggctgttatt 8460
ggatacgata cggatacaat gtctctgcgc gcgtacgact ggagactgtc ctacctgat 8520
ttggagcacc gagacggata cttctccaag ctcaaagctt caatcgaaga gactaagcgt 8580
atgacagggtg agaagacagt tctgacgggc cattccatgg gctcccaggt catcttctac 8640
ttcatgaagt gggctgaggc cgagggatat ggaggaggag gtcccaactg ggtcaatgac 8700
catattgaat cctttgtcga catttccggc tccatgctgg gtactccaa gacctggtt 8760
gctcttctgt ctggagaaat gaaggatacc gtgcagctga acgcgatggc tgtgtatgga 8820

ctggagcagt tcttctctcg acgagagcga gccgatctgc tgcgaacatg gggaggaatt 8880
gcttccatga ttccaaggg tggtaaggct atctgggggtg atcattctgg agccctgat 8940
gacgagcccg gccagaatgt cacccttggc aacttcatca agttcaagga gtccttgacc 9000
gagtactctg ctaagaacct caccatggat gaaaccgttg acttctgtga ttctcagtct 9060
cccagtggtt ttgtgaaccg aaccgagggt gcttactcct ttggaattgc caagactcga 9120
aagcaggttg agcagaatga gaagcgacct tctacctgga gcaaccctct ggaagctgct 9180
ctcccaatg ccccgatct caagatctac tgcttctatg gagtccgtaa ggataccgag 9240

cgagcctact actaccagga tgagcccaat cccgagcaga ccaacttgaa cgtcagatc 9300
gctggaaacg accctgatgg tgtgcttatg ggtcagggcg atggaaccgt ctcccttgtg 9360
acccatacca tgtgtcaccg atggaaggac gagaattcca agttcaacc tggtaacgcc 9420
caggtcaagg ttgtggagat gttgcaccag cctgatcgac ttgatattcg aggcggtgct 9480
cagactgccg agcatgtgga cattctgggg cgttctgagt tgaacgagat ggttctgaag 9540
gtggctagtg gaaagggaaa tgagattgaa gagagagtca tctccaacat tgatgagtgg 9600
gtgtggaaga ttgatctcgg cagcaattag gc 9632

<210> 85

<211> 15262

<212> DNA

<213> Artificial Sequence

<220><223> Plasmid pZK16-ML8N

<400> 85

gtacgcattc ttgacttgct cccagttgc cctttttggt cataactcag ccccgactct 60
ctcccaaag aggagatgga cttacaatag atacgagaca acagatattt tcggccacct 120
gggtgcaattg gtctgtttta ttaccccaaa taatatccac agcgtccgaa tcagctcgac 180
aagctctttc tcagcatact gatcttacc cataccccc agacagcggg gccatatcca 240
taccgcact acctgagaca gtaaacctag ggcaaggagc ggtcttggtc catcgcgacg 300

ctgtaagctt tgcctttgcg ttttgtgac caccacacgg cagtcaaggt ctcggttgaa 360

agatgggtct ggcaatgtgg attgctgaat cattgacata taaatacgag tgatcctgcc 420

atttcccaat ctttagaaat catcttttct tcacacatac aaaaaacaca catcacaatg 480

cagttctccg ccgccgccgt ccttgccctc gccgccttcg tctctgccca ggccgcctcc 540

gctgcttctg aagctgggtg tgccgttgct tccggtgcct ccggtgccgc tgccgccgct 600

tcttctgctg tcattggctc cgagctcag gccgcttctt ctgctgcctc ttctgctgct 660

gagtctgctt ctccgctgc ctctgagtc gccctctccg cttccgcctc cgagtccaag 720

gcctctgagt ccgcttccaa gtccgcctcc aaggttctg agtctgcttc caaggccgag 780

tcttctgctg ccaaggtctc ctccgctgct ggttctgctt cttccgccgc tgcctctgcc 840

accaagtctg gtgatgcctc cggtcttcc aagccttccg gctctgcctc ctctaccggt 900

gacaagtccg gtgctggtat tgttgccccc gctgttggtg ccatcggtgt cgggtgctctt 960

gctcagctca tctaagcgtt ctttgtacaa ctaccactt aatatttata ttatgatata 1020

gactgcaata ctattgcatg tgtaatgctt cgtttgaaag tcgtagtccg gccattcatg 1080

gcagtgtatg gcagtattgc aaccaattgg ctaataaca tatacaaagt ggcaattctg 1140

tgttcgaacc atcacacaac ccaacatatt tgacattgct acaagtattc gacttgttct 1200

tggccgcac gacatgcatt gctagcagca cacttgcaag agcttgcca atggttaggt 1260

tgacctttaa ccagtaaac gcaatatagg ttgccaatta acaacaaca cagcccagaa 1320

ctggtgtaaa gggcttggtt tttgtgcaat caactcattg actcacctcc cccaatgctc 1380

catgaacca tgaatccata gtacctccat cccaattcag agaggctcag aaacacgcaa 1440

ttggcaaaag taacaaagg tctcgaacc ttaaggggtg cagtgttgca aagttaaaat 1500

tacactgttg gatggtgttc cactcaact gtggatggat cagacggttg cactaaacca 1560

agattcaata tgctctaca gtagcgtgcc atgaggtgta taatatcaca catgatcaga 1620

gttaaggcct ggaaaaaag tgtccgcggc tatgtaacac tgctgaatac agtgtgtgcc 1680

gtttccacgg actggaatga atgcactcaa cagtttggtg gctaacttcc ccaactattg 1740

acttcacttt actgtagcta ctctaagtac gataatgaca tcatgaatca ccaactttat 1800

cttccccacc gaaagtaaga gatgaaagag aggctaacag cttttgtaag aaatcacaa 1860

aatctactta ccgtattaca aaagtacaag tatcagtaac caggcgcgcc agctgcatta 1920

atgaatcggc caacgcgcgg ggagaggcgg ttgctgattt gggcgctctt ccgcttctc 1980

gctcactgac tcgtgcgct cggtcgttcg gctgcgcga gcggtatcag ctactcaaa 2040

ggcggtaata cggttatcca cagaatcagg ggataacgca ggaaagaaca tgtgagcaaa	2100
aggccagcaa aaggccagga accgtaaaaa ggccgcgttg ctggcgtttt tccataggct	2160
ccgccccct gacgagcatc acaaaaatcg acgtcgaagt cagagggtggc gaaacccgac	2220
aggactataa agataaccagg cgtttcccc tggaaagctcc ctctgtcgct ctctgttcc	2280
gacctgccc cttaccggat acctgtccgc ctttctccct tcgggaagcg tggcgctttc	2340
tcatagctca cgtgttaggt atctcagttc ggtgtaggtc gttcgctcca agctgggctg	2400
tgtgcacgaa cccccgttc agcccgaccg ctgcgcctta tccggtaact atcgtcttga	2460
gtccaacccg gtaagacacg acttatcgcc actggcagca gccactggta acaggattag	2520
cagagcgagg tatgtaggcg gtgctacaga gttcttgaag tggtagccta actacggcta	2580
cactagaaga acagtatttg gtatctgcgc tctgtgaag ccagttacct tcggaaaaag	2640
agttggtagc tcttgatccg gcaaacaaac caccgctggt agcgggtggt tttttgtttg	2700
caagcagcag attacgcgca gaaaaaaagg atctcaagaa gatcctttga tcttttctac	2760
ggggctgac gctcagtga acgaaaactc acgttaaggg attttgtca tgagattatc	2820
aaaaaggatc ttacactaga tctttttaa ttaaaaatga agttttaaat caatctaaag	2880
tatatatgag taaacttggc ctgacagtta ccaatgctta atcagtgagg cacctatctc	2940
agcgatctgt ctatttcgtt catccatagt tgcctgactc cccgtcgtgt agataactac	3000
gatacgggag ggcttaccat ctggccccag tcttgcaatg ataccgcgag acccacgctc	3060
accggctcca gatttatcag caataacca gccagccgga agggccgagc gcagaagtgg	3120
tcttgcgaact ttatccgct ccatccagtc tattaattgt tgccgggaag ctagagtaag	3180
tagttcgcca gttaatagtt tgcgcaacgt tgttgccatt gctacaggca tcgtggtgtc	3240
acgtcgtcgc ttgggtatgg cttcattcag ctccggttcc caacgatcaa ggcgagttac	3300
atgatcccc atgttgtca aaaaagcggc tagctccttc ggtcctccga tcgttgtcag	3360
aagtaagttg gccgcagtgt tatcactcat ggttatggca gcaactgcata attctcttac	3420
tgtcatgcca tccgtaagat gcttttctgt gactggtgag tactcaacca agtcattctg	3480
agaatagtgt atgcggcgac cgagttgtc ttgcccggcg tcaatacggg ataataccgc	3540
gccacatagc agaactttaa aagtgtcat catttgaaaa cgttcttcgg ggcgaaaact	3600
ctcaaggatc ttaccgtgt tgagatccag ttcgatgtaa cccactcgt caccctaactg	3660
atcttcagca tcttttactt tcaccagcgt ttttgggtga gcaaaaacag gaaggcaaaa	3720
tgccgaaaa aagggaataa gggcgacacg gaaatgttga atactcatal tcttctttt	3780
tcaatattat tgaagcattt atcagggtta ttgtctcatg agcggataca tatitgaatg	3840
tatttagaaa aataaaciaa taggggttcc gcgcacattt cccgaaaag tgccacctga	3900

tcgggtgtga aataccgcac agatgcgtaa ggagaaaata ccgcatcagg aaattgttaag	3960
cgtaataatt ttgttaaaat tcgcgttaaa tttttgttaa atcagctcat tttttaacca	4020
ataggccgaa atcggcaaaa tcccttataa atcaaaagaa tagaccgaga tagggttgag	4080
tgttgttcca gtttgaaca agagtcact attaaagaac gtggactcca acgtcaaagg	4140
gcgaaaaacc gtctatcagg gcgatggccc actacgtgaa ccatcacct aatcaagttt	4200
tttggggtcg aggtgccgta aagcactaaa tcggaaccct aaaggagacc cccgatttag	4260
agcttgacgg ggaagccgg cgaacgtggc gagaaaggaa gggaagaaag cgaaaggagc	4320
gggcgctagg gcgctggcaa gtgtagcggc cacgctgcgc gtaaccacca cccccccgc	4380
gcttaatgcg ccgtacagg gcgctccat tcgccattca ggctgcgcaa ctgttgggaa	4440
gggcgatcgg tcggggcctc ttcgtatta cgccagctgg cgaaaggggg atgtgctgca	4500
aggcgattaa gttgggtaac gccagggttt tcccagtcac gacgttgtaa aacgacggcc	4560
agtgaattgt aatcagactc actatagggc gaattgggcc cgacgtcgca tgctatgtac	4620
actgtgtgag ttagagggc aatggtggtg acaactgagc tactatacac acacagcttg	4680
tactgttata agttgtgagc tacagtacaa gtaatactac aaaattcagt atgtacaagt	4740
atgccgttac ggcagtacga tgtatcaaag agagacagaa agaaggctgg aaagcgaagg	4800
ccagctgagc gtgggaaccg cagtacgtgg aagtgcctaaa cctgcgacag aatgacatgg	4860
tcttgcactt ttaccggta ccggtacagt acgataccgg tacgcacttg ctacagtaga	4920
tcatagacca ccagaataac cctgcattag cgccacgttg ccaagccctc aactttcttt	4980
gcagctacac cttttacacc ttttaggctg ccaggaagag cataaatgaa ttgtacttgg	5040
gggcacagaa ttgggttga agtgttcctg ggagagaatg tgcttggctg tatagtggg	5100
tggctctcct tcatagtaga tgtgaggta cagcacttgt agattgtgta cgcgttctgt	5160
ttaagtggt accagtatta tgggttgacg tgaagaaaag ggaggaggca ggggtacagc	5220
aagagtgtgg gcgcacactc acacatgaag caagacctgt ttcgtcattt ctatttgttt	5280
aacaaggacg acataatata cttaggacag catgggccga ccggcaggaa taatctgcaa	5340
acacaatcgt attgtggcgt cccctctcga tcacgatagt tgctaagttt agcagctga	5400
gagtttggct gtgacgcaag accgagagag agacagagag ttgtcccagt ttgagtactc	5460
gtaccgtact acagagtaaa tcgaaatcga gcgctcactt tctgcacttt gtgaaggcgc	5520
agcattcaac atttgtgtcc agcccccttc cgtttgaagc gacttgacca cgaaatgttg	5580
gaataaagct gtgttacata cgagacctac aagtagagtg tcggatatgc caactgaaca	5640
ggttagatga aagttaacta acaacaggag agcggacacc aacgtgcgaa aatgtacgga	5700
tgaaccatt gtctgtccgc atagcaatct gttttcgcca gcagtgcttg tgtcgttga	5760

ttgctgccat tgiactatit gtactttgtt ccattaacac tcgcatcttg ctacccgatt	5820
agcagcactc ctgttgccctc actgggcttg tctgtttcag ggtggaatgc atgagcaatg	5880
cgatcgtcaa ggggaggttg aaagcgcgta tggataaact gtggttctat agctgagggg	5940
aggatatcga ggacagccac ggttcagtgg gggatttcgg gagctactca gtaacatgta	6000
ctcttgtccg aacgccctgc tgatggcaaa tatgacaaaa cgcaacaatg tggagcaaca	6060
tttccatttc gagagatatt acatgggttg cattggggcc tgggcatccc gtgcaataac	6120
cgagaccatg tgaaggtgca cgatcgacac ttatccagcc agtcatacgt ggtgtcgcaa	6180
gctaagcaac aatgggattc acgaatctat gtgggggaac gtgcacaggg gaaatgaaac	6240
aaaacggagg aggggggtctt tgaaaatgia gagaaatata acaaaaaatc acaaaaacgc	6300
accaaaaata tgccaaaatc gtcaacagtg gaaatcacgc gcccttggtc cgacttccca	6360
ctgacatacc cccctcaac caatatcata tccacctag ccaatcgctt aattaagttg	6420
cgacacatgt ctgatagta tcttgaattc tctctcttga gcttttccat aacaagttct	6480
tctgcctcca ggaagtccat ggggtggttg atcatggttt tgggtgtagtg gtagtgcagt	6540
gggtggtattg tgactgggga ttagttgag aataagtcac acacaagtca gctttcttcg	6600
agcctcatat aagtataagt agttcaacgt attagcactg taccagcat ctccgtatcg	6660
agaaacacaa caacatgccc cattggacag atcatgcgga tacacaggtt gtgcagtatc	6720
atacatactc gatcagacag gtcgtctgac catcatacaa gctgaacaag cgctccatac	6780
ttgcacgctc tctatataca cagttaaatt acatatccat agtctaacct ctaacagtta	6840
atcttctggt aagcctccca gccagccttc tggatatcgt tggcctctc aataggatct	6900
cggttctggc cgtacagacc tcggccgaca attatgatat ccgttccggt agacatgaca	6960
tcctcaacag ttcggtactg ctgtccgaga gcgtctcctt tgcgtcaag acccaccg	7020
ggggtcagaa taagccagtc ctacagatcg ccttaggtc ggttctgggc aatgaagcca	7080
accacaaact cggggtcgga tcgggcaagc tcaatggtct gcttgagta ctgccagtg	7140
gccagagagc ctttgcaaga cagctcgcc agcatgagca gaccttggc cagcttctcg	7200
ttgggagagg ggactaggaa ctccctgtac tgggagttct cgtagtcaga gacgtcctcc	7260
ttcttctgtt cagagacagt ttctcggca ccagctcgca ggccagcaat gattccggtt	7320
ccgggtacac cgtgggcgtt ggtgatatcg gaccactcgg cgattcgggt acaccgtac	7380
tgggtcttga cagtgttgcc aatatctgcg aactttctgt cctcgaacag gaagaaccg	7440
tgcttaagag caagttcctt gagggggagc acagtgccg cgtaggtgaa gtcgtcaatg	7500

atgtcgatat gggttttgat catgcacaca taaggtccga ccttatcggc aagctcaatg	7560
agctccttgg tggtaggaac atccagagaa gcacacaggt tggttttctt ggctgccacg	7620
agcttgagca ctcgagcggc aaaggcggac ttgtggacgt tagctcgagc ttcgtaggag	7680
ggcattttgg tggtagaagag gagactgaaa taaatttagt ctgcagaact ttttatcgga	7740
accttatctg gggcagtgaa gtatatgtta tggtaaatgt tacgagttag ttgaacttat	7800
agatagactg gactatacgg ctatcgggcc aaattagaaa gaacgtcaat ggctctctgg	7860
gcgtcgccctt tgccgacaaa aatgtgatca tgatgaaagc cagcaatgac gttgcagctg	7920
atattgttgt cgccaaccg cgccgaaaac gcagctgtca gaccacagc ctccaacgaa	7980
gaatgtatcg tcaaagtgat ccaagcacac tcatagtgg agtcgtactc caaaggcggc	8040
aatgacgagt cagacagata ctgctcgacc ttttccttgg gaaccaccac cgtcagccct	8100
tctgactcac gtattgtagc caccgacaca ggcaacagtc cgtggatagc agaatatgtc	8160
ttgtcgggcc atttctcacc aacttttagc gtcaagttaa tgttcagaa gaagtatgtg	8220
ccttcattga gaatcgggtg tgctgatttc aataaagtct tgagatcagt ttggccagtc	8280
atgttgtggg gggtaattgg attgagttat cgcctacagt ctgtacaggt atactcgctg	8340
cccactttat actttttgat tccgctgcac ttgaagcaat gtcgtttacc aaaagtgaga	8400
atgtccaca gaacacaccc cagggtatgg ttgagcaaaa aataaacact ccgatacggg	8460
gaatcgaacc ccggtctcca cggttctcaa gaagtattct tgatgagagc gtatcgatag	8520
ttggagcaag ggagaaatgt agagtgtgaa agactcacta tgggtccgggc ttatctcgac	8580
caatagccaa agtctggagt ttctgagaga aaaaggcaag atacgtatgt aacaaagcga	8640
cgcatggtac aataataacc gaggcattga tcatagagag ttagtgggtc gatgatggca	8700
ctggtgcctg gtatgacttt atacggctga ctacatatct gtcctcagac atacaattac	8760
agtcaagcac ttacccttgg acatctgtag gtacccccg gccaaagacga tctcagcgtg	8820
tcgtatgtcg gattggcgta gtcctctgc tcgtcaattg gtcctcatct actttcttct	8880
gcttggtctac acccagcatg tctgctatgg ctggttttcg tgccttatct atcctcccag	8940
tattaccaac tctaaatgac atgatgtgat tgggtctaca ctttcatac agagataagg	9000
agtagcacag ttgcataaaa agcccaactc taatcagctt cttcctttct tgtaattagt	9060
acaaagtgta ttagcgaat ctggaagctt agttggccct aaaaaaatca aaaaaagcaa	9120
aaaacgaaaa acgaaaaacc acagttttga gaacaggag gtaacgaagg atcgatatata	9180
tatatatata tatatatacc cacggatccc gagaccggcc tttgattctt ccctacaacc	9240
aaccattctc accacctaa ttcaacaacca tgtctattgg ttctccaac cccgtgctct	9300
tggctgcgat tcccttcgtc tacctgtttg tctcccacg agtcctggct ttcctgcctc	9360

agaaggctca gttcctggcc aaatgtattg tggctctgat tgccacgctt atcatgtccg	9420
ttgcaggctg ctcatctcg atcgtgtgcg ctcttctgga caagagatac gtcatacaatt	9480
acgttgtgtc gcgattgttc tcccttccttg ccgctcgacc gtgtggtgtg acctataaga	9540
ttgttgggtga ggaacacctc gataagtacc ctgctatcgt ggtctgtaac catcaatcct	9600
ctatggatat gatggttttg ggacgagttt ttccaaagca ctgctgtgtc atggcgaaga	9660
aggaactcct gtactttccc tttttgggaa tgtttatgaa actgagcaac gctatcttca	9720
tcgaccggaa gaaccacaag aaagccatcg agtctaccac ccaagccgtg gcggacatga	9780
agaagcacia ctctggaate tggattttcc cagagggcac ccggtctaga ctggacaagg	9840
cagacctgct gcccttcaag aaaggtgcct ttcattctgc aattcaggcc cagctcccta	9900
ttctcccat tatctcgag ggctattccc atatctacga ctcttcgaag cggtaacttc	9960
ccggtggaga gctcgagatc agagtcctgg agcccatcc tacaactggc ctcaactactg	10020
atgatgtgaa cgacctgatg gacaagacac gaaaccttat gctcaagcac ttgaaggaga	10080
tggattccca gtattcgtcg agcactgctg aaaatggatc cacgcacatc gacgccgata	10140
ttgccaagtc tacagccacc agcattggca aactgacga cgcaattaca aaacgtcgt	10200
cccctaagga ataagcggcc gcattgatga ttggaaacac acacatgggt tatatctagg	10260
tgagagttag ttggacagtt atataataaa tcagctatgc caacggtaac ttcattcatg	10320
tcaacgagga accagtact gcaagtaata tagaattga ccaccttgcc attctcttgc	10380
actcctttac tatatctcat ttatttctta tatacaaatc acttcttctt cccagcatcg	10440
agctcggaaa cctcatgagc aataacatcg tggatctcgt caatagaggg ctttttggac	10500
tccttgctgt tggccacctt gtccttgctg tttaaacat catctaagg cctcaaaact	10560
acctcggaa tctgtcgctg atctggacac cacagaggtt ccgagcactt taggttgac	10620
caaatgtccc accaggtgca ggcagaaaac gctggaacag cgtgtacagt ttgtcttaac	10680
aaaaagttag ggcgtgagg tcgagcaggg tgggtgtgact tgttatagcc tttagagctg	10740
cgaaagcgcg tatggatttg gctcatcagg ccagattgag ggtctgtgga cacatgtcat	10800
gttagtgtac ttcaatcgcc ccctggatat agcccgaca ataggccgtg gcctcatitt	10860
tttgccttcc gcacatttcc attgctcggg acccacacct tgcttctcct gcacttgcca	10920
accttaatac tggttttacat tgaccaacat cttaacaagc gggggcttgt ctagggtata	10980
tataaacagt ggctctccca atcggttgcc agtctctttt ttcctttctt tccccacaga	11040
ttcgaaatct aaactacaca tcacaccatg gtctccaacc acctgttcga cgccatgcga	11100
gtgcccgtc ccggagacgc acctttcatt cgaatcgaca acgtcggac ctggacttac	11160
gatgacgcca ttgtctttc cggtcgaatc gctggagcta tggacgcact cggcattcga	11220

cccggagaca gagggtgccg gcaggtcgag aagtcgtccg aggcgttgat tctctacctg	11280
gcctgtcttc gaaccggagc tgtctacctg cctctcaaca ctgcctacac cctggccgag	11340
ctcgactact tcatcggcga tgccgaaccg cgtctggtgg tcgttgctcc cgcagctcga	11400
ggtggcgtgg agacaattgc caagcgacac ggtgctatcg tcgaaaccct cgacgccgat	11460
ggacgaggct ctttgctgga ctttgctaga gatgagcctg ccgactttgt cgatgcttcg	11520
cgatctgccg acgatctggc tgctattctc tacacttccg gtacaaccgg acgatcgaag	11580
ggtgccatgc ttactcatgg caatctgctc tccaacgctc tcaccttgcg agactattgg	11640
agagttaccg cagacgatcg actcatccat gccttgccaa tctttcacac tcatggtctg	11700
ttcgttgcta cgaacgtcac actgcttgca ggagcctcga tgtttctgct ctccaagttc	11760
gatgccgacg aggtcgtttc tctcatgcca caggccacca tgcttatggg cgtgcccaca	11820
ttctacgttc gattgctgca gaggctcga ctcgagaagg gtgctgtggc cagcatcaga	11880
ctgttcattt ctggatcagc tcccttgctt gccgaaacc acgccgagtt tcatgctcgt	11940
actggtcacg ccattctcga gcgatacggc atgacggaaa ccaacatgaa tacttccaac	12000
ccctacgagg gcaagcgtat tgccggaacc gttggttttc ctctgccga cgtcactgtg	12060
cgagtcaccg atcccgccac cggctctgct cttccacctg aagagactgg catgatcgag	12120
atcaaggac ccaacgtctt caagggtat tgccgaatgc ccgaaaagac cgctgccgag	12180
tttaccgcag acggtttctt tatctctgga gatctcggca agatcgaccg agaaggttac	12240
gttcacattg tgggacgagg caaggacctg gtcatttccg gtggctacaa catctatccc	12300
aaagaggctg aaggcgagat cgaccagatc gaggggtggt tcgagctgc tgtcattggt	12360
gttcctcatc ccgatttcgg agaagggtgc accgctgttg tcgtgtgcaa acctggtgcc	12420
gttctcgacg aaaagaccat cgtgtctgct ctgcaggacc gtcttgccc atacaagcaa	12480
cccaagcgga ttatctttgc cgacgatctg cctcgaaaca ctatgggaaa ggttcagaag	12540
aacattcttc gacagcaata cgccgatctc tacaccagac gataagcggc cgcattgagaa	12600
gataaatata taaatacatt gagatattaa atgcgctaga ttagagagcc tcatactgct	12660
cggagagaag ccaagacgag tactcaaagg ggattacacc atccatatcc acagacacaa	12720
gctggggaaa ggttctatct acactttccg gaataccgta gtttccgatg ttatcaatgg	12780
gggcagccag gatttcaggc acttcggtgt ctcggggtga aatggcgttc ttggcctcca	12840
tcaagtcgta ccatgtcttc atttgctgtg caaagtaaaa cagaagcaga tgaagaatga	12900
acttgaagtg aaggaattta aatagttgga gcaagggaga aatgtagagt gtgaaagact	12960

cactatggtc cgggcttata tcgaccaata gccaaagtct ggagtttctg agagaaaaag	13020
gcaagatacg tatgtaacaa agcgacgcat ggtacaataa taccggaggc atgtatcata	13080
gagagttagt ggttcgatga tggcactggc gcctggtagt actttatacg gctgactaca	13140
tatttgcctt cagacataca attacagtca agcacttacc cttggacatc tgtaggtacc	13200
ccccggccaa gacgatctca gcgtgtcgta tgtcggattg gcgtagctcc ctgctcgtc	13260
aattggctcc catctacttt cttctgcttg gctacacca gcatgtctgc tatggctcgt	13320
tttcgtgcct taictatcct ccagtatia ccaactctaa atgacatgat gtgattgggt	13380
ctacactttc atatcagaga taaggagtag cacagttgca taaaaagccc aactctaate	13440
agcttcttcc ttctttgtaa ttagtataaa ggtgattagc gaaatctgga agcttagttg	13500
gccctaaaaa aatcaaaaaa agcaaaaaac gaaaaacgaa aaaccacagt ttgagaaca	13560
gggaggtaac gaaggatcgt atatatatat atatatatat ataccacgg atcccgagac	13620
cgccctttga ttcttccta caaccaacca ttctcaccac cctaattcac aacctgggtg	13680
aaggcttctc gacaggctct gccctcgtc atcgacggaa aggtgtacga cgtctccgt	13740
tgggtgaact tccacctgg tggagctgaa atcattgaga actaccaggg acgagatgct	13800
actgacgctt tcatggttat gcactctcag gaagccttcg acaagctcaa gcgaatgccc	13860
aagatcaacc aggttccga gctgcctccc caggctgccg tcaacgaagc tcaggaggat	13920
ttccgaaagc tccgagaaga gctgatcgcc actggcatgt ttgacgctc tcccctctgg	13980
tactgtaca agatcttgac caccctgggt cttggcgtgc ttgccttctt catgctggtc	14040
cagtaccacc tgtacttcat tggctctctc gtgctcggta tgactacca gcaaatggga	14100
tggctgtctc atgacatctg ccaccaccag accttcaaga accgaaactg gaataacgtc	14160
ctgggtctgg tctttggcaa cggactccag ggcttctccg tgacctggtg gaaggacaga	14220
cacaacgccc atcattctgc taccaacgtt cagggtcacg atcccgacat tgataacctg	14280
cctctgctcg cctgggtccga ggacgatgtc actcgagctt ctcccatctc ccgaaagctc	14340
attcagttcc aacagtacta ttctctggtc atctgtattc tctgcatg catctgggtg	14400
ttccagtctg tgetgacctg tcatccctc aaggaccgag acaaccagtt ctaccgatct	14460
cagtacaaga aagaggccat tggactcgtc ctgcactgga ctctcaagac cctgttccac	14520
ctcttcttta tgcctccat cctgacctcg atgctgggtg tctttgttgc cgagctcgtc	14580
gggtggcttcg gaattgcat cgtgggtctc atgaaccact acctctgga gaagatcgg	14640
gattccgtct gggacggaca tggcttctct gtgggtcaga tccatgagac catgaacatt	14700
cgacgaggca tcaatctga ctgggtcttt ggaggcctga actaccagat cgagcaccat	14760
ctctggccca cctgcctcg acacaacctc actgccgttt cctaccaggt ggaacagctg	14820

tgccagaagc acaacctccc ctaccgaaac cctctgcccc atgaaggtct cgtcatcctg	14880
ctccgatacc tgtcccagtt cgctcgaatg gccgagaagc agcccgggtgc caaggctcag	14940
taagcggccg caagtgtgga tggggaagtg agtgcccgggt tctgtgtgca caattggcaa	15000
tccaagatgg atggattcaa cacagggata tagcgagcta cgtgggtggtg cgaggatata	15060
gcaacggata tttatgtttg acacttgaga atgtacgata caagcactgt ccaagtacaa	15120
tactaaacat actgtacata ctcatactcg taccggggca acggtttcac ttgagtgcag	15180
tggctagtgc tcttactcgt acagtgtgca atactgcgta tcatagtctt tgatgtatat	15240
cgtattcatt catgttagtt gc	15262
<210> 86	
<211> 15181	
<212> DNA	
<213> Artificial Sequence	
<220><223> Plasmid pZK16-MyL8N	
<400> 86	
aaaccatcat ctaagggcct caaaactacc tcggaactgc tgcgtgatc tggacaccac	60
agaggttccg agcactttag gttgcaccaa atgtcccacc aggtgcaggc agaaaacgct	120
ggaacagcgt gtacagtttg tcttaacaaa aagtgagggc gctgaggtcg agcagggtgg	180
tgtgacttgt tatagccttt agagctgcga aagcgcgtat ggatttggt catcaggcca	240
gattgagggt ctgtggacac atgtcatgtt agtgtacttc aatcgcccc tggatatagc	300
cccgacaata ggccgtggcc tcattttttt gccttcgca catttccatt gctcgtacc	360
cacaccttgc ttctcctgca cttgccaacc ttaatactgg ttacattga ccaacatctt	420
acaagcgggg ggcttgtcta gggatatatat aaacagtggc tctccaatc ggttgccagt	480
ctcttttttc ctttctttcc ccacagattc gaaatctaaa ctacacatca cccatggtc	540
tccaaccacc tgttcgacgc catgcgagct gccgtcccg gagacgcacc tttcattcga	600
atcgacaacg ctcggaactg gacttacgat gacgccattg ctttttccgg tcgaatcgct	660
ggagctatgg acgcactcgg cattcgaccg ggagacagag ttgccgtgca ggtcgagaag	720
tctgccgagg cgttgattct ctacctggcc tgtcttcgaa ccggagctgt ctacctgcct	780
ctcaacactg cctacacct ggccgagctc gactacttca tcggcgatgc cgaaccgcgt	840
ctggtggtcg ttgctccgc agctcgaggt ggcgtggaga caattgcaa gcgacacggt	900
gctatcgteg aaacctcga cgccgatgga cgaggctcct tgctggacct tgctagagat	960
gagcctgccg actttgtcga tgcttcgca tctgccgacg atctggctgc tattctctac	1020

acttccggtg caaccggacg atcgaagggt gccatgctta ctcatggcaa tctgctctcc 1080

aacgctctca ccttgcgaga ctattggaga gttaccgcag acgatcgact catccatgcc 1140

ttgccaatct ttcacactca tggctctgtc gttgctacga acgtcacact gcttgcagga 1200

gcctcgatgt ttctgctctc caagttcgat gccgacgagg tcgtttctct catgccacag 1260

gccaccatgc ttatgggcgt gccacattc tacgttcgat tgcgcagag tcctcgactc 1320

gagaagggtg ctgtggccag catcagactg ttcatttctg gatcagctcc ctgtcttgcc 1380

gaaaccacg ccgagtttca tgctcgact ggtcacgcca ttctcgagcg atacggcatg 1440

acggaacca acatgaatac ttccaacccc tacgagggca agcgtattgc cggaaccgtt 1500

ggttttcttc tgcccagct cactgtgcga gtcaccgatc ccgccaccgg tctcgttctt 1560

ccacctgaag agactggcat gatcgagatc aagggaccca acgtcttcaa gggtattgg 1620

cgaatgcccc aaaagaccgc tgccgagttt accgcagacg gtttctttat ctctggagat 1680

ctcggcaaga tcgaccgaga aggttacgtt cacattgtgg gacgaggcaa ggacctggtc 1740

atttccggtg gctacaacat ctatccaaa gaggtcgaag gcgagatcga ccagatcgag 1800

gggtgtggtc agtctgctgt cattgggtgt cctcatcccg atttcggaga aggtgtcacc 1860

gctgttgtc tgtgcaaacc tgggtccgtt ctgcacgaaa agaccatcgt gtctgctctg 1920

caggaccgtc ttgcccata caagcaacc aagcggatta tctttgccga cgatctgcct 1980

cgaaacacta tgggaaaggt tcagaagaac attcttcgac agcaatacgc cgatctctac 2040

accagacgat aagcggccgc atgagaagat aaatatataa atacattgag atattaaatg 2100

cgctagatta gagagcctca tactgctcgg agagaagcca agacgagtac tcaaagggga 2160

ttacaccatc catatccaca gacacaagct ggggaaaggt tctatataca ctttccggaa 2220

taccgtagtt tccgatgta tcaatggggg cagccaggat ttcaggcact tcggtgtctc 2280

ggggtgaaat ggcgttcttg gcctccatca agtcgtacca tgtcttcatt tgcctgtcaa 2340

agtaaacacg aagcagatga agaataact tgaagtgaag gaatttaaat agttggagca 2400

aggagaaaat gtagagtgtg aaagactcac tatggtccgg gcttatctcg accaatagcc 2460

aaagtctgga gtttctgaga gaaaaaggca agatacgtat gtaacaaagc gacgcattgt 2520

acaataatac cggaggcatg tatcatagag agttagtgg tcatgatgg cactggtgcc 2580

tggtatgact ttatacggct gactacatat ttgtctcag acatacaatt acagtcaagc 2640

acttaccctt ggacatctgt aggtaccccc cgccaagac gatctcagcg tgcgtatgt 2700

cggattggcg tagctccctc gctcgtcaat tggtcccat ctactttctt ctgcttggt 2760

acaccagca tgtctgctat ggctcgtttt cgtgccttat ctatcctccc agtattacca 2820

actctaatac acatgatgtg attgggtcta cactttcata tcagagataa ggagtagcac 2880

agttgcataa aaagcccaac tctaatacagc ttcttccctt cttgtaatta gtacaaaggt	2940
gattagcgaa atctggaagc ttagttggcc ctaaaaaat caaaaaagc aaaaaacgaa	3000
aaacgaaaaa ccacagtttt gagaacaggg aggtaacgaa ggatcgtata tatatatata	3060
tatatatata cccacggatc ccgagaccgg cctttgattc ticcctacaa ccaaccattc	3120
tcaccacctt aattcacaac catggtgaag gcttctcgac aggcctcgcc cctcgtcatc	3180
gacggaaagg tgtacgacgt ctccgcttgg gtgaacttcc accctgggtg agctgaaatc	3240
attgagaact accagggacg agatgctact gacgccttca tggttatgca ctctcaggaa	3300
gccttcgaca agctcaagcg aatgcccaag atcaaccagg cttccgagct gcctccccag	3360
gctgccgtca acgaagctca ggaggatttc cgaaagctcc gagaagagct gatcgccact	3420
ggcatgtttg acgcctctcc cctctgggtac tcgtacaaga tcttgaccac cctgggtctt	3480
ggcgtgcttg ccttcttcat gctgggtccag taccacctgt attcatttgg tgctctcgtg	3540
ctcgglatgc actaccagca aatgggatgg ctgtctcatg acatctgcca ccaccagacc	3600
ttcaagaacc gaaactggaa taacgtcctg ggtctggtct ttggcaacgg actccagggc	3660
ttctccgtga cctgggtggaa ggacagacac aacgcccatc attctgtac caacgttcag	3720
ggtcacgata ccgacattga taacctgcct ctgctcgctt ggtccgagga cgatgtcact	3780
cgagcttctc ccattctccg aaagctcatt cagttccaac agtactatct cctggtcatac	3840
tgtattctcc tgcgattcat ctgggtgttc cagtctgtgc tgaccgttcg atccctcaag	3900
gaccgagaca accagttcta ccgatctcag tacaagaaag aggccattgg actcgtcttg	3960
cactggactc tcaagacctt gtccacctc ttctttatgc cctccatcct gacctcgatg	4020
ctggtgttct ttgtttccga gctcgtcggg ggcttcggaa ttgccatcgt ggtcttcata	4080
aaccactacc ctctggagaa gatcgggtgat tccgtctggg acggacatgg cttctctgtg	4140
ggtcagatcc atgagacat gaacattcga cgaggcatca ttactgactg gttctttgga	4200
ggcctgaact accagatcga gcacatctc tggcccaccc tgcctcgaca caacctcact	4260
gccgtttcct accaggtgga acagctgtgc cagaagcaca acctcccta ccgaaacctt	4320
ctgccccatg aaggtctcgt catcctgctc cgatacctgt ccagtttcgc tcgaatggcc	4380
gagaagcagc ccggtgcca ggctcagtaa gcggccgcaa gtgtggatgg ggaagtgagt	4440
gccccgttct gtgtgcacaa ttggcaatcc aagatggatg gattcaacac agggatatag	4500
cgagctacgt ggtgggtcga ggatatagca acggatatct atgtttgaca cttgagaatg	4560
tacgatataa gcaactgtcca agtacaatac taaacatact gtacatactc atactcgtac	4620
ccgggcaacg gtttcaattg agtgcagtgg ctagtctctt tactcgtaca gtgtgcaata	4680

ctgcgtatca tagtctttga tgtatatcgt attcattcat gttagttagcg tacgcattct	4740
tgacttgctc cccagttgcc ctttttggtc ataactcagc cccgactctc tcccaaaga	4800
ggagatggac ttacaataga tacgagacaa cagatatatt cgccacctg gtgcaattgg	4860
tctgttttat taccccaat aatatccaca gcgtccgaat cagctcgaca agctctttct	4920
cagcatactg atcttaccce atacccecca gacagcggtg ccatatccat accgcacta	4980
cctgagacag taaactcagg gcaaggagcg gtcttggtcc atcgcgacgc tgtaagcttt	5040
gcctttgcgt ttgtgatcc accacacggc agtcaaggtc tcggttgaaa gatgggtctg	5100
gcaatgtgga ttgctgaatc attgacatat aaatacgagt gatcctgcc tttccaatc	5160
tttagaaatc atcttttctt cacacataca aaaaacacac atcacaatgc agttctccgc	5220
cgcgcctgc cttgccctcg ccgccttcgt ctctgccag gccgcctccg ctgcttctga	5280
agctgggtgt gccgttgtct ccggtgcctc cggtgccgt gccgcctctt cttctgctgt	5340
cattggctcc gagtctcagg ccgttcttct tgctgcctct tctgctgctg agtctgcttc	5400
ttccgtgcc tctgagtcg cctcctccgc ttccgcctcc gagtccaagg cctctgagtc	5460
cgttccaag tccgcctcca aggttcttga gtctgcttcc aaggccgagt cttctgctgc	5520
caaggcttcc tccgtgctg gttctgcttc ttccgcctgc cctctgcc ccaagtctgg	5580
tgatgcctcc ggctcttcca agccttccgg ctctgcctcc tctaccggtg acaagtccgg	5640
tgctgggtatt gtigccccg ctgttggtgc catcgttgtc ggtgctcttg ctgagctcat	5700
ctaagcggtc ttgtacaac taccactta atatttatat tatgatatag actgcaatac	5760
tattgcatgt gtaatgttc gtttgaaagt cgtagtccgg ccattcatgg cagtgtatgg	5820
cagtattgca accaattggc ttaataacat atacaaagt gcaattctgt gttcgaacca	5880
tcacacaacc caacatatat gacattgcta caagtattcg acttgcttct cgccgcacg	5940
acatgcattg ctgacgacac acttgcagaa gcttggccaa tggtaggtt gaccttaac	6000
cagtaaaacg caatataggt tgccaattaa caaacaacac agcccagaac tgggtgaaag	6060
ggcttgggtt ttgtgcaatc aactcattga ctacactccc ccaatgctcc atgaacccat	6120
gaatccatag tacttccatc ccaattcaga gaggtcaga aacacgcaat tggcaaaagt	6180
aaacaaaggt ctcgaaacct taaggggtgc agtgttgcaa agttaaaatt acactgttgg	6240
atggtgttcc actcaaactg tggatggatc agacggttgc actaaacca gattcaatat	6300
gctctacaag tagcgtgcc tgaggtgtat aatatcacac atgatcagag ttaaggcctg	6360
gaaaaaaagt gtccgcggct atgtaacact gctgaataga gtgtgtgccg tttccacgga	6420

ctggaatgaa tgcactcaac acgttggtag ctaacttccc caactattga cttcacttta	6480
ctgtagtctac tctaagtagc ataagacat catgaatcac caactttatc ttccccaccg	6540
aaagtaagag atgaaagaga ggctaacagc ttttctaaga aaatcacaaa atctacttac	6600
cgtattacaa aagtacaagt atcagtaacc aggcgcgcca gctgcattaa tgaatcggcc	6660
aacgcgcggg gagaggcggg ttgcgtattg ggcgctcttc cgcttctctg ctcactgact	6720
cgctgcgctc ggtcgttcgg ctgcggcgag cggtatcagc tcaactcaaag gcggtaatac	6780
ggttatccac agaatacagg gataacgcag gaaagaacat gtgagcaaaa ggccagcaaa	6840
aggccaggaa ccgtaaaaag gccgcgttgc tggcgttttt ccataggctc cgccccctg	6900
acgagcatca caaaaatcga cgctcaagtc agaggtaggcg aaacccgaca ggactataaa	6960
gataccaggc gtttccccct ggaagctccc tcgtgcgctc tcctgttccg accctgccgc	7020
ttaccggata cctgtccgcc ttttccccct cggaagcgt ggcgctttct catagctcac	7080
gctgtaggta ttcagttcg gtgtaggctc ttcgctccaa gctgggctgt gtgcacgaac	7140
ccccgttca gcccgaccgc tgcgccttat ccgctaacta tcgtcttgag tccaacccgg	7200
taagacacga cttatcgcca ctggcagcag ccaactggtaa caggattagc agagcgaggt	7260
atgtaggcgg tgctacagag ttcttgaagt ggtggcctaa ctacggctac actagaagaa	7320
cagtatttgg tatctgcgct ctgctgaagc cagttacctt cggaaaaaga gttggtagct	7380
cttgatccgg caaacaacc accgctggta gcgggtggtt ttttgtttgc aagcagcaga	7440
ttacgcgcag aaaaaaagga tctcaagaag atcctttgat cttttctacg gggtctgacg	7500
ctcagtggaa cgaaaactca cgttaaggga ttttggctat gagattatca aaaaggatct	7560
tcacctagat ctttttaaat taaaaatgaa gttttaaatc aatctaaagt atatatgagt	7620
aaacttggtc tgacagttac caatgcttaa tcagtaggc acctatctca gcgatctgtc	7680
tatttcgttc atccatagtt gcctgactcc ccgtcgtgta gataactacg atacgggagg	7740
gcttaccatc tggccccagt gctgcaatga taccgcgaga cccacgctca ccggctccag	7800
atttatcagc aataaaccag ccagccggaa gggccgagcg cagaagtggc cctgcaactt	7860
tatccgcctc catccagctt attaatgtt gccgggaagc tagagtaagt agttcgccag	7920
ttaatagttt gcgcaacgtt gttgccattg ctacagccat cgtgggtgtc cgctcgtcgt	7980
ttggtatggc ttcatcagc tccggttccc aacgatcaag gcgagttaca tgatcccca	8040
tgttgtgcaa aaaagcggtt agctccttcg gtcctccgat cgttgtcaga agtaagtgg	8100
ccgcagtgtt atcactcatg gttatggcag cactgcataa ttctcttact gtcatgccat	8160
ccgtaagatg cttttctgtg actggtgagt actcaacaa gtcattctga gaatagtgt	8220
tgccggcacc gattgtctt tgcccggcgt caatacggga taataccgcg ccacatagca	8280

gaactttaaa agtgcctcatc attggaaaac gttcttcggg gcgaaaactc tcaaggatct	8340
taccgtgtt gagatccagt tcgatgtaac cactcgtgc acccaactga tcttcagcat	8400
cttttacttt caccagcgtt tctgggtgag caaaaacagg aaggcaaat gccgcaaaaa	8460
aggaataag ggcgacacgg aaatgttgaa tactcatact cttccttttt caatattatt	8520
gaagcattta tcagggttat tgtctcatga gcggatacat atttgaatgt atttagaaaa	8580
ataaacaat aggggttccg cgcacatttc cccgaaaagt gccacctgat gcggtgtgaa	8640
ataccgcaca gatgcgtaag gagaaaatac cgcacagga aattgtaagc gttaatattt	8700
tgttaaaatt cgcgttaaat ttttgttaaa tcagctcatt ttttaacaa taggccgaaa	8760
tcggcaaat cccttataaa tcaaaagaat agaccgagat aggggtgagt gttgttccag	8820
tttgaacaa gattccacta ttaaagaacg tggactccaa cgtcaaaggc gaaaaaccg	8880
tctatcaggc cgatggccca ctacgtgaac catcaccta atcaagtttt ttggggtcga	8940
ggtgccgtaa agcactaaat cggaacccta aaggagagccc ccgatttaga gcttgacggg	9000
gaaagccggc gaacgtggcg agaaaggaag ggaagaaagc gaaaggagcg ggcgctaggc	9060
cgctggcaag tgtagcggtc acgctgcgcg taaccaccac acccgccgcg cttaatgcgc	9120
cgctacaggc cgcgtccatt cgccattcag gctgcgcaac tgttgggaag ggcgatcggc	9180
gcgggcctct tcgctattac gccagctggc gaaaggggga tgtgctgcaa ggcgattaag	9240
ttgggtaacg ccagggtttt ccagctcacg acgttgtaaa acgacggcca gtgaattgta	9300
atagactca ctataggcg aattgggccc gacgtcgcat gctatgtaca ctgtgtgagt	9360
gtagagggca atgggtgtga caactgagct actatacaca cacagcttgt actggtataa	9420
gttgtgagct acagtacaag taatactaca aaattcagta tgtacaagta tgccgttacg	9480
gcagtagcat gtatcaaaga gagacagaaa gaaggctgga aagcgaagc cagctgagcg	9540
tgggaaccgc agtacgtgga agtgccaaac ctgcgacaga atgacatggt ctgacattt	9600
ttaccgttac cggtacagta cgataccggt acgcacttgc tacagtagat catagaccac	9660
cagaataacc ctgcattagc gccacgttgc caagccctca actttctttg cagctacacc	9720
ttttacacct tttaggttgc caggaagagc ataatgaat tgtacttggg ggcacagaat	9780
ttgggttgaa gtgttcctgg gagagaatgt gcttggctgt atagtgggt ggctctcctt	9840
catagtagat gtgaggttac agcacttgta gattgtgtac gcgttctgtt taaggtggt	9900
ccagtattat ggggtgacgt gaaagaaagg gaggaggcag gggtagagca agagtgtggg	9960
cgcacactca cacatgaagc aagacctgtt tcgtcatttc tatttgttta acaaggacga	10020
cataatatac ttaggacagc atgggccgac cggcaggaat aatctgcaa cacaatcgta	10080
ttgtggcgct ccctctcgat cacgatagtt gctaagttaa gcagctgag agtttggctg	10140

tgacgcaaga ccgagagaga gacagagagt tgtcccagtt tgagtactcg taccgtacta	10200
cagagtaaata cgaaatcgag cgctcacttt ctgcactttg tgaaggcgca gcattcaaca	10260
tttgtgtcca gcccccttcc gtttgaagcg acttgaccac gaaatgttgg aataaagctg	10320
tgttacatac gagacctaca agtagagtgt cggatatgcc aactgaacag gttagatgaa	10380
agttactcaa caacaggaga gcggacacca acgtgcgaaa atgtacggat gaaccattg	10440
tctgtccgca tagcaatctg ttttcgccag cagtgccttg gtctgttgat tgctgccatt	10500
gtactatttg tactttgttc cattaacact cgcactctgc tcaccgatta gcagcactcc	10560
tgttgccctca ctgggcttgt ctgtttcagg gtggaatgca tgagcaatgc gatcgtcaag	10620
gggaggttga aagcgcgtat ggaataactg tggttctata gctgagggga ggatatcgag	10680
gacagccacg gttcagtggg ggatttcggg agctactcag taacatgtac tcttgtccga	10740
acgcctgct gatggcaaat atgacaaaac gcaacaatgt ggagcaacat ttccatttcg	10800
agagatatta catgggttgc attggggcct gggcatcccg tgcaataacc gagacatgt	10860
gaaggtgcac gatcgacact tatccagcca gtcatacgtg gtgtcgcaag ctaagcaaca	10920
atgggattca cgaatctatg tgggggaacg tgcacagggg aatgaaaca aaacggagga	10980
gggggtcttt gaaaatgtag agaaatataa caaaaaatca caaaaacgca ccaaaaatat	11040
gccaaaatcg tcaacagtgg aaatcacgcg cccctggctc gacttccac tgacataccc	11100
cccctcaacc aatatcatat ccacctagc caatcgctta attaatgtgc gacacatgtc	11160
ttgatagtat cttaattct ctctcttgag cttttccata acaagttctt ctgcctccag	11220
gaagtccatg ggtggtttga tcatggtttt ggtgtagtgg tagtgcatg gtggtattgt	11280
gactggggat gtagttgaga ataagtcata cacaagtcag ctttcttcga gcctcatata	11340
agtataagta gttcaacgta ttagcactgt acccagcatc tccgtatcga gaaacacaac	11400
aacatgcccc attggacaga tcatcggtat acacaggttg tgcagtatca tacatactcg	11460
atcagacagg tcgtctgacc atcatacaag ctgaacaagc gtcatact tgcacgtct	11520
ctatatcac agttaaatata catatccata gtctaacctc taacagttaa tcttctggta	11580
agcctcccag ccagccttct ggtatcgctt ggcctcctca ataggatctc ggttctggcc	11640
gtacagacct cggccgacaa ttatgatatc cgttccggtg gacatgacat cctcaacagt	11700
tcggtactgc tgtccgagag cgtctccctt gtcgtcaaga cccaccccg gggtcagaat	11760
aagccagtcc tcagagtcgc ccttaggtcg gttctgggca atgaagcaa ccacaaactc	11820
ggggtcggat cgggcaagct caatggtctg cttggagtac tcgccagtgg ccagagagcc	11880

cttgcaagac agctcggcca gcatgagcag acctctggcc agcttctcgt tgggagaggg	11940
gactaggaac tccttgact gggagtctc gtagtcagag acgtcctcct tcttctgttc	12000
agagacagtt tcctcggcac cagctcgcag gccagcaatg attccggttc cgggtacacc	12060
gtgggcgttg gtgatatcgg accactcggc gattcgggtga caccgggtact ggtgcttgac	12120
agtgttgcca atatctgcga actttctgtc ctggaacagg aagaaaccgt gcttaagagc	12180
aagttccttg agggggagca cagtccggc gtaggtgaag tcgtcaatga tgtcgatatg	12240
ggttttgatc atgcacacat aaggccgac cttatcggca agctcaatga gtccttggt	12300
ggtggaaca tcagagaag cacacaggtt ggttttcttg gctgccacga gcttgagcac	12360
tcgagcggca aaggcggact tgtggacgtt agctcgagct tcgtaggagg gcattttggt	12420
ggtgaagagg agactgaaat aaatttagtc tgcagaactt tttatcggaa cttatctgg	12480
ggcagtgaag tatatgttat ggtaatagt acgagttagt tgaacttata gatagactgg	12540
actatacggc tatcggcca aattagaaag aacgtcaatg gctctctggg cgtcgcttt	12600
gccgacaaaa atgtgatcat gatgaaagcc agcaatgacg ttgcagctga tattgtgtc	12660
ggccaaccgc gccgaaaacg cagctgtcag acccacagcc tccaacgaag aatgtatcgt	12720
caaagtgatc caagcacact catagttgga gtcgtactcc aaaggcggca atgacgagtc	12780
agacagatac tcgtcgacct tttccttggg aaccaccacc gtcagccctt ctgactcacg	12840
tatttgatgc accgacacag gcaacagtcc gtggatagca gaatatgtct tgtcggcca	12900
tttctacca actttagcgc tcaagtgaat gttgcagaag aagtatgtgc cttcattgag	12960
aatcgggtgt gctgatttca ataaagtctt gagatcagtt tggccagtca tgttgtggg	13020
ggtaattgga ttgagttatc gcctacagtc tgtacaggta tactcgctgc ccactttata	13080
ctttttgatt ccgtgcact tgaagcaatg tcgtttacca aaagtgagaa tgctccacag	13140
aacacacccc agggatatgt tgagcaaaaa ataaacactc cgatacgggg aatcgaaccc	13200
cggctctccac ggttctcaag aagtattctt gatgagagcg tatcgatagt tggagcaagg	13260
gagaaatgta gagtgtgaaa gactcactat ggtccgggct tatctcgacc aatagccaaa	13320
gltcggagtt tctgagagaa aaaggcaaga tacgtatgta acaaagcgac gcatggtaca	13380
ataataccgg aggcattgat catagagagt tagtggttcg atgatggcac tggcctgg	13440
tatgacttta tacggctgac tacatatattg tcctcagaca tacaattaca gtcaagcact	13500
tacccttgga catctgtagg taccctccgg ccaagacgat ctacgcgtgt cgtatgtcgg	13560
attggcgtag ctccctcgt cgtcaattgg ctccatcta ctttctctg cttggctaca	13620
cccagcatgt ctgctatggc tcgttttcgt gccttatcta tcctccagc attaccaact	13680
ctaaatgaca tgatgtgatt gggctacac ttcatatca gagataagga gtagcacagt	13740

tgataaaaa gcccaactct aatcagcttc ttcctttctt gtaattagta caaaggtgat	13800
tagcgaaatc tggaagctta gttggcccta aaaaaatcaa aaaaagcaaa aaacgaaaaa	13860
cgaaaaacca cagttttgag aacagggagg taacgaagga tcgtatatat atatatatat	13920
atataatccc acggatcccg agaccggcct ttgattcttc cctacaacca accattctca	13980
ccacctaat tcacaacat gtccgttgca tccaagctcg tttctacgt ccgcgccgcc	14040
atcgccgtgg tcattcttgc cgctgtgcc acctacggcg tgcctggctc caccattctc	14100
accgccatcg gcaagcaggg cctggcccaa tggaccgttg ccagagcctt ctactactcg	14160
gtgcgcattc tcttgggtat cagcatcaag ctgcgtagcc ggcaggtgac cggaaccgcc	14220
ggctctggatg cctccaagat ccaggtcgcc aacaccacca agccattga cgacatcacc	14280
aaacacctgc cccgaccatg cattctgatt tccaaccacc agaacgaaat ggacattctg	14340
gtgctcggtc gcatcttccc ccagtactgc tccgtcaccg ccaaaaaggc cctcaagtgg	14400
taccctctgc tgggccagtt catggcgctg tccggcacca tcttcttgga ccgaaaggac	14460
cgaaccaagt ccgtgcagac cctcggcggc gccgtcaaga ccatccagag cggcaacgga	14520
ggcaagggcc agagcgtctt catgttcccc gagggaaacc gatcctactc caaggacgtc	14580
ggcatcatgc cctcaagaa gggctgtttc cacctggcgg tccagtcggg cgctccatt	14640
gtccccgtgg tggccagaa cacctccga atgttttctt tcggccgagg caagctggac	14700
gccggagaga tcttgtcga cgtcttgagc ccattgaga ccaagggtct ggacgccagc	14760
aacgtcgacg ctctcatggc caccattat aaggccatgt gcgagactgc cgaccagatt	14820
ggctacgtg gccagaagac tcagtaggcg gccgcatgag aagataata tataaatata	14880
ttgagatatt aaatgcgcta gattagagag cctcactatg ctcgagaga agccaagacg	14940
agtactcaaa ggggattaca ccatccatat ccacagacac aagctgggga aaggttctat	15000
atacatttc cggaataacc tagtttccga tgttatcaat gggggcagcc aggatttcag	15060
gcatttcggt gtctcgggt gaaatggcgt tcttggcctc catcaagtcg taccatgtct	15120
tcatttcct gtcaaagtaa aacagaagca gatgaagaat gaactgaag tgaaggaatt	15180
t	15181

<210> 87

<211> 1272

<212> DNA

<213> Artificial Sequence

<220><223> mutant EgD8S-23 delta-8 desaturase (also "EgD8M")

<220><221> CDS

<222> (2)..(1270)

<300><302> MUTANT DELTA-8 DESATURASE GENES ENGINEERED BY TARGETED MUTAGENESIS
AND THEIR USE IN MAKING POLYUNSATURATED FATTY ACIDS

<310> US 7709239

<311> 2006-12-07

<312> 2010-05-04

<313> (1)..(1272)

<400> 87

c atg gtg aag gct tct cga cag gct ctg ccc ctc gtc atc gac gga aag 49

Met Val Lys Ala Ser Arg Gln Ala Leu Pro Leu Val Ile Asp Gly Lys

1 5 10 15

gtg tac gac gtc tcc gct tgg gtg aac ttc cac cct ggt gga gct gaa 97

Val Tyr Asp Val Ser Ala Trp Val Asn Phe His Pro Gly Gly Ala Glu

20 25 30

atc att gag aac tac cag gga cga gat gct act gac gcc ttc atg gtt 145

Ile Ile Glu Asn Tyr Gln Gly Arg Asp Ala Thr Asp Ala Phe Met Val

35 40 45

atg cac tct cag gaa gcc ttc gac aag ctc aag cga atg ccc aag atc 193

Met His Ser Gln Glu Ala Phe Asp Lys Leu Lys Arg Met Pro Lys Ile

50 55 60

aac cag gct tcc gag ctg cct ccc cag gct gcc gtc aac gaa gct cag 241

Asn Gln Ala Ser Glu Leu Pro Pro Gln Ala Ala Val Asn Glu Ala Gln

65 70 75 80

gag gat ttc cga aag ctc cga gaa gag ctg atc gcc act ggc atg ttt 289

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

85 90 95

gac gcc tct ccc ctc tgg tac tcg tac aag atc ttg acc acc ctg ggt 337

Asp Ala Ser Pro Leu Trp Tyr Ser Tyr Lys Ile Leu Thr Thr Leu Gly

100 105 110

ctt ggc gtg ctt gcc ttc ttc atg ctg gtc cag tac cac ctg tac ttc 385

Leu Gly Val Leu Ala Phe Phe Met Leu Val Gln Tyr His Leu Tyr Phe

115	120	125	
att ggt gct ctc gtg ctc ggt atg cac tac cag caa atg gga tgg ctg			433
Ile Gly Ala Leu Val Leu Gly Met His Tyr Gln Gln Met Gly Trp Leu			
130	135	140	
tct cat gac atc tgc cac cac cag acc ttc aag aac cga aac tgg aat			481
Ser His Asp Ile Cys His His Gln Thr Phe Lys Asn Arg Asn Trp Asn			
145	150	155	160
aac gtc ctg ggt ctg gtc ttt ggc aac gga ctc cag ggc ttc tcc gtg			529
Asn Val Leu Gly Leu Val Phe Gly Asn Gly Leu Gln Gly Phe Ser Val			
165	170	175	
acc tgg tgg aag gac aga cac aac gcc cat cat tct gct acc aac gtt			577
Thr Trp Trp Lys Asp Arg His Asn Ala His His Ser Ala Thr Asn Val			
180	185	190	
cag ggt cac gat ccc gac att gat aac ctg cct ctg ctc gcc tgg tcc			625
Gln Gly His Asp Pro Asp Ile Asp Asn Leu Pro Leu Leu Ala Trp Ser			
195	200	205	
gag gac gat gtc act cga gct tct ccc atc tcc cga aag ctc att cag			673
Glu Asp Asp Val Thr Arg Ala Ser Pro Ile Ser Arg Lys Leu Ile Gln			
210	215	220	
ttc caa cag tac tat ttc ctg gtc atc tgt att ctc ctg cga ttc atc			721
Phe Gln Gln Tyr Tyr Phe Leu Val Ile Cys Ile Leu Leu Arg Phe Ile			
225	230	235	240
tgg tgt ttc cag tct gtg ctg acc gtt cga tcc ctc aag gac cga gac			769
Trp Cys Phe Gln Ser Val Leu Thr Val Arg Ser Leu Lys Asp Arg Asp			
245	250	255	
aac cag ttc tac cga tct cag tac aag aaa gag gcc att gga ctc gct			817
Asn Gln Phe Tyr Arg Ser Gln Tyr Lys Lys Glu Ala Ile Gly Leu Ala			
260	265	270	
ctg cac tgg act ctc aag acc ctg ttc cac ctc ttc ttt atg ccc tcc			865
Leu His Trp Thr Leu Lys Thr Leu Phe His Leu Phe Phe Met Pro Ser			
275	280	285	

atc ctg acc tcg atg ctg gtg ttc ttt gtt tcc gag ctc gtc ggt ggc 913
Ile Leu Thr Ser Met Leu Val Phe Phe Val Ser Glu Leu Val Gly Gly
290 295 300
ttc gga att gcc atc gtg gtc ttc atg aac cac tac cct ctg gag aag 961
Phe Gly Ile Ala Ile Val Val Phe Met Asn His Tyr Pro Leu Glu Lys
305 310 315 320
atc ggt gat tcc gtc tgg gac gga cat ggc ttc tct gtg ggt cag atc 1009
Ile Gly Asp Ser Val Trp Asp Gly His Gly Phe Ser Val Gly Gln Ile
325 330 335
cat gag acc atg aac att cga cga ggc atc att act gac tgg ttc ttt 1057
His Glu Thr Met Asn Ile Arg Arg Gly Ile Ile Thr Asp Trp Phe Phe
340 345 350
gga ggc ctg aac tac cag atc gag cac cat ctc tgg ccc acc ctg cct 1105
Gly Gly Leu Asn Tyr Gln Ile Glu His His Leu Trp Pro Thr Leu Pro
355 360 365
cga cac aac ctc act gcc gtt tcc tac cag gtg gaa cag ctg tgc cag 1153
Arg His Asn Leu Thr Ala Val Ser Tyr Gln Val Glu Gln Leu Cys Gln
370 375 380
aag cac aac ctc ccc tac cga aac cct ctg ccc cat gaa ggt ctc gtc 1201
Lys His Asn Leu Pro Tyr Arg Asn Pro Leu Pro His Glu Gly Leu Val
385 390 395 400
atc ctg ctc cga tac ctg tcc cag ttc gct cga atg gcc gag aag cag 1249
Ile Leu Leu Arg Tyr Leu Ser Gln Phe Ala Arg Met Ala Glu Lys Gln
405 410 415
ccc ggt gcc aag gct cag taa gc 1272
Pro Gly Ala Lys Ala Gln
420
<210> 88
<211> 422
<212> PRT
<213> Artificial Sequence

<220><223> Synthetic Construct

<400> 88

Met Val Lys Ala Ser Arg Gln Ala Leu Pro Leu Val Ile Asp Gly Lys

1 5 10 15

Val Tyr Asp Val Ser Ala Trp Val Asn Phe His Pro Gly Gly Ala Glu

20 25 30

Ile Ile Glu Asn Tyr Gln Gly Arg Asp Ala Thr Asp Ala Phe Met Val

35 40 45

Met His Ser Gln Glu Ala Phe Asp Lys Leu Lys Arg Met Pro Lys Ile

50 55 60

Asn Gln Ala Ser Glu Leu Pro Pro Gln Ala Ala Val Asn Glu Ala Gln

65 70 75 80

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

85 90 95

Asp Ala Ser Pro Leu Trp Tyr Ser Tyr Lys Ile Leu Thr Thr Leu Gly

100 105 110

Leu Gly Val Leu Ala Phe Phe Met Leu Val Gln Tyr His Leu Tyr Phe

115 120 125

Ile Gly Ala Leu Val Leu Gly Met His Tyr Gln Gln Met Gly Trp Leu

130 135 140

Ser His Asp Ile Cys His His Gln Thr Phe Lys Asn Arg Asn Trp Asn

145 150 155 160

Asn Val Leu Gly Leu Val Phe Gly Asn Gly Leu Gln Gly Phe Ser Val

165 170 175

Thr Trp Trp Lys Asp Arg His Asn Ala His His Ser Ala Thr Asn Val

180 185 190

Gln Gly His Asp Pro Asp Ile Asp Asn Leu Pro Leu Leu Ala Trp Ser

195 200 205

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

210 215 220

Phe Gln Gln Tyr Tyr Phe Leu Val Ile Cys Ile Leu Leu Arg Phe Ile

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

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

Gly Gly Leu Asn Tyr Gln Ile Glu His His Leu Trp Pro Thr Leu Pro
 355 360 365
 Arg His Asn Leu Thr Ala Val Ser Tyr Gln Val Glu Gln Leu Cys Gln
 370 375 380
 Lys His Asn Leu Pro Tyr Arg Asn Pro Leu Pro His Glu Gly Leu Val
 385 390 395 400
 Ile Leu Leu Arg Tyr Leu Ser Gln Phe Ala Arg Met Ala Glu Lys Gln
 405 410 415

Pro Gly Ala Lys Ala Gln
 420

<210> 89

<211> 15105

<212> DNA

<213> Artificial Sequence

<220><223> Plasmid pZK16-ML3

<400> 89

gtacgcattc ttgacttgct cccagttgc cctttttggt cataactcag ccccgactct

60

ctccccaag aggagatgga cttacaatag atacgagaca acagatattt tcggccacct 120

ggtgcaattg gtctgtttta ttaccccaaa taatatccac agcgtccgaa tcagctcgac 180

aagctctttc tcagcatact gatcttacc cataccccc agacagcggg gccatatcca 240

taccgcact acctgagaca gtaaacctcag ggcaaggagc ggtcttggtc catcgcgacg 300

ctgtaagctt tgcctttgcg ttttgtgatc caccacacgg cagtcaaggt ctcggttgaa 360

agatgggtct ggcaatgtgg attgctgaat cattgacata taaatcagag tgatcctgcc 420

atttcccaat ctttagaaat catcttttct tcacacatac aaaaaacaca catcacaatg 480

cagttctccg ccgcccgct ccttgccctc gccgccttcg tctctgccca ggccgcctcc 540

gtctgttctg aagctgggtg tgccgttgct tccggtgcct ccggtgccgc tgccgcgct 600

tcttctgctg tcattggctc cgagctcag gccgttctt ctgctgcctc ttctgctgct 660

gagctctgct cttccgctgc ctctgagtc gccctcctcc cttccgcctc cgagtccaag 720

gcctctgagt ccgcttccaa gtccgcctcc aaggcttctg agtctgctc caaggccgag 780

tcttctgctg ccaaggtctc ctccgctgct ggttctgctt cttccgccgc tgcctctgcc 840

accaagtctg gtgatgcctc cggtcttctc aagccttccg gctctgcctc ctctaccggt 900

gacaagtccg gtgctggat tgttgcccc gctgttggtg ccatcggtgt cgggtgctctt 960

gtcagctca tctaagcggg ctttgtacaa ctaccactt aatatttata ttatgatata 1020

gactgcaata ctattgcatg tgtaatgctt cgtttgaaag tcgtagtccg gccattcatg 1080

gcagtgtatg gcagtattgc aaccaattgg cttaataaca tatacaaagt ggcaattctg 1140

tgttcgaacc atcacacaac ccaacatatt tgacattgct acaagtattc gacttgttct 1200

tcggccgcac gacatgcatt gctagcagca cacttgcaag agcttgcca atggttaggt 1260

tgacctttaa ccagtaaaac gcaatatagg ttgccaatta acaacaaca cagcccagaa 1320

ctggtgtaaa gggttgggt tttgtgcaat caactcattg actcacctcc cccaatgctc 1380

catgaacca tgaatccata gtacctccat cccaattcag agaggctcag aaacacgcaa 1440

ttggcaaaag taacaaaagg tctcgaaacc ttaaggggtg cagtgttgca aagttaaaat 1500

tacactgttg gatggtgtc cactcaaaact gtggatggat cagacggttg cactaaacca 1560

agattcaata tgctctacaa gtagcgtgcc atgagggtga taatatcaca catgatcaga 1620

gttaaggcct ggaaaaaaag tgctcgccgc tatgtaacac tgctgaatac agtgtgtgcc 1680

gtttccacgg actggaatga atgcactcaa caggttggtg gctaacttcc ccaactattg 1740

acttacttt actgtagcta ctctaagtag gataatgaca tcatgaatca ccaactttat 1800

cttccccacc gaaagtaaga gatgaaagag aggctaacag cttttgtaag aaaatcacia 1860

aatctactta ccgtattaca aaagtacaag tatcagtaac caggcgccgc agctgcatta 1920

atgaatcggc caacgcgcgg ggagaggcgg ttgctgtatt gggcgctctt ccgcttcctc	1980
gctcactgac tcgctgcgct cggctcgttcg gctgcgcgca gcggtatcag ctactcaaa	2040
ggcggtaata cggttatcca cagaatcagg ggataacgca ggaaagaaca tgtgagcaaa	2100
agggcagcaa aaggccagga accgtaaaaa ggccgcgttg ctggcgtttt tccataggct	2160
ccgccccct gacgagcatc acaaaaatcg acgtcgaagt cagaggtggc gaaacccgac	2220
aggactataa agataccagg cgtttcccc tggaagctcc ctctgcgct ctctgttcc	2280
gacctgccc cttaccgat acctgtccgc ctttctccct tcgggaagcg tggcgctttc	2340
tcatagctca cgctgtaggt atctcagttc ggtgtaggtc gttcgctcca agctgggctg	2400
tgtgcacgaa cccccgttc agcccgaccg ctgcgcctta tccgtaact atcgtcttga	2460
gtccaacccg gtaagacacg acttatcgcc actggcagca gccactggtg acagattag	2520
cagagcgagg tatgtaggag gtgctacaga gttcttgaag tggtagccta actacggcta	2580
cactagaaga acagtatttg gtatctgcgc tctgctgaag ccagttacct tcggaaaaag	2640
agttgtagc tcttgatccg gcaacaaac caccgtggg agcggtggtt tttttgttg	2700
caagcagcag attacgcgca gaaaaaagg atctcaagaa gatcctttga tctttctac	2760
ggggtctgac gtcagtggg acgaaaactc acgttaaggg attttggtca tgagattatc	2820
aaaaaggatc ttcacctaga tccttttaaa ttaaaaatga agttttaaat caatctaaag	2880
tatatatgag taaacttggg ctgacagtta ccaatgctta atcagtgagg cacctatctc	2940
agcgatctgt ctatttcgtt catccatagt tgctgactc ccgctcgtgt agataactac	3000
gatacgggag ggttaccat ctggccccag tctgcaatg ataccgcgag acccacgctc	3060
accggctcca gatattatcag caataaacca gccagccgga agggccgagc gcagaagtgg	3120
tcctgcaact ttatccgct ccatccagtc tattaattgt tgccgggaag ctagagtaag	3180
tagttcgcca gttaatagtt tgcgcaacgt tgttgccatt gctacaggca tcgtggtgtc	3240
acgtcgtcg tttggtatgg cttcattcag ctccggttcc caacgatcaa ggcgagttac	3300
atgatcccc atgttgtgca aaaaagcggg tagctccttc ggtcctccga tcgttgtcag	3360
aagtaagtgt gccgcagtgt tatcactcat ggttatggca gcaactgcata attctcttac	3420
tgtcatgcca tccgtaagat gcttttctgt gactggtgag tactcaacca agtcattctg	3480
agaatagtgt atcgggcgac cgagttgctc ttgcccggcg tcaatcggg ataataccgc	3540
gccacatagc agaactttaa aagtgtcat cattggaaaa cgttcttcgg ggcgaaaact	3600
ctcaaggatc ttaccgctgt tgagatccag ttcatgtaa cccactcgtg cacccaactg	3660

atcttcagca tcttttactt tcaccagcgt ttctgggtga gcaaaaacag gaaggcaaaa	3720
tgccgcaaaa aagggaataa gggcgacacg gaaatgtga atactcatac tcttcctttt	3780
tcaatattat tgaagcatit atcagggtta ttgtctcatg agcggataca tatttgaatg	3840
tatttagaaa aataaacaat taggggttcc ggcacacatit ccccgaaaag tgccacctga	3900
tgcggtgtga aataccgcac agatgcgtaa ggagaaaata ccgcatcagg aaattgtaag	3960
cgttaatatt ttgttaaaat tcgcgttaaa tttttgttaa atcagctcat tttttaacca	4020
ataggccgaa atcggaataa tcccttataa atcaaaagaa tagaccgaga taggggttag	4080
tggtgttcca gtttgaaca agagtccact attaaagaac gtggactcca acgtcaaagg	4140
gcgaaaaacc gtctatcagg gcgatggccc actacgtgaa ccatcaccct aatcaagttt	4200
tttggggtcg aggtgccgta aagcactaaa tcggaaccct aaaggggagcc cccgatttag	4260
agcttgacgg ggaagccgg cgaacgtggc gagaaaaggaa gggaagaaag cgaaaggagc	4320
gggcgctagg gcgctggcaa gtgtagcggc cacgtgcgc gtaaccacca caccgcccgc	4380
gcttaatcgc ccgtacagg gcgcgtccat tcgccattca ggctgcgcaa ctgttgggaa	4440
gggcgatcgg tgcgggcctc ttcgtatta cgccagctgg cgaaaggggg atgtgtgca	4500
aggcgattaa gtgggtaac gccagggttt tcccagtcac gacgttgtaa aacgacggcc	4560
agtgaattgt aatacgactc actatagggc gaattgggcc cgacgtcgca tgctatgtac	4620
actgtgtgag tglagagggc aatgggtgtg acaactgagc tactatacac acacagcttg	4680
tactgttata agttgtgagc tacagtacaa gtaatactac aaaattcagt atgtacaagt	4740
atgccgttac ggcagtagca tgtatcaaag agagacagaa agaaggctgg aaagcgaagg	4800
ccagctgagc gtgggaaccg cagtacgtgg aagtgcgcaa cctgcgacag aatgacatgg	4860
tcttgcaatt ttaccggta ccggtacagt acgataccgg tacgcacttg ctacagtaga	4920
tcatagacca ccagaataac cctgcattag cgccacgttg ccaagccctc aactttcttt	4980
gcagctacac cttttacacc ttttaggttg ccaggaagag cataaatgaa ttgtacttgg	5040
gggcacagaa tttgggttga agtgttctg ggagagaatg tgcttggctg tatagttggg	5100
tggtctcct tcatagtaga tgtgaggtta cagcacttgt agatttgtga cgcgttctgt	5160
ttaaggtgt accagtatta tgggttgacg tgaaagaaag ggaggaggca ggggtacagc	5220
aagagtgtgg gcgcacactc acacatgaag caagacctgt ttcgtcattt ctatttgttt	5280
aacaaggacg acataatata cttaggacag catgggccga ccggcaggaa taatctgcaa	5340
acacaatcgt attgtggcgt cccctctcga tcacgatagt tgctaagttt agcagctcga	5400
gagtttggct gtgacgcaag accgagagag agacagagag ttgtcccagt ttgagtactc	5460
gtaccgtact acagagtaaa tcgaaatcga gcgctcactt tctgcacttt gtgaaggcgc	5520

agcattcaac atttgtgtcc agcccccttc cgtttgaagc gacttgacca cgaaatgttg 5580
 gaataaagct gtgttacata cgagacctac aagtagagtg tcggatatgc caactgaaca 5640
 ggtagatga aagtactca acaacaggag agcggacacc aacgtgcgaa aatgtacgga 5700
 tgaaccatt gtctgtccgc atagcaatct gttttcgcca gcagtgttg tgtcgttgga 5760

 ttgctgcat tgtactatgt gtactttgtt ccattaacac tcgcatcttg ctaccgatt 5820
 agcagcactc ctgttgctc actgggcttg tctgtttcag ggtggaatgc atgagcaatg 5880
 cgatcgtcaa ggggaggttg aaagcgcgta tgaataact gtggttctat agctgagggg 5940
 aggatatcga ggacagccac ggttcagtgg gggatttcgg gagctactca gtaacatgta 6000
 ctcttgtccg aacgcccctgc tgatggcaaa tatgacaaaa cgcaacaatg tggagcaaca 6060
 tttccatttc gagagatatt acatgggttg cattggggcc tgggcatccc gtgcaataac 6120
 cgagaccatg tgaaggtgca cgatcgacac ttatccagcc agtcatacgt ggtgtcgcaa 6180

 gctaagcaac aatgggattc acgaatctat gtgggggaac gtgcacaggg gaaatgaaac 6240
 aaaacggagg aggggggtctt tgaaaatgta gagaaatata acaaaaaatc acaaaaacgc 6300
 accaaaaata tgccaaaatc gtcaacagtg gaaatcacgc gcccttggt cgacttccca 6360
 ctgacatacc cccctcaac caatatcata tccaccctag ccaatcgctt aattaagttg 6420
 cgacacatgt cttgatagta tcttgaattc tctctcttga gcttttccat aacaagttct 6480
 tctgcctcca ggaagtccat ggggtggttg atcatggttt tgggttagtg gtagtgcagt 6540
 ggtggtattg tgactgggga ttagttgag aataagtcac acacaagtca gctttcttcg 6600

 agcctcatat aagtataagt agttcaacgt attagcactg taccagcat ctccgtatcg 6660
 agaaacacaa caacatgccc cattggacag atcatgcgga tacacaggtt gtgcagtatc 6720
 atacatactc gatcagacag gtcgtctgac catcatacaa gctgaacaag cgctccatac 6780
 ttgcacgctc tctatataca cagttaaatt acatatccat agtctaacct ctaacagtta 6840
 atcttctggt aagcctccca gccagccttc tggatcgtt tggcctcctc aataggatct 6900
 cggttctggc cgtacagacc tcggccgaca attatgatat ccgttcgggt agacatgaca 6960
 tcctcaacag ttcggtactg ctgtccgaga gcgtctccct tgtcgtcaag acccaccg 7020

 ggggtcagaa taagccagtc ctgagagtcg cccttaggtc ggttctgggc aatgaagcca 7080
 accacaaact cggggtcgga tcgggcaagc tcaatgttct gcttgagta ctgccagt 7140
 gccagagagc ctttgaaga cagctcgcc agcatgagca gacctctggc cagcttctcg 7200
 ttgggagagg ggactaggaa ctcttgtac tgggagttct cgtagtcaga gacgtcctcc 7260
 ttcttctgtt cagagacagt ttctcggca ccagctcgca ggccagcaat gattccggtt 7320
 ccgggtacac cgtgggcgtt ggtgatatcg gaccactcgg cgattcgtg acaccgtac 7380

tggtgcttga cagtgttgcc aatatctgcg aactttctgt cctcgaacag gaagaaaccg	7440
tgcttaagag caagttcctt gaggggggagc acagtgccgg cgtaggtgaa gtcgtcaatg	7500
atgtcgatat gggttttgat catgcacaca taaggccga ccttatcggc aagctcaatg	7560
agctccttgg tggtagtaac atccagagaa gcacacaggt tggttttctt ggctgccacg	7620
agcttgagca ctcgagcggc aaaggcggac ttgtggacgt tagctcgagc ttcgtaggag	7680
ggcattttgg tggtagaagag gagactgaaa taaatttagt ctgcagaact ttttatcgga	7740
accttatctg gggcagtgaa gtatatgtta tggtaaatgt tacgagttag ttgaacttat	7800
agatagactg gactatacgg ctatcgggtcc aaattagaaa gaacgtcaat ggctctctgg	7860
gcgtgcctt tgccgacaaa aatgtgatca tgatgaaagc cagcaatgac gttgcagctg	7920
atattgttgt cgccaaccg cgccgaaaac gcagctgtca gaccacagc ctccaacgaa	7980
gaatgtatcg tcaaagtgat ccaagcacac tcatagttag agtcgtactc caaaggcggc	8040
aatgacgagt cagacagata ctgctcgacc ttttccttgg gaaccaccac cgtcagccct	8100
tctgactcac gtattgtagc caccgacaca ggcaacagtc cgtggatagc agaatatgtc	8160
ttgtcgggtc atttctcacc aactttaggc gtcaagttaa tgttcagaa gaagtatgtg	8220
ccttcattga gaatcgggtg tgctgatttc aataaagtct tgagatcagt ttggccagtc	8280
atgttgtggg gggtaattgg attgagttat cgcctacagt ctgtacaggt atactcgtg	8340
cccactttat actttttgat tccgtgcac ttgaagcaat gtcgtttacc aaaagtgaga	8400
atgtccaca gaacacaccc cagggtatgg ttgagcaaaa aataaacact ccgatacggg	8460
gaatcgaacc ccggtctcca cggttctcaa gaagtattct tgatgagagc gtatcgatag	8520
ttggagcaag ggagaaatgt agagtgtgaa agactcacta tggccgggc ttatctcgac	8580
caatagccaa agtctggagt ttctgagaga aaaaggcaag atactgtatgt aacaaagcga	8640
cgcatgggtac aataataaccg gaggcattga tcatagagag ttagtggttc gatgatggca	8700
ctggtgcctg gtatgacttt atacggctga ctacatattt gtcctcagac atacaattac	8760
agtcaagcac ttacccttgg acatctgtag gtacccccg gccaaagcga tctcagcgtg	8820
tcgtatgtcg gattggcgta gctccctcgc tcgtcaattg gctcccatct actttcttct	8880
gcttggctac acccagcatg tctgtatgg ctggttttcg tgccttatct atcctcccag	8940
tattaccaac tctaaatgac atgatgtgat tgggtctaca ctttcatac agagataagg	9000
agtagcacag ttgcataaaa agcccaactc taatcagctt cttcctttct tgtaattagt	9060
acaaaggtga ttagcgaaat ctggaagctt agttggccct aaaaaaatca aaaaaagcaa	9120

aaaacgaaaa acgaaaaaac acagttttga gaacaggag gtaacgaagg atcgtatata	9180
tatatatata tatatatacc caggatccc gagaccggcc ttgtattctt ccctacaacc	9240
aaccattctc accaccctaa ttcacaacca tgtctattgg ttcgtccaac cccgtgctct	9300
tggtgcgat tcccttcgtc taccgtttg tctcccacg agtcctggct ttcctgcctc	9360
agaaggtca gttcctggcc aaatgtattg tggctctgat tgccacgctt atcatgtccg	9420
ttgcaggctg cttcatctcg atcgtgtgcg ctcttctgga caagagatac gtcataaatt	9480
acgttgtgtc gcgattgttc tcttccttg ccgctcgacc gtgtggtgtg acctataaga	9540
ttgttggatga ggaacacctc gataagtacc ctgctatcgt ggtctgtaac catcaatcct	9600
ctatggatat gatggttttg ggacgagttt ttccaaagca ctgctgttc atggcgaaga	9660
aggaaactct gtactttccc tttttgggaa tgtttatgaa actgagcaac gctatcttca	9720
tcgaccggaa gaaccacaag aaagccatcg agtctaccac ccaagccgtg gcggacatga	9780
agaagcacia ctctggaatc tggattttcc cagagggcac ccggtctaga ctggacaagg	9840
cagacctgct gcccttcaag aaaggtgcct ttcattctgc aattcaggcc cagctcccta	9900
ttctcccat tatctcgag ggctattccc atatctacga ctcttgaag cggctacttc	9960
ccggtggaga gctcgagatc agagtcctgg agccattcc tacaactggc ctactactg	10020
atgatgtgaa cgacctgatg gacaagacac gaaaccttat gctcaagcac ttgaaggaga	10080
tggtattcca gtattcgtc agcactgctg aaaatggatc cagcacatc gacgccgata	10140
ttgccaagtc tacagccacc agcattggca aactgacga cgcaattaca aaacgtcgta	10200
cccctaagga ataagcgcc gcattgatga ttggaacac acacatgggt tatatctagg	10260
tgagagttag ttggacagtt atatattaaa tcagctatgc caacggtaac ttcattcatg	10320
tcaacgagga accagtgact gcaagtaata tagaatttga ccaccttgcc attctcttgc	10380
actcctttac tatatctcat ttatttctta tataaaaac acttcttctt cccagcatcg	10440
agctcggaac cctcatgagc aataacatcg tggatctcgt caatagaggg ctttttggac	10500
tccttgcgtg tgccacctt gtccttgcgt tttaaaccat catctaaggg cctcaaaact	10560
acctcggaac tgcctgcgtg atctggacac cacagagggt ccgagcactt taggttgac	10620
caaatgtccc accaggtgca ggcagaaaac gctggaacag cgtgtacagt ttgtctaac	10680
aaaaagttag ggcgctgagg tcgagcaggg tgggtgtgact tgttatagcc tttagagctg	10740
cgaaagcgcg tatggatttg gctcatcagg ccagattgag ggtctgtgga cacatgtcat	10800
gttagtgtac ttcaatcgcc ccttgatata agccccgaca ataggccgtg gcctcatttt	10860
tttgcttcc gcacatttcc attgtcgggt acccacacct tgccttctct gcacttgcca	10920
accttaatac tggtttacct tgaccaacat cttacaagcg gggggcttgt ctagggtata	10980

tataaacagt ggctctccca atcggttgcc agtctctttt ttcttttctt tccccacaga	11040
ttcgaaatct aaactacaca tcacacatg gtctccaacc acctgttcga cgccatgcga	11100
gctgccgctc ccggagacgc acctttcatt cgaatcgaca acgctcggac ctggacttac	11160
gatgacgcca ttgtcttttc cggtcgaatc gctggagcta tggacgcact cggcattcga	11220
cccggagaca gatttgccgt gcaggctgag aagtctgccg aggcgttgat tctctacctg	11280
gcctgtcttc gaaccggagc tgtctacctg cctctcaaca ctgcctacac cctggccgag	11340
ctcgactact tcatcggcga tgccgaaccg cgtctggtgg tcgttgctcc cgcagctcga	11400
ggtggcgtgg agacaattgc caagcgacac ggtgctatcg tcgaaacct cgacgccgat	11460
ggacgaggct ccttgctgga ccttgctaga gatgagcctg ccgactttgt cgatgcttcg	11520
cgatctgccg acgatctggc tgctattctc tacacttccg gtacaaccgg acgatcgaag	11580
ggtgccatgc ttactcatgg caatctgctc tccaacgctc tcaccttgcg agactattgg	11640
agagttaccg cagacgatcg actcatccat gccttgccaa tctttcacac tcatggtctg	11700
ttcgttgcta cgaacgtcac actgcttgca ggagcctcga tgtttctgct ctccaagttc	11760
gatgccgacg aggtcgttct tctcatgcca caggccacca tgcttatggg cgtgccacac	11820
ttctacgttc gattgctgca gagtctcga ctcgagaagg gtgctgtggc cagcatcaga	11880
ctgttcattt ctggatcagc tcccttgctt gccgaaacct acgccgagtt tcatgctcgt	11940
actggtcacg ccattctcga gcgatacggc atgacggaaa ccaacatgaa tacttccaac	12000
ccctacgagg gcaagcgtat tgccggaacc gttggttttc ctctgcccgga cgtcactgtg	12060
cgagtcaccg atcccgccac cggctctggt cttccacctg aagagactgg catgatcgag	12120
atcaagggac ccaacgtctt caagggtat tggcgaatgc ccgaaaagac cgctgccgag	12180
tttaccgcag acggtttctt tatctctgga gatctcggca agatcgaccg agaaggttac	12240
gttcacattg tgggacgagg caaggacctg gtcatttccg gtggtacaaa catctatccc	12300
aaagaggtcg aaggcgagat cgaccagatc gaggggtgtg tcgagtctgc tgtcattggt	12360
gttctctatc ccgatttcgg agaaggtgtc accgctgttg tcgtgtgcaa acctggtgcc	12420
gttctcgacg aaaagacat cgtgtctgct ctgcaggacc gtcttgcccc atacaagcaa	12480
cccaagcgga ttatctttgc cgacgatctg cctcgaaaca ctatgggaaa ggttcagaag	12540
aacattcttc gacagcaata cgccgatctc tacaccagac gataagcggc cgcatgagaa	12600
gataaatata taatacatt gagatattaa atgcgctaga ttagagagcc tcatactgct	12660
cggagagaag ccaagacgag tactcaaagg ggattacacc atccatatcc acagacacaa	12720
gctggggaaa ggttctatat acattttccg gaataccgta gtttccgatg ttatcaatgg	12780
gggcagccag gatttcaggc acttcggtgt ctcggggtga aatggcgttc ttggcctcca	12840

tcaagtcgta ccatgtcttc atttgcctgt caaagtaaaa cagaagcaga tgaagaatga	12900
acttgaagtg aaggaattta aatgtagcta acggtagcag gcgaactact ggtacatacc	12960
tcccccgaa tatgtacagg cataatgcgt atctgtggga catgtggtcg ttgcgccatt	13020
atgtaagcag cgtgtactcc tctgactgtc catatggttt gctccatctc accctcatcg	13080
ttttcattgt tcacaggcgg ccacaaaaaa actgtcttct ctcttctct ctctgcctta	13140
gtctactcgg accagtttta gtttagcttg gcgccactgg ataaatgaga cctcaggcct	13200
tgtgatgagg aggtcactta tgaagcatgt taggaggtgc ttgtatggat agagaagcac	13260
ccaaaataat aagaataata ataaaacagg gggcgttgtc atttcataac gtgttttcac	13320
catcaataca cctccaaaca atgcccttca tgtggccagc cccaatattg tcctgtagtt	13380
caactctatg cagctcgtat cttattgagc aagtaaaact ctgtcagccg atattgcccg	13440
accgcgcaca agggteaaca aggtgggtgta aggccttcgc agaagtcaaa actgtgcca	13500
acaaacatct agagtctctt tgggtgttct cgcataatatt twatcggctg tcttacgtat	13560
ttgcgcctcg gtaccggact aatttcggat catccccaat acgcttttct ttcgcagctg	13620
tcaacagtgt ccatgatcta tccacctaaa tgggtcatat gaggcgtata atttcgtggt	13680
gctgataata attcccatat atttgacaca aaacttcccc ccctagacat acatctcaca	13740
atctcacttc ttgtgcttct gtcacacatc tcctccagct gacttcaact cacacctctg	13800
ccccagttgg tctacagcgg tataagggtt ctccgcatag aggtgcacca ctctcccga	13860
tacttgtttg tgtgacttgt gggtcacgac atatatatct acacacattg cgccaccctt	13920
tggttcttcc agcacaacaa aaacacgaca cgctaaccat ggagtctgga cccatgcctg	13980
ctggcattcc ctccctgag tactatgact tctttatgga ctggaagact cccctggcca	14040
tcgtgccac ctacactgtc gccgtcggtc tcttcaacct caaggttggc aaggtctccc	14100
gagtggttgc caagtcggct aacgcaaagc ctgccgagcg aaccagctcc ggagctgcca	14160
tgactgcctt cgtctttgtg cacaacctca ttctgtgtgt ctactctggc atcaccttct	14220
actacatgtt tcctgctatg gtcaagaact tccgaacca cactgcac gaagcctact	14280
gcgacacgga tcagtccctc tggaacaacg cacttggtta ctggggttac ctcttctacc	14340
tgtccaagtt ctacgaggtc attgacacca tcatcatcat cctgaaggga cgacggctct	14400
cgtctgttca gacctaccac catgctggag ccatgattac catgtggtct ggcatcaact	14460
accaagccac tccatttgg atctttgtgg tcttcaactc cttcattcac accatcatgt	14520
actgttacta tgccttcacc tctatcggtat tccatcctcc tggcaaaaag tacctgactt	14580

cgatgcagat tactcagttt ctggtcggta tcaccattgc cgtgtcctac ctcttcgttc 14640
ctggctgcat ccgaacaccc ggtgtctaga tggctgtctg gatcaacgtc ggctacctgt 14700
ttcccttgac ctatctgttc gtggactttg ccaagcgaac ctactccaag cgatctgcca 14760
ttgccgtca gaaaaaggct cagtaagcgg ccgcaagtgt ggatggggaa gtgagtgtccc 14820
ggttctgtgt gcacaattgg caatccaaga tggatggatt caacacaggg atatagcgag 14880
ctacgtggtg gtgcgaggat atagcaacgg atatttatgt ttgacacttg agaatgtacg 14940
atacaagcac tgtccaagta caatactaaa catactgtac atactcatac tcgtacccgg 15000

gcaacggttt cacttgagtg cagtggctag tgcctctact cgtacagtgt gcaatactgc 15060
glatcatagt ctttgatgta tatcgtattc attcatgtta gttgc 15105

<210> 90

<211> 828

<212> DNA

<213> Mortierella alpina

<220><221> CDS

<222> (1)..(828)

<223> synthetic C16/18 elongase (codon-optimized for Yarrowia
lipolytica)

<300><302> A MORTIERELLA ALPINA C16/18 FATTY ACID ELONGASE

<310> U.S. 7,470,532

<311> 2005-10-19

<312> 2008-12-30

<313> (1)..(828)

<400> 90

atg gag tct gga ccc atg cct gct ggc att ccc ttc cct gag tac tat 48

Met Glu Ser Gly Pro Met Pro Ala Gly Ile Pro Phe Pro Glu Tyr Tyr

1 5 10 15

gac ttc ttt atg gac tgg aag act ccc ctg gcc atc gct gcc acc tac 96

Asp Phe Phe Met Asp Trp Lys Thr Pro Leu Ala Ile Ala Ala Thr Tyr

20 25 30

act gct gcc gtc ggt ctc ttc aac ccc aag gtt ggc aag gtc tcc cga 144

Thr Ala Ala Val Gly Leu Phe Asn Pro Lys Val Gly Lys Val Ser Arg

gtg gtt gcc aag tcg gct aac gca aag cct gcc gag cga acc cag tcc	192
Val Val Ala Lys Ser Ala Asn Ala Lys Pro Ala Glu Arg Thr Gln Ser	
50 55 60	
gga gct gcc atg act gcc ttc gtc ttt gtg cac aac ctc att ctg tgt	240
Gly Ala Ala Met Thr Ala Phe Val Phe Val His Asn Leu Ile Leu Cys	
65 70 75 80	
gtc tac tct ggc atc acc ttc tac tac atg ttt cct gct atg gtc aag	288
Val Tyr Ser Gly Ile Thr Phe Tyr Tyr Met Phe Pro Ala Met Val Lys	
85 90 95	
aac ttc cga acc cac aca ctg cac gaa gcc tac tgc gac acg gat cag	336
Asn Phe Arg Thr His Thr Leu His Glu Ala Tyr Cys Asp Thr Asp Gln	
100 105 110	
tcc ctc tgg aac aac gca ctt ggc tac tgg ggt tac ctc ttc tac ctg	384
Ser Leu Trp Asn Asn Ala Leu Gly Tyr Trp Gly Tyr Leu Phe Tyr Leu	
115 120 125	
tcc aag ttc tac gag gtc att gac acc atc atc atc atc ctg aag gga	432
Ser Lys Phe Tyr Glu Val Ile Asp Thr Ile Ile Ile Ile Leu Lys Gly	
130 135 140	
cga cgg tcc tcg ctg ctt cag acc tac cac cat gct gga gcc atg att	480
Arg Arg Ser Ser Leu Leu Gln Thr Tyr His His Ala Gly Ala Met Ile	
145 150 155 160	
acc atg tgg tct ggc atc aac tac caa gcc act ccc att tgg atc ttt	528
Thr Met Trp Ser Gly Ile Asn Tyr Gln Ala Thr Pro Ile Trp Ile Phe	
165 170 175	
gtg gtc ttc aac tcc ttc att cac acc atc atg tac tgt tac tat gcc	576
Val Val Phe Asn Ser Phe Ile His Thr Ile Met Tyr Cys Tyr Tyr Ala	
180 185 190	
ttc acc tct atc gga ttc cat cct cct ggc aaa aag tac ctg act tcg	624
Phe Thr Ser Ile Gly Phe His Pro Pro Gly Lys Lys Tyr Leu Thr Ser	
195 200 205	
atg cag att act cag ttt ctg gtc ggt atc acc att gcc gtg tcc tac	672

Met Gln Ile Thr Gln Phe Leu Val Gly Ile Thr Ile Ala Val Ser Tyr
 210 215 220
 ctc ttc gtt cct ggc tgc atc cga aca ccc ggt gct cag atg gct gtc 720

Leu Phe Val Pro Gly Cys Ile Arg Thr Pro Gly Ala Gln Met Ala Val
 225 230 235 240
 tgg atc aac gtc ggc tac ctg ttt ccc ttg acc tat ctg ttc gtg gac 768
 Trp Ile Asn Val Gly Tyr Leu Phe Pro Leu Thr Tyr Leu Phe Val Asp
 245 250 255
 ttt gcc aag cga acc tac tcc aag cga tct gcc att gcc gct cag aaa 816
 Phe Ala Lys Arg Thr Tyr Ser Lys Arg Ser Ala Ile Ala Ala Gln Lys

260 265 270
 aag gct cag taa 828
 Lys Ala Gln
 275

<210> 91
 <211> 275
 <212> PRT
 <213> Mortierella alpina
 <400> 91

Met Glu Ser Gly Pro Met Pro Ala Gly Ile Pro Phe Pro Glu Tyr Tyr
 1 5 10 15

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

Val Tyr Ser Gly Ile Thr Phe Tyr Tyr Met Phe Pro Ala Met Val Lys
 85 90 95

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

Arg Arg Ser Ser Leu Leu Gln Thr Tyr His His Ala Gly Ala Met Ile
 145 150 155 160
 Thr Met Trp Ser Gly Ile Asn Tyr Gln Ala Thr Pro Ile Trp Ile Phe
 165 170 175
 Val Val Phe Asn Ser Phe Ile His Thr Ile Met Tyr Cys Tyr Tyr Ala
 180 185 190
 Phe Thr Ser Ile Gly Phe His Pro Pro Gly Lys Lys Tyr Leu Thr Ser
 195 200 205

Met Gln Ile Thr Gln Phe Leu Val Gly Ile Thr Ile Ala Val Ser Tyr
 210 215 220
 Leu Phe Val Pro Gly Cys Ile Arg Thr Pro Gly Ala Gln Met Ala Val
 225 230 235 240
 Trp Ile Asn Val Gly Tyr Leu Phe Pro Leu Thr Tyr Leu Phe Val Asp
 245 250 255
 Phe Ala Lys Arg Thr Tyr Ser Lys Arg Ser Ala Ile Ala Ala Gln Lys
 260 265 270

Lys Ala Gln
 275

<210> 92

<211> 13237

<212> DNA

<213> Artificial Sequence

<220><223> Plasmid pZKMP-ML9DP

<400> 92

gtacgactga tgtgtatgta agctcaatga gcagcgtgct cctcgggtcta ctatatggcg 60

acatatctct tcgcttctgt ttacccttta catacgggtac agctcttaca atgacattta 120

tccactggcg tcgatccata aaccacacga accctgtttt gttagtcacc atgaccgggg	180
cgtgtcgtgt tacgtttccac cgtttcacct cagccggtta cgattcaact tgccgcgtca	240
ttctgcgttg ctacgggagc gactacgagt aactagactt tcgataagct gaatgacttc	300
agggtgcatg agaggcgagc atcggatatt tcggacttgt tcttctagaa gactggtttg	360
agagcaatcg cggaagagtt gggaaccgct ggagagcaga atgggttagt acagagacta	420
tctgaictcc cccgcttggt tctagcccca ctttatactc tatttggcag ttgttgcttg	480
ttattcaagc acagcatgtc tgctcgtatc actttggaga catccctaac ctctaacct	540
ttatagatgc cttaccttct gagcttgagc agagcattcc acagcgaacc cagttgacat	600
tatcagttga catcttacgc tcacaccaca ttacgcatct cacacacaat cacacgcaca	660
agtacacacg caagcacatc atacaatggt caaggcgcgc cagctgcatt aatgaatcgg	720
ccaacgcgcg gggagaggcg gtttgcgtat tgggcgtctt tccgcttcct cgctcactga	780
ctcgtcgcgc tcggtcgttc ggctgcggcg agcgggtatca gctcactcaa aggcggtaat	840
acggttatcc acagaatcag gggataacgc aggaagaac atgtgagcaa aaggccagca	900
aaaggccagg aaccgtaaaa aggcgcggtt gctggcggtt tccataggc tccgcccccc	960
tgacgagcat caaaaaatc gacgtcaag tcagaggtgg cgaaaccga caggactata	1020
aagataccag gcgtttcccc ctggaagctc cctcgtgcgc tctcctgttc cgacctgcc	1080
gcttaccgga tacctgtccg cctttctccc ttccggaagc gtggcgcttt ctcatagctc	1140
acgtgtagg tatctcagtt cgggttaggt cgttcgtcc aagctgggct gtgtgcacga	1200
acccccgtt cagcccgacc gctgcgcctt atccggtaac tatcgtcttg agtccaacct	1260
ggtaagacac gacttatcgc cactggcagc agccactggg aacaggatta gcagagcgag	1320
gtatgtaggc ggtgctacag agttcttgaa gtggtggcct aactacggct aactagaag	1380
aacagtatth ggtatctcgc ctctgctgaa gccagttacc ttccgaaaaa gagttggtag	1440
ctcttgatcc ggcaaaaaa ccaccgtgg tagcgggtgt tttttgttt gcaagcagca	1500
gattacgcgc agaaaaaag gatctcaaga agatcctttg atcttttcta cggggtctga	1560
cgctcagtgg aacgaaaact cacgttaagg gatthtggc atgagattat caaaaaggat	1620
cttcacctag atcctthtaa attaaaaatg aagthttaa tcaatctaaa gtatatatga	1680
gtaaacttgg tctgacagtt accaatgctt aatcagtgag gcacctatct cagcgatctg	1740
tctatthcgt tcatccatag ttgcctgact ccccgctgtg tagataacta cgatacggga	1800
gggcttacca tctggcccca gtgctgcaat gataccgca gaccacgct caccgctcc	1860
agatttatca gcaataaacc agccagccgg aaggcccgag cgcagaagtg gtcctgcaac	1920

tttatccgcc tccatccagt ctattaattg ttgccgggaa gctagagtaa gtagttcgcc	1980
agttaatagt ttgcgcaacg ttgttgccat tgctacaggc atcgttggtg cagctcgtc	2040
gtttggtatg gcttcattca gctccggttc ccaacgatca aggcgagtta catgatcccc	2100
catgttgtgc aaaaaagcgg ttagctcctt cggtcctccg atcgttgtca gaagtaagtt	2160
ggccgcagtg ttatcactca tggttatggc agcactgcat aattctctta ctgtcatgcc	2220
atccgtaaga tgcttttctg tgactgggtga gtactcaacc aagtcattct gagaatagtg	2280
tatgcggcga ccgagttgct cttgcccggc gtcaatacgg gataataccg cgccacatag	2340
cagaacttta aaagtgtca tcattggaaa acgttcttcg gggcgaaaac tctcaaggat	2400
cttaccgctg ttgagatcca gttcgatgia acccactcgt gcaccaact gatcttcagc	2460
atcttttact ttcaccagcg tttctgggtg agcaaaaaca ggaaggcaaa atgccgcaaa	2520
aaagggaata agggcgacac ggaaatgttg aatactcata ctcttctttt ttcaatatta	2580
ttgaagcatt tatcagggtt attgtctcat gagcggatac atatttgaat gtatttagaa	2640
aaataaacia ataggggttc cgcgacatt tccccgaaa gtgccacctg atgcggtgtg	2700
aaataccgca cagatgcgta aggagaaaat accgcatcag gaaattgtaa gcgttaatat	2760
tttgttaaaa ttcgcgttaa atttttgtta aatcagctca tttttaacc aataggccga	2820
aatcgcaaa atcccttata aatcaaaaga atagaccgag atagggttga gtgttgttcc	2880
agtttgaac aagagtccac tattaagaa cgtggactcc aacgtcaaag ggcgaaaaac	2940
cgtctatcag ggcatggcc cactacgtga accatcccc taatcaagtt ttttggggtc	3000
gaggtgccgt aaagcaacta atcggaacct taaagggagc ccccgattta gagcttgacg	3060
gggaaagccg gcgaacgtgg cgagaaagga agggaagaaa gcgaaaggag cgggcgctag	3120
ggcgttgcca agttagtcgg tcacgtcgcg cgtaaccacc acaccgccg cgcttaatgc	3180
gccgctacag ggcgctcca ttcccatc aggctgcgca actgttggga agggcgatcg	3240
gtgcgggctt ctctctatt acgccagctg gcgaaagggg gatgtgctgc aaggcgatta	3300
agttgggtaa cgccagggtt tteccagtca cgacgttgta aaacgacggc cagtgaattg	3360
taatacgact cactataggc cgaattgggc ccgacgtcgc atgctaggag agtctgtcgt	3420
tcgcgagtgg cgcgccaatg tcatgagtta cacctgattc aaatagtga taataatata	3480
acagccacaa atgaaagatg taccacggg tagatatggc ttaattact gatgattgaa	3540
tgattatctg agtgctacag ttgtaacgag tacgagttat tacctgctac tattgtactc	3600
ctaaatcaaa gtactcgtac atgcaatgaa ttacttgtac ttgcgccaat cagatctatc	3660

ggttgatttg tatataatac tagtctattc ttttgcttac cagatggtgt aacctccatc 3720
 gacaatgaga tcgtggccgg tcatgtagct ggaggcgtcg gaagccatga agacaatagg 3780
 gcctccaaac tctcgggct cagccattcg tcggaagga attcgctcgt cccagtcgtt 3840
 tcggagctct ccgtcagct cgatgatgtg acgggtcaga ggagtaagaa tgtagccggg 3900
 agacagagtg ttgactcgga tgttgtactg ggcccactcg gcagccagag acttgccat 3960
 gtggatgaca cctgccttgg acatgttgta gggggtctgg ggctgggggt cgttgacaat 4020
 ggagccactc atggatgcaa ccagaatcag agatccagga gactcgtcca ggacaagcga 4080

 cttgccacc tcggtagcga agtttagga tcccatgagg ttgacgtcca gcagtcgctt 4140
 gacctggggg ttgggataat caaaggcagt catattctcg cagaatccgg cacagttaat 4200
 taagttgcga cacatgtctt gatagtatct tgaattctct ctcttgagct tttccataac 4260
 aagttcttct gcctccagga agtccatcgg tggtttgatc atggttttgg tgtagtggta 4320
 gtgcagtggg ggtatttgta ctggggatgt agttgagaat aagtcataca caagtcagct 4380
 ttcttcgagc ctcatataag tataagtagt tcaacgtatt agcactgtac ccagcatctc 4440
 cgtatcgaga aacacaacaa catgccccat tggacagatc atgcggatac acaggttgtg 4500

 cagtatcata catactcgat cagacaggtc gtctgaccat catacaagct gaacaagcgc 4560
 tccatacttg cacgtctct atatacacag ttaaattaca tatccatagt ctaacctcta 4620
 acagttaate ttctggtaag cctccagcc agccttctgg tatcgcttgg cctcctcaat 4680
 aggatctcgg ttctggccgt acagacctcg gccgacaatt atgatatccg ttccggtaga 4740
 catgacatcc tcaacagttc ggtactgtg tccgagagcg tctcccttgt cgtcaagacc 4800
 caccgccggg gtcagaataa gccagtcctc agagtcgccc ttaggtcggg tctgggcaat 4860
 gaagccaacc acaaactcgg ggtcggatcg ggcaagctca atggtctgct tggagtactc 4920

 gccagtggcc agagagccct tgcaagacag ctggccagc atgagcagac ctctggccag 4980
 ctctcgttg ggagagggga ctaggaaactc ctgtactgg gagttctcgt agtcagagac 5040
 gtctcttc ttctgttcag agacagtctc ctggcacca gctcgaggc cagcaatgat 5100
 tccggttccg ggtacacgt gggcgttgg gatatcggac cactcggcga ttcggtgaca 5160
 ccggtactgg tgcttgacag tgttgccaat atctgcgaac tttctgtcct cgaacaggaa 5220
 gaaacctgtc ttaagagcaa gtctcttgag ggggagcaca gtgccggcgt aggtgaagtc 5280
 gtcaatgatg tcgatatggg ttttgatcat gcacacataa ggtccgacct tatcggcaag 5340

 ctcaatgagc tccttggttg tggtaacatc cagagaagca cacaggttgg tttcttggc 5400
 tgccacgagc ttgagcactc gagcggcaaa ggcggacttg tggacgttag ctcgagcttc 5460
 gtaggagggc attttggttg tgaagaggag actgaaataa attagtctg cagaactttt 5520

tatcggaacc ttatctgggg cagtgaagta tatgttatgg taatagttac gagttagtgtg	5580
aacttataga tagactggac tatacggcta tcggtccaaa ttgaaagaa cgtcaatggc	5640
tctctgggcg tcgcctttgc cgacaaaaat gtgatcatga tgaaagccag caatgacgtt	5700
gcagctgata ttgttgctcg ccaaccgcgc cgaaaacgca gctgtcagac ccacagcctc	5760
caacgaagaa tgtatcgta aagtgatcca agcacactca tagttggagt cgtactccaa	5820
aggcggcaat gacgagtcag acagatactc gtcgaccttt tccttgggaa ccaccaccgt	5880
cagcccttct gactcacgta ttgtagccac cgacacaggc aacagtccgt ggatagcaga	5940
atatgtcttg tcggtccatt tctcaccaac tttaggcgtc aagtgaatgt tgcagaagaa	6000
glatgtgcct tcattgagaa tcggtgttgc tgatttcaat aaagtcctga gatcagtttg	6060
gccagtcatg ttgtgggggg taattggatt gagttatcgc ctacagtctg tacaggtata	6120
ctcgtgccc actttatact ttttgattcc gctgcacttg aagcaatgtc gtttaccaaa	6180
agtgagaatg ctccacagaa cacaccccag gglatgggtg agcaaaaaat aaacactccg	6240
atacggggaa tcgaaccccg gtctccacgg ttctcaagaa gtattcttga tgagagcgta	6300
tcgatggtta atgctgctgt gtgctgtgtg tgtgtgtgt ttggcgctca ttgttgcgtt	6360
atgcagcgta caccacaata ttggaagctt attagccttt ctatttttgc gtttgcaagg	6420
cttaacaaca ttgctgtgga gagggatggg gatatggagg ccgctggagg gagtccgaga	6480
ggcgttttgg agcggcttgg cctggcgccc agctcgcgaa acgcacctag gaccctttgg	6540
cacgccgaaa tgtgccactt ttcagtctag taacgcctta cctacgtcat tccatgcgtg	6600
catgtttgcg ccttttttcc cttgcccttg atcgccacac agtacagtgc actgtacagt	6660
ggaggttttg ggggggtctt agatgggagc taaaagcggc ctagcggtag actagtggga	6720
ttgtatggag tggcatggag cctaggtgga gcctgacagg acgcacgacc ggctagcccc	6780
tgacagacga tgggtggctc ctgttgtcca ccgctacaa atgtttgggc caaagtcttg	6840
tcagccttgc ttgcgaacct aattccaat ttgtcactt cgcaccccca ttgatcgagc	6900
cctaaccctt gccatcagg caatccaatt aagctcgcct tgtctgcctt gtttagtttg	6960
gtccttgcgc gtttcggcgt ccacttgcac aaacacaaac aagcattata tataaggctc	7020
gtctctccct cccaaccaca ctacttttt tgcccgtctt cctttgctaa cacaaaagtc	7080
aagaacacaa acaaccacce caacccctt acacacaaga catactaca gcaatggcca	7140
tggccaaaag caaacgacgg tcggaggctg tggaagagca cgtgaccggc tcggacgagg	7200
gcttgaccga tacttcgggt cacgtgagcc ctgccgcaa gaagcagaag aactcggaga	7260
ttcatttcac caccaggtt gccacgagt tggatcggga gcgcaaggag gagtatctgg	7320
actcgtgat cgacaacaag gactatctca agtaccgtcc tcgaggctgg aagctcaaca	7380

accgcctac cgaccgacct gtgcgaatct acgccgatgg agtgtttgat ttgttccatc 7440

 tgggacacat gcgtcagctg gagcagtcga agaaggcctt cccaacgca gtgttgattg 7500
 tgggcattcc cagcgacaag gagacccaca agcgggaaggg attgaccgtg ctgagtgacg 7560
 tccagcggta cgagacggtg cgacactgca agtgggtgga cgaggtgggtg gaggatgctc 7620
 cctgggtgtgt caccatggac tttctggaaa aacacaaaat cgactacgtg gcccatgacg 7680
 atctgcccta cgcttccggc aacgacgatg atatctacaa gcccatcaag gagaaggga 7740
 tgtttctggc caccagcga accgagggca tttccacctc ggacatcatc accaagatta 7800
 tccgagacta cgacaagtat ttaatgcgaa actttgcccg gggtgctaac cgaaaggatc 7860

 tcaacgtctc gtggctcaag aagaacgagc tggacttcaa gcgtcatgtg gccgagttcc 7920
 gaaactcgtt caagcgaaag aaggctcggtg aggatctcta cggcgagatt cgcggtctgc 7980
 tgcagaatgt gctcatttgg aacggcgaca actccggcac ttcactccc cagcgaaaga 8040
 cgctgcagac caacgccaag aagatgtaca tgaacgtgct caagactctg caggctcctg 8100
 acgtgttga cgtggactcc tcggagaacg tgtctgagaa cgtcactgat gaggaggagg 8160
 aagacgacga cgaggttgat gaggacgaag aagccgacga cgacgacgaa gacgacgaag 8220
 acgaggaaga cgacgagtag gcggccgcat tgatgattgg aaacacacac atgggttata 8280

 tctaggtgag agttagttgg acagttatat attaaatcag ctatgccaac ggtaacttca 8340
 ttcatgtcaa cgaggaacca gtgactgcaa gtaatataga atttgaccac cttgccattc 8400
 tcttgactc ctttactata tctcatttat ttcttatata caaatcactt cttcttccca 8460
 gcatcgagct cggaaacctc atgagcaata acatcgtgga tctcgtcaat agagggtttt 8520
 ttggactcct tgctgttggc caccttgtcc ttgctgttta aaccacgcca actgatatcc 8580
 ccttacgtta cccctcatc acctggtgag gcgaaactgt aagtgaaag ctaaaaatga 8640
 catctcagct gcacgaagga cgggggctta aaagacgggc tgggtgcttgat gatitaaaac 8700

 tggacaaatc tcagcttgct tgaaattttg gtctccaact gtttccgagc gaatcgacaa 8760
 caaacgggc tttctctctc agaccacgcc cccgaaactc tttctccac caccaccaac 8820
 actcccttc cattccaca ccgttctctc ctctgtcgtc atccttgctc aatcatcttc 8880
 gtctgcgaca tattgtacga catacagtag cacggaacgt ttcagaccgt cgacgtgaac 8940
 acatcttagg aacagcaacc tgagctacag aaatctatct ataggcggt aaaaaaacgc 9000
 accactgct cgtcctcctt gctcctcgaa accgactcct ctacacacgt caaatccgag 9060
 gttgaaatct tccccacatt tggcagccaa accagcacat cccagcaacc tcgcacagcg 9120

ccgaaatcga cctgtcgact tggccacaaa aaaaagcacc ggctctgcaa cagttctcac	9180
gaccaattac gtacaagtac gaaatcggtt gtggaccgtg actgataagc tcccactttt	9240
tcttctaaca acaggcaaca gacaagtcac aaaaaacaaa agccatggtg aaaaacgtgg	9300
accaagtggg tctctcgagc gtcgacacca ttgcctccgg ccgagatgtc aactacaagg	9360
tcaagtacac ctccggcggt aagatgagcc agggcgcccta cgacgacaag ggccgccaca	9420
tttccgagca gcccttcacc tgggccaact ggccaccagca catcaactgg ctcaacttca	9480
ttctgggtgat tgcgctgcct ctgtcgtcct ttgctgccgc tcccttcgtc tccttcaact	9540
ggaagaccgc cgcgttttct gtccgctatt acatgtgcac cggctcctgg atcaccgccg	9600
gctaccaccg aatgtgggcc catcgagcct acaaggccgc tctgcccgtt cgaatcatcc	9660
ttgctctgtt tggaggagga gctgtcgagg gctccatccg atggtgggcc tcgtctcacc	9720
gagtccacca ccgatggacc gactccaaca aggaccctta cgacgcccga aagggtattct	9780
ggttctccca ctttggctgg atgctgcttg tgcccaaccc caagaacaag ggccgaactg	9840
acatttttga cctcaacaac gactgggttg tccgactcca gcacaagtac tacgtttacg	9900
ttctcgtctt catggccatt gtcttgccca cctcgtctg tggctttggc tggggcgact	9960
ggaagggagg tctgtctac gccggtatca tgcgatacac ctttgtgcag caggtgactt	10020
tctgtgtcaa ctcccttgcc cactggattg gagagcagcc cttcgacgac cgacgaactc	10080
cccagacca cgtctttacc gccctgggtca ctttggaga gggctaccac aacttcacc	10140
acgagttccc ctccgactac cgaaacgccc tcactctggtta ccagtcagac cccaccaagt	10200
ggctcatctg gacctcaag caggttggtc tcgctggga cctccagacc ttctccaga	10260
acgccatcga gcagggtctc gtgcagcagc gacagaagaa gctggacaag tggcgaaaca	10320
acctcaactg gggatatccc attgagcagc tgcctgtcat tgagtttgag gatttccaag	10380
agcaggccaa gacccgagat ctggttctca tttctggcat tgtccacgac gtgtctgcct	10440
ttgtcgagca ccacctgggt ggaaaggccc tcattatgag cgccgtcggc aaggacggta	10500
ccgtgtctt caacggaggt gtctaccgac actccaacgc tggccacaac ctgcttgcca	10560
ccatgcgagt ttcggtcatt cgaggcggca tggaggttga ggtgtggaag actgccaga	10620
acgaaaagaa ggaccagaac attgtctccg atgagagtgg aaaccgaatc caccgagctg	10680
gtctccaggc caccgggtc gagaaccccg gtatgtctgg catggctgct taggcggccg	10740
catgagaaga taaatatata aatacattga gatattaat gcgctagatt agagagcctc	10800
atactgctcg gagagaagcc aagacgagta ctcaaagggg attacacat ccatatccac	10860
agacacaagc tggggaaagg ttctatatac actttccgga ataccgtagt ttccgatgtt	10920
atcaatgggg gcagccagga tttcaggcac ttcggtgtct cggggtgaaa tggcgttctt	10980

ggcctccatc aagtcgtacc atgtcttcat ttgcctgtca aagtaaaaca gaagcagatg	11040
aagaatgaac ttgaagtga ggaatttaaa ttttcgagat tttacagata tttctcgag	11100
tttttcacgt ccccttgcc ttgtcctatt gtttcaaata aactctcgtc tactgatttc	11160
acatggaacc ttigtatatt cggggataac ccccttgcc attgcacgat ggacgtggca	11220
aaagaaagat cgccctgcgg ggatacttat catgtggtca catgctgtga ttagaaataa	11280
agaaaaaggt gcttttttgg cgctgtgatt aacatctcgt ctgccgtgct ctactagtcg	11340
caatagcaaa aactcgctta atagtgtgca tagtgcgggg tagcaggata ctgaactaca	11400
gtacgatttg cttgctactg cttgtagcaa ttacctttac tgtagggacc acacctcctg	11460
gtttcaatgt ctttctcgcc ctgcacaaag caaaactgtc acccaatcac accttgttca	11520
tattcattag tgcatccgtt aaccttgaca tgacacttct catactagtg atagggtgt	11580
agttgagaca agttgattca caggataca tacaagcct cagagagcaa atgttatata	11640
ctcagggacc gaccaatcaa aaaaacacac tcctaataac caccatttcc atctacgct	11700
actcactctg tcagctgccc cacattgccc aatgcacaat gcacaatgat gtgtgcaaac	11760
aacgcaatca aaagtctatg gatgtgacc aaactctgat caccaagttg cgaacatgaa	11820
aaagaagacc tgtgtatata taagtaaggg ggagagccct aactagatct ttcgaaaacc	11880
ccccgacctt caccttcac aacctgatc atcttatacg ttttgccgt tgcggtctcc	11940
ttctcatct tcaagagagt cacctacacc atggcttcta ttggttcgt caaccccggtg	12000
ctcttggtg cgattccctt cgtctacctg ttgtctctcc cagagtcct ggctttcctg	12060
cctcagaagg ctgatttctt ggccaaatgt attgtgttcc tgattgccac gcttatcatg	12120
tccgttgtag gctgcttcat ctgatcgtg tgcgtcttc tggacaagag atacgtcatc	12180
aattacgttg tgtcgcgatt gttctcttc ctggccgtc gaccgtgtgg tgtgacctat	12240
aagattgttg gtgaggaaca cctcgataag taccctgcta tcgttgtctg taacctcaa	12300
tcctctatgg atatgatgtt ttgggacga gtttttcaa agcactgcgt tgtcatggcg	12360
aagaaggaac tctgtactt tccctttttg ggaatgttta tgaaactgag caacgtatc	12420
ttcatcgacc ggaagaacca caagaaagcc atcgagtcta ccaccaagc cgtggcggac	12480
atgaagaagc acaactctgg aatctggatt ttcccagagg gcacccggtc tagactggac	12540
aaggcagacc tctgtccctt caagaaaggt gcctttcacc ttgcaattca ggcccagctc	12600
cctattctcc ccattatctc gcagggtat tcccatatct acgactcttc gaagcggtac	12660
ttccccggtg gagagctcga gatcagagtc ctggagccca ttcttacaac tggcctcact	12720
actgatgatg tgaacgacct gatggacaag acacgaaacc ttatgtctaa gcacttgaag	12780
gagatggatt cccagtatc gtcgagcact gctgaaaatg gatccacgca catcgacgcc	12840

gatattgcc a gcttacagc caccagcatt ggcaacactg acgacgcaat taaaaacgt 12900

cgtacccta aggaataagc ggccgcaagt gtggatgggg aagtgagtgc ccggttctgt 12960

gtgcacaatt ggcaatccaa gatggatgga ttcaacacag ggatatagcg agctacgtgg 13020

tgggtcgagg atatagcaac ggatatattat gtttgacact tgagaatgta cgatacaagc 13080

actgtccaag tacaatacta aacatactgt acatactcat actcgtaacc gggcaacggt 13140

ttcacttgag tgcagtggct agtgccttta ctctacagt gtgcaatact gcgtatcata 13200

gtctttgatg tatatcgtat tcattcatgt tagttgc 13237

<210> 93

<211> 899

<212> DNA

<213> Yarrowia lipolytica

<400> 93

tttcgagatt ttacagatat ttctcgcagt ttttcacgtc ccttctcct tgcctattg 60

tttcaataa actctcgtct actgatttca catggaacct ttgctatttc ggggataacc 120

ccctttgcc ttgcacgatg gacgtggcaa aagaaagatc gccctgcggg gatactatc 180

atgtggtcac atgctgtgat tagaaataa gaaaaagggtg cttttttggc gctgtgatta 240

acatctcgtc tgccgtgctc tactagtgc aatagcaaaa actcgcttaa tagtgtcat 300

agtgcggggt agcaggatag tgaactacag tacgatttgc ttgctactgc ttgtagcaat 360

tacctttact gtagggacca cacctcctgg tttcaatgtc tttcctcgcc tcgacaaagc 420

aaaactgtca ccaatcaca ccttgttcat attcattagt gcatccgtta accttgacat 480

gacacttctc atactagtga tagggctgta gttgagacaa gttgattcac acggatacat 540

acaaagcctc agagagcaaa tggttatatac tcagggaccg accaatcaaa aaaacacact 600

cctaataacc accatttcca tctacgcgtc ctactctgt cagctgcccc acattgccca 660

atgcacaatg cacaatgatg tgtgcaaaac acgcaatcaa aagtctatgg atgctgacca 720

aactctgac accaagttgc gaacatgaaa aagaagacct gtgtatatat aagtaagggg 780

gagagcccta actagatctt tcgaaaaccc cccgaccttc accttcaca accatgatca 840

tcttatacgt tttggccgtt gcggctctct tctctatctt caagagagtc acctacacc 899

<210> 94

<211> 722

<212> DNA

<213> Yarrowia lipolytica

<400> 94

ccacgccaac tgatatcccc ttacgttacc cctcatcac ctggtgaggc gaaactgtaa 60
 ggtgaaagct aaaaatgaca tctcagctgc acgaaggacc ggggcttaaa agacgggctg 120
 gtgcttgtga tttaaaactg gacaaatctc agcttgcttg aaattttggt ctccaactgt 180
 ttccgagcga atcgcacaca aaccgggctt ctctctgcag accacgcccc cgaaactctt 240
 tctcccacca ccaccaaac tccctttcca tcccacacc gttcctctct cgtgcgtcat 300

ccttgcgcaa tcatcttctg ctgcgacata ttgtacgaca tacagtacca cggaacgttt 360
 cagaccgtcg acgtgaacac atcttaggaa cagcaacctg agctacagaa atctatctat 420
 aggcgataa aaaaacgcac ccactgctcg tctccttgc tctcgaaac cgactcctct 480
 acacacgtca aatccgaggt tgaaatcttc cccacatttg gcagccaaac cagcacatcc 540
 cagcaacctc gcacagcgcc gaaatcgacc tgtcgacttg gccacaaaa aaagcaccgg 600
 ctctgcaaca gtctcacga ccaattacgt acaagtacga aatcgttctg ggaccgtgac 660
 tgataagctc ccactttttc ttctaacaac aggcaacaga caagtcacac aaaacaaaag 720

cc 722

<210> 95

<211> 13320

<212> DNA

<213> Artificial Sequence

<220><223> Plasmid pZKMP-ML9DCB

<400> 95

cgatggttaa tgctgctgtg tgctgtgtgt gtgtgttgtt tggcgctcat tgttgcgtta 60
 tgcagcgtag accacaatat tggaagctta ttagcctttc tattttttcg ttgcaaggc 120
 ttaacaacat tgctgtggag agggatgggg atatggaggc cgctggaggg agtcggagag 180
 gcgttttgga ggggcttggc ctggcgccca gctcgcaaaa cgcacctagg accctttggc 240
 acgccgaaat gtgccacttt tcagtctagt aacgccttac ctacgtcatt ccatgcgtgc 300

atgtttgcgc cttttttccc ttgcccttga tcgccacaca gtacagtga ctgtacagtg 360
 gaggttttgg gggggtctta gatgggagct aaaagcggcc tagcggtaga ctagtgggat 420
 tgtatggagt ggcatggagc ctaggtggag cctgacagga cgcacgaccg gctagcccg 480
 gacagacgat ggggtggctcc tgttgtccac cgcgtacaaa tgtttgggcc aaagtcttgt 540
 cagccttget tgcgaacctt attcccatt ttgtcacttc gcaccccat tgatcgagcc 600
 ctaaccctg cccatcaggc aatccaatta agctcgcat gtctgccttg tttagtttgg 660

ctcctgcccc ttctggcgtc cacttgcaca aacacaaaca agcattatat ataaggctcg	720
tctctccctc ccaaccacac tcactttttt gcccgctctc ccttgctaac acaaaagtca	780
agaacacaaa caaccacccc aaccccccta cacacaagac atatctacag caatggccat	840
gggcgtattc attaaacagg agcagcttcc ggctctcaag aagtacaagt actccgccga	900
ggatcactcg ttcatctcca acaacattct gcgccccctc tggcgacagt ttgtcaaat	960
cttcctctg tggatggccc ccaacatggt gactctgctg ggcttcttct ttgtcattgt	1020
gaacttcac accatgctca ttgtgatcc caccacgac cgcgagcctc ccagatgggt	1080
ctacctcacc tacgtctcgg gtctgttctt ttaccagaca ttgatgcct gtgacggatc	1140
ccatgccccga cgaactggcc agatgggacc ccttgagag ctgtttgacc actgtgtcga	1200
cgccatgaat acctctctga ttctcacggt ggtgggtgtcc accaccata tgggatataa	1260
catgaagcta ctgattgtgc agattgccgc tctcggaac ttctacctgt cgacctggga	1320
gacctacat accggaactc tgtacctttc tggcttctct ggtcctgttg aaggtactt	1380
gattctgggt gctcttttcg tctcacctt cttcactggt cccaacgtgt acgctctgac	1440
cgtctacgag gctcttcccg agtccatcac ttcgctgctg cctgccagct tctggacgt	1500
caccatcacc cagatctaca ttggattcgg agtgcctggc atggtgttca acatctacgg	1560
cgctcgcca aacgtgatca agtactacaa caacaagggc aagagcgctc tccccccat	1620
tctcggaatc gcccccttg gcattctta cgtcgcgctc ttgcctggg cccatgttgc	1680
tcctctgctt ctctccaagt acgccatcgt ctatctgttt gccattgggg ctgcctttgc	1740
catgcaagtc ggccagatga ttcttgccca tctcgtgctt gctcccttc cccactggaa	1800
cgtgctgctc ttcttccct ttgtgggact ggcatgacac tacattgcac ccgtgtttgg	1860
ctgggacgcc gatatcgtg cggtaaacac tctcttcacc tgttttggcg ccacctctc	1920
catttacgcc ttctttgtgc ttgagatcat cgacgagatc accaactacc tcgatatctg	1980
gtgtctgca atcaagtacc ctgaggagaa gaagaccgaa taagcgccg catgtacata	2040
caagattatt tatagaaatg aatcgcatc gaacaaagag tacgagtgt cagtagggg	2100
atgatgataa aagtgaaga agttccgcat ctttgattt atcaacgtgt aggacgatac	2160
ttcctgtaaa aatgcaatgt ctttaccata ggttctgctg tagatgttat taactacat	2220
taacatgtct actgtacag ttgcagacca gttggagtat agaatggtac acttacaaa	2280
aagtgtgat ggtgtgaact acgatataa aaactgtga cgggatctgc gtacactgtt	2340
taaaccacgc caactgatat ccccttacgt taccctca tcacctggtg aggcgaaact	2400

gtaaggtgaa agctaaaaat gacatctcag ctgcacgaag gaccggggct taaaagacgg 2460
 gctggtgctt gtgattttaa actggacaaa tctcagcttg cttgaaattt tggcttccaa 2520
 ctgtttccga gcgaatcgca cacaaccgg gcttctctct gcagaccacg cccccgaaac 2580
 tctttctccc accaccacca acatccctt tccattccca caccgttctt ctctcgtgcg 2640
 tcatccttgc gcaatcatct tcgtctgcga catattgtac gacatacagt accacggaac 2700
 gtttcagacc gtcgacgtga acacatctta ggaacagcaa cctgagctac agaaatctat 2760
 ctataggcgg ataaaaaac gcacccactg ctgctctcc ttgctcctcg aaaccgactc 2820

 ctctacacac gtcaaatccg aggttgaaat ctccccaca tttggcagcc aaaccagcac 2880
 atcccagcaa cctcgacag cgccgaaatc gacctgtcga cttggccaca aaaaaagca 2940
 ccggctctgc aacagtctc acgaccaatt acgtacaagt acgaaatcgt tcgtggaccg 3000
 tgactgataa gctcccactt tttcttctaa caacaggcaa cagacaagtc acacaaaaca 3060
 aaagccatgg tgaaaaacgt ggaccaagt gactctctgc aggtcgacac cattgcctcc 3120
 ggccgagatg tcaactacaa ggtcaagtac acctccggcg ttaagatgag ccagggcgcc 3180
 tacgacgaca agggccgcca catttccgag cagcccttca cctgggcca ctggcaccag 3240

 cacatcaact ggctcaactt cattctggtg attgcgctgc ctctgtctc ctttgcctgc 3300
 gctcccttcg tctccttcaa ctggaagacc gccgcgttg ctgtcgcta ttacatgtgc 3360
 accggtctcg gtatcacgc cggtaccac cgaatgtggg cccatcgagc ctacaaggcc 3420
 gctctgcccc ttcaatcat ccttgcctcg tttggaggag gagctgtcga gggctccatc 3480
 cgatgggtgg cctcgtctca ccgagtcac caccgatgga ccgactcaa caaggacct 3540
 tacgacgcc gaaagggtt ctggttctcc cactttggct ggatgctgct tgtgcccaac 3600
 cccaagaaca agggccgaac tgacatttct gacctcaaca acgactgggt tgtccgactc 3660

 cagcacaagt actacgttta cgttctctc ttcattggcca ttgttctgcc caccctctc 3720
 tgtggctttg gctggggcga ctggaaggga ggtcttgtct acgccggtat catgcgatac 3780
 acctttgtgc agcaggtgac tttctgtgtc aactcccttg cccactggat tggagagcag 3840
 cccttcgacg accgacgaac tccccgagac cacgtctta ccgccctggt cacctttgga 3900
 gagggtacc acaacttcca ccacagttc cctcggact accgaaacgc cctcatctgg 3960
 taccgtacg accccacca gtggctcatc tggacctca agcaggttgg tctcgctgg 4020
 gacctcaga ctttctccca gaacccatc gagcagggtc tcgtgcagca gcgacagaag 4080

 aagctggaca agtggcgaaa caacctcaac tggggtatcc ccattgagca gctgcctgtc 4140
 attgagttt aggagttcca agagcaggcc aagaccgag atctggttct cttttctggc 4200
 attgtccacg acgtgtctgc ctttctcgag caccacctg gtggaaaggc cctcattatg 4260

agcgccgtcg gcaaggacgg taccgtgtc ttcaacggag gtgtctaccg acactccaac 4320
 gctggccaca acctgcttgc caccatgcga gtttcgggtca ttcgagggcg catggaggtt 4380
 gaggtgtgga agactgcccc gaacgaaaag aaggaccaga acattgtctc cgatgagagt 4440
 ggaaaccgaa tccaccgagc tgggtctccag gccacccggg tcgagaaccc cggtatgtct 4500

 ggcatggctg cttaggcggc cgcatgagaa gataaatata taaatacatt gagatattaa 4560
 atgcgctaga ttagagagcc tcatactgct cggagagaag ccaagacgag tactcaaagg 4620
 ggattacacc atccatatcc acagacacaa gctggggaaa ggttctatat acactttccg 4680
 gaataccgta gtttccgatg ttatcaatgg gggcagccag gatttcaggc acttcggtgt 4740
 ctgggggtga aatggcggtc ttggcctcca tcaagtcgta ccatgtcttc atttgctgt 4800
 caaagtaaaa cagaagcaga tgaagaatga acttgaagtg aaggaattta aattttcgag 4860
 attttacaga tatttctcgc agtttttcac gtcccttgt ccttgctcta ttgtttcaaa 4920

 taaactctcg tctactgatt tcacatggaa cctttgctat ttcggggata accccctttg 4980
 ccattgcacg atggacgtgg caaaagaaag atcgccctgc ggggatactt atcatgtggt 5040
 cacatgctgt gattagaaa aaagaaaaag gtgctttttt ggcgctgtga ttaacatctc 5100
 gtctgccgtg ctctactagt cgcaatagca aaaactcgct taatagtgtg catagtgcgg 5160
 ggtagcagga tactgaacta cagtacgatt tgcttgctac tgcttgtagc aattaccttt 5220
 actgtaggga ccacacctcc tggtttcaat gtctttctc gcctcgacaa agcaaaactg 5280
 tcaccaatc acaccttggt catattcatt agtgcacccg ttaacctga catgacactt 5340

 ctcatactag tgatagggt gtagttgaga caagttgatt cacacggata catacaaagc 5400
 ctgagagagc aaatgttata tactcaggga ccgaccaatc aaaaaaacac actcctaata 5460
 accaccattt ccactctacg gtactcactc tgcagctgc cccacattgc ccaatgcaca 5520
 atgcacaatg atgtgtgcaa acaacgcaat caaaagtcta tggatgctga ccaaactctg 5580
 atcaccaagt tgcaacatg aaaaagaaga cctgtgtata tataagtaag ggggagagcc 5640
 ctaactagat ctttcgaaaa cccccgacc ttcacctcc acaacctga tcatttata 5700
 cgttttggcc gttgcggct ctttctcat cttcaagaga gtcacctaca ccatggcttc 5760

 tattggttcg tccaaccccg tgctcttggc tgcgattccc ttcgtctacc tgtttgtcct 5820
 cccacgagtc ctggctttcc tgctcagaa ggctcagttc ctggccaaat gtattgtggt 5880
 cctgattgcc acgcttatca tgctcggttc aggtctgttc atctcgatcg tgtgcgtctt 5940
 tctggacaag agatactca tcaattacgt tgtgtcgcga ttgttctcct tccttgccgc 6000
 tcgaccgtgt ggtgtgacct ataagattgt tggtaggaa cacctcgata agtacacctgc 6060
 tatcgtggtc tgtaaccatc aatctctat ggatatgatg gttttgggac gagtttttcc 6120

aaagcactgc gttgtcatgg cgaagaagga actcctgtac tttccctttt tgggaatgtt	6180
tatgaaactg agcaacgcta tcttcatcga ccggaagaac cacaagaaag ccatcgagtc	6240
taccacccaa gccgtggcgg acatgaagaa gcacaactct ggaatctgga ttttccaga	6300
gggcacccgg tctagactgg acaaggcaga cctgctgccc ttcaagaaag gtgcctttca	6360
tcttgcaatt caggccccagc tccctattct ccccatattc tcgcagggct attcccatat	6420
ctacgactct tcgaagcggg acttccccgg tggagagctc gagatcagag tcctggagcc	6480
cattcctaca actggcctca ctactgatga tgtgaacgac ctgatggaca agacacgaaa	6540
ccttatgctc aagcacttga aggagatgga ttcccagtat tcgtcgagca ctgctgaaaa	6600
tggatccacg cacatcgacg ccgatattgc caagtctaca gccaccagca ttggcaacac	6660
tgacgacgca attacaaaac gtcgtacccc taaggaataa gcggccgcaa gtgtggatgg	6720
ggaagtgagt gcccggttct gtgtgcacaa ttggcaatcc aagatggatg gattcaacac	6780
agggatatag cgagctactg ggtgtgtcga ggatatagca acggatatat atgtttgaca	6840
cttgagaatg tacgatacaa gcactgtcca agtacaatac taaacatact gtacatactc	6900
atactcgtac ccgggcaacg gtttcacttg agtgcagtgg ctagtgtctt tactcgtaca	6960
gtgtgcaata ctgcgtatca tagtctttga tgtatatcgt attcattcat gttagtgtcg	7020
tacgactgat gtgtatgtaa gctcaatgag cagcgtgctc ctcggtctac tatatggcga	7080
catactctct cgcttctgtt taccctttac atacggtaca gctcttaca tgacatttat	7140
ccactggcgt cgatccataa accacacgaa ccctgttttg ttagtcacca tgaccggggc	7200
gtgtcgtgtt acgttccacc gtttcacctc agccggttac gattcaactt gccgcgtcat	7260
tctgcgttgc tagcggagcg agtacgagta actagacttt cgataagctg aatgacttca	7320
gggtgcatga gaggcgcaga tcggtatttt cggacttgtt cttctagaag actggtttga	7380
gagcaatcgc ggaagagtig ggaaccgctg gagagcagaa tgggttagta cagagactat	7440
ctgatctccc ccgcttgtgt ctagccccac ttataactct atttggcagt tgttgcttgt	7500
tattcaagca cagcatgtct gctcgtatca ctttgagac atccctaacc tctaaacctt	7560
tatagatgcc ttaccttctg agcttgcaga gagcattcca cagcgaaccc agttgacatt	7620
atcagttgac atcttacgt cacaccacat tacgcatctc acacacaatc acacgcacaa	7680
gtacacacgc aagcacatca tacaatggtc aaggcgcgcc agctgcatta atgaatcggc	7740
caacgcgcgg ggagaggcgg tttgcgtatt gggcgctctt ccgcttcctc gctcactgac	7800
tcgctgcgct cggtcgttcg gctgcggcga gcggtatcag ctcactcaaa ggcggttaata	7860

cggttatcca cagaatcagg ggataacgca ggaaagaaca tgtgagcaaa aggccagcaa	7920
aaggccagga accgtaaaaa ggccgcgttg ctggcgtttt tccataggct ccgccccct	7980
gacgagcatc acaaaaaatcg acgctcaagt cagaggtggc gaaacccgac aggactataa	8040
agataccagg cgtttcccc tggaagctcc ctctgtcgct ctctgtttcc gacctgccc	8100
cttaccggat acctgtccgc ttttctccct tcgggaagcg tggcgctttc tcatagctca	8160
cgctgtaggt atctcagttc ggtgtaggtc gttcgctcca agctgggctg tgtgcacgaa	8220
cccccgttc agcccgaccg ctgcgcctta tccggtaact atcgtcttga gtccaaccg	8280
gtaagacacg acttatcgcc actggcagca gccactggta acaggattag cagagcgagg	8340
tatgtaggcg gtgctacaga gttcttgaag tggtagccta actacggcta cactagaaga	8400
acagtatttg gtatctgcgc tctgtgaag ccagttacct tcgaaaaag agttggtagc	8460
tcttgatccg gcaaacaaac caccgtggt agcgggtggt tttttgttg caagcagcag	8520
attacgcga gaaaaaagg atctcaagaa gatccttga tctttctac ggggtctgac	8580
gctcagtgga acgaaaactc acgttaaggg attttggtca tgagattatc aaaaaggatc	8640
ttcacctaga tctttttaa ttaaaaatga agttttaaat caatctaaag tatatatgag	8700
taaacttggc ctgacagtta ccaatgctta atcagtgagg cacctatctc agcgatctgt	8760
ctatttcgtt catccatagt tgcctgactc cccgtcgtgt agataactac gatacgggag	8820
ggcttaccat ctggccccg tgctgcaatg ataccgcgag acccagctc accggtcca	8880
gatttatcag caataaacca gccagccgga agggccgagc gcagaagtgg tctgcaact	8940
ttatccgctt ccatccagtc tattaattgt tgccgggaag ctagagtaag tagttcgcca	9000
gttaatatgt tgcgcaacgt tgttgccatt gctacaggca tcgtgggtgc acgctcgtcg	9060
tttggtatgg cttcattcag ctccggttcc caacgatcaa ggcgagttac atgatcccc	9120
atgttgtgca aaaaagcggc tagctccttc ggtcctccga tcgttgtcag aagtaagttg	9180
gccgcagtgt tatcactcat ggttatggca gcactgcata attctcttac tgtcatgcca	9240
tccgtaagat gcttttctgt gactgggtgag tactcaacca agtcattctg agaatagttg	9300
atgcggcgac cgagttgctc ttgcccggcg tcaatacggg ataataccgc gccacatagc	9360
agaactttaa aagtgtcat cattggaaaa cgttcttcgg ggcgaaaact ctcaaggatc	9420
ttaccgtgt tgagatccag ttcatgtaa cccactctg caccctactg atcttcagca	9480
tcttttactt tcaccagctt ttctgggtga gcaaaaacag gaaggcaaaa tgccgcaaaa	9540
aagggaataa gggcgacacg gaaatgtga atactcatc tcttctttt tcaatattat	9600
tgaagcattt atcagggtta ttgtctcatg agcggataca tatttgaatg tatitagaaa	9660
aataaacaaa taggggttcc gcgcacattt ccccgaaaag tgccacctga tgcggtgtga	9720

aataccgcac agatgcgtaa ggagaaaata ccgcatcagg aaattgtaag cgtaaatt	9780
ttgttaaaat tcgcgttaaa tttttgttaa atcagctcat ttttaacca ataggccgaa	9840
atcggcaaaa tcccttataa atcaaaagaa tagaccgaga tagggttag tgttgtcca	9900
gtttggaaca agagtccact attaaagaac gtggactcca acgtcaaagg gcgaaaaacc	9960
gtctatcagg gcgatggccc actacgtgaa ccatcacct aatcaagttt tttgggtcg	10020
aggtgccgta aagcactaaa tcggaacct aaaggagcc cccgatttag agcttgacgg	10080
ggaaagccgg cgaacgtggc gagaaaggaa gggaagaaag cgaaaggagc gggcgctagg	10140
gcgctggcaa gtgtagcgt cagctgcgc gtaaccacca caccgccgc gcttaatgcg	10200
ccgtacagg gcgctccat tcgccattca ggctgcgcaa ctgttgggaa gggcgatcgg	10260
tgcggcctc ttgctatta cgccagctgg cgaaagggg atgtgctgca aggcgattaa	10320
gttgggtaac gccagggtt tccagtcac gacgttgtaa aacgacggcc agtgaattgt	10380
aatacgaact actatagggc gaattgggcc cgacgtcgca tgctaggaga gtctgtcgtt	10440
cgcgagtggc gcgccaatgt catgagttac acctgattca aatagtgaat aataatataa	10500
cagccacaaa tgaaagatgt atccacgggt agatatggct ttaattactg atgattgaat	10560
gattatctga gtgtacagt tgtaacgagt acgagttatt acctgctact attgtactcc	10620
taaatcaaag tactcgtaca tgcaatgaat tacttgtact tgcgccaatc agatctatcg	10680
gttgtattgt atataact agtctattct ttgtcttacc agatgggtga acctccatcg	10740
acaatgagat cgtggccggt catgtagctg gaggcgtcgg aagccatgaa gacaataggg	10800
cctccaaact cctcgggctc agccattcgt cgaaaggaa ttcgtcgtc ccagtcgttt	10860
cggagctctc cgtcagctc gatgatgtga cgggtcagag gagtaagaat gtagccggga	10920
gacagagtgt tgactcggat gttgtactgg gccactcgg cagccagaga ctggccatg	10980
tggatgacac ctgccttga catgtttag ggggtctggg gctgggggtc gttgacaatg	11040
gagccactca tggatgcaac cagaatcaga gatccaggag actcgtccag gacaagcgac	11100
ttggccacct cggtagcgaa gttgtaggat cccatgaggt tgacgtccag cagtcgcttg	11160
acctgggggt tgggataatc aaaggcagtc atattctgc agaaccggc acagttaatt	11220
aagttgcgac acatgtcttg atagtatctt gaattctctc tcttagctt ttccataaca	11280
agttcttctg cctccaggaa gtccatcggg ggtttgatca tggttttggt gtagtggtag	11340
tgcagtgggt gtattgtgac tgggatgta gttgagaata agtcatacac aagtcagctt	11400
tcttcgagcc tcatataagt ataagtagtt caacgtatta gcactgtacc cagcatctcc	11460
gtatcgagaa acacaacaac atgccccatt ggacagatca tgcggataca caggttgtgc	11520
agtatcatac atactcgatc agacaggctg tctgaccatc atacaagctg aacaagcgct	11580

ccatacttgc acgctctcta tatacacagt taaattacat atccatagtc taacctctaa	11640
cagttaatct tctggtaagc ctcccagcca gccttctggt atcgcttgge ctectcaata	11700
ggatctcggt tctggccgta cagacctcgg ccgacaatta tgatatccgt tccggtagac	11760
atgacatcct caacagttcg gtactgctgt ccgagagcgt ctcccttgtc gtcaagaccc	11820
accccggggg tcagaataag ccagtcctca gaggcgccct taggtcggtt ctgggcaatg	11880
aagccaacca caaactcggg gtcggatcgg gcaagctcaa tggctgctt ggagtactcg	11940
ccagtggcca gagagccctt gcaagacagc tcggccagca tgagcagacc tctggccagc	12000
ttctcgttgg gagaggggac taggaactcc ttgtactggg agttctcgta gtcagagacg	12060
tcctcttct tctgttcaga gacagtttcc tcggcaccag ctgcagggcc agcaatgatt	12120
ccggttccgg gtacaccgtg ggcgttggtg atatcgacc actcggcgat tcggtgacac	12180
cggtactggt gcttgacagt gttgccaata tctgcgaact ttctgtctc gaacaggaag	12240
aaaccgtgct taagagcaag ttccttgagg gggagcacag tgccggcgta ggtgaagtcg	12300
tcaatgatgt cgatatgggt ttgatcatg cacacataag gtccgacctt atcggaagc	12360
tcaatgagct cttgggtgt ggtaacatcc agagaagcac acaggttggt ttctttggct	12420
gccacgagct tgagcactcg agcggcaaag gcggacttgt ggacgttagc tcgagcttcg	12480
taggagggca ttttggtggt gaagaggaga ctgaaataaa tttagtctgc agaacttttt	12540
atcggaacct tatctggggc agtgaagtat atgttatggt aatagttacg agttagtga	12600
acttatagat agactggact atacggctat cggtcctaat tagaagaac gtcaatggct	12660
ctctggcggt cgcctttgcc gacaaaaatg tgatcatgat gaaagccagc aatgacgttg	12720
cagctgatat tgttgtcggc caaccgcgcc gaaaacgcag ctgtcagacc cacagcctcc	12780
aacgaagaat gtatcgtcaa agtgatcaa gcacactcat agttggagtc gtactccaaa	12840
ggcggcaatg acgagtcaga cagatactcg tcgacctttt ccttggaac caccaccgtc	12900
agcccttctg actcacglat ttagccacc gacacaggca acagtcctgt gatagcagaa	12960
tatgtcttgt cggtcatttt ctccacaaact ttaggcgtca agtgaatgtt gcagaagaag	13020
tatgtgcctt cattgagaat cgggtgtgct gatttcaata aagcttgag atcagtttgg	13080
ccagtcatgt tgtggggggt aattggattg agttatcgcc tacagtctgt acaggtatac	13140
tcgtgcccc ctttatactt ttgatccg ctgcactga agcaatgtcg tttaacaaaa	13200
gtgagaatgc tccacagaac acacccagg gtatggttga gcaaaaaata aacactccga	13260
tacggggaat cgaaccccg tctccacggt tctcaagaag tattcttgat gagagcgtat	13320

<210> 96

<211> 13975

<212> DNA

<213> Artificial Sequence

<220><223> Plasmid pZKSL-5S5A5

<400> 96

cgatggttaa tgctgctgtg tgctgtgtgt gtgtgttgtt tggcgctcat tgttgcgtta 60
 tgcagcgta accacaatat tggaaagctta ttagcctttc ttttttttcg tttgcaaggc 120
 ttaacaacat tgctgtggag agggatgggg atatggaggc cgctggaggg agtcggagag 180
 gcgttttga gcggcttggc ctggcgccca gctcgcaaaa cgcacctagg accctttggc 240
 acgccgaaat gtgccacttt tcagcttagt aacgccttac ctacgtcatt ccatgcgtgc 300
 atgtttgcgc cttttttccc ttgcccttga tcgccacaca gtacagtga ctgtacagtg 360

gaggttttgg gggggtctta gatgggagct aaaagcggcc tagcggta caagtgggat 420
 tgtatggagt ggcatggagc ctaggtggag cctgacagga cgcacgaccg gctagcccg 480
 gacagacgat ggggtggctcc tgtttgccac cgcgtacaaa tgtttgggcc aaagtcttgt 540
 cagccttgc tgcgaaccta attcccaatt ttgtcacttc gcaccccat tgatcgagcc 600
 ctaaccctg ccatcagc aatccaatta agctcgcat gtctgccttg tttagtttgg 660
 ctctgcccc tttcgcgctc cacttgcaca aacacaaaca agcattatat ataaggctcg 720
 tctctccctc ccaaccacac tcactttttt gcccgctctt ccttgctaac acaaaagtca 780

agaacacaaa caaccacccc aacccctt caacaaagac atatctacag caatggccat 840
 ggctctcagt cttaccacag aacagctgtt agaacgcct gatttgggtg cgattgatgg 900
 catcctctac gaccttgaag ggcttgccaa agttcatcca ggatccgatt tgattctcgc 960
 ttctgggtgc tctgatgctt cccctctctt ttattcaatg catccatacg tcaaaccgga 1020
 gaactccaaa ttgcttcaac agttcgctcg agggaagcat gaccgcacct cgaaggacat 1080
 tgtctacacg tatgattctc ctttcgcaca agacgttaag cggacaatgc gcgaggtgat 1140
 gaaaggagg aactggtagc caaccctgg cttctggctg cgcaccgttg ggatcatcgc 1200

cgtagcggcc ttttgcgagt ggcaactggc taccacgggg atggtgctgt ggggcctgtt 1260
 gactggattc atgcacatgc agatcggtt atccatccag catgatgcgt cccacggggc 1320
 catcagcaag aagccttggg tcaacgcctt cttcgctac ggcatcgac tcatcgatc 1380
 gtcccggtg atttggctgc agtcgcacat catcggcac cacacctaca ccaaccagca 1440
 cggcctcgac ctggatgcgg agtcggcaga gccgttctg gtgttcaca actacccgc 1500
 cgcaaacacc gcccgaaagt ggttcaccg cttccaggct tggtacatgt accttgtgt 1560

gggggcatac ggggtatcgc tgggtgtacaa cccgtcttac attttccgga tgcagcacia	1620
tgcaccatc ccagagtctg tcacggccat gcgggaaaat ggctttctgc ggcgctaccg	1680
cacacttgca ttctgtatgc gagctttctt catcttccgg accgcattct tgccttggtg	1740
cctcactggg acctcattgc tgatcacat tctcttggtg cccaccgcaa ctggtgcctt	1800
cttgacgttc ttcttcattt tgtcccacaa ttttgatggc tccgaacgga tccccgacaa	1860
gaactgcaag gtttaagcat ctgagaagga cgttgaggct gaccaaattg actggtatcg	1920
ggcgcagggt gagacgtcct ccacatacgg tggcccatc gccatgttct tactggcgg	1980
tctcaatttc cagatcgagc accacctctt tccccgatg tcgtcttggc actaccctt	2040
cgctccagcag gcggtccggg agtgttgca acgacatgga gtgcgatatg tttctaccc	2100
taccatcgtc ggcaacatca tctccacct gaagtacatg cataaggtgg gtgtcgtcca	2160
ctgcgtgaag gacgcacagg attcctaagc ggccgcatg atgattggaa acacacacat	2220
gggttatatc taggtgagag ttagttggac agttatata taaatcagct atgccaacgg	2280
taacttcatt catgtcaacg aggaaccagt gactgcaagt aatatagaat ttgaccacct	2340
tgccattctc ttgactcct ttactatac tcatatttt cttatataca aatcattct	2400
tcttcccagc atcgagctcg gaaacctcat gagcaataac atcgtggatc tcgtcaatag	2460
agggttttt ggactccttg ctgttgcca cctgtcctt gctgtctgga tcattctgtt	2520
tcaacgcctt ttaattaaat ctctgcagc ttctcagcg agacactgct cctgtctgga	2580
cccgagctaa ggctctgttc gacaaacacg ttctgcgaat tggcgagtaa ttctattaat	2640
gcaaatagac gtgtatttga aaaggaggag atgatgtac cagtaattta cgggttttga	2700
atagacaaat tatatatata aaattaacct gatgaatgag tgtatgaatg gattgcttaa	2760
tataccgagg gagagccgc attatcaata ttattgtcc taagaagcta aaatatgctg	2820
ttccgttgat agcgatgtct gtttgagagt caatggcaga acgcggagga gtattgttag	2880
ttggtgatcg gtttactga catcgagtag ggggcgagac agaagatac ccgaatccat	2940
tgcattgttt attaggatgt tcacaacaca actccctcta gactctgggg atgtgcgtta	3000
gaagaatgac ctggagcaag agtgttcaga ttgatccgtt caaattttca agattactgt	3060
tgggtgtgtt ttgaaatcca cattagctgg acgactattg catctgagcg ctcggaacgt	3120
ctccctttcc cttcagcttt atacgagtcg attccataag cgcgacctct cgaattgcct	3180
tcggttgatga agccatagc aaaggtgtgg ctatggaatg catgcgacgt cgggcccaat	3240
tcgcctata gtgagtcgta ttacaattca ctggccgtcg ttttacaacg tcgtgactgg	3300

gaaaaccctg gcgttaccca acttaatcgc cttgcagcac atcccccttt cgccagctgg 3360
 cgtaatatgcg aagaggcccg caccgatcgc ccttccaac agttgcgcag cctgaatggc 3420
 gaatggacgc gccctgtagc ggcgcatata gcgcggcggg tgtggtggtt acgcgcagcg 3480
 tgaccgctac acttgccagc gccctagcgc ccgctccttt cgctttcttc ccttcctttc 3540
 tcgccacgtt cgcggctttt ccccgtaag ctctaaatcg ggggctccct ttagggttcc 3600
 gatttagtgc ttacggcac ctgcaccca aaaaacttga ttaggggat ggttcacgta 3660
 gtgggccatc gccctgatag acggtttttc gccctttgac gttggagtcc acgtttctta 3720

 atagtggact cttgttccaa actggaacaa cactcaacc tatctcggtc tattcttttg 3780
 attataagg gattttgccg atttcggcct attggttaaa aaatgagctg atttaacaaa 3840
 aatttaacgc gaattttaac aaaatattaa cgcttacaat ttcctgatgc ggtattttct 3900
 ccttacgcat ctgtgcgga tttcacaccg catcaggtgg cacttttcgg ggaatgtgc 3960
 gcggaacccc tatttgttta tttttctaaa tacattcaaa tatgtatccg ctcatgagac 4020
 aataaccctg ataatgctt caataatatt gaaaaaggaa gagtatgagt attcaacatt 4080
 tccgtgtcgc cttattccc tttttgcgg cattttgct tctgttttt gctcaccag 4140

 aaacgctggt gaaagtaaaa gatgctgaag atcagttggg tgcacgagt ggttacatcg 4200
 aactggatct caacagcggg aagatccttg agagttttcg cccgaagaa cgttttccaa 4260
 tgatgagcac ttttaaagtt ctgctatgtg gcgcggtatt atccgtatt gacgccgggc 4320
 aagagcaact cggtcgccg atacactatt ctcagaatga cttggttag tactcaccag 4380
 tcacagaaaa gcatcttacg gatggcatga cagtaagaga attatgcagt gctgccataa 4440
 ccatgagtga taacactgcg gccaaacttac tctgacaac gatcggagga ccgaaggagc 4500
 taaccgcttt ttgcacaac atgggggatc atgtaactcg ccttgatcgt tgggaaccgg 4560

 agctgaatga agccatacca aacgacgagc gtgacaccac gatgcctgta gcaatggcaa 4620
 caacgttgcg caaactatta actggcgaac tacttactct agcttcccgg caacaattaa 4680
 tagactggat ggaggcggat aaagtgcag gaccattct gcgctcgcc cttccggctg 4740
 gctggtttat tgctgataaa tctggagccg gtgagcgtgg gtctcgcggt atcattgcag 4800
 cactggggcc agatggtaag ccctcccgtc tcgtagtat ctacacgacg gggagtcagg 4860
 caactatgga tgaacgaat agacagatcg ctgagatagg tgcctcactg attaagcatt 4920
 ggtaactgtc agaccaagtt tactcataata tacttttagat tgatttaaaa cttcattttt 4980

 aatttaaaag gatctagggtg aagatccttt ttgataatct catgacaaa atccctaac 5040
 gtgagttttc gttccactga gcgtcagacc ccgtagaaaa gatcaaagga tcttcttgag 5100
 atcccttttt tctgcgcgta atctgctgct tgcaaacaaa aaaaccaccg ctaccagcgg 5160

tggtttgttt gccgatcaa gagctaccaa ctctttttcc gaaggtaact ggcttcagca 5220
 gagcgcagat accaaatact gttcttctag tgtagccgta gttaggccac cacttcaaga 5280
 actctgtagc accgcctaca tacctcgctc tgctaatect gttaccagtg gctgctgcca 5340
 gtggcgataa gtcgtgtctt accgggttgg actcaagacg atagttaccg gataaggcgc 5400

 agcggtcggg ctgaacgggg gggttcgtgca cacagcccag cttggagcga acgacctaca 5460
 ccgaactgag atacctacag cgtgagctat gagaaagcgc cacgcttccc gaaggagaaa 5520
 aggcggacag gtatccggtg agcggcaggg tcggaacagg agagcgcacg agggagcttc 5580
 cagggggaaa cgcttggtat ttttatagtc ctgtcgggtt tcgccacctc tgacttgagc 5640
 gtcgattttt gtgatgctcg tcaggggggc ggagcctatg gaaaaacgcc agcaacgcgg 5700
 cttttttacg gttcctggcc ttttgctggc cttttgctca catgttcttt cctgcgttat 5760
 cccctgattc tgttgataac cgtattaccg cttttgagtg agctgatacc gctcgccgca 5820

 gccgaacgac cgagcgcagc gagtcagtga gcgaggaagc ggaagagcgc ccaatacgca 5880
 aaccgcctct ccccgcgctt tggccgattc attaatgcag ctggcgcgcc gaggtttcca 5940
 acgagataac atcgtggcag ctgccacatg gaccgatatg cctgaaataa gtgaattgac 6000
 catcagaagt tctgcatcca aatagaacta tatcatattc ggctcagttt tttcaataat 6060
 agtccaatcc ctaagtctcc tatccaaaat gggtcctgac cagccacatc catgatcatg 6120
 actgctgac aggaacagtc attccgtgga tgaacgactt tacgctcagg tactgtaaat 6180
 atctgtaaaag ggcagacaac caaccaattg agtaacctgt gagacttgaa acgtaagatg 6240

 acttcacaca caaagtcact tgactcaacc gtggctctca attgcacaaa atcactctgc 6300
 actaatctat tgcaggagtc aggctatgaa caactagacg acagctactt atgtgttata 6360
 tagaggaata ttaaaaaatc taagaataat cataaagtta caaaataatt atcagatttc 6420
 gagccacagg tcacccetaa catgtgttat tgcacacca caatcctcag ctgatgtca 6480
 tttaatcttt ccagccacca tctctctctc caaccetaat ggcaaacttt attttggtgg 6540
 agcgatgact cttactcaac tgcagcatac ttaagcacia ttgttcccca gcctgatacg 6600
 acacaccatc cattgtcaag cttcaccaca tacaacaaca cagcgtacgc aactaacatg 6660

 aatgaatacg atatacatca aagactatga tacgcagtat tgcacactgt acgagtaaga 6720
 gcactagcca ctgcactcaa gtgaaaccgt tgcccgggta cgagtatgag tatgtacagt 6780
 atgtttagta ttgtacttgg acagtgttgg tatcgtacat tctcaagtgt caaacataaa 6840
 tatccgttgc tatatcctcg caccaccacg tagctcgcta tatccctgtg ttgaatccat 6900
 ccatcttggg ttgccaattg tgcacacaga accgggcact cacttcccca tccacacttg 6960
 cggccgctta ggaatcctga gcgtccttga cacagtgaac cacaccgact ttgtgcatgt 7020

acttgagggt ggaaatgatg ttgccacaaa tggtagggta gaagacgtac cgaactccgt 7080

gtcgttcgca acactctcgg acagcttgct gcacgaaggg atagtgccaa gacgacattc 7140

gaggaaagag gtgatgctcg atctggaagt tgagaccgcc agtaaagaac atggcaatgg 7200

gtccaccgta ggtggaagag gtctccacct gagctctgta ccagtcgac tgatcggtt 7260

caacgtcctt ctcgagctc ttgaccttgc agttcttgc ggggattcgc tccgagccat 7320

cgaagtgtg agacaagatg aaaaagaagg tgaggaaggc accggtagca gtgggcacca 7380

gaggaatggt gatgagcagg gaggttccag tgagatacca gggcaagaag gcggttcgaa 7440

agatgaagaa agctcgcata acgaatgcaa gggttcggta ccgtcgaga aagccgttct 7500

ctcgcatggc tggacagac tcgggaatgg tgcgttgtg ctgcattcgg aagatgtaga 7560

gagggttgta caccagcgaa acgccgtagg ctccaagcac gaggtacatg taccagcct 7620

ggaatcgggt aaaccacttt cgagcagtgt tggcagcagg gtagttgtgg aacacaagga 7680

atggttctgc ggactcggca tccaggtcga gaccatgctg attggtgtag gtgtgatgtc 7740

gcatgatgtg agactgcagc cagatccatc tggacgatcc aatgacgtcg atgccgtagg 7800

caaagagagc gttgaccagg ggctttttgc tgatggcacc atgagaggca tcgtgctgaa 7860

tggacaggcc gatctgcatg tgcatgaatc cagtcaagag accccacagc accattccgg 7920

tagtagccca gtgccactcg caaaaggcgg tgacagcaat gatgccaacg gttcgcagcc 7980

agaatccagg tgggcatac cagttccgac ctttcatgac ctctcgata gttcgcttga 8040

cgctctgtgc aaaggagag tcgtaggtgt agacaatgtc cttggaggtt cggctgtgct 8100

tgcctcgac gaactgtga agcagcttcg agttctcggg cttgacgtaa ggggtgcatgg 8160

agtagaacag aggagaagca tcggaggcac cagaagcgag gatcaagtgc gatccgggat 8220

ggaccttggc aagaccttc agatcgtaga gaatccgtc gatggcaacc aggtcgggtc 8280

gtcgcagcag ctgctcggtg gtaagggaga gagccatgga gagctgggtt agtttgtgta 8340

gagagtgtgt gttgctagcg actttcggat tgtgtcatta caaaaacgc gtcgtctcga 8400

cactgatctt gtcgtggata ctacggctc ggacatcgtc gccgacgatg acaccggact 8460

ttcgcttaag gacgtcagta acaggcattg tgtgatgtgt agtttagatt tcgaatctgt 8520

ggggaaagaa aggaaaaaag agactggcaa ccgattggga gagccactgt ttatatatac 8580

cctagacaag cccccgctt gtaagatgtt ggtcaatgta aaccagtatt aaggttggca 8640

agtgaggag aagcaaggtg tgggtaccga gcaatggaaa tgtgcggaag gcaaaaaaat 8700

gaggccacgg cctattgtcg gggctatatc cagggggcga ttgaagtaca ctaacatgac 8760

atgtgtccac agaccctcaa tctggcctga tgagccaaat ccatacgcgc tttcgagct	8820
ctaaaggcta taacaagta caccacctg ctgacctca ggcacctcac tttttgttaa	8880
gacaaactgt acacgtgtt ccagcgtttt ctgctgcac ctggtgggac atttggtgca	8940
acctaaagt ctggaacct ctgtggtgtc cagatcagcg cagcagttcc gaggtagttt	9000
tgaggccctt agatgatgca atggtgtcag tcgctggatc acgagtctta atggcagtat	9060
tcgttcttat ttgtgccatt gagccccgtt atcctcgtat cttctacccc ccatcccatc	9120
cccttgttgg tgcaacccta cccatttatt gttgggtgca gcccaaccga cgtggagagc	9180
ttggcttggc catataaaaa ggcccccccc tagtggcaat ggcagaaaagt cagctgtgag	9240
ttgttgaatt tgcattctag gcggcctggc cgtcttctcc ggggcaattg gggctgtttt	9300
ttgggacaca aatagccgc caaccgggtc tctcctgaat tctgcagatg ggctgcagga	9360
attccgtcgt cgctgagtc gactccaact ttccacctg agcgtaaaat gtggagaaga	9420
aatcggcact aaaaagtcag gtagactgga aaatgcgcca tgaaatgaat atctcttgct	9480
acagtaatgc ccagcatcga ggggtattgt gtcaccaaca ctatagtggc agctgaagcg	9540
ctcgtgattg tagtatgagt ctttatttgt gatgggaaga gttcactcaa tattctcgtt	9600
actgcaaaaa caccacggta atcggccaga caccatggat gtagatcacc aagcctgtga	9660
atgttattcg agctaaaatg cacatggttg gtgaaaggag tagttgctgt cgaattccgt	9720
cgctgcctga gtcattttt atttaccagt tggccacaaa cccttgacga tctcgtatgt	9780
cccctccgac atactcccgg ccggctgggg tacgttcgat agcgtatcg gcatcgacaa	9840
ggtttgggtc cctagccgat accgcactac ctgagtcaca atcttcggag gtttagtctt	9900
ccacatagca cgggcaaaag tgcgtatata tacaagagcg tttgccagcc acagattttc	9960
actccacaca ccacatcaca catacaacca cacacatcca caatggaacc cgaaactaag	10020
aagaccaaga ctgactcaa gaagattgtt cttctcggcg gcgacttctg tggccccgag	10080
gtgattgccg aggccgtcaa ggtgtcaag tctgttctg aggctccgg caccgagttt	10140
gtgtttgagg accgactcat tggaggagct gccattgaga aggaggcgca gcccatcacc	10200
gacgctactc tcgacatctg ccgaaaggct gactctatta tgctcggtgc gtctggaggc	10260
gctgccaaca ccgtatggac cactcccgc ggacgaaccg acgtgcgacc cgagcagggt	10320
ctctcaagc tcgaaagga cctgaacctg tacgccaacc tgcgacctg ccagctgctg	10380
tcgccaagc tcgccatct ctcceccatc cgaaacgttg agggcaccga cttcatcatt	10440
gtccgagagc tcgtcggagg tatctacttt ggagagcgaa aggaggatga cggatctggc	10500
gtcgttccg acaccgagac ctactccgtt cctgaggttg agcgaattgc ccgaatggcc	10560
gccttctgg cccttcagca caacccccct ctcccggtgt ggtctcttga caaggccaac	10620

gtgctggcct ccctcgcact ttggcgaaaag actgtcactc gagtcctcaa ggacgaattc	10680
ccccagctcg agctcaacca ccagctgatc gactcgcccg ccatgatcct catcaagcag	10740
ccctccaaga tgaatggat catcatcacc accaacaatgt ttggcgatat catctccgac	10800
gaggcctccg tcatccccgg ttctctgggt ctgctgccct ccgcctctct ggcttctctg	10860
ccccacacca acgaggcggt cggtctgtac gagccctgtc acggatctgc ccccgatctc	10920
ggcaagcaga aggtcaaccc cattgccacc attctgtctg ccgcatgat gctcaagttc	10980
tctcttaaca tgaagccgc cggtgacgct gttgaggctg ccgtcaagga gtccgtcgag	11040
gctggatca ctaccgccga tatcgaggc tcttctcca cctccgaggt cggagacttg	11100
ttgccaacaa ggtcaaggag ctgctcaaga aggagtaagt cgtttctacg acgcattgat	11160
ggaaggagca aactgacgc cctgcgggtt ggtctaccgg cagggtccgc tagtgtataa	11220
gactctataa aaagggccct gccctgctaa tgaatgatg atttataatt taccggtgta	11280
gcaaccttga ctagaagaag cagattgggt gtgtttgtag tggaggacag tggtagttt	11340
tggaaacagt cttcttgaaa gtgtctgtc tacagtatat tcactcataa cctcaatagc	11400
caagggtgta gtcggtttat taaaggaagg gattgtggc tgatgtggat agatatctt	11460
aagctggcga ctgcacccaa cgagtgtgtt gtagctgtt ttaaacagag tgtgaaagac	11520
tcactatggt cgggcttat ctcgaccaat agccaaagtc tggagtttct gagagaaaaa	11580
ggcaagatac gtaigtataa aagcgacgca tggtaacaata ataccggagg catgtatcat	11640
agagagttag tggttcgatg atggcactgg tgcctggtat gactttatag ggctgactac	11700
atatttgtcc tcagacatac aattacagtc aagcattac ccttggacat ctgtaggtac	11760
ccccggcca agacgatctc agcgtgtcgt atgtcggatt ggcttagctc cctcgtcgt	11820
caattggctc ccatctactt tcttctgctt ggctacaccc agcatgtctg ctatggctcg	11880
ttttcgtgcc ttatctatcc tcccagtatt accaactcta aatgacatga tgtgattggg	11940
tctacacttt catatcagag ataaggagta gcacagttgc ataaaaagcc caactcta	12000
cagcttcttc ctttcttgta attagtacaa aggtgattag cgaaatctgg aagcttagtt	12060
ggccctaaaa aaatcaaaaa aagcaaaaaa cgaaaaacga aaaaccacag ttttgagaac	12120
agggaggtaa cgaaggatcg tatatatata tatatatata tatacccacg gatcccgaga	12180
ccggcctttg attcttccct acaaccaacc attctacca ccctaattca caaccatggc	12240
caccatctcc ctgactaccg agcagctcct ggaacacccc gagctcgttg ccatcgacgg	12300
agtcctgtac gatctcttcg gtctggccaa ggtccatgcc ggaggcaacc tcatcgaagc	12360
tgccggtgca tccgacggaa ccgctctgtt ctactccatg catcctggag tcaagccaga	12420
gaactcgaag ctctgcagc aatttgcggc aggcaagcac gaacgaagct ccaaggatcc	12480

cggtgtacacc ttctgactctc cctttgtctca ggacgtcaag cagtccgttc gagaggtcat 12540

gaagggtcga aactgggtacg ccactcctgg ctctctggctg agaaccgcac tcatcatcgc 12600

ttgtactgcc attggcgagt ggtactggat cacaaccgga gcagtgatgt ggggtatctt 12660

tactggatac ttccactcgc agattggctt ggccattcaa cacgatgctt ctacaggagc 12720

catcagcaaa aagccctggg tcaacgcctt ttctgcttat ggcatcgacg ccattgggtc 12780

ctctcggttg atctggctgc agtccacat tatgcgacat cacacttaca ccaaccagca 12840

tggcctcgac ctggatgctg cctcggcaga gccgttcac tttgttccact cctatcctgc 12900

taccaacgcc tctcgaagt ggtaccacg atttcaggcg tggtagatgt acatcgttct 12960

gggaatgtat ggtgtctcga tgggttaciaa tcccatgtac ctcttcacaa tgcagcacia 13020

cgacaccatt cccgaggcca ctctctcag accaggcagc tttttcaatc ggcagcgagc 13080

tttcgccgtt tcccttcgac tgctcttcat ctctcgaaac gcctttcttc cctggtagat 13140

tgctgggtgc tctcctcgc tcaccattct tctgggtccc acggtcacag gcatcttctt 13200

cacctttgtg ttctgtctgt cccataactt cgaggagacc gaacggaccc cagagaagaa 13260

ctgcaaggcc aaacgagcta aggaaggcaa ggaggtcaga gacgtggaag aggatcgagt 13320

cgactggtac cgagcacagg ccgagactgc tgccacctac ggtggcagcg tgggaatgat 13380

gcttacaggc ggtctcaacc tgcagatcga gcatcacttg tttccccgaa tgtcctcttg 13440

gcactatccc ttattcaag acaccgttcg ggagtgttg aagcgacatg gcgtccgtta 13500

cacatactat cctaccattc tcgagaacat catgtccact ctctgataca tgcagaaggt 13560

gggtgttgct cacaccattc aggatgccca ggagttctaa gcggccgcat gtacatacaa 13620

gattatttat agaaatgaat cgcgatcga caaagagtac gagtgtacga gtaggggatg 13680

atgataaaag tggaagaagt tccgcatctt tggatttacc aacgtgtagg acgatacttc 13740

ctgtaaaaat gcaatgtctt taccataggt tctgctgtag atgttattaa ctaccattaa 13800

catgtctact tgtacagttg cagaccagtt ggagtataga atggtacact taccaaaaag 13860

tggtgatggt tgtaactacg atatataaaa ctgttgacgg gatccccgt gatatgccta 13920

aggaacaatc aaagaggaag atattaattc agaatgctag tatacagtta gggat 13975

<210> 97

<211> 14619

<212> DNA

<213> Artificial Sequence

<220><223> Plasmid pZP2-85m98F

<400> 97

cgatggaagc cggtagaacc gggctgcttg tgcttggaga tggaagccgg tagaaccggg	60
ctgcttgggg ggatttgggg ccgctgggct ccaaagaggg gtaggcattt cgttgggggtt	120
acgtaattgc ggcatttggg tctgcgcgc atgtccatt ggtcagaatt agtccggata	180
ggagacttat cagccaatca cagcgccgga tccacctgta ggttgggttg ggtgggagca	240
ccctccaca gactagagtc aaacagcagc agcaacatga tagttggggg tgtgcgtgtt	300
aaaggaaaaa aaagaagctt gggttatatt cccgctctat ttagaggttg cgggatagac	360
gccgacggag ggcaatggcg ctatggaacc ttgcggatat ccatcgccg cggcggactg	420
cgtcgaacc agtccagca gcgttttttc cgggccattg agccgactgc gaccccgcca	480
acgtgtcttg gccacgcac tcatgtcatg ttggtgttg gaggccactt ttaagtagc	540
acaaggcacc tagctcgag caaggtgtcc gaaccaaaga agcggctgca gtggtgcaaa	600
cggggcggaa acggcgggaa aaagccacgg gggcacgaat tgaggcacgc cctcgaattt	660
gagacgagtc acggcccat tgcgccgcgc aatggctcgc caacggccgg tctttgcac	720
cacatcaggt tacccaagc caaacctttg tgttaaaaag cttacatat tataccgaac	780
gtaggttttg ggggcttg tccgtctgc caaggcaaca tttatataag ggtctgcatc	840
gccggctcaa ttgaatcttt tttctcttc tcttctctat attcattctt gaattaaaca	900
cacatcaacc atggtcaagc gaccgctct gcctctcacc gtggacggtg tcacctacga	960
cgtttctgcc tggtcaacc accatcccgg aggtgccgac attatcgaga actaccgagg	1020
tcgggatgct accgacgtct tcatggttat gcactccgag aacgccgtgt ccaaactcag	1080
acgaatgccc atcatggaac ctctctctcc cctgactcca acacctcca agccaaactc	1140
cgacgaacct caggaggatt tccgaaagct gcgagacgag ctcatgtctg caggcatgtt	1200
cgatgcctct ccatgttgt acgtttacaa gacctgtcg actctcgac tgggtgtcct	1260
tgcctgtctg ttgatgacc agtggcactg gtacctggtt ggtgtatcg tctcggcat	1320
tcactttcaa cagatgggat ggctctcgca cgacatttg catcaccagc tgttcaagga	1380
ccgatccatc aacaatgcca ttggcctgct ctccgaaac gtgcttcagg gcttttctgt	1440
cacttggtagg aaggaccgac acaacgtca tcaactcgcc accaacgtgc agggtcacga	1500
tcccacatc gacaacctgc ctctctggc gtggtccaag gaggacgtcg agcgagctgg	1560
cccgtttct cgacgatga tcaagtacca acagtattac ttctttttca tctgtccct	1620
tctgcgattc atctggtgct ttcagtccat tcatactgcc acgggtctca aggatcgaag	1680
caatcagtag tatcgaagac agtacgagaa ggagtcgctc ggtctggcac tccactgggg	1740
tctcaaggcc ttgttctact atttctacat gccctcggtt ctcaccggac tcatggtgtt	1800

ctttgtctcc gagctgcttg gtggcttcgg aattgccatc gttgtcttca tgaaccacta	1860
ccctctggag aagattcagg actccgtgtg ggatggtcac ggcttctgtg ctggacagat	1920
tcacgagacc atgaacgttc agcgaggcct cgtcacagac tggtttttcg gtggcctcaa	1980
ctaccagatc gaacatcacc tgtggcctac tcttcccaga cacaacctca ccgctgcctc	2040
catcaaatgt gagcagctgt gcaagaagca caacctgccc taccgatctc ctcccatgct	2100
cgaaggtgtc ggcatcttta tctctacct gggcaccttc gctcgaatgg ttgccaaggc	2160
agacaaggcc taagcggccg cattgatgat tggaaacaca cacatgggtt atatctaggt	2220
gagagttagt tggacagtta tatattaaat cagctatgcc aacggtaact tcattcatgt	2280
caacgaggaa ccagtactg caagtaatat agaatttgac caccttgcca ttctcttgca	2340
ctcctttact atatctcatt tatttcttat atacaaatca cttcttcttc ccagcatcga	2400
gctcgaaac ctcatgagca ataacatcgt ggatctcgtc aatagaggcg tttttggact	2460
ccttgctgtt ggccaccttg tcttgctgt ttaaaccatcg tggtaaatgc tgctgtgtgc	2520
tgtgtgtgtg tgtgtttgg cgctcattgt tgcgttatgc agcgtacacc acaatatgg	2580
aagcttatta gcctttctat tttttcgttt gcaaggctta acaacattgc tgtggagagg	2640
gatggggata tggaggccgc tggagggagc cggagaggcg ttttggagcg gcttggcctg	2700
gcgccagct cgcgaaacgc acctaggacc ctttggcacg ccgaaatgtg ccacttttca	2760
gtctagtaac gccttaccta cgtcattcca tgcgtgcatg tttgcgcctt ttttccttg	2820
cccttgatcg ccacacagta cagtgcactg tacagtggag gttttggggg ggtcttagat	2880
gggagctaaa agcggcctag cggtaacta gtgggatgt atggagtggc atggagccta	2940
ggtggagcct gacaggacgc acgaccggct agcccgtgac agacgatggg tggctcctgt	3000
tgtccaccgc gtacaaatgt ttgggccaaa gtcttgcag ccttgcttgc gaacctaat	3060
cccaattttg tcaattcgca cccccattga tcgagcccta acccctgccc atcaggcaat	3120
ccaattaagc tcgattgtc tgcttgttt agtttggtc ctgcccgttt cggcgtccac	3180
ttgcacaaac acaacaagc attatatata aggtcgtct ctcctccca accacactca	3240
cttttttgcc cgtcttccct tgctaacaca aaagtcaaga acacaaaca ccacccaac	3300
ccccttacac acaagacata tctacagca tggccatggc tctctccctt actaccgagc	3360
agctgtctga gcgaccgac ctggttgcca tcgacggcat tctctacgat ctggaaggtc	3420
ttgccaaggt ccatcccga tccgacttga tctcgtctt tggcgcctcc gatgtctctc	3480
ctctgttcta ctccatgcac ccttacgtca agcccagaa ctcgaagctg cttcaacagt	3540

tcgtgcgagg caagcacgac cgaacctcca aggacattgt ctacacctac gactctccct	3600
ttgcacagga cgtcaagcga actatgcgag aggtcatgaa aggtcggaa tggtatgcca	3660
cacctggatt ctggctgcga accgttggca tcattgctgt caccgccttt tgcgagtggc	3720
actgggctac taccggaatg gtgctgtggg gtctcttgac tggattcatg cacatgcaga	3780
tcggcctgtc cattcagcac gatgcctctc atggtgccat cagcaaaaag ccctgggtca	3840
acgctctctt tgcctacggc atcgacgtca ttggatcgtc cagatggatc tggctgcagt	3900
ctcacatcat gcgacatcac acctacacca atcagcatgg tctcgacctg gatgccgagt	3960
ccgcagaacc attccttgtg ttccacaact accctgtctc caacactgct cgaaagtgg	4020
ttcaccgatt ccaggcctgg tacatgtacc tcgtgcttgg agcctacggc gtttcgctgg	4080
tgtacaacc tctctacatc ttccgaatgc agcacaacga caccattccc gagtctgtca	4140
cagccatgcg agagaacggc tttctgcgac ggtaccgaac ccttgcattc gttatgcgag	4200
ctttcttcat ctttcgaacc gccttcttgc cctggatatc cactggaacc tccctgtca	4260
tcaccattcc tctgggtccc actgctaccg gtgccttctt cacttcttt ttcatttgt	4320
ctcacaactt cgatggctcg gagcgaatcc ccgacaagaa ctgcaaggtc aagagctccg	4380
agaaggacgt tgaagccgat cagatcgact ggtacagagc tcaggtggag acctcttcca	4440
cctacgggtg acccattgcc atgttcttta ctggcggtct caacttcag atcgagcatc	4500
acctctttcc tcgaatgtcg tcttggcact atcccttctg gcagcaagct gtccgagagt	4560
gttgcgaacg acacggagt ttggtacgtct tctacctac cattgtgggc aacatcattt	4620
ccacctcaa gtacatgcac aaagtcggtg tggttcactg tgtcaaggac gctcaggatt	4680
cctaagcggc cgcatgagaa gataaatata taaatacatt gagatattaa atgcgctaga	4740
ttagagagcc tcatactgct cggagagaag ccaagacgag tactcaaagg ggattacacc	4800
atccatatcc acagacacaa gctggggaaa ggttctatat aactttccg gaataccgta	4860
gtttccgatg ttatcaatgg gggcagccag gatttcaggc acttcggtgt ctcggggtga	4920
aatggcgctt ttggcctcca tcaagtcgta ccatgtcttc atttgctgt caaagtaaaa	4980
cagaagcaga tgaagaatga acttgaagtg aaggaattta aatgtaacga aactgaaatt	5040
tgaccagata ttgtgtccgc ggtggagctc cagcttttgt tccctttagt gagggttaat	5100
ttcgagcttg gcgtaatcat ggtcatagct gtttctgtg tgaaattgtt atccgctcac	5160
aagcttccac acaacgtacg ggcgtcgttg ctgtgtgat ttttaggac ccatcccttt	5220
ggatatataag tatactctgg ggttaagggt gcccggttag tctaggttat agttttcatg	5280
tgaaataccg agagccgagg gagaataaac gggggtatit ggacttggtt ttttcgcgga	5340
aaagcgtcga atcaaccctg cgggccttgc accatgtcca cgacgtgttt ctgccecaa	5400

ttcgccccctt gcacgtcaaa attaggccctc catctagacc cctccataac atgtgactgt	5460
ggggaaaagt ataagggaaa ccatgcaacc atagacgacg tgaaagacgg ggaggaacca	5520
atggaggcca aagaaatggg gtagcaacag tccaggagac agacaaggag acaaggagag	5580
ggcgccccgaa agatcgga aaacaacatg tccaattggg gcagtgcagg aaacgacacg	5640
gacacttcag tacaatggac cgaccatctc caagccaggg ttattccggt atcaccttgg	5700
ccgtaacctc ccgtgtgtac ctgatattgt acacgttcac attcaatata ctttcagcta	5760
caataagaga ggctgtttgt cgggcatgtg tgtccgtcgt atggggtgat gtccgagggc	5820
gaaattcgct acaagcttaa ctctggcgct tgtccagtat gaatagacaa gtcaagacca	5880
gtggtgccat gattgacagg gaggtacaag acttcgatac tcgagcatta ctcgacttg	5940
tggcgattga acagacgggc gatcgcttct cccccgtatt gccggcgcgc cagctgcatt	6000
aatgaatcgg ccaacgcgcg gggagaggcg gtttgcgtat tgggcgctct tccgcttct	6060
cgctcactga ctgcgtgcgc tcggtcgttc ggctgcggcg agcggtatca gctcactcaa	6120
aggcggtaat acggttatcc acagaatcag gggataacgc aggaaagaac atgtgagcaa	6180
aaggccagca aaaggccagg aaccgtaaaa aggccgcgtt gctggcgttt ttccataggc	6240
tccgcccccc tgacgagcat caaaaaatc gacgtcaag tcagagggtg cgaaaccga	6300
caggactata aagataccag gcgtttcccc ctggaagctc cctcgtgcgc tctcctgttc	6360
cgacctgcc gcttaccgga tacctgtccg cttttctccc ttcgggaagc gtggcgcttt	6420
ctcatagctc acgtgtagg tatctcagtt cgggttaggt cgttcgctcc aagctgggct	6480
gtgtgcacga accccccgtt cageccgacc gctgcgcctt atccggtaac tatcgtcttg	6540
agtccaaccc ggtaagacac gacttatcgc cactggcagc agccactggt aacaggatta	6600
gcagagcgag gtatgtaggc ggtgtctacag agttcttgaa gtggtggcct aactacggct	6660
acactagaag aacagtatit ggtatctgcg ctctgctgaa gccagttacc ttcggaaaaa	6720
gagtgtgtag ctcttgatcc ggcaaaaaa ccaccgtgg tagcgggtgt tttttgttt	6780
gcaagcagca gattacgcgc agaaaaaaag gatctcaaga agatccttg atcttttcta	6840
cggggctga cgctcagtg aacgaaaact cacgttaagg gattttggct atgagattat	6900
caaaaaggat cttcacctag atccttttaa attaaaaatg aagtttttaa tcaatctaaa	6960
gtatatatga gtaaaccttg tctgacagtt accaatgctt aatcagtgag gcacctatct	7020
cagcgatctg tctatttcgt tcatccatag ttgcctgact ccccgctgtg tagataacta	7080
cgatacggga gggcttacca tctggcccca gtgctgcaat gataccgcga gaccacgct	7140
caccgctcc agatttatca gcaataaacc agccagccgg aagggccgag cgcagaagtg	7200
gtcctgcaac ttatccgcc tccatccagt ctattaattg ttgccgggaa gctagagtaa	7260

gtagttcgcc agttaatagt ttgcgcaacg ttgttgccat tgctacaggc atcgtggtgt 7320

cacgctcgtc gtttggatag gcttcattca gctccggttc ccaacgatca aggcgagtta 7380

catgatcccc catgtttgtc aaaaaagcgg ttagctcctt cggtcctccg atcgtttgtca 7440

gaagtaagtt ggccgcagtg ttatcactca tggttatggc agcactgcat aattctctta 7500

ctgtcatgcc atccgtaaga tgcttttctg tgaactggta gtactcaacc aagtcattct 7560

gagaatagtg tatgcccga cagagtgtct cttgcccggc gtcaatacgg gataatacgg 7620

cgccacatag cagaacttta aaagtgtca tcattggaaa acgttcttcg gggcgaaaac 7680

tctcaaggat cttaccgctg ttgagatcca gttcgatgta acccactcgt gcaccaact 7740

gatcttcagc atcttttact ttcaccagcg tttctgggtg agcaaaaaca ggaaggcaaa 7800

atgccgcaaa aaaggaata agggcgacac ggaaatgttg aatactcata ctcttccttt 7860

ttcaatatta ttgaagcatt tatcaggggtt attgtctcat gagcggtac atatttgaat 7920

gtatttagaa aaataaaca ataggggttc cgcgcacatt tccccgaaa gtgccacctg 7980

atgcggtgtg aaataccgca cagatgcgtg aggagaaaat accgcatcag gaaattgtaa 8040

gcgttaatat tttgttaaaa ttgcggttaa atttttgtta aatcagctca ttttttaacc 8100

aataggccga aatcggaaca atcccttata aatcaaaaga atagaccgag atagggttga 8160

gtgttgttcc agtttggaa aagagtcac tatataagaa cgtggactcc aacgtcaaag 8220

ggcgaaaaac cgctatcag ggcatggcc cactacgtga accatcacc taatcaagtt 8280

ttttgggtc gaggtgccgt aaagcactaa atcggaaccc taaaggagc ccccattta 8340

gagcttgacg gggaaagccg gcgaacgtgg cgagaaagga aggaagaaa gcgaaaggag 8400

cgggcgctag ggcgctggca agtgiagcgg tcacgtcgcg cgtaaccacc acaccgccg 8460

cgcttaatgc gccgtacag ggcgcgtcca ttgccattc aggtcgcga actgttggga 8520

agggcgatcg gtgcgggcct cttcgtatt acgccagctg gcgaaaggg gatgtgctgc 8580

aaggcgatta agttgggtaa cgccaggggt tcccagtcg cgacgttgta aaacgacggc 8640

cagtgaattg taatacgaact cactataggg cgaattgggc ccgacgtcgc atgcgtgat 8700

gacactttgg tctgaaagag atgcattttg aatcccaaac ttgcagtgcc caagtgacat 8760

acatctccgc gttttgaaa atgttcagaa acagttgatt gtgttggaa ggggaatggg 8820

gaatgaaaa atgactcaag tatcaattcc aaaaacttct ctggttgga gtacctactg 8880

tccatactac tgcattttct ccagtcaggc cactctatac tcgacgacac agtagtaaaa 8940

cccagataat ttgcacataa acaagaaaac agaccaata atatttatat atagtcagcc 9000

gtttgtccag ttcagactgt aatagccgaa aaaaaatcca aagtttctat tctaggaaaa	9060
tatatccaa tatttttaat tcttaatctc atttatttta ttctagcgaa atacatttca	9120
gctacttgag acatgtgata cccacaaatc ggattcggac tcggttggtc agaagagcat	9180
atggcattcg tgctcgcttg ttcacgtatt ctctctgttc catctcttgg ccgacaatca	9240
cacaaaaatg gggttttttt ttttaattcta atgattcatt acagcaaaat tgagatatag	9300
cagaccacgt attccataat caccaaggaa gtctttgggc gtcttaatta actcacctgc	9360
aggattgaga ctatgaatgg attcccgtgc ccgtattact ctactaattt gatcttggaa	9420
cgcgaaaata cgtttctagg actccaaaga atctcaactc ttgtccttac taaatatact	9480
acccatagtt gatggtttac ttgaacagag aggacatgtt cacttgacct aaagtttctc	9540
gcatctcttg gatatttgaa caacggcgtc cactgaccgt cagttatcca gtcacaaaac	9600
ccccacattc atacattccc atgtacgttt acaaagttct caattccatc gtgcaaatca	9660
aaatcacatc tattcattca tcatatataa acccatcatg tctactaaca ctcaaacctc	9720
catagaaaac atcgactcag aacacacgct ccatgcggcc gcttactgag ccttggcacc	9780
gggctgcttc tcggccattc gagcgaactg ggacaggtat cggagcagga tgacgagacc	9840
ttcatggggc agaggggttc ggtaggggag gttgtgcttc tggcacagct gttccacctg	9900
gtaggaaacg gcagtgaggt tgtgtcgagg cagggtgggc cagagatggt gctcgatctg	9960
gtagttcagg cctccaaaga accagtcagt aatgatgcct cgtcgaatgt tcatggtctc	10020
atggatctga cccacagaga agccatgtcc gtcccagacg gaatcaccga tcttctccag	10080
agggtagtgg ttcatgaaga ccacgatggc aattccgaag ccaccgacga gctcggaaac	10140
aaagaacacc agcatcgagg tcaggatgga gggcataaag aagaggtgga acagggctctt	10200
gagagtccag tgcagagcga gtccaatggc ctctttcttg tactgagatc ggtagaactg	10260
gttgtctcgg tcttgaggg atcgaacggt cagcacagac tggaaacacc agatgaatcg	10320
caggagaata cagatgacca ggaaatagta ctgttggaac tgaatgagct ttcgggagat	10380
gggagaagct cgagtacatc cgtcctcgga ccaggcgagc agaggcaggt tatcaatgtc	10440
gggatcgtga cctgaacgt tggtagcaga atgatggcg tttgtgtgt ccttcacca	10500
ggtcacggag aagccctgga gtccgttgcc aaagaccaga cccaggacgt tattccagtt	10560
tcggttcttg aaggtctggt ggtggcagat gtcatgagac agccatccca tttgctggta	10620
gtgcataccg agcacgagag caccaatgaa gtacaggtgg tactggacca gcatgaagaa	10680
ggcaagcacg ccaagaccga ggggtggtaa gatcttgtac gattaccaga ggggagaggc	10740
gtcaaacatg ccagtggcga tcagctcttc tcggagcttt cggaaatcct cctgagcttc	10800
gttgacggca gcctggggag gcagctcgga agcctggttg atcttggca ttcgcttgag	10860

cttgtcgaag gcttcctgag agtgcataac catgaaggcg tcagtagcat ctcgccctg	10920
gtagttctca atgatttcag ctccaccagg gtggaagttc acccaagcgg agacgtcgta	10980
cacctttccg tcgatgacga ggggcagagc ctgtcgagaa gccttcacgg atcccatgac	11040
ggccagagag tcgtagtagg tagcgggagg aagtccggca ggtcgagcgg gaccggcgcc	11100
ctgaatcttt ttggctccct tgtgctttcg gacgatgtag gtctgcacgt agaagttgag	11160
gaacagacac aggacagtac caacgtagaa gtagttagaa aaccagccaa acattctcat	11220
tccatcttgt cggtagcagg gaatgttccg gtacttccag acgatgtaga agccaacgtt	11280
gaactgaatg atctgcatag aagtaatcag ggacttgggc atagggaact tgagcttgat	11340
cagtcgggtc caatagtagc cgtacatgat ccagtgaatg aagccgttga gcagcacaaa	11400
gatccaaacg gcttcgtttc ggtagttgta gaacagccac atgtccatag gagctccgag	11460
atggtgaaag aactgcaacc aggtcagagg ctgccccatg aggggcagat agaaggagtc	11520
aatgtactcg aggaacttgc tgaggtagaa cagctgagtg gtgattcgga agacattgtt	11580
gtcgaaagcc ttctcgagtg gtctggacat gacaccaatg gtgtacatgg ctagggccat	11640
agagaggaag gagcccagcg agtagatgga catgagcagg ttgtagttagg tgaacacaaa	11700
cttcattcga gactgacct tgggtccgag aggaccaagg gtgaacttca ggatgacgaa	11760
ggcgatggag aggtacagca cctcgagtg cgaggcatca gaccagagct gagcatagtc	11820
gaccttggga agaacctcct ggccaatgga gacgatttcg ttcacgacct ccatggttgt	11880
gaattagggt ggtgagaatg gttggttgta gggaagaatc aaaggccggt ctcgggatcc	11940
gtgggtatat atatataat atatataac gatccttcgt tacctccctg ttctcaaaac	12000
tgtggttttt cgtttttcgt tttttgcttt ttttgatttt tttaggcca actaagcttc	12060
cagatttcgc taatcacctt tgtactaatt acaagaaagg aagaagctga ttagagttgg	12120
gctttttatg caactgtgct actccttatt tctgatatga aagtgtagac ccaatcacat	12180
catgtcattt agagttaggta atactgggag gatagataag gcacgaaaac gagccatagc	12240
agacatgctg ggtgtagcca agcagaagaa agtagatggg agccaattga cgagcgaggg	12300
agctacgcca atccgacata cgacacgtg agatcgtctt ggccggggggg tacctacaga	12360
tgtccaaggg taagtgttg actgtaattg tatgtctgag gacaaatatg tagtcagccg	12420
tataaagtea taccaggcac cagtccatc atcgaccac taactctcta tgatacatgc	12480
ctccgtattt atgtacat gcgtcgcttt gttacatag tatcttgctt ttttctctca	12540
gaaactccag aattctctct cttgagcttt tccataacaa gttcttctgc ctccaggaag	12600
tccatgggtg gtttgatcat ggttttgggt tagtggtagt gcagtgggtg tattgtgact	12660
ggggatgtag ttgagaataa gtcatacaca agtcagcttt cttcgagcct catataagta	12720

taagtagttc aacgtattag cactgtaccc agcatctccg tatcgagaaa cacaacaaca	12780
tgccccattg gacagatcat gcggatacac aggttgtgca gtatcataca tactcgatca	12840
gacaggtcgt ctgacatca tacaagctga acaagcgctc catacttgca cgctctctat	12900
atacacagtt aaattacata tccatagtct aacctctaac agttaatctt ctggtaagcc	12960
tcccagccag cttcttggtg tcgcttggtc tcctcaatag gatctcggtt ctggccgtac	13020
agacctgggc cgacaattat gatatccgtt ccggtagaca tgacatcctc aacagttcgg	13080
tactgtgtgc cgagagcgct tcccttgctg tcaagacca ccccgggggt cagaataagc	13140
cagtcctcag agtcgccctt aggtcgggtt tgggcaatga agccaaccac aaactcgggg	13200
tcggatcggg caagctcaat ggtctgcttg gactactcgc cagtggccag agagcccttg	13260
caagacagct cggccagcat gagcagacct ctggccagct tctcgttggg agaggggact	13320
aggaactcct tgtactggga gttctcgtag tcagagacgt cctccttctt ctgttcagag	13380
acagtttctt cggcaccagc tcgcaggcca gcaatgattc cggttccggg tacaccgtgg	13440
gcgttggtga tatcggaaca ctcgcgatt cggtagacacc ggtactggtg ctgacagtg	13500
ttgccaatat ctgcgaactt tctgtcctcg aacaggaaga aaccgtgctt aagagcaagt	13560
tccttgaggg ggagcacagt gccggcgtag gtgaagtcgt caatgatgtc gatatgggtt	13620
ttgatcatgc acacataagg tccgacctta tcggcaagct caatgagctc ctgggtggtg	13680
glaacatcca gagaagcaca caggttggtt ttcttggtg ccacgagctt gagcactcga	13740
gcggcaaagg cggacttggt gacgttagct cgagcttcgt aggagggcat ttgtgtggtg	13800
aagaggagac tgaataaat ttagtctgca gaacttttta tcggaacctt atctggggca	13860
gtgaagtata tgttatggtg atagttacga gttagttaa cttatagata gactggacta	13920
tacggctatc ggtccaaatt agaaagaacg tcaatggctc tctgggcgtc gcctttgccg	13980
acaaaaatgt gatcatgatg aaagccagca atgacgttgc agctgatatt gttgtcggcc	14040
aaccgcgcg aaaacgcagc tgtcagacc acagcctcca acgaagaatg tatcgtcaaa	14100
gtgatccaag cacactcata gttggagtcg tactccaaag gcggcaatga cgagtcagac	14160
agatactcgt cgaccttttc cttgggaacc accaccgtca gcccttctga ctcacgtatt	14220
gtagccaccg acacaggcaa cagtcctggt atagcagaat atgtcttgct ggtccatttc	14280
tcaccaactt taggcgtcaa gtgaatgttg cagaagaagt atgtgccttc attgagaatc	14340
ggtgttgctg atttcaataa agtcttgaga tcagtttggc cagtcagtgt gtggggggtg	14400
attggattga gttatgcct acagtctgta caggtatact cgctgccac ttatacttt	14460

ttgattccgc tgcacttgaa gcaatgtcgt ttacaaaag tgagaatgct ccacagaaca 14520
 caccgccggg tatgggtgag caaaaaataa acactccgat acggggaatc gaaccccggt 14580
 ctccacggtt ctcaagaagt attcttgatg agagcgat 14619

<210> 98

<211> 993

<212> DNA

<213> Artificial Sequence

<220><223> 993 bp stuffer fragment

<400> 98

ctctatccaa tcacacagtc aactccgact accacgtcc ggccggcagt ggagccttct 60
 actctacca gaatggctct gagcgcgaca tgccaccgcc cgccacgtcc taccgaggat 120
 accagcctcc gggatattga aaaaaaagtg taaatatatta ttatcgtcaa tcgatcacga 180

attatctgaa cggaatgaat gcagccgtgt gtgagctgtg tagatcatgt aacgtaaacg 240
 gatacaatgg ctccatcctt tcttgactcg tttccacttc tgcaattgag tcgtaggagg 300
 tgatctcact catgatcatg acttgtgcac taacgttggc tctacagaat agtcagaata 360
 gtccagctcg tgccagtcgc gtgatttgaa acgcgaccaa caatgagaat gtttcataag 420
 ctgcgctaag ttggtgatgt ctacattca cattcgtttc atgacgcgat caaacttgtc 480
 tacagttacc tattgcgaaa caaaactgcg ggacagtttc ttctatcatc tataaatcat 540
 taatttatac aatccgtatc cgacggattc gtttttccca ttctttttgt ctgcgtctcc 600

gtctaagcga gagaagtgaa gaacttgtgg gcgagctcga gcggccgcat tctagatgca 660
 gacgcacctt gaccagatca aacggatgac caaccagaac ggccgagata cctcaaagc 720
 caccggcggc gaaggactcg gcctgcgaca ggaacgactt gagagcggaa gggggagcga 780
 tggccttaac gccagagtca atttcgggag cgtcactcat tgtgtggggt gtgtggaagg 840
 gatagttgag tgcagttagg agatggcaac gtccgagtta taagaagtgt ttagagttgt 900
 tcacacggga aggttacctc ggcatgtgaa tatggaggat tcggtatagg ggggtgtagc 960
 ggccgggttta gggggcgatt gtgtatatat atg 993

<210> 99

<211> 7338

<212> DNA

<213> Artificial Sequence

<220><223> Plasmid pYPS234

<400> 99

aaatgagtat ctgtctgact cgtcattgcc gcctttggag tacgactcca actatgagtg	60
tgcttgatc actttgacga tacattcttc gttaggagct gtgggtctga cagctgcgtt	120
ttcggcgagg ttggccgaca acaatatcag ctgcaacgac attgctggct ttcacatga	180
tcacatcttt gtgcgcaaag ggcagcccca gagagccatt gacgttcttt ctaatttga	240
ccgatagccg tatagtccag tctatctata agttcaacta actcgtaact attaccataa	300
catatacttc actgccccag ataagggtcc gataaaaagt tctgcagact aaatttattt	360
cagtctcttc ttcaccacca aaatgccctc ctacgaagct cgagctaacg tccacaagtc	420
cgctttgcc gctcgagtgc tcaagctcgt ggcagccaag aaaaccaacc tgtgtgcttc	480
tctggatgtt accaccacca aggagctcat tgagcttgcc gataaggctg gaccttatgt	540
gtgcatgac aaaaccata tcgacatcat tgacgacttc acctacgccg gactgtgct	600
ccccctaag gaacttgctc ttaagcacgg tttcttctg ttcgaggaca gaaagttcgc	660
agatatggc aacactgtca agcaccagta ccggtgtcac cgaatcgccg agtggccga	720
tatcaccaac gccacgggtg taccggaac cggaatcatt gctggcctgc gagctgggtc	780
cgaggaaact gtctctgaac agaagaagga ggacgtctct gactacgaga actcccagta	840
caaggagttc ctagtccct ctccaacga gaagctggcc agaggtctgc tcatgtggc	900
cgagctgtct tgcaagggtc ctctggccac tggcgagtac tccaagcaga ccattgagct	960
tgcccgatcc gaccccgagt ttgtggttgg ctctattgcc cagaaccgac ctaaggcgca	1020
ctctgaggac tggcttatcc tgacccccgg ggtgggtctt gacgacaagg gagacgtct	1080
cggacagcag taccgaactg ttgaggatgt catgtctacc ggaacggata tcataattgt	1140
cgcccgaggt ctgtacggcc agaaccgaga tcctattgag gaggccaagc gataccagaa	1200
ggctggctgg gaggttacc agaagattaa ctgttagagg ttagactatg gatatgtaat	1260
ttaactgtgt atatagagag cgtgcaagta tggagcgctt gttcagcttg tatgatggc	1320
agacgacctg tctgatcgag tatgtatgat actgcacaac ctgtgtatcc gcatgatctg	1380
tccaatgggg catgttgttg tgtttctcga tacggagatg ctgggtacag tgctaatacg	1440
ttgaactact tatacttata tgaggctcga agaaagctga cttgtgtatg acttaattaa	1500
cgcggcgcg cagctgcatt aatgaatcgg ccaacgcgcg gggagaggcg gtttgcgtat	1560
tgggcgctct tccgcttctt cgtcactga ctgcgtgcgc tcggtcgttc ggctgcggcg	1620
agcgglatca gtcactcaa aggcggtaat acggttatcc acagaatcag gggataacgc	1680
aggaaagaac atgtgagcaa aaggccagca aaaggccagg aaccgtaaaa aggccgcgtt	1740
gtggcgcttt ttccatagcg tccgcccccc tgacgagcat cacaaaaatc gacgtcaag	1800
tcagaggtag cgaaaccga caggactata aagataccag gcgtttcccc ctggaagctc	1860

cctcgtgcgc tctcctgttc cgacctgcc gcttacccga tacctgtccg cttttctccc	1920
ttcgggaagc gtggcgcttt ctcatagctc acgctgtagg tatctcagtt cgggtgtaggt	1980
cgttcgtcc aagctgggct gtgtgcacga accccccgtt cagcccgacc gctgcgcctt	2040
atccggtaac tatcgtcttg agtccaaccc ggtaagacac gacttatcgc cactggcagc	2100
agccactggt aacaggatta gcagagcgag gtatgtaggc ggtgtacag agttcttgaa	2160
gtggtggcct aactacggct acactagaag aacagtatit ggtaictgcg ctctgctgaa	2220
gccagttacc ttcggaaaaa gaggtagtag ctcttgatcc ggcaaaaaa ccaccgctgg	2280
tagcggtagt tttttgttt gcaagcagca gattacgcgc agaaaaaag gatctcaaga	2340
agatcctttg atcttttcta cggggtctga cgctcagtag aacgaaaact cacgttaagg	2400
gattttggtc atgagattat caaaaaggat cttcacctag atccttttaa attaaaaatg	2460
aagttttaaa tcaatctaaa gtatatatga gtaaacttgg tctgacagtt accaatgctt	2520
aatcagtgag gcacctatct cagcgatctg tctatttctg tcatccatag ttgcctgact	2580
ccccgtcgtg tagataacta cgatacggga gggtttacca tctggcccca gtgctgcaat	2640
gataccgga gaccacgct caccggctcc agattttatca gcaataaacc agccagccgg	2700
aagggccgag cgagaaagt gtcttgcaac ttatccgcc tccatccagt ctattaattg	2760
ttgccgggaa gctagagtaa gtagttgcc agttaatagt ttgcgcaac ttgttgccat	2820
tgctacaggc atcgtgggtg cacgctcgtc gtttggtagt gcttcattca gctccggttc	2880
ccaacgatca aggcgagtta catgatcccc catgttgtgc aaaaaagcgg ttagctcctt	2940
cggctctccg atcgttgta gaagtaagt ggccgcagtg ttatcactca tggttatggc	3000
agcactgcat aattctctta ctgtcatgcc atccgtaaga tgcttttctg tgactggtga	3060
gtactcaacc aagtcattct gagaatagt tatgcgcga ccgagttgct ctgcccggc	3120
gtcaatacgg gataataccg cgccacatag cagaacttta aaagtgtca tcattggaaa	3180
acgttcttcg gggcgaaaac tctcaaggat cttaccgctg ttgagatcca gttcgatga	3240
accactcgt gcaccaact gatcttcagc atcttttact ttcaccagcg tttctgggtg	3300
agcaaaaaa ggaaggcaaa atgccgcaaa aaagggaata agggcgacac ggaaatgttg	3360
aatactcata ctcttccttt ttcaatatta ttgaagcatt tatcagggtt attgtctcat	3420
gagcggatac atatttgaat gtatttagaa aaataaaca atagggttc cgcgcacatt	3480
tccccgaaaa gtgccacctg atcggtgtg aaataccgca cagatgcgt aggagaaaat	3540
accgcatcag gaaattgtaa gcgttaatat ttgttataa ttcgcgttaa atttttgtta	3600
aatcagctca ttttttaacc aataggccga aatcgcaaa atcccttata aatcaaaaga	3660

atagaccgag ataggggtga gtgttgttcc agtttggaac aagagtcac tattaagaa	3720
cgtaggactcc aacgtcaaag ggcgaaaaac cgtctatcag ggcgatggcc cactacgtga	3780
accatcaccc taatcaagtt ttttggggtc gaggtgccgt aaagcactaa atcggaaccc	3840
taaaggagc ccccgattta gagcttgacg gggaaagccg gcgaacgtgg cgagaaagga	3900
agggaagaaa gcgaaaggag cgggcgctag ggcgctggca agtgtagcgg tcacgtcgcg	3960
cgtaaccacc acaccgccg cgcttaatgc gccgctacag ggcgctcca ttcgccattc	4020
aggctgcga acgtttggga agggcgatcg gtgcgggcct cttcgctatt acgccagctg	4080
gcgaaagggg gatgtgctgc aaggcgatta agttgggtaa cgccagggtt tcccagtc	4140
cgacgttgta aaacgacggc cagtgaattg taatacgact cactataggg cgaattgggc	4200
ccgacgtcgc atgcatcaat cgatgactgt cgactcacgt acgggtgggt agaattggtg	4260
gttgtaggga agaataaaag gccgtctcg ggatccgtgg gtatatatat atatatatat	4320
atatacgatc cttcgctacc tccctgttct caaaactgtg gtttttcgtt tttcgttttt	4380
tgcttttttt gattttttta gggccaacta agcttccaga tttcgctaata cacctttgta	4440
ctaattacaa gaaaggaaga agctgattag agttgggctt tttatgcaac tgtgctactc	4500
cttatctctg atatgaaagt gtagacccaa tcacatcatg tcatttagag ttggtataac	4560
tgggaggata gataaggcac gaaaacgagc catagcagac atgctgggtg tagccaagca	4620
gaagaaagta gatgggagcc aattgacgag cgaggagct acgccaatcc gacatacgac	4680
acgctgagat cgtcttgccc ggggggtacc tacagatgtc caagggtgag tgcttgactg	4740
taattgtatg tctgaggaca aatatgtagt cagccgtata aagtcatacc aggcaccagt	4800
gccatcatcg aaccactaac tctctatgat acatgcctcc ggtattattg taccatgcgt	4860
cgctttgtta catacgtatc ttgccttttt ctctcagaaa ctccagactt tggctattgg	4920
tcgagataag cccggacat agtgagtctt tcacactctg tttaaacaag ctaccaccac	4980
actcgttggg tgcagtcgcc agcttaaaga tatctatcca catcagccac aactcccttc	5040
ctttaataaa ccgactacac ccttggctat tgaggttatg agtgaatata ctgtagacaa	5100
gacactttca agaagactgt ttccaaaacg taccactgtc ctccactaca aacacaccca	5160
atctgcttct tctagtcaag gttgctacac cggtaaatta taaatcatca tttcattagc	5220
agggcagggc cttttttata gagtcttata cactagcgga cctgcccgt agaccaaccc	5280
gcaggcgcgt cagtttgctc ctccatcaa tgcgtcgtag actagtctct atccaatcac	5340
acagtcaact ccgactacca cgctccggcg ggcagtggag ctttctactc tcaccagaat	5400

ggctctgagc gcgacatgcc accgcccgcc acgtcctacc gaggatacca gcctccggga 5460
 tattgaaaa aaagtgtaaa tatttattat cgtcaatcga tcacgaatta tctgaacgga 5520
 atgaatgcag ccgtgtgtga gctgtgtaga tcatgtaacg taaacggata caatggctcc 5580
 atcctttctt gactcgtttc cacttctgca attgagtcgt aggaggatgat ctcactcatg 5640
 atcatgactt gtgcactaac gttggctcta cagaatagtc agaatagtcc agctcgtgcc 5700
 agtcgcgtga ttgaaacgc gaccaacaat gagaatgttt cataagctgc gctaagttgg 5760
 tgatgtctca cattcacatt cgtttcatga cgcgatcaaa cttgtctaca gttacctatt 5820

 gcgaaacaaa actgcgggac agtttcttct atcatctata aatcattaat ttatacaatc 5880
 cgtatccgac ggattcgitt ttccattct tttgtctgc gtctccgtct aagcgagaga 5940
 agtgaagaac ttgtggcgga gctcgagcgg ccgcattcta gatgcagacg caccttgacc 6000
 agatcaaacg gatgaccaac cagaacggcg cagatacttc caaagccacc ggcggcgaag 6060
 gactcggcct gcgacaggaa cgacttgaga gcggaagggg gagcgatggc cttaacgcca 6120
 gagtcaattt cgggagcgtc actcattgtg tggggtgtgt ggaaggata gttgagtga 6180
 gttaggagat ggcaacgtcc gagttataag aagtgtttag agttgttcac acgggaaggt 6240

 tacctcggca tgtgaatatg gaggattcgg tatagggggg tgtagcggcg ggtttagggg 6300
 gcgatttgtt atatatatgg gatcccggtt ctgtgtgcac aattggcaat ccaagatgga 6360
 tggattcaac acagggatat agcgagctac gtggtggtgc gaggatatag caacggatat 6420
 ttatgtttga cacttgagaa tgtacgatac aagcactgtc caagtacaat actaaacata 6480
 ctgtacatac tcatactcgt acccgggcaa cggtttcact tgagtgcagt ggctagtgtc 6540
 cttaactcga cagtgtgcaa tactgcgtat catagtcttt gatgtatac gtattcattc 6600
 atgttagttg cgtacgtgt gttgttgtat gtggtgaagc ttgacaatgg atggtgtgtc 6660

 gtatcaggct ggggaacaat tgtgcttaag tatgtgcag ttgagtaaga gtcacgctc 6720
 cacaaaaata aagtttgcca ttagggttgg agagagagat ggtggctgga agaattaaat 6780
 gacatcaagc tgaggattgt ggggtgtgcaa taacacatgt taggggtgac ctgtggctcg 6840
 aaatctgata attattttgt aactttatga ttattcttag attttttaat attcctctat 6900
 ataacacata agtagctgtc gtctagtgtt tcatagcctg actcctgcaa tagattagtg 6960
 cagagtgatt ttgtgcaatt gagagccacg gttgagtcaa gtgactttgt gtgtgaagtc 7020
 atcttacgtt tcaagtctca caggttactc aattgggttg ttgtctgccc ttacagata 7080

 ttacagtac ctgagcgtaa agtcgttcat ccacggaatg actgttcttg tcacgcagtc 7140
 atgatcatgg atgtggctgg tcaggaacca ttttgatag gagacttagg gattggacta 7200
 ttattgaaaa aactgagccg aatatgatat agttctatct gaatgcagaa cttctgatgg 7260

tcaattcact tatttcaggc atatcgggtca tggtaggcagc tgccacgatg ttatctcgtt	7320
ggaaaagatc ttcgattt	7338
<210> 100	
<211> 1019	
<212> DNA	
<213> Artificial Sequence	
<220><223> 1019 bp stuffer fragment	
<400> 100	
actcactctg tcagctgccc cacattgccc aatgcacaat gcacaatgat gtgtgcaaac	60
aacgcaatca aaagtctatg catgctgacc aaactctgat caccaagttg cgaacatgaa	120
aaagaagacc tgtgtatata taagtaaggg ggagagccct aactagatct ttcgaaaacc	180
ccccgacctt caccttccac aacctgac atcttatacg ttttgccgtg tcggtctcc	240
ttcctcatct tcaagagagt cacctacacg atgcgaagcc gagagctcgc caagaagtgg	300
cactgtgagg agcctcaca cctgaatgat ctgatatagc gccgccatct taacggtgtg	360
ttcgtaaga tgacccggga cgaggctgcc ttgcccaga ccgagaagct cattaacgca	420
taaagggatg tatacttgta ctgtagatgg attaataaat ttttatgat gaacaaacat	480
tgactgagaa aggcaccata ttgtagcttc aagtactgta ctggtagaag taagcctgtc	540
gttaccact tatggagatc caatattgtt tgatgcggcc atcgactgc ttggagcgtg	600
cagagaattg gtaaaatgtg gtggctgggt tggtttcgat tgaaattatc aagatctctg	660
aggtattctt ggtgtgaag ggatgttctt ggactactgc ttgtaacata gatgggatgt	720
agggatggca tggggtataa tcgagacaac cattgtggag gtaggcaaga gctcttaca	780
gtaccatacg agtacaacta ctgaggcata atgtctggaa ttcccgtta agtccattac	840
agttgatatc attcaaactg gacagatatt ctccgaggtg atacatctct gtccccgaga	900
gaagctgatt tggagcacat ccaggttacg gggggagaga aatttatgta atgtgaggat	960
tttttgagc tgaataataa agtctgttg atctagcttc aattgagtca acatccgcc	1019
<210> 101	
<211> 7364	
<212> DNA	
<213> Artificial Sequence	
<220><223> Plasmid pYPS233	
<400> 101	
aatgagtat ctgtctgact cgtcattgcc gcctttggag tacgactcca actatgagt	60

tgcttggatc acittgacga tacattcttc gttggaggct gtgggtctga cagctgcgtt	120
ttcggcgcggt ttggccgaca acaatatcag ctgcaacgtc attgctggct ttcattcatga	180
tcacatTTTT gtcggcaaag gcgacgcca gagagccatt gacgttcttt ctaatttga	240
ccgatagccg tatagtccag tctatctata agttcaacta actcgtaact attaccataa	300
catatacttc actgccccag ataaggttcc gataaaaagt tctgcagact aaatttattt	360
cagtctcttc ttcaccacca aaatgccctc ctacgaagct cgagctaacg tccacaagtc	420
cgcctttgcc gctcgagttc tcaagctcgt ggacgccaag aaaaccaacc tgtgtgcttc	480
tctggatgtt accaccacca aggagctcat tgagcttgcc gataaggtcg gaccttatgt	540
gtgcatgac aaaaccata tcgacatcat tgacgacttc acctacgccc gcactgtgct	600
ccccctcaag gaacttgctc ttaagcacgg tttcttctg ttcgaggaca gaaagttcgc	660
agatatggc aacactgtca agcaccagta ccggtgtcac cgaatcgccg agtgggtccga	720
tatcaccaac gccacgggtg taccggaac cgggaatcatt gctggcctgc gagctgggtgc	780
cgaggaaact gtctctgaac agaagaagga ggacgtctct gactacgaga actcccagta	840
caaggagttc ctagtccctt ctcccacga gaagctggcc agaggtctgc tcatgtggc	900
cgagctgtct tgcaagggtt ctctggccac tggcgagtac tccaagcaga ccattgagct	960
tgcccgatcc gaccccgagt ttgtggttgg cttcattgcc cagaaccgac ctaaggcgca	1020
ctctgaggac tggcttattc tgacccccgg ggtgggtctt gacgacaagg gagacgtctt	1080
cggacagcag taccgaactg ttgaggatgt catgtctacc ggaacggata tcataattgt	1140
cggccgaggt ctgtacggcc agaaccgaga tcctattgag gaggccaagc gataccagaa	1200
ggctggctgg gaggttacc agaagattaa ctgttagagg ttagactatg gatatgtaat	1260
ttaactgtgt atatagagag cgtgcaagta tggagcgtt gttcagctt tatgatggtc	1320
agacgacctg tctgatcgag tatgtatgat actgcacaac ctgtgtatcc gcatgatctg	1380
tccaatgggg catgttgttg tgtttctcga tacggagatg ctgggtacag tgctaatacg	1440
ttgaactact tatacttata tgaggtcga agaaagctga cttgtgtatg acttaattaa	1500
cgcggcgcg cagctgcatt aatgaatcgg ccaacgcgcg gggagaggcg gtttgcgtat	1560
tggcgctct tccgttctt cgtcactga ctgcgtgcgc tcggtcgttc ggctgcggcg	1620
agcggatca gctcactcaa aggcggtaat acggttatcc acagaatcag gggataacgc	1680
aggaaagaac atgtgagcaa aaggccagca aaaggccagg aaccgtaaaa aggccgcgtt	1740
gctggcgttt ttccatagc tccgcccc tgacgagcat cacaaaaatc gacgtcaag	1800
tcagaggtgg cgaaaccga caggactata aagataccag gcgtttcccc ctggaagctc	1860

cctcgtgcgc tctcctgttc cgaccctgcc gcttacggga tacctgtccg cttttctccc	1920
ttcgggaagc gtggcgcttt ctcatagctc acgctgtagg tatctcagtt cgggtgtaggt	1980
cgttcgtcc aagctgggct gtgtgcacga acccccgtt cagcccgacc gctgcgcctt	2040
atccggtaac tatcgtcttg agtccaacc ggtaagacac gacttatcgc cactggcagc	2100
agccactggt aacaggatta gcagagcgag gtatgtaggc ggtgtacag agttcttgaa	2160
gtggtggcct aactacggct acactagaag aacagtatit ggtatctgcg ctctgctgaa	2220
gccagttacc ttcggaaaaa gatttggtag ctcttgatcc ggcaaaaaa ccaccgctgg	2280
tagcggtggt tttttgttt gcaagcagca gattacgcgc agaaaaaaag gatctcaaga	2340
agatcctttg atcttttcta cggggtctga cgctcagtgg aacgaaaact cacgttaagg	2400
gattttggtc atgagattat caaaaaggat cttcacctag atccttttaa attaaaaatg	2460
aagtttttaa tcaatctaaa gtatatatga gtaaaactgg tctgacagtt accaatgctt	2520
aatcagttag gcacctatct cagcgatctg tctatttctg tcatccatag ttgcctgact	2580
ccccgtcgtg tagataacta cgatacggga gggtttacca tctggcccca gtgctgcaat	2640
gataccgga gaccacgct caccggctcc agatttatca gcaataaacc agccagccgg	2700
aagggccgag cgcagaagtg gtcttgaac ttatccgcc tccatccagt ctattaattg	2760
ttgccgggaa gctagagtaa gtagttgcc agttaatagt ttgcgcaac ttgttgccat	2820
tgctacaggc atcgtgggtg cagctcgtc gtttggtatg gcttcattca gctccggttc	2880
ccaacgatca aggcgagtta catgatcccc catgttgtgc aaaaaagcgg ttagctcctt	2940
cggctctccg atcgttgtca gaagtaagt ggccgcagtg ttatcactca tggttatggc	3000
agcactgcat aattctctta ctgtcatgcc atccgtaaga tgcttttctg tgactggtga	3060
gtactcaacc aagtcattct gagaatagt tatgcggcga ccgagttgct ctgcccggc	3120
gtcaatacgg gataataccg cgccacatag cagaacttta aaagtgtca tcattggaaa	3180
acgttcttcg gggcgaaaac tctcaaggat cttaccgtg ttgagatcca gttcgtatga	3240
accactcgt gcacccaact gatcttcagc atcttttact ttaccagcg tttctgggtg	3300
agcaaaaaca ggaaggcaaa atgccgcaaa aaagggaata agggcgacac ggaaatgttg	3360
aatactcata ctcttccttt ttcaatatta ttgaagcatt tatcagggtt attgtctcat	3420
gagcggatag atatttgaat gtatttagaa aaataaacia atagggttc cgcgcacatt	3480
tccccgaaaa gtgccacctg atgcggtgtg aaataccgca cagatgcgt aggagaaaat	3540
accgcatcag gaaattgtaa gcgttaatat ttgttataa ttcgcgttaa atttttgtta	3600

aatcagctca ttttttaacc aataggccga aatcggcaaa atcccttata aatcaaaaga	3660
atagaccgag atagggttga gtgtgttcc agtttggaac aagagtccac tattaagaa	3720
cgtggactcc aacgtcaaag ggcgaaaaac cgtctatcag ggcatggcc cactacgtga	3780
accatcacc taatcaagtt ttttgggtc gaggtgccgt aaagcactaa atcggaaccc	3840
taaaggagc ccccgattta gagcttgacg gggaagccg gcgaacgtgg cgagaaagga	3900
agggaagaaa gcgaaggag cgggcgctag ggcgtggca agttagcgg tcacgtgcg	3960
cgtaaccacc acaccgccg cgcttaatgc gccgtacag ggcgcgtcca ttcgccattc	4020
aggctgcga actgttggga agggcgatcg gtgcggcct cttcgtatt acgccagt	4080
gcgaaaggg gatgtgtgc aaggcgatta agttgggtaa cgccagggtt tcccagtca	4140
cgacgttga aaacgacggc cagtgaattg taatcact cactatagg cgaattgggc	4200
ccgacgtcg atgcattcca tagccacacc ttgcctatg gcttcacaac cgaaggcaat	4260
tcgagagtc gcgttatgg aatcgactcg tataaagctg aagggaagg gagacgttc	4320
gagcgtcag atgcaatagt cgtccagcta atgtggattc aaaaacaacc ccaacagtaa	4380
tcttgaaaat ttgaacggat caatctgaac actctgtc caggtcattc ttctaacgca	4440
catccccaga gtctagagg agttgtgtg tgaacatcct aataaacaat gcaatggatt	4500
cgggatatct tctgtctcg cccctactcg atgtcgagta aaccgatcac caactaaca	4560
tactctccg cgttctgcca ttgactctca aacagacatc gctatcaac gaacagcata	4620
ttttagcttc ttaggacaat aaatatgat aatgccggt ctcctcgggt atattaagca	4680
atccattcat acactcattc atcaggttaa ttttatatat ataattgtc tattcaaca	4740
ccgtaaatca ctggtaccat catctctcc tttcaaata cagctctatt tgcattaatg	4800
aaattactcg ccaattcgca gaacgtgtt gtgcaacaga gccttagctc gggtccagac	4860
aggagcagtg tctcgtgag gaagctgcag gagagttaat taaaaggcgt tgaaacagaa	4920
tgagccagac agcaaggaca aggtggccaa cagcaaggag tccaaaagc cctctattga	4980
cgagatccac gatgttattg ctcatgaggt ttccgagctc gatgtggga agaagaagt	5040
attgtatat aagaaataaa tgagatatag taaaggagt caagagaatg gcaagtggt	5100
caaattctat attacttga gtcactggt cctcgttgac atgaatgaag ttaccgttg	5160
catagctgat ttaatatata actgtccaac taactctcac ctatagataa cccatgtgtg	5220
tgttccacg cgtactcact ctgtcagctg cccacattg cccaatgcac aatgcacaat	5280
gatgtgtgca aacaacgcaa tcaaaagtct atgcatgctg accaaactct gatcaccaag	5340
ttgcgaacat gaaaagaag acctgtgtat atataagtaa gggggagagc ctaactaga	5400
tctttcgaaa accccccgac cttcacctc cacaacatg atcatctat acgttttggc	5460

cgttgcggtc tctttcctca tcttcaagag agtcacctac acgatgcgaa gccgagagct 5520
cgccaagaag tggcactgtg aggagcctca caacctgaat gatctagata gcggccgcca 5580
tcttaacggt gtgttcgtta agatgacccg ggacgaggct gcctttgccg agaccgagaa 5640
gtcattaac gcataaaggg atgtatactt gtactgtaga tggattaata aattatttat 5700

gatgaacaaa cattgactga gaaaggcacc atattgtagc ttcaagtact gtactggtag 5760
aagtaagcct gtcgttacc accatggag atccaatatt gttgatgcg gccatcggac 5820
tgcttgagc gtacagagaa ttggtaaaat gtggtggctg ggttggtttc gattgaaatt 5880
atcaagatct ctgaggtatt cttggtgttg aagggatgtt cttggactac tgcttgaac 5940
atagatggga ttaggggatg gcatggggta taatcgagac aaccattgtg gaggtaggca 6000
agagctctta caagtacat acgagtacaa ctactgagc ataagtctg gaattcccgg 6060
ttaagtccat tacagttgat atcattcaaa tcggacagat attctccgag gtgatacatc 6120

tctgtccccg agagaagctg atttggagca catccaggtt acggggggag agaaatttat 6180
gtaatgtgag gattttttgg agctgaataa taaagtcctg ttgatctagc ttcaattgag 6240
tcaacatccg ccgtcgactg acgtacgggt ggtgagaatg gttggttgta gggaagaatc 6300
aaaggccggt ctgggatcc gtgggtatat atatatatat atatatatac gatccttcgt 6360
tacctccctg ttctcaaaac tgtggttttt cgtttttcgt tttttgcttt ttttgatttt 6420
tttagggcca actaagcttc cagatttcgc taatcacctt tgtactaatt acaagaaagg 6480
aagaagctga ttagagttgg gctttttatg caactgtgct actccttacc tctgatatga 6540

aagtgtagac ccaatcacat catgtcattt agagttggta atactgggag gatagataag 6600
gcacgaaaac gagccatagc agacatgctg ggtgtagcca agcagaagaa agtagatggg 6660
agccaattga cgagcgaggg agctacgcca atccgacata cgacacgctg agatcgtctt 6720
ggccggggggg tacctacaga tgtccaaggg taagtgtctg actgtaatg tatgtctgag 6780
gacaaatatg tagtcagccg tataaagtca taccaggcac cagtgccatc atcgaaccac 6840
taactctcta tgatacatgc ctccggtatt attgtacat gcgtcgtttt gttacatacg 6900
tatcttgcct ttttctctca gaaactccag actttggcta ttggtcgaga taagcccgga 6960

ccatagttag tctttcacac tctgtttaaa caagctacca ccacactcgt tgggtgcagt 7020
cgccagctta aagatatcta tccacatcag ccacaactcc cttcttttaa taaaccgact 7080
acacccttgg ctattgaggt tatgagtga tatactgtag acaagacact ttcaagaaga 7140
ctgtttccaa aacgtaccac tgtcctccac tacaacaca cccaatctgc ttcttctagt 7200
caaggttgc acaccggtaa attataaatc atcatttcat tagcagggca gggccctttt 7260
tatagagtct tatacactag cggaccctgc cggtagacca acccgaggc gcgtcagttt 7320

gtctcttcca tcaatgcgtc gtagactagt tcagatctcg attt 7364

<210> 102

<211> 9211

<212> DNA

<213> Artificial Sequence

<220><223> Plasmid pYPS241

<400> 102

cgatgactgt cgactccaac ttttcacact gagcgtaaaa tgtggagaag aaatcggcac 60
taaaaagtca ggtagactgg aaaatgcgcc atgaaatgaa tatctcttgc tacagtaatg 120
cccagcatcg aggggtattg tgtcaccaac actatagtgg cagctgaagc gtcctgatt 180
gtagtatgag tctttattgg tgatgggaag agttcactca atattctcgt tactgccaaa 240
acaccacggt aatcggccag acaccatgga ttagatcac caagcctgtg aatgttattc 300
gagctaaaat gcacatggtt ggtgaaagga gtagttgctg tcgaattccg tcgtcgctg 360

agtcacatt tatttaccag ttggccacaa acccttgacg atctcgatg tccccccga 420
catactcccg gccggctggg gtacgttcga tagcgtatc ggcatcgaca aggtttgggt 480
ccctagccga taccgacta cctgagtcac aatcttcgga ggtttagtct tccacatagc 540
acgggcaaaa gtgcgtatat atacaagagc gtttgccagc cacagatttt cactccacac 600
accacatcac acatacaacc acacacatcc acaatggaac ccgaaactaa gaagaccaag 660
actgactcca agaagattgt tcttctcggc ggcgacttct gtggccccga ggtgattgcc 720
gaggccgtca aggtgtcaa gtctgttctg gaggcctccg gcaccgagtt tgtgtttgag 780

gaccgactca ttggaggagc tgccattgag aaggaggggc agcccatcac cgacgtact 840
ctcgacatct gccgaaagcg tgactctatt atgctcgggtg ctgtcggagg cgctgccaac 900
accgtatgga cactcccga cggacgaacc gacgtgagc cagagcaggg tctcctcaag 960
ctgcgaaagg acctgaacct gtacgccaac ctgcgacct gccagctgct gtgcaccaag 1020
ctgcgcgac tctccccat ccgaaacgtt gagggcaccg attcatcat tgtcgcgag 1080
ctcgtcggag gtatctactt tggagagcga aaggaggatg acgcatctgg cgtcgcttcc 1140
gacaccgaga cctactccgt tctgaggtt gagcgaattg cccgaatggc cgccttctg 1200

gcccttcagc acaaccccc tcttcccgtg tggctctctg acaaggccaa cgtgctggcc 1260
tcctctcgac ttggcgaaa gactgtcact cgagtcctca aggacgaatt cccccagctc 1320
gagctcaacc accagctgat cgactcggcc gccatgatcc tcatcaagca gccctccaag 1380
atgaatggta tcatcatcac caccaacatg ttggcgata tcatctccga cgaggcctcc 1440

gtcatcccg gtctctggg tctgtgccc tccgcctctc tggcttctct gcccgacacc	1500
aacgaggcgt tcggtctgta cgagccctgt cacggatctg ccccgatct cggcaagcag	1560
aaggtcaacc ccattgccac cattctgtct gccgcatga tgcacaagt ctctcttaac	1620
atgaagcccg ccggtgacgc tgttgaggct gccgtcaagg agtccgtcga ggctggatc	1680
actaccgccg atatcgagg ctcttctctc acctccgagg tcggagactt gtgccaaca	1740
aggtcaagga gctgctcaag aaggagtaag tcgtttctac gacgcattga tggaggagc	1800
aaactgacgc gcctgcgggt tggcttaccg gcagggtccg ctagtgtata agactctata	1860
aaaaggcccc tgcctgcta atgaaatgat gatattataat ttaccggtgt agcaaccttg	1920
actagaagaa gcagattggg tgtgtttgta gtggaggaca gtgtacgtt ttggaaacag	1980
tcttcttgaa agtgtctgt ctacagtata ttactcata acctcaatag ccaagggtgt	2040
agtcggttta ttaaaggaag ggagttgtgg ctgatgtgga tagatatctt taagctggcg	2100
actgcaccca acgagtgagg tggtagcttg tttaaacggg aggttagtgc ttggtatatg	2160
agttgtaggc atgacaattt ggaaaggggt ggactttggg aatattgtgg gatttcaata	2220
ccttagtttg tacgcaacta acatgaatga atacgatata catcaaagac tatgatacgc	2280
agtattgcac actgtacgag taagagcact agccactgca ctcaagtga accgttgccc	2340
gggtacgagt atgagtatgt acagtatgt tagtattgta cttggacagt gcttgatcg	2400
tacattctca agtgtcaaac ataaatatcc gttgtatat cctcgacca ccacgtagct	2460
cgctatatcc ctgtgttgaa tccatccatc ttggattgcc aattgtcac acagaaccgg	2520
gcactcactt ccccatccac acttcggcc gcttaggcct tgtctgcctt ggcaaccatt	2580
cgagcgaagg tgcccaggta ggagataaga atccgcacac cttcgagcat gggaggagat	2640
cggtaggcca ggttgtgctt cttgcacagc tgtccactt tgatggagc agcgggtgagg	2700
ttgtgtctgg gaagagtagg ccacaggtga tgttcgatct ggtagttgag gccaccgaaa	2760
aaccagtctg tgacgaggcc tcgtgaacg ttcattgtct cgtgaatctg tccagcacag	2820
aagccatgac catcccacac ggagtcctga atcttctcca gagggtagtg gttcatgaag	2880
acaacgatgg caattccgaa gccaccaagc agctcggaga caaagaacac catgagtccg	2940
gtgagaaacg agggcatgta gaaatagtag aacaaggcct tgagaccca gtggagtgcc	3000
agaccgacgg actccttctc gtactgtctt cgatagtact gattgcttcg atccttgaga	3060
cccgtggcag tatgaatgga ctgaaagcac cagatgaatc gcagaagggc acagatgaaa	3120
aagaagtaat actgttggtta cttgatcatc cgtcgagaaa acgggccagc tcgctcgacg	3180
tcctccttgg accacgccag gagaggcagg ttgtcgatgt cgggatcgtg acctgcacg	3240

ttggtggcgg agt gatgagc gttgtgtcgg tccttcacc aagtacaga aaagccctga	3300
agcacgtttc cgaagagcag gccaatggca ttgttgatgg atcggtcctt gaacagctgg	3360
tgatggcaaa tgtcgtgcga gagccatccc atctgttgaa agtgaatgcc gaggacgata	3420
gcaccaacca ggtaccagtg ccactgggtc atcaacagca cggcaaggac acccagtcgg	3480
agagtcgaca gggctcttgta agcgtaccac atgggagagg catcgaacat gcctgcagca	3540
atgagctcgt ctgcagctt tcggaatcc tcctgagggt cgtcggagtt tggcttggga	3600
gggttggag tcaggggaga ggaaggttcc atgatgggca ttcgtctgag ttgacacg	3660
gcgttctcgg agtcataac catgaagacg tcggtagcat cccgacctcg gtagttctcg	3720
ataatgtcgg cacctccggg atggtggttg agccaggcag aaacgtcgta ggtgacaccg	3780
tccacggtga gaggcagagc gggctcgttg accatggttg atgtgtgtt aattcaagaa	3840
tgaatataga gaagagaaga agaaaaaga ttcaattgag ccaaaccgat ttaaatgagt	3900
atctgtctga ctgtcattg ccgcctttgg agtacactc caactatgag tgtgcttgg	3960
tcactttgac gatacattct tcgttggagg ctgtgggtct gacagctgcg ttttcggcgc	4020
ggttggccga caacaatata agctgcaacg tcattgtctgg cttcatcat gatcacattt	4080
ttgtcggcaa aggcgacgcc cagagagcca ttgacgttct ttctaatttg gaccgatagc	4140
cgtatagtec agtctatcta taagttcaac taactcgtaa ctattacat aacataact	4200
tcactgcccc agataagggt ccgataaaaa gtctgcaga ctaatttat ttcagctcc	4260
tcttcaccac caaatgccc tcctacgaag ctgcagctaa cgtccacaag tccgcctttg	4320
ccgtcagat gctcaagctc gtggcagcca agaaaaccaa cctgtgtgct tctctggatg	4380
ttaccaccac caaggagctc attgagcttg ccgataagggt cggaccttat gtgtgcatga	4440
tcaaaacca tatgcacatc attgacgact tcacctacgc cggcactgtg cccccctca	4500
aggaacttgc tcttaagcac gggttcttcc tgttcagga cagaaagttc gcagatattg	4560
gcaacactgt caagcaccag taccggtgtc accgaatcg cgagtgttc gatcacca	4620
acgccacgg tgtaccgga accggaatca ttgtggcct gcgagctggt gccgaggaaa	4680
ctgtctctga acagaagaag gaggacgtct ctgactacga gaactcccag tacaaggagt	4740
tcctagtccc ctctcccaac gagaagctgg ccagaggtct gctcatgctg gccgagctgt	4800
cttgcaagg ctctctggcc actggcagat actccaagca gaccattgag ctgcccgat	4860
ccgacccga gtttgtgtt ggcttcattg ccagaaccg acctaaaggc gactctgagg	4920
actggcttat tctgaccccc ggggtgggtc ttgacgaaa gggagacgt ctcggacagc	4980

agtaccgaac tgttgaggat gtcattgtcta ccggaacgga tatcataatt gtcggccgag 5040
 gtctgtaccg ccagaaccga gatcctattg aggaggccaa gcgataccag aaggctggct 5100
 gggaggctta ccagaagatt aactgttaga ggtagacta tggatatgta atttaactgt 5160
 gtatatagag agcgtgcaag tatggagcgc ttgttcagct tgtatgatgg tcagacgacc 5220
 tgtctgatcg agtatgtatg atactgcaca acctgtgtat ccgcatgatc tgtccaatgg 5280
 ggcatgttgt tgtgtttctc gatacggaga tgcgggtac agtgctaata cgttgaacta 5340
 cttatactta tatgaggctc gaagaaagct gacttgtgta tgacttaatt aacgcggcgc 5400

 gccagctgca ttaatgaatc ggccaacgcg cggggagagg cgttttgcgt attgggcgct 5460
 ctccgcttc ctgctcact gactcgctgc gctcggctgt tcggctgcgg cgagcggtat 5520
 cagctcactc aaaggcggtg atacggttat ccacagaatc aggggataac gcaggaaaga 5580
 acatgtgagc aaaaggccag caaaaggcca ggaaccgtaa aaaggccgcg ttgctggcgt 5640
 tttccatag gctccgccc cctgacgagc atcacaaaaa tcgacgtca agtcagaggt 5700
 ggcgaaacc gacaggacta taaagatacc aggcgtttcc ccctggaagc tcctcgtgc 5760
 gctctcctgt tccgacctg ccgcttaccg gatacctgac cgcctttctc cttcgggaa 5820

 gcgtggcgct ttctcatagc tcacgtgta ggtatctcag ttcggtgtag gtcgttcgt 5880
 ccaagctggg ctgtgtgcac gaacccccg ttcagccga ccgtgcgc ttatccggtg 5940
 actatcgtct tgagtccaac ccggtgaagc acgacttacc gccactggca gcagccactg 6000
 glaacaggat tagcagagcg aggtatgtag gcggtgctac agagtcttg aagtgtggc 6060
 ctaactacgg ctacactaga agaacagtat ttggtatctg cgctctgctg aagccagtta 6120
 cttcggaaa aagagttggt agctcttgat ccggcaaaaca aaccaccgct ggtagcgggtg 6180
 gttttttgt ttgcaagcag cagattacgc gcagaaaaaa aggatctcaa gaagatcctt 6240

 tgatcttttc tacggggtct gacgtcagt ggaacgaaaa ctcacgttaa gggattttgg 6300
 tcatgagatt atcaaaaagg atcttcacct agatcctttt aaattaaaaa tgaagtttta 6360
 aatcaatcta aagtatatat gagtaaacct ggtctgacag ttaccaatgc ttaatcagtg 6420
 aggcacctat ctacgcgac tgtctatttc gttcatccat agttgcctga ctccccgtcg 6480
 tgtagataac tacgatacgg gagggcttac catctggccc cagtgtgca atgataccgc 6540
 gagaccacg ctaccggct ccagatttat cagcaataaa ccagccagcc ggaagggccg 6600
 agcgcagaag tggctctgca actttatccg cctccatcca gtctattaat tgttgccggg 6660

 aagctagagt aagtagttcg ccagttaata gtttgcgcaa cgttgttgcc attgctacag 6720
 gcatcgttgt gtcacgtcg tcgtttggta tggcttcatt cagctccggt tcccaacgat 6780
 caaggcgagt tacatgatcc ccatgttgt gcaaaaaagc ggtagctcc ttcggtcctc 6840

cgatcgttgt cagaagtaag ttggccgcag tgttatcact catggttatg gcagcactgc 6900
ataattctct tactgtcatg ccatccgtaa gatgcttttc tgtgactggt gactactcaa 6960
ccaagtcatt ctgagaatag tgtatgcggc gaccgagttg ctcttgcccg gcgtcaatac 7020
gggataatac cgcgccacat agcagaactt taaaagtgt catcattgga aaacgttctt 7080

cggggcgaaa actctcaagg atcttaccgc tgttgagatc cagttcgatg taaccactc 7140
gtgcacccaa ctgatcttca gcatctttta ctttcaccag cgtttctggg tgagcaaaaa 7200
caggaaggca aaatgccga aaaaaggga taaggcgac acggaaatgt tgaatactca 7260
tactcttct ttttcaatat tattgaagca tttatcaggg ttattgtctc atgagcggat 7320
acatatttga atgtatttag aaaaataaac aaataggggt tccgcgcaca tttccccgaa 7380
aagtgccacc tgatgcggtg tgaaataacc cacagatgcg taaggagaaa ataccgcatc 7440
aggaaattgt aagcgtaaat attttgtaa aattcgcgtt aaattttgt taaatcagct 7500

cattttttaa ccaataggcc gaaatcggca aaatccctta taaatcaaaa gaatagaccg 7560
agatagggtt gagtgttgtt ccagtttga acaagagtcc actattaaag aacgtggact 7620
ccaacgtcaa agggcgaaaa accgtctatc agggcgatgg cccactacgt gaaccatcac 7680
cctaatacag ttttttggg tgcaggtgcc gtaaagcact aaatcggaac ctaaaggga 7740
gccccgatt tagagcttga cggggaaagc cggcgaacgt ggcgagaaag gaagggaaga 7800
aagcgaaagg agcgggcgtt agggcgctgg caagtgtagc ggtcacgctg cgcgtaacca 7860
ccacaccgc cgcgcttaat gcgccgtac agggcgctc cattcgccat tcaggctgcg 7920

caactgttgg gaaggcgat cggtcgggc ctcttcgcta ttacgccagc tggcgaaagg 7980
gggatgtgct gcaaggcgt taagtgggt aacgccaggg ttttccagt cacgacgttg 8040
taaaacgacg gccagtgaat tgtaatacga ctactatag gggaatttg gcccacgtc 8100
gcatgccact tagtctacag attacgtgtg gtgtactggt acctgtggtg tagcgacagt 8160
attgtacata ctgtactcc cactctacac acccgccgcc accacttacc ctctgagaa 8220
ccccctcaga aactcacaac cgtcacctta gcagaagggt ttagccgtt ggagccgaca 8280
gggatgtcag gcaatacagc tgtcaccgag gggacaccgg ggtgtcaaca acagtgtcat 8340

ttcggtgctt gtggagttag tgacatcacc gagaaagttag tagtatgtag gggtttttgg 8400
aatggttgtt tcatgatcca ttgggtactgc tcggcaaaaa tagcctgatg gaatcacgtt 8460
tcgttctggg gttagcggc atctcagctt tctccatcca cgagtgtctc agagataaga 8520
cctgcatatc ttttatctgt tttcaaagt taagtctat agagtgccg gagggggtca 8580
aggatgtgca tggacgtgac tacagtgaga gtttttggag gaagtaaga ccgagtatcg 8640
accatccaga aaccccatat cttctggaat tccacaaatg gccaaactg gcttctttta 8700

tccacttttt gataactaaaa atgtctaatt tgacctctct tccactaaca aaaacagtct 8760

cgccgttctc caaccgcac caccaactac aagtacagta gctataattt ttggcacca 8820

ctcgttttct tctttttagg cgaaacttgt ctcaaactta tcatgccata ggcattaatt 8880

tccgggtctt gcagagctac tgcaactggg ggggggggtg tacaacgcac gaccgagtgc 8940

tatgtaatgt acacaatggc cacaacaca actgttactg atgtaccgat gactcagcac 9000

ctacagtagt ccacgcgttt gacttccggg gctcttcggg cccaattttt ttcttcgtct 9060

gcttcccaaa tgccagattt atgcaagggt catgtgacgt cccgggtccga ggattgctat 9120

aacgggcgaa ataaggacac aaacaactca aaaataatac caaaatatag ctgcaacaa 9180

gggcttcaga ttgatcccat tattggtcca t 9211

<210> 103

<211> 1260

<212> DNA

<213> Euglena anabaena UTEX 373

<220><221> CDS

<222> (1)..(1260)

<223> synthetic delta-8 desaturase (codon-optimized for Yarrowia lipolytica)

<400> 103

atg gtc aag cga ccc gct ctg cct ctc acc gtg gac ggt gtc acc tac 48

Met Val Lys Arg Pro Ala Leu Pro Leu Thr Val Asp Gly Val Thr Tyr

1 5 10 15

gac gtt tct gcc tgg ctc aac cac cat ccc gga ggt gcc gac att atc 96

Asp Val Ser Ala Trp Leu Asn His His Pro Gly Gly Ala Asp Ile Ile

20 25 30

gag aac tac cga ggt cgg gat gct acc gac gtc ttc atg gtt atg cac 144

Glu Asn Tyr Arg Gly Arg Asp Ala Thr Asp Val Phe Met Val Met His

35 40 45

tcc gag aac gcc gtg tcc aaa ctc aga cga atg ccc atc atg gaa cct 192

Ser Glu Asn Ala Val Ser Lys Leu Arg Arg Met Pro Ile Met Glu Pro

50 55 60

tcc tct ccc ctg act cca aca cct ccc aag cca aac tcc gac gaa cct 240

225	230	235	240	
atc tgg tgc ttt cag tcc att cat act gcc acg ggt ctc aag gat cga				768
Ile Trp Cys Phe Gln Ser Ile His Thr Ala Thr Gly Leu Lys Asp Arg				
245	250	255		
agc aat cag tac tat cga aga cag tac gag aag gag tcc gtc ggt ctg				816
Ser Asn Gln Tyr Tyr Arg Arg Gln Tyr Glu Lys Glu Ser Val Gly Leu				
260	265	270		
gca ctc cac tgg ggt ctc aag gcc ttg ttc tac tat ttc tac atg ccc				864
Ala Leu His Trp Gly Leu Lys Ala Leu Phe Tyr Tyr Phe Tyr Met Pro				
275	280	285		
tcg ttt ctc acc gga ctc atg gtg ttc ttt gtc tcc gag ctg ctt ggt				912
Ser Phe Leu Thr Gly Leu Met Val Phe Phe Val Ser Glu Leu Leu Gly				
290	295	300		
ggc ttc gga att gcc atc gtt gtc ttc atg aac cac tac cct ctg gag				960
Gly Phe Gly Ile Ala Ile Val Val Phe Met Asn His Tyr Pro Leu Glu				
305	310	315	320	
aag att cag gac tcc gtg tgg gat ggt cat ggc ttc tgt gct gga cag				1008
Lys Ile Gln Asp Ser Val Trp Asp Gly His Gly Phe Cys Ala Gly Gln				
325	330	335		
att cac gag acc atg aac gtt cag cga ggc ctc gtc aca gac tgg ttt				1056
Ile His Glu Thr Met Asn Val Gln Arg Gly Leu Val Thr Asp Trp Phe				
340	345	350		
ttc ggt ggc ctc aac tac cag atc gaa cat cac ctg tgg cct act ctt				1104
Phe Gly Gly Leu Asn Tyr Gln Ile Glu His His Leu Trp Pro Thr Leu				
355	360	365		
ccc aga cac aac ctc acc gct gcc tcc atc aaa gtg gag cag ctg tgc				1152
Pro Arg His Asn Leu Thr Ala Ala Ser Ile Lys Val Glu Gln Leu Cys				
370	375	380		
aag aag cac aac ctg ccc tac cga tct cct ccc atg ctc gaa ggt gtc				1200

Lys Lys His Asn Leu Pro Tyr Arg Ser Pro Pro Met Leu Glu Gly Val
 385 390 395 400
 ggc att ctt atc tcc tac ctg ggc acc ttc gct cga atg gtt gcc aag 1248
 Gly Ile Leu Ile Ser Tyr Leu Gly Thr Phe Ala Arg Met Val Ala Lys
 405 410 415
 gca gac aag gcc 1260
 Ala Asp Lys Ala
 420
 <210> 104
 <211> 420
 <212> PRT
 <213> Euglena anabaena UTEX 373
 <400> 104
 Met Val Lys Arg Pro Ala Leu Pro Leu Thr Val Asp Gly Val Thr Tyr
 1 5 10 15
 Asp Val Ser Ala Trp Leu Asn His His Pro Gly Gly Ala Asp Ile Ile
 20 25 30
 Glu Asn Tyr Arg Gly Arg Asp Ala Thr Asp Val Phe Met Val Met His
 35 40 45
 Ser Glu Asn Ala Val Ser Lys Leu Arg Arg Met Pro Ile Met Glu Pro
 50 55 60
 Ser Ser Pro Leu Thr Pro Thr Pro Pro Lys Pro Asn Ser Asp Glu Pro
 65 70 75 80
 Gln Glu Asp Phe Arg Lys Leu Arg Asp Glu Leu Ile Ala Ala Gly Met
 85 90 95
 Phe Asp Ala Ser Pro Met Trp Tyr Ala Tyr Lys Thr Leu Ser Thr Leu
 100 105 110
 Gly Leu Gly Val Leu Ala Val Leu Leu Met Thr Gln Trp His Trp Tyr
 115 120 125
 Leu Val Gly Ala Ile Val Leu Gly Ile His Phe Gln Gln Met Gly Trp
 130 135 140
 Leu Ser His Asp Ile Cys His His Gln Leu Phe Lys Asp Arg Ser Ile

145 150 155 160
 Asn Asn Ala Ile Gly Leu Leu Phe Gly Asn Val Leu Gln Gly Phe Ser
 165 170 175

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

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

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

 Pro Arg His Asn Leu Thr Ala Ala Ser Ile Lys Val Glu Gln Leu Cys
 370 375 380
 Lys Lys His Asn Leu Pro Tyr Arg Ser Pro Pro Met Leu Glu Gly Val
 385 390 395 400

Gly Ile Leu Ile Ser Tyr Leu Gly Thr Phe Ala Arg Met Val Ala Lys
405 410 415

Ala Asp Lys Ala
420

<210> 105

<211> 1350

<212> DNA

<213> *Euglena gracilis*

<220><221> CDS

<222> (1)..(1350)

<223> EgD5M (i.e., EgD5R*-34g158g) mutant delta-5 desaturase

<400> 105

atg gcc ctg tct ctc act acc gaa cag ctc ctg gag cga cct gat ctc 48

Met Ala Leu Ser Leu Thr Thr Glu Gln Leu Leu Glu Arg Pro Asp Leu

1 5 10 15

gtc gct atc gat ggt atc ctg tac gac ctc gag ggc ctg gcc aag gtg 96

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

20 25 30

cat ggt ggt gga gac ctc att ctg gcc tct gga gcc tcc gac gcc tct 144

His Gly Gly Gly Asp Leu Ile Leu Ala Ser Gly Ala Ser Asp Ala Ser

35 40 45

ccc ctc ttc tac tct atg cat ccc tac gtc aag ccc gag aac tcc aag 192

Pro Leu Phe Tyr Ser Met His Pro Tyr Val Lys Pro Glu Asn Ser Lys

50 55 60

ctc ctg cag caa ttc gtc cga ggg aag cat gac cgc acc tcg aag gac 240

Leu Leu Gln Gln Phe Val Arg Gly Lys His Asp Arg Thr Ser Lys Asp

65 70 75 80

att gtc tac acg tat gat tct ccc ttc gca caa gac gtt aag cgg aca 288

Ile Val Tyr Thr Tyr Asp Ser Pro Phe Ala Gln Asp Val Lys Arg Thr

85 90 95

atg cgc gag gtg atg aaa ggg agg aac tgg tac gca acc cct ggc ttc 336

Met Arg Glu Val Met Lys Gly Arg Asn Trp Tyr Ala Thr Pro Gly Phe

100 105 110

tgg ctg cgc acc gtt ggg atc atc gcc gtg acg gcc ttt tgc gag tgg 384

Trp Leu Arg Thr Val Gly Ile Ile Ala Val Thr Ala Phe Cys Glu Trp

115 120 125

cac tgg gct acc acg ggg atg gtg ctg tgg ggc ctg ttg act gga ttc 432

His Trp Ala Thr Thr Gly Met Val Leu Trp Gly Leu Leu Thr Gly Phe

130 135 140

atg cac atg cag atc ggc tta tcc atc cag cat gat gcg ggt cac ggg 480

Met His Met Gln Ile Gly Leu Ser Ile Gln His Asp Ala Gly His Gly

145 150 155 160

gcc atc agc aag aag cct tgg gtc aac gcc ctc ttc gcc tac ggc att 528

Ala Ile Ser Lys Lys Pro Trp Val Asn Ala Leu Phe Ala Tyr Gly Ile

165 170 175

gac gtc atc gga tcg tcc cgg tgg att tgg ctg cag tcg cac atc atg 576

Asp Val Ile Gly Ser Ser Arg Trp Ile Trp Leu Gln Ser His Ile Met

180 185 190

cgg cac cac acc tac acc aac cag cac ggc ctc gac ctg gat gcg gag 624

Arg His His Thr Tyr Thr Asn Gln His Gly Leu Asp Leu Asp Ala Glu

195 200 205

tcg gca gag ccg ttc ctg gtg ttc cac aac tac ccc gcc gca aac acc 672

Ser Ala Glu Pro Phe Leu Val Phe His Asn Tyr Pro Ala Ala Asn Thr

210 215 220

gcc cga aag tgg ttc cac cgc ttc cag gct tgg tac atg tac ctt gtg 720

Ala Arg Lys Trp Phe His Arg Phe Gln Ala Trp Tyr Met Tyr Leu Val

225 230 235 240

ctg ggg gca tac ggg gta tcg ctg gtg tac aac ccg ctc tac att ttc 768

Leu Gly Ala Tyr Gly Val Ser Leu Val Tyr Asn Pro Leu Tyr Ile Phe

245 250 255

cgg atg cag cac aat gac acc atc cca gag tct gtc acg gcc atg cgg	816
Arg Met Gln His Asn Asp Thr Ile Pro Glu Ser Val Thr Ala Met Arg	
260 265 270	
gaa aat ggc ttt ctg cgg cgc tac cgc aca ctt gca ttc gtg atg cga	864
Glu Asn Gly Phe Leu Arg Arg Tyr Arg Thr Leu Ala Phe Val Met Arg	
275 280 285	
gct ttc ttc atc ttc cgg acc gca ttc ttg ccc tgg tac ctc act ggg	912
Ala Phe Phe Ile Phe Arg Thr Ala Phe Leu Pro Trp Tyr Leu Thr Gly	
290 295 300	
acc tca ttg ctg atc acc att cct ctg gtg ccc acc gca act ggt gcc	960
Thr Ser Leu Leu Ile Thr Ile Pro Leu Val Pro Thr Ala Thr Gly Ala	
305 310 315 320	
ttc ttg acg ttc ttc ttc att ttg tcc cac aat ttt gat ggc tcc gaa	1008
Phe Leu Thr Phe Phe Phe Ile Leu Ser His Asn Phe Asp Gly Ser Glu	
325 330 335	
cgg atc ccc gac aag aac tgc aag gtt aag cga tct gag aag gac gtt	1056
Arg Ile Pro Asp Lys Asn Cys Lys Val Lys Arg Ser Glu Lys Asp Val	
340 345 350	
gag gct gac caa att gac tgg tat cgg gcg cag gtg gag acg tcc tcc	1104
Glu Ala Asp Gln Ile Asp Trp Tyr Arg Ala Gln Val Glu Thr Ser Ser	
355 360 365	
aca tac ggt ggc ccc atc gcc atg ttc ttc act ggc ggt ctc aat ttc	1152
Thr Tyr Gly Gly Pro Ile Ala Met Phe Phe Thr Gly Gly Leu Asn Phe	
370 375 380	
cag atc gag cac cac ctc ttt ccc cgg atg tcg tct tgg cac tac ccc	1200
Gln Ile Glu His His Leu Phe Pro Arg Met Ser Ser Trp His Tyr Pro	
385 390 395 400	
ttc gtc cag cag gcg gtc cgg gag tgt tgc gaa cga cat gga gtg cga	1248
Phe Val Gln Gln Ala Val Arg Glu Cys Cys Glu Arg His Gly Val Arg	
405 410 415	
tat gtt ttc tac cct acc atc gtc ggc aac atc atc tcc acc ctg aag	1296

Tyr Val Phe Tyr Pro Thr Ile Val Gly Asn Ile Ile Ser Thr Leu Lys
420 425 430

tac atg cat aag gtg ggt gtc gtc cac tgc gtg aag gac gca cag gat 1344

Tyr Met His Lys Val Gly Val Val His Cys Val Lys Asp Ala Gln Asp

435 440 445

tcc taa 1350

Ser

<210> 106

<211> 449

<212> PRT

<213> Euglena gracilis

<400> 106

Met Ala Leu Ser Leu Thr Thr Glu Gln Leu Leu Glu Arg Pro Asp Leu

1 5 10 15

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

20 25 30

His Gly Gly Gly Asp Leu Ile Leu Ala Ser Gly Ala Ser Asp Ala Ser

35 40 45

Pro Leu Phe Tyr Ser Met His Pro Tyr Val Lys Pro Glu Asn Ser Lys

50 55 60

Leu Leu Gln Gln Phe Val Arg Gly Lys His Asp Arg Thr Ser Lys Asp

65 70 75 80

Ile Val Tyr Thr Tyr Asp Ser Pro Phe Ala Gln Asp Val Lys Arg Thr

85 90 95

Met Arg Glu Val Met Lys Gly Arg Asn Trp Tyr Ala Thr Pro Gly Phe

100 105 110

Trp Leu Arg Thr Val Gly Ile Ile Ala Val Thr Ala Phe Cys Glu Trp

115 120 125

His Trp Ala Thr Thr Gly Met Val Leu Trp Gly Leu Leu Thr Gly Phe

130 135 140

Met His Met Gln Ile Gly Leu Ser Ile Gln His Asp Ala Gly His Gly
145 150 155 160
Ala Ile Ser Lys Lys Pro Trp Val Asn Ala Leu Phe Ala Tyr Gly Ile
165 170 175
Asp Val Ile Gly Ser Ser Arg Trp Ile Trp Leu Gln Ser His Ile Met
180 185 190
Arg His His Thr Tyr Thr Asn Gln His Gly Leu Asp Leu Asp Ala Glu
195 200 205

Ser Ala Glu Pro Phe Leu Val Phe His Asn Tyr Pro Ala Ala Asn Thr
210 215 220
Ala Arg Lys Trp Phe His Arg Phe Gln Ala Trp Tyr Met Tyr Leu Val
225 230 235 240
Leu Gly Ala Tyr Gly Val Ser Leu Val Tyr Asn Pro Leu Tyr Ile Phe
245 250 255
Arg Met Gln His Asn Asp Thr Ile Pro Glu Ser Val Thr Ala Met Arg
260 265 270

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

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

385 390 395 400

Phe Val Gln Gln Ala Val Arg Glu Cys Cys Glu Arg His Gly Val Arg

405 410 415

Tyr Val Phe Tyr Pro Thr Ile Val Gly Asn Ile Ile Ser Thr Leu Lys

420 425 430

Tyr Met His Lys Val Gly Val Val His Cys Val Lys Asp Ala Gln Asp

435 440 445

Ser

<210> 107

<211> 1350

<212> DNA

<213> *Euglena gracilis*

<220><221> CDS

<222> (1)..(1350)

<223> EgD5M1 (i.e., EgD5R*-34g158g347s) mutant delta-5 desaturase

<400> 107

atg gcc ctg tct ctc act acc gaa cag ctc ctg gag cga cct gat ctc 48

Met Ala Leu Ser Leu Thr Thr Glu Gln Leu Leu Glu Arg Pro Asp Leu

1 5 10 15

gtc gct atc gat ggt atc ctg tac gac ctc gag ggc ctg gcc aag gtg 96

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

20 25 30

cat ggt ggt gga gac ctc att ctg gcc tct gga gcc tcc gac gcc tct 144

His Gly Gly Gly Asp Leu Ile Leu Ala Ser Gly Ala Ser Asp Ala Ser

35 40 45

ccc ctc ttc tac tct atg cat ccc tac gtc aag ccc gag aac tcc aag 192

Pro Leu Phe Tyr Ser Met His Pro Tyr Val Lys Pro Glu Asn Ser Lys

50 55 60

ctc ctg cag caa ttc gtc cga ggg aag cat gac cgc acc tcg aag gac 240

Leu Leu Gln Gln Phe Val Arg Gly Lys His Asp Arg Thr Ser Lys Asp

65	70	75	80	
att gtc tac acg tat gat tct ccc ttc gca caa gac gtt aag cgg aca				288
Ile Val Tyr Thr Tyr Asp Ser Pro Phe Ala Gln Asp Val Lys Arg Thr				
	85	90	95	
atg cgc gag gtg atg aaa ggg agg aac tgg tac gca acc cct ggc ttc				336
Met Arg Glu Val Met Lys Gly Arg Asn Trp Tyr Ala Thr Pro Gly Phe				
	100	105	110	
tgg ctg cgc acc gtt ggg atc atc gcc gtg acg gcc ttt tgc gag tgg				384
Trp Leu Arg Thr Val Gly Ile Ile Ala Val Thr Ala Phe Cys Glu Trp				
	115	120	125	
cac tgg gct acc acg ggg atg gtg ctg tgg ggc ctg ttg act gga ttc				432
His Trp Ala Thr Thr Gly Met Val Leu Trp Gly Leu Leu Thr Gly Phe				
	130	135	140	
atg cac atg cag atc ggc tta tcc atc cag cat gat gcg ggt cac ggg				480
Met His Met Gln Ile Gly Leu Ser Ile Gln His Asp Ala Gly His Gly				
	145	150	155	160
gcc atc agc aag aag cct tgg gtc aac gcc ctc ttc gcc tac ggc att				528
Ala Ile Ser Lys Lys Pro Trp Val Asn Ala Leu Phe Ala Tyr Gly Ile				
	165	170	175	
gac gtc atc gga tcg tcc cgg tgg att tgg ctg cag tcg cac atc atg				576
Asp Val Ile Gly Ser Ser Arg Trp Ile Trp Leu Gln Ser His Ile Met				
	180	185	190	
cgg cac cac acc tac acc aac cag cac ggc ctc gac ctg gat gcg gag				624
Arg His His Thr Tyr Thr Asn Gln His Gly Leu Asp Leu Asp Ala Glu				
	195	200	205	
tcg gca gag ccg ttc ctg gtg ttc cac aac tac ccc gcc gca aac acc				672
Ser Ala Glu Pro Phe Leu Val Phe His Asn Tyr Pro Ala Ala Asn Thr				
	210	215	220	
gcc cga aag tgg ttc cac cgc ttc cag gct tgg tac atg tac ctt gtg				720
Ala Arg Lys Trp Phe His Arg Phe Gln Ala Trp Tyr Met Tyr Leu Val				
	225	230	235	240

ctg ggg gca tac ggg gta tcg ctg gtg tac aac ccg ctc tac att ttc	768
Leu Gly Ala Tyr Gly Val Ser Leu Val Tyr Asn Pro Leu Tyr Ile Phe	
245 250 255	
cgg atg cag cac aat gac acc atc cca gag tct gtc acg gcc atg cgg	816
Arg Met Gln His Asn Asp Thr Ile Pro Glu Ser Val Thr Ala Met Arg	
260 265 270	
gaa aat ggc ttt ctg cgg cgc tac cgc aca ctt gca ttc gtg atg cga	864
Glu Asn Gly Phe Leu Arg Arg Tyr Arg Thr Leu Ala Phe Val Met Arg	
275 280 285	
gct ttc ttc atc ttc cgg acc gca ttc ttg ccc tgg tac ctc act ggg	912
Ala Phe Phe Ile Phe Arg Thr Ala Phe Leu Pro Trp Tyr Leu Thr Gly	
290 295 300	
acc tca ttg ctg atc acc att cct ctg gtg ccc acc gca act ggt gcc	960
Thr Ser Leu Leu Ile Thr Ile Pro Leu Val Pro Thr Ala Thr Gly Ala	
305 310 315 320	
ttc ttg acg ttc ttc ttc att ttg tcc cac aat ttt gat ggc tcc gaa	1008
Phe Leu Thr Phe Phe Phe Ile Leu Ser His Asn Phe Asp Gly Ser Glu	
325 330 335	
cgg atc ccc gac aag aac tgc aag gtt aag agc tct gag aag gac gtt	1056
Arg Ile Pro Asp Lys Asn Cys Lys Val Lys Ser Ser Glu Lys Asp Val	
340 345 350	
gag gct gac caa att gac tgg tat cgg gcg cag gtg gag acg tcc tcc	1104
Glu Ala Asp Gln Ile Asp Trp Tyr Arg Ala Gln Val Glu Thr Ser Ser	
355 360 365	
aca tac ggt ggc ccc atc gcc atg ttc ttc act ggc ggt ctc aat ttc	1152
Thr Tyr Gly Gly Pro Ile Ala Met Phe Phe Thr Gly Gly Leu Asn Phe	
370 375 380	
cag atc gag cac cac ctc ttt ccc cgg atg tcg tct tgg cac tac ccc	1200
Gln Ile Glu His His Leu Phe Pro Arg Met Ser Ser Trp His Tyr Pro	
385 390 395 400	

ttc gtc cag cag gcg gtc cgg gag tgt tgc gaa cga cat gga gtg cga 1248

Phe Val Gln Gln Ala Val Arg Glu Cys Cys Glu Arg His Gly Val Arg

405 410 415

tat gtt ttc tac cct acc atc gtc ggc aac atc atc tcc acc ctg aag 1296

Tyr Val Phe Tyr Pro Thr Ile Val Gly Asn Ile Ile Ser Thr Leu Lys

420 425 430

tac atg cat aag gtg ggt gtc gtc cac tgc gtg aag gac gca cag gat 1344

Tyr Met His Lys Val Gly Val Val His Cys Val Lys Asp Ala Gln Asp

435 440 445

tcc taa 1350

Ser

<210> 108

<211> 449

<212> PRT

<213> Euglena gracilis

<400> 108

Met Ala Leu Ser Leu Thr Thr Glu Gln Leu Leu Glu Arg Pro Asp Leu

1 5 10 15

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

20 25 30

His Gly Gly Gly Asp Leu Ile Leu Ala Ser Gly Ala Ser Asp Ala Ser

35 40 45

Pro Leu Phe Tyr Ser Met His Pro Tyr Val Lys Pro Glu Asn Ser Lys

50 55 60

Leu Leu Gln Gln Phe Val Arg Gly Lys His Asp Arg Thr Ser Lys Asp

65 70 75 80

Ile Val Tyr Thr Tyr Asp Ser Pro Phe Ala Gln Asp Val Lys Arg Thr

85 90 95

Met Arg Glu Val Met Lys Gly Arg Asn Trp Tyr Ala Thr Pro Gly Phe

100 105 110

Trp Leu Arg Thr Val Gly Ile Ile Ala Val Thr Ala Phe Cys Glu Trp
115 120 125

His Trp Ala Thr Thr Gly Met Val Leu Trp Gly Leu Leu Thr Gly Phe
130 135 140

Met His Met Gln Ile Gly Leu Ser Ile Gln His Asp Ala Gly His Gly
145 150 155 160

Ala Ile Ser Lys Lys Pro Trp Val Asn Ala Leu Phe Ala Tyr Gly Ile
165 170 175

Asp Val Ile Gly Ser Ser Arg Trp Ile Trp Leu Gln Ser His Ile Met
180 185 190

Arg His His Thr Tyr Thr Asn Gln His Gly Leu Asp Leu Asp Ala Glu
195 200 205

Ser Ala Glu Pro Phe Leu Val Phe His Asn Tyr Pro Ala Ala Asn Thr
210 215 220

Ala Arg Lys Trp Phe His Arg Phe Gln Ala Trp Tyr Met Tyr Leu Val
225 230 235 240

Leu Gly Ala Tyr Gly Val Ser Leu Val Tyr Asn Pro Leu Tyr Ile Phe
245 250 255

Arg Met Gln His Asn Asp Thr Ile Pro Glu Ser Val Thr Ala Met Arg
260 265 270

Glu Asn Gly Phe Leu Arg Arg Tyr Arg Thr Leu Ala Phe Val Met Arg
275 280 285

Ala Phe Phe Ile Phe Arg Thr Ala Phe Leu Pro Trp Tyr Leu Thr Gly
290 295 300

Thr Ser Leu Leu Ile Thr Ile Pro Leu Val Pro Thr Ala Thr Gly Ala
305 310 315 320

Phe Leu Thr Phe Phe Phe Ile Leu Ser His Asn Phe Asp Gly Ser Glu
325 330 335

Arg Ile Pro Asp Lys Asn Cys Lys Val Lys Ser Ser Glu Lys Asp Val
340 345 350

Glu Ala Asp Gln Ile Asp Trp Tyr Arg Ala Gln Val Glu Thr Ser Ser

355 360 365
 Thr Tyr Gly Gly Pro Ile Ala Met Phe Phe Thr Gly Gly Leu Asn Phe
 370 375 380
 Gln Ile Glu His His Leu Phe Pro Arg Met Ser Ser Trp His Tyr Pro
 385 390 395 400

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

<210> 109

<211> 1350

<212> DNA

<213> *Euglena gracilis*

<220><221> CDS

<222> (1)..(1350)

<223> EgD5S-36s157g mutant delta-5 desaturase

<400> 109

atg gct ctc tcc ctt act acc gag cag ctg ctc gag cga ccc gac ctg	48
Met Ala Leu Ser Leu Thr Thr Glu Gln Leu Leu Glu Arg Pro Asp Leu	
1 5 10 15	
gtt gcc atc gac ggc att ctc tac gat ctg gaa ggt ctt gcc aag gtc	96
Val Ala Ile Asp Gly Ile Leu Tyr Asp Leu Glu Gly Leu Ala Lys Val	
20 25 30	
cat ccc gga tcc gac ttg atc ctc gct tct ggt gcc tcc gat gct tct	144
His Pro Gly Ser Asp Leu Ile Leu Ala Ser Gly Ala Ser Asp Ala Ser	
35 40 45	
cct ctg ttc tac tcc atg cac cct tac gtc aag ccc gag aac tcg aag	192
Pro Leu Phe Tyr Ser Met His Pro Tyr Val Lys Pro Glu Asn Ser Lys	

50	55	60	
ctg ctt caa cag ttc gtg cga ggc aag cac gac cga acc tcc aag gac			240
Leu Leu Gln Gln Phe Val Arg Gly Lys His Asp Arg Thr Ser Lys Asp			
65	70	75	80
att gtc tac acc tac gac tct ccc ttt gca cag gac gtc aag cga act			288
Ile Val Tyr Thr Tyr Asp Ser Pro Phe Ala Gln Asp Val Lys Arg Thr			
85	90	95	
atg cga gag gtc atg aaa ggt cgg aac tgg tat gcc aca cct gga ttc			336
Met Arg Glu Val Met Lys Gly Arg Asn Trp Tyr Ala Thr Pro Gly Phe			
100	105	110	
tgg ctg cga acc gtt ggc atc att gct gtc acc gcc ttt tgc gag tgg			384
Trp Leu Arg Thr Val Gly Ile Ile Ala Val Thr Ala Phe Cys Glu Trp			
115	120	125	
cac tgg gct act acc gga atg gtg ctg tgg ggt ctc ttg act gga ttc			432
His Trp Ala Thr Thr Gly Met Val Leu Trp Gly Leu Leu Thr Gly Phe			
130	135	140	
atg cac atg cag atc ggc ctg tcc att cag cac gat ggt tct cat ggt			480
Met His Met Gln Ile Gly Leu Ser Ile Gln His Asp Gly Ser His Gly			
145	150	155	160
gcc atc agc aaa aag ccc tgg gtc aac gct ctc ttt gcc tac ggc atc			528
Ala Ile Ser Lys Lys Pro Trp Val Asn Ala Leu Phe Ala Tyr Gly Ile			
165	170	175	
gac gtc att gga tcg tcc aga tgg atc tgg ctg cag tct cac atc atg			576
Asp Val Ile Gly Ser Ser Arg Trp Ile Trp Leu Gln Ser His Ile Met			
180	185	190	
cga cat cac acc tac acc aat cag cat ggt ctc gac ctg gat gcc gag			624
Arg His His Thr Tyr Thr Asn Gln His Gly Leu Asp Leu Asp Ala Glu			
195	200	205	
tcc gca gaa cca ttc ctt gtg ttc cac aac tac cct gct gcc aac act			672

Ser Ala Glu Pro Phe Leu Val Phe His Asn Tyr Pro Ala Ala Asn Thr

210	215	220	
gct cga aag tgg ttt cac cga ttc cag gcc tgg tac atg tac ctc gtg			720
Ala Arg Lys Trp Phe His Arg Phe Gln Ala Trp Tyr Met Tyr Leu Val			
225	230	235	240
ctt gga gcc tac ggc gtt tcg ctg gtg tac aac cct ctc tac atc ttc			768
Leu Gly Ala Tyr Gly Val Ser Leu Val Tyr Asn Pro Leu Tyr Ile Phe			
245	250	255	
cga atg cag cac aac gac acc att ccc gag tct gtc aca gcc atg cga			816
Arg Met Gln His Asn Asp Thr Ile Pro Glu Ser Val Thr Ala Met Arg			
260	265	270	
gag aac ggc ttt ctg cga cgg tac cga acc ctt gca ttc gtt atg cga			864
Glu Asn Gly Phe Leu Arg Arg Tyr Arg Thr Leu Ala Phe Val Met Arg			
275	280	285	
gct ttc ttc atc ttt cga acc gcc ttc ttg ccc tgg tat ctc act gga			912
Ala Phe Phe Ile Phe Arg Thr Ala Phe Leu Pro Trp Tyr Leu Thr Gly			
290	295	300	
acc tcc ctg ctc atc acc att cct ctg gtg ccc act gct acc ggt gcc			960
Thr Ser Leu Leu Ile Thr Ile Pro Leu Val Pro Thr Ala Thr Gly Ala			
305	310	315	320
ttc ctc acc ttc ttt ttc atc ttg tct cac aac ttc gat ggc tcg gag			1008
Phe Leu Thr Phe Phe Phe Ile Leu Ser His Asn Phe Asp Gly Ser Glu			
325	330	335	
cga atc ccc gac aag aac tgc aag gtc aag agc tcc gag aag gac gtt			1056
Arg Ile Pro Asp Lys Asn Cys Lys Val Lys Ser Ser Glu Lys Asp Val			
340	345	350	
gaa gcc gat cag atc gac tgg tac aga gct cag gtg gag acc tct tcc			1104
Glu Ala Asp Gln Ile Asp Trp Tyr Arg Ala Gln Val Glu Thr Ser Ser			
355	360	365	

acc tac ggt gga ccc att gcc atg ttc ttt act ggc ggt ctc aac ttc 1152
 Thr Tyr Gly Gly Pro Ile Ala Met Phe Phe Thr Gly Gly Leu Asn Phe
 370 375 380

cag atc gag cat cac ctc ttt cct cga atg tcg tct tgg cac tat ccc 1200
 Gln Ile Glu His His Leu Phe Pro Arg Met Ser Ser Trp His Tyr Pro
 385 390 395 400

ttc gtg cag caa gct gtc cga gag tgt tgc gaa cga cac gga gtt cgg 1248

Phe Val Gln Gln Ala Val Arg Glu Cys Cys Glu Arg His Gly Val Arg
 405 410 415

tac gtc ttc tac cct acc att gtg ggc aac atc att tcc acc ctc aag 1296
 Tyr Val Phe Tyr Pro Thr Ile Val Gly Asn Ile Ile Ser Thr Leu Lys
 420 425 430

tac atg cac aaa gtc ggt gtg gtt cac tgt gtc aag gac gct cag gat 1344
 Tyr Met His Lys Val Gly Val Val His Cys Val Lys Asp Ala Gln Asp

435 440 445

tcc taa 1350
 Ser

<210> 110
 <211> 449
 <212> PRT
 <213> Euglena gracilis
 <400> 110

Met Ala Leu Ser Leu Thr Thr Glu Gln Leu Leu Glu Arg Pro Asp Leu
 1 5 10 15

Val Ala Ile Asp Gly Ile Leu Tyr Asp Leu Glu Gly Leu Ala Lys Val
 20 25 30

His Pro Gly Ser Asp Leu Ile Leu Ala Ser Gly Ala Ser Asp Ala Ser
 35 40 45

Pro Leu Phe Tyr Ser Met His Pro Tyr Val Lys Pro Glu Asn Ser Lys
 50 55 60

Leu Leu Gln Gln Phe Val Arg Gly Lys His Asp Arg Thr Ser Lys Asp

65				70				75				80			
Ile	Val	Tyr	Thr	Tyr	Asp	Ser	Pro	Phe	Ala	Gln	Asp	Val	Lys	Arg	Thr
				85				90				95			
Met	Arg	Glu	Val	Met	Lys	Gly	Arg	Asn	Trp	Tyr	Ala	Thr	Pro	Gly	Phe
100								105				110			
Trp	Leu	Arg	Thr	Val	Gly	Ile	Ile	Ala	Val	Thr	Ala	Phe	Cys	Glu	Trp
115								120				125			
His	Trp	Ala	Thr	Thr	Gly	Met	Val	Leu	Trp	Gly	Leu	Leu	Thr	Gly	Phe
130								135				140			
Met	His	Met	Gln	Ile	Gly	Leu	Ser	Ile	Gln	His	Asp	Gly	Ser	His	Gly
145				150				155				160			
Ala	Ile	Ser	Lys	Lys	Pro	Trp	Val	Asn	Ala	Leu	Phe	Ala	Tyr	Gly	Ile
				165				170				175			
Asp	Val	Ile	Gly	Ser	Ser	Arg	Trp	Ile	Trp	Leu	Gln	Ser	His	Ile	Met
180								185				190			
Arg	His	His	Thr	Tyr	Thr	Asn	Gln	His	Gly	Leu	Asp	Leu	Asp	Ala	Glu
195								200				205			
Ser	Ala	Glu	Pro	Phe	Leu	Val	Phe	His	Asn	Tyr	Pro	Ala	Ala	Asn	Thr
210								215				220			
Ala	Arg	Lys	Trp	Phe	His	Arg	Phe	Gln	Ala	Trp	Tyr	Met	Tyr	Leu	Val
225				230				235				240			
Leu	Gly	Ala	Tyr	Gly	Val	Ser	Leu	Val	Tyr	Asn	Pro	Leu	Tyr	Ile	Phe
				245				250				255			
Arg	Met	Gln	His	Asn	Asp	Thr	Ile	Pro	Glu	Ser	Val	Thr	Ala	Met	Arg
260								265				270			
Glu	Asn	Gly	Phe	Leu	Arg	Arg	Tyr	Arg	Thr	Leu	Ala	Phe	Val	Met	Arg
275								280				285			
Ala	Phe	Phe	Ile	Phe	Arg	Thr	Ala	Phe	Leu	Pro	Trp	Tyr	Leu	Thr	Gly
290								295				300			
Thr	Ser	Leu	Leu	Ile	Thr	Ile	Pro	Leu	Val	Pro	Thr	Ala	Thr	Gly	Ala
305				310				315				320			

Phe Leu Thr Phe Phe Phe Ile Leu Ser His Asn Phe Asp Gly Ser Glu
325 330 335

Arg Ile Pro Asp Lys Asn Cys Lys Val Lys Ser Ser Glu Lys Asp Val
340 345 350

Glu Ala Asp Gln Ile Asp Trp Tyr Arg Ala Gln Val Glu Thr Ser Ser
355 360 365

Thr Tyr Gly Gly Pro Ile Ala Met Phe Phe Thr Gly Gly Leu Asn Phe
370 375 380

Gln Ile Glu His His Leu Phe Pro Arg Met Ser Ser Trp His Tyr Pro
385 390 395 400

Phe Val Gln Gln Ala Val Arg Glu Cys Cys Glu Arg His Gly Val Arg
405 410 415

Tyr Val Phe Tyr Pro Thr Ile Val Gly Asn Ile Ile Ser Thr Leu Lys
420 425 430

Tyr Met His Lys Val Gly Val Val His Cys Val Lys Asp Ala Gln Asp
435 440 445

Ser

<210> 111

<211> 1365

<212> DNA

<213> Euglena anabaena

<220><221> CDS

<222> (1)..(1365)

<223> EaD5S-35a158g mutant delta-5 desaturase

<400> 111

atg gcc acc atc tcc ctg act acc gag cag ctc ctg gaa cac ccc gag	48
Met Ala Thr Ile Ser Leu Thr Thr Glu Gln Leu Leu Glu His Pro Glu	
1 5 10 15	
ctc gtt gcc atc gac gga gtc ctg tac gat ctc ttc ggt ctg gcc aag	96
Leu Val Ala Ile Asp Gly Val Leu Tyr Asp Leu Phe Gly Leu Ala Lys	

20	25	30	
gtc cat gcc gga ggc aac ctc atc gaa gct gcc ggt gca tcc gac gga			144
Val His Ala Gly Gly Asn Leu Ile Glu Ala Ala Gly Ala Ser Asp Gly			
35	40	45	
acc gct ctg ttc tac tcc atg cat cct gga gtc aag cca gag aac tcg			192
Thr Ala Leu Phe Tyr Ser Met His Pro Gly Val Lys Pro Glu Asn Ser			
50	55	60	
aag ctt ctg cag caa ttt gcc cga ggc aag cac gaa cga agc tcc aag			240
Lys Leu Leu Gln Gln Phe Ala Arg Gly Lys His Glu Arg Ser Ser Lys			
65	70	75	80
gat ccc gtg tac acc ttc gac tct ccc ttt gct cag gac gtc aag cag			288
Asp Pro Val Tyr Thr Phe Asp Ser Pro Phe Ala Gln Asp Val Lys Gln			
85	90	95	
tcc gtt cga gag gtc atg aag ggt cga aac tgg tac gcc act cct ggc			336
Ser Val Arg Glu Val Met Lys Gly Arg Asn Trp Tyr Ala Thr Pro Gly			
100	105	110	
ttc tgg ctg aga acc gca ctc atc atc gct tgt act gcc att ggc gag			384
Phe Trp Leu Arg Thr Ala Leu Ile Ile Ala Cys Thr Ala Ile Gly Glu			
115	120	125	
tgg tac tgg atc aca acc gga gca gtg atg tgg ggt atc ttt act gga			432
Trp Tyr Trp Ile Thr Thr Gly Ala Val Met Trp Gly Ile Phe Thr Gly			
130	135	140	
tac ttc cac tcg cag att ggc ttg gcc att caa cac gat ggt tct cac			480
Tyr Phe His Ser Gln Ile Gly Leu Ala Ile Gln His Asp Gly Ser His			
145	150	155	160
gga gcc atc agc aaa aag ccc tgg gtc aac gcc ttt ttc gct tat ggc			528
Gly Ala Ile Ser Lys Lys Pro Trp Val Asn Ala Phe Phe Ala Tyr Gly			
165	170	175	
atc gac gcc att ggt tcc tct cgt tgg atc tgg ctg cag tcc cac att			576

Ile Asp Ala Ile Gly Ser Ser Arg Trp Ile Trp Leu Gln Ser His Ile	
180 185 190	
atg cga cat cac act tac acc aac cag cat ggc ctc gac ctg gat gct	624
Met Arg His His Thr Tyr Thr Asn Gln His Gly Leu Asp Leu Asp Ala	
195 200 205	
gcc tcg gca gag ccg ttc atc ttg ttc cac tcc tat cct gct acc aac	672
Ala Ser Ala Glu Pro Phe Ile Leu Phe His Ser Tyr Pro Ala Thr Asn	
210 215 220	
gcc tct cga aag tgg tac cac cga ttt cag gcg tgg tac atg tac atc	720
Ala Ser Arg Lys Trp Tyr His Arg Phe Gln Ala Trp Tyr Met Tyr Ile	
225 230 235 240	
gtt ctg gga atg tat ggt gtc tcg atg gtg tac aat ccc atg tac ctc	768
Val Leu Gly Met Tyr Gly Val Ser Met Val Tyr Asn Pro Met Tyr Leu	
245 250 255	
ttc aca atg cag cac aac gac acc att ccc gag gcc act tct ctc aga	816
Phe Thr Met Gln His Asn Asp Thr Ile Pro Glu Ala Thr Ser Leu Arg	
260 265 270	
cca ggc agc ttt ttc aat cgg cag cga gct ttc gcc gtt tcc ctt cga	864
Pro Gly Ser Phe Phe Asn Arg Gln Arg Ala Phe Ala Val Ser Leu Arg	
275 280 285	
ctg ctc ttc atc ttc cga aac gcc ttt ctt ccc tgg tac att gct ggt	912
Leu Leu Phe Ile Phe Arg Asn Ala Phe Leu Pro Trp Tyr Ile Ala Gly	
290 295 300	
gcc tct cct ctg ctc acc att ctt ctg gtg ccc acg gtc aca ggc atc	960
Ala Ser Pro Leu Leu Thr Ile Leu Leu Val Pro Thr Val Thr Gly Ile	
305 310 315 320	
ttc ctc acc ttt gtg ttc gtt ctg tcc cat aac ttc gag gga gcc gaa	1008
Phe Leu Thr Phe Val Phe Val Leu Ser His Asn Phe Glu Gly Ala Glu	
325 330 335	
cgg acc cca gag aag aac tgc aag gcc aaa cga gct aag gaa ggc aag	1056
Arg Thr Pro Glu Lys Asn Cys Lys Ala Lys Arg Ala Lys Glu Gly Lys	

340	345	350	
gag gtc aga gac gtg gaa gag gat cga gtc gac tgg tac cga gca cag			1104
Glu Val Arg Asp Val Glu Glu Asp Arg Val Asp Trp Tyr Arg Ala Gln			
355	360	365	
gcc gag act gct gcc acc tac ggt ggc agc gtg gga atg atg ctt aca			1152
Ala Glu Thr Ala Ala Thr Tyr Gly Gly Ser Val Gly Met Met Leu Thr			
370	375	380	
ggc ggt ctc aac ctg cag atc gag cat cac ttg ttt ccc cga atg tcc			1200
Gly Gly Leu Asn Leu Gln Ile Glu His His Leu Phe Pro Arg Met Ser			
385	390	395	400
tct tgg cac tat ccc ttc att caa gac acc gtt cgg gag tgt tgc aag			1248
Ser Trp His Tyr Pro Phe Ile Gln Asp Thr Val Arg Glu Cys Cys Lys			
405	410	415	
cga cat ggc gtc cgt tac aca tac tat cct acc att ctc gag aac atc			1296
Arg His Gly Val Arg Tyr Thr Tyr Tyr Pro Thr Ile Leu Glu Asn Ile			
420	425	430	
atg tcc act ctt cga tac atg cag aag gtg ggt gtt gct cac acc att			1344
Met Ser Thr Leu Arg Tyr Met Gln Lys Val Gly Val Ala His Thr Ile			
435	440	445	
cag gat gcc cag gag ttc taa			1365
Gln Asp Ala Gln Glu Phe			
450			
<210> 112			
<211> 454			
<212> PRT			
<213> Euglena anabaena			
<400> 112			
Met Ala Thr Ile Ser Leu Thr Thr Glu Gln Leu Leu Glu His Pro Glu			
1	5	10	15
Leu Val Ala Ile Asp Gly Val Leu Tyr Asp Leu Phe Gly Leu Ala Lys			
20	25	30	

Val His Ala Gly Gly Asn Leu Ile Glu Ala Ala Gly Ala Ser Asp Gly
35 40 45

Thr Ala Leu Phe Tyr Ser Met His Pro Gly Val Lys Pro Glu Asn Ser
50 55 60

Lys Leu Leu Gln Gln Phe Ala Arg Gly Lys His Glu Arg Ser Ser Lys
65 70 75 80

Asp Pro Val Tyr Thr Phe Asp Ser Pro Phe Ala Gln Asp Val Lys Gln
85 90 95

Ser Val Arg Glu Val Met Lys Gly Arg Asn Trp Tyr Ala Thr Pro Gly
100 105 110

Phe Trp Leu Arg Thr Ala Leu Ile Ile Ala Cys Thr Ala Ile Gly Glu
115 120 125

Trp Tyr Trp Ile Thr Thr Gly Ala Val Met Trp Gly Ile Phe Thr Gly
130 135 140

Tyr Phe His Ser Gln Ile Gly Leu Ala Ile Gln His Asp Gly Ser His
145 150 155 160

Gly Ala Ile Ser Lys Lys Pro Trp Val Asn Ala Phe Phe Ala Tyr Gly
165 170 175

Ile Asp Ala Ile Gly Ser Ser Arg Trp Ile Trp Leu Gln Ser His Ile
180 185 190

Met Arg His His Thr Tyr Thr Asn Gln His Gly Leu Asp Leu Asp Ala
195 200 205

Ala Ser Ala Glu Pro Phe Ile Leu Phe His Ser Tyr Pro Ala Thr Asn
210 215 220

Ala Ser Arg Lys Trp Tyr His Arg Phe Gln Ala Trp Tyr Met Tyr Ile
225 230 235 240

Val Leu Gly Met Tyr Gly Val Ser Met Val Tyr Asn Pro Met Tyr Leu
245 250 255

Phe Thr Met Gln His Asn Asp Thr Ile Pro Glu Ala Thr Ser Leu Arg
260 265 270

Pro Gly Ser Phe Phe Asn Arg Gln Arg Ala Phe Ala Val Ser Leu Arg

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

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

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

<210> 113

<211> 13926

<212> DNA

<213> Artificial Sequence

<220><223> Plasmid pZR5AU-555

<400> 113

cgatccctaa ctgtatacta gcattctgaa ttaatatctt cctctttgat tgttccttag	60
gcatatcagc ggggatcccg tcaacagttt tatatatcgt agttacaacc atcaacactt	120
tttgtaagt gtaccattct atactccaac tggctctgcaa ctgtacaagt agacatgtta	180

atggtagtta ataacatcta cagcagaacc tatggtaaag acattgcatt ttacaggaa	240
gtatcgtcct acacgttgat aaatccaaag atgcggaact tcttcactt ttatcatcat	300
cccctactcg tacactcgta ctctttgttc gatcgcgatt ctttctata aataatcttg	360
tatgtacatg cggccgctta gaactcctgg gcatcctgaa tgggtgagc aacacccacc	420
ttctgcatgt atcgaagagt ggacatgatg ttctcgagaa tggtaggata gtatgtgtaa	480
cggacgccat gtcgcttgca acactcccga acggtgtctt gaatgaagg atagtccaa	540
gaggacattc ggggaaacaa gtgatgctcg atctgcaggt tgagaccgcc tgtaagcatc	600
attcccacgc tgccaccgta ggtggcagca gtctcgccct gtgctcgta ccagtgcact	660
cgatcctctt ccacgtctct gacctccttg ccttccttag ctgctttggc ctgacgttc	720
ttctctgggg tccgttcggc tccctcgaag ttatgggaca gaacgaacac aaaggtgagg	780
aagatgcctg tgaccgtggg caccagaaga atggtgagca gaggagaggc accagcaatg	840
taccaggga gaaaggcgtt tcggaagatg aagagcagtc gaagggaac ggcgaaagct	900
cgctgccgat tgaanaagct gcctggcttg agagaagtgg cctcggaat ggtgtcgttg	960
tgctgcattg tgaagaggta catgggattg tacaccatcg agacaccata cattccaga	1020
acgatgtaca tgiaccacgc ctgaaatcgg tggtaaccact ttcgagaggc gttggtagca	1080
ggataggagt ggaacaagat gaacggctct gccgaggcag catccaggtc gaggccatgc	1140
tggttgggtg aagtgtgatg tcgcataatg tgggactgca gccagatcca acgagaggaa	1200
ccaatggcgt cgatgccata agcgaaaaag gcgttgacct agggcttttt gctgatggct	1260
ccgtgagaac catcgtgttg aatggccaag ccaatctgcg agtggaagta tccagtaaag	1320
ataccacaca tactgtctcc ggtttgtgatc cagtaccact cgccaatggc agtacaagcg	1380
atgatgagt cggttctcag ccagaagcca ggagtggcgt accagtttcg acccttcagt	1440
acctctcgaa cggactgctt gacgtcctga gcaaaggag agtcgaaggt gtacacggga	1500
tccttggagc ttcgttcgtg cttgcctcgg gcaaattgct gcagaagctt cgagttctct	1560
ggcttgactc caggatgcat ggagtagaac agagcggctt cgctcgatgc accggcagct	1620
tcgatgaggt tgcctccggc atggaccttg gccagaccga agagatcgt caggactccg	1680
tcgatggcaa cgagctcggg gtgttccagg agctgctcgg tagtcaggga gatggtggcc	1740
atggttgtga attagggtgg tgagaatggt tggttgtagg gaagaatcaa aggccggtct	1800
cgggatccgt gggatataat atatatatat atatataga tccttcgtta cctccctgtt	1860
ctcaaaactg tggtttttcg ttttctgtt tttgttttt ttgatttttt tagggccaac	1920
taagcttcca gatttcgcta atcacctttg tactaattac aagaaaggaa gaagctgatt	1980
agagtgggc tttttatgca actgtgctac tccttatctc tgatatgaaa gtgtagacct	2040

aatcacatca tgcatttag agttggtaat actgggagga tagataaggc acgaaaacga	2100
gccatagcag acatgctggg tgtagccaag cagaagaaag tagatgggag ccaattgacg	2160
agcgaggag ctacgccaat ccgacatacg acacgctgag atcgtcttgg ccggggggta	2220
cctacagatg tccaagggtg agtgcttgac tgtaattgta tgtctgagga caaatatgta	2280
gtcagccgta taaagtcata ccaggcacca gtgccatcat cgaaccacta actctctatg	2340
atacatgcct ccggtattat tgtaccatgc gtcgctttgt tacatacgta tcttgccctt	2400
ttctctcaga aactccagac ttggctatt ggtcgagata agcccgacc atagtgagtc	2460
tttcacactc tgtttaaaca agctaccacc acactcgttg ggtgcagtcg acgagtatct	2520
gtctgactcg tcattgccgc ctttgagta cgactccaac tatgagtgtg ctggatcac	2580
tttgacgata cattcttctg tggaggctgt gggtctgaca gctgcgtttt cggcgcggtt	2640
ggcgcacaac aatatcagct gcaacgtcat tcttggtttt catcatgac acatttttgt	2700
cggcaaaggc gacgcccaga gagccattga cgttctttct aatttggacc gatagccgta	2760
tagtccagtc tatctataag ttcaactaac tcgtaactat taccataaca tatacttcac	2820
tgccccagat aaggttccga taaaaagttc tgcagactaa atttatttca gtctctctt	2880
caccacaaa atgccctctt acgaagctcg agctaacgtc cacaagtcg ctttgccgc	2940
tcgagtgtc aagctcgttg cagccaagaa aaccaacctg tgtgtttctc tggatgttac	3000
caccaccaag gagtcattg agcttgccga taaggtcgga ccttatgtgt gcatgatcaa	3060
aaccatata gacatattg acgacttcac ctacgccggc actgtgtcc ccctcaagga	3120
acttgtctt aagcacggtt tcttctgtt cgaggacaga aagttcgag atattggcaa	3180
caactgtcaag caccagtacc ggtgtcaccg aatcgccgag tggccgata tcaccaacgc	3240
ccacggtgta cccggaaccg gaatcattgc tggcctcgga gctggtgccg aggaaactgt	3300
ctctgaacag aagaaggagg acgtctctga ctacgagaac tcccagtaca aggagtctct	3360
agtccctct cccaacgaga agctggccag aggtctgtc atgctggccg agctgtcttg	3420
caagggtct ctggccactg gcgagtactc caagcagacc attgagcttg cccgatccga	3480
ccccgagttt gtggttggct tcatigccca gaaccgacct aaggcgact ctgaggactg	3540
gcttattctg acccccgagg tgggtcttga cgacaaggga gacgtctcg gacagcagta	3600
ccgaactgtt gaggatgtca tgtctaccgg aacgatata ataattgtcg gccgaggtct	3660
gtacggccag aaccgagatc ctattgagga ggccaagcga taccagaagg ctggctggga	3720
ggcttaccag aagattaact gttagaggtt agactatgga tatgtaat t aactgtgtat	3780

atagagagcg tgcaagtatg gagcgcttgt tcagcttgta tgatggtcag acgacctgtc	3840
tgatcgagta tgtatgatac tgcacaacct gtgtatccgc atgatctgtc caatggggca	3900
tgttgttgtg tttctcgata cggagatgct gggtacagtg ctaatacgtt gaactactta	3960
tacttatatg aggcctgaag aaagctgact tgtgtatgac ttattctcaa ctacatcccc	4020
agtcacaata ccaccactgc actaccacta caccaaaacc atgatcaaac caccgatgga	4080
cttctctggag gcagaagaac ttgttatgga aaagctcaag agagagaatt caggagagac	4140
cgggttggcg gcgtatttgt gtcccaaaaa acagccccaa ttgccccgga gaagacggcc	4200
aggccgccta gatgacaaat tcaacaactc acagctgact ttctgccatt gccactaggg	4260
gggggccttt ttatatggcc aagccaagct ctccacgtcg gttgggctgc acccaacaat	4320
aatgggtag gggtgcacca acaaagggat gggtatggggg gtagaagata cgaggataac	4380
ggggctcaat ggcacaaata agaacgaata ctgccattaa gactcgtgat ccagcgactg	4440
acaccattgc atcatctaag ggcctcaaaa ctacctcgga actgctgcgc tgatctggac	4500
accacagagg ttccgagcac tttaggttgc accaaatgtc ccaccaggtg caggcagaaa	4560
acgtctggaac agcgtgtaca gtttgtctta acaaaaagtg agggcgctga ggtcgagcag	4620
gggtgtgtga ctgtttatag cctttagagc tgcgaaagcg cgtatggatt tggtcatca	4680
ggccagattg agggctctgtg gacacatgtc atgttagtgt acttcaatcg cccctggat	4740
atagccccga caataggccg tggcctcatt tttttgcctt ccgcacattt ccattgctcg	4800
glaccacac ctgtcttctc ctgcacttgc caaccttaat actggtttac attgaccaac	4860
atcttacaag cggggggctt gtctagggtg tatataaaca gtggctctcc caatcggttg	4920
ccagtctctt ttttctttt tttcccccaca gattcgaaat ctaaactaca catcacaaa	4980
tgctgttac tgacgtcctt aagcgaagat ccggtgtcat cgtcggcgac gatgtccgag	5040
ccgtgagtat ccacgacaag atcagtgtcg agacgacgcg ttttgtgtaa tgacacaatc	5100
cgaagtctgc tagcaacaca cactctctac aaaaactaac ccagctctcc atggctctct	5160
cccttactac cgagcagctg ctcgagcgac ccgacctggt tgccatcgac ggcattctct	5220
acgatctgga aggtcttgcc aaggctcatc ccggatccga cttgatectc gcttctgggtg	5280
cctccgatgc ttctctctg ttctactcca tgcaccctta cgtcaagccc gagaactcga	5340
agctgttca acagtctgtg cgaggcaagc acgaccgaac ctccaaggac attgtctaca	5400
cctacgactc tccctttgca caggacgtca agcgaactat gcgagaggtc atgaaaggtc	5460
ggaactggta tgccacacct ggattctggc tgcgaaccgt tggcatcatt gctgtcaccg	5520
ccttttgcga gtggcactgg gctactaccg gaatgggtgct gtggggctctc ttgactggat	5580
tcatgcacat gcagatcggc ctgtccattc agcacgatgg ttctcatggt gccatcagca	5640

aaaagccctg ggtaacgct ctctttgcct acggcatcga cgtcattgga tcgtccagat	5700
ggatctggct gcagtctcac atcatcgac atcacaccta caccaatcag catggctcgc	5760
acctggatgc cgagtcgca gaaccattcc ttgtgttcca caactacct gctgccaaca	5820
ctgctcgaaa gtggtttcac cgattccagg cctggtacat gtacctctg ctgggagcct	5880
acggcgtttc gctgggttac aacctctct acatcttcg aatgcagcac aacgacacca	5940
ttcccagtc tgcacagcc atgcgagaga acggctttct gcgacgtac cgaaccttg	6000
cattcgttat gcgagcttc ttcatcttc gaaccgcctt cttgccctgg tatctcactg	6060
gaacctccct gctcatcacc attctctgg tgcctctgc taccggtgcc ttctcactt	6120
tctttttcat ctgtctcac aacttcgat gctcggagcg aatccccgac aagaactgca	6180
aggtcaagag ctccgagaag gacgttgaag ccgacagat cgactggtac agagctcagg	6240
tggagacctc ttccacctac ggtggacca ttgccatgt ctttactggc ggtctcaact	6300
tccagatcga gcatcacctc ttctctgaa tgcgtcttg gcactatccc ttcgtgcagc	6360
aagctgtccg agagtgtgc gaacgacacg gagttcggta cgtctttac cctaccattg	6420
tgggcaacat catttcacc ctcaagtaca tgcacaaagt cgggtgtgtt cactgtgtca	6480
aggacgtca ggattcctaa gcggccgcaa gtgtggatgg ggaagtgagt gcccgttct	6540
gtgtgcacaa ttggcaatcc aagatggatg gattcaacac agggatatag cgagctacgt	6600
ggtgggtcga ggatatagca acggatattt atgtttgaca cttgagaatg tacgatacaa	6660
gcactgtcca agtacaatac taaacatact gtacatactc atactctac ccgggcaacg	6720
gtttcacttg agtgcagtgg ctagtctct tactctaca gtgtgcaata ctgcgtatca	6780
tagtctttga tgtatatcgt attcattcat gttagtgcg tacgacagac ggacaatgac	6840
gagccaggtt ccccaactg acggaatct ctatagctt ctcatcaagc gcataaatat	6900
agtatataaa gaattacatg gtgggccaca gcttggcgat agtatgacgg tctgggctcc	6960
agatcacacc catgagggat gccgtgcgag tagatgtaga ggtgggtgctt ccagggtgtc	7020
ccaacaaca atcgctactg gaaaataccg tgaaagtacg tacaccaac tctgtgactc	7080
gcgttgactt tcactttca atttggtcgt gtcttacggt gcatgtctg tattggacat	7140
gaagaagggg ttggcaagct gaaccatgcc tgaggtgctg gagtctgata ctgtctagca	7200
gagggcgac aaaagatcac catecacag ttgtttcca tattcaatct ccaatttcca	7260
acatctctt ggctaatagc cactttgaga gatatttgg atcgcgtatt cacttgtctc	7320
atatcccggt tttctgttc accaaggctt ggacgatatt atcgcgagaa aaaccgtatt	7380
gaggtattta ttagagtca gagaggagga acaggtattc ggaagatgga aaagctgtac	7440
agagcgagaa acgatcccat gtcgacgtc ctgatggctt atactgagac atggaatcag	7500

gatcttgccc ttcttcacag ctctacttgc agctacagaa ttgctttttc tcgtatccga	7560
tcataattcc gctgctatga taggcagcca ccgatacaat ataggagaga gtataggtat	7620
tgactctata gacactgtat gatgctgaca aagctgactg agggattcga ctcagcttcg	7680
ccgttgcttg atctccaatt aattatcgta ggcgcgccag ctgcattaat gaatcggcca	7740
acgcgcgggg agaggcggtt tgcgtattgg gcgctcttcc gcttcctcgc tcaactgactc	7800
gctgcgctcg gtcgttcggc tgcggcgagc ggtatcagct cactcaaagg cggaataacg	7860
gttatccaca gaatcagggg ataacgcagg aaagaacatg tgagcaaaag gccagcaaaa	7920
ggccaggaac cgtaaaaagg ccgcgttgct ggcggttttc cataggtcc gcccccctga	7980
cgagcatcac aaaaatcgac gctcaagtca gaggtggcga aaccgcagag gactataaag	8040
ataccaggcg tttcccctg gaagctccct cgtgcgtctt cctgttccga ccctgccgt	8100
taccggatac ctgtccgctt ttctcccttc gggaagcgtg gcgctttctc atagctcacg	8160
ctgtaggtat ctcagttcgg tgtaggtcgt tcgctccaag ctgggctgtg tgcacgaacc	8220
ccccgttcag ccgaccgtt gcgccttatt cggttaactat cgtcttgagt ccaaccgggt	8280
aagacacgac ttatcgccac tggcagcagc cactggtaac aggattagca gagcgaggta	8340
tgtaggcgggt gctacagagt tcttgaagtg gtggcctaac tacggctaca ctagaagaac	8400
agtatttgggt atctgcgctc tgctgaagcc agttaccttc ggaaaaagag ttggtagctc	8460
ttgatccggc aaacaaacca ccgctggtag cggtaggtttt ttgttttgca agcagcagat	8520
tacgcgcaga aaaaaaggat ctcaagaaga tcctttgatc ttttctacgg ggtctgacgc	8580
tcagtggaaac gaaaactcac gtttaagggtt tttggtcatt agattatcaa aaaggatctt	8640
cacctagatc cttttaaat aaaaatgaag ttttaaatca atctaaagta tatatgagta	8700
aacttggctt gacagttacc aatgcttaat cagtgaggca cctatctcag cgatctgtct	8760
atttcgttca tccatagtig cctgactccc cgtcgtgtag ataactacga tacgggaggg	8820
cttaccatct ggccccagtg ctgcaatgat accgcgagac ccacgctcac cggtccaga	8880
tttatcagca ataaaccagc cagccggaag ggccgagcgc agaagtggtc ctgcaacttt	8940
atccgcctcc atccagtcta ttaattgttg ccgggaagct agagtaagta gttcgccagt	9000
taatagtttg cgcaacgttg ttgccattgc tacaggcatc gtgggtgtcac gctcgtcgtt	9060
tggtatggct tcaattcagct ccggttccca acgatcaagg cgagttacat gatccccat	9120
gttgtgcaaa aaagcggtta gctccttcgg tcctccgacg gttgtcagaa gtaagttggc	9180
cgcagtgtaa tcaactatgg ttatggcagc actgcataat tctcttactg tcatgccatc	9240

cgtaagatgc ttttctgtga ctgggtgagta ctcaaccaag tcattctgag aatagtgtat	9300
gcggcgaccg agttgctctt gcccggcgtc aatacgggat aataccgcgc cacatagcag	9360
aactttaaaa gtgctcatca ttggaaaacg ttcttcgggg cgaaaactct caaggatctt	9420
accgctgttg agatccagtt cgatgtaacc cactcgtgca cccaactgat ctacagcatc	9480
ttttactttt accagcgttt ctgggtgagc aaaaacagga aggcaaatg ccgcaaaaaa	9540
gggaataagg gcgacacgga aatgttgaat actcatactc ttctttttt aatattattg	9600
aagcatttat cagggttatt gtctcatgag cggatacata ttgaaatgta tttagaaaaa	9660
taaacaaata ggggttccgc gcacatttcc ccgaaaagt ccacctgat cggtgtgaaa	9720
taccgcacag atgcgtaagg agaaaatacc gcatcaggaa attgtaagcg ttaatatatt	9780
gttaaaattc gcgttaaat tttgttaaat cagctcattt tttaaccaat aggccgaaat	9840
cggcaaaatc cttataaat caaaagaata gaccgagata ggggtgagt ttgttccagt	9900
ttggaacaag agtccactat taaagaacgt ggactccaac gtcaaaggcg gaaaaaccgt	9960
ctatcaggcg gatggccac tacgtgaacc atcacctaa tcaagttttt tggggtcgag	10020
gtgccgtaaa gcactaaatc ggaaccctaa agggagcccc cgatttagag cttagcgggg	10080
aaagccggcg aacgtggcga gaaaggaagg gaagaaagcg aaaggagcgg gcgctaggcg	10140
gctggcaagt gtagcggta cgctgcgctg aaccaccaca cccgccgcgc ttaatgcgcc	10200
gtacacggcg gcgtccatt gccattcagg ctgcgcaact gttgggaagg gcgatcggtg	10260
cgggcctctt cgctattacg ccagctggcg aaagggggat gtgctgcaag gcgattaagt	10320
tgggtaacgc cagggttttc ccagtcacga cgttgtaaaa cgacggccag tgaattgtaa	10380
tacgactcac tataggcgga attgggcccg acgtcgcatg cattaccaac gttggcgcg	10440
cacaacgtca gaaactctt tccatccgat aagccgtgtc taactatacg accctcaaac	10500
aaggtcagcc agaccaact ttagcactcg caaaacacac acccaaacc aggccgcgac	10560
gcagccggag cccgtctact tatactgttt cgccgcagac gccaagtgc aaaaaagcg	10620
cacaaacct cccccctg ggcggcgac atccacaaca aacgacccat aaacggcact	10680
cccacacgga aacaaaacgt ttgtcgggtt actcctcacg ccaagacaca ctacacaca	10740
caacctccca ctttcatca ccactcaact acgggtacag tagtggtaga aaacacagag	10800
tccagagtc acacagacac gcacagagt ctgtcacaca agcaccatgc caacaagcag	10860
accgttcatt tcgtctgtgt ttggaagctt ctccacgtcg tatcaacgca acaccacc	10920
acagacctcc gcatcgacag ccaccgtcac caccggcccc acgtcgaga tgtacaagat	10980
gtcgcacaac cactcgata ctgagcagcc ccagcagccg acctgactc agacacagcc	11040
agcgggtgcag acggcgcaga cggcgagac ggcgagcag gcaagtggga ctgcggacct	11100

cagcgaatcg ccacacgaca agctgtggat tggaggccga tcaaacgacg gacgggaaaa	11160
gttctaccgg ctggaaccgg ccaagcgaag ggcatccttc gatcgattt ctctggatca	11220
ggtgtctatt tgatgaggcg ccaaacagtg gccagaccaa gtctggtatc tcaaccctc	11280
aaaattgata accaagtttc tgggtggagt caaaccaagg ggagctacgg agccgttcac	11340
aaggtgtcca cagaggttca aagcacattg caccaagacc gccatcatcc ttaattaaaa	11400
ggcgttgaaa cagaatgagc cagacagcaa ggacaagggt gccaacagca aggagtccaa	11460
aaagccctct attgacgaga tccacgatgt tattgtctcat gaggtttccg agctcgatgc	11520
tgggaagaag aagtgatttg tatataagaa ataaatgaga tatagtaaag gagtgcgaaga	11580
gaatggcaag gtggtcaaat tctatattac ttgcagtcac tggttcctcg ttgacatgaa	11640
tgaagtacc gtggcatag ctgatttaat atataactgt ccaactaact ctcacctaga	11700
tataacccat gtgtgtgttt ccaatcatca atcgggccgc ttaggaatcc tgtgcgtcct	11760
tcacgcagtg gacgacacc acccttatgca tgtacttcag ggtggagatg atgttgccga	11820
cgatggtagg gtagaaaaca tatcgactc catgtcgttc gcaacactcc cggaccgcct	11880
gctggacgaa ggggtagtgc caagacgaca tccggggaaa gagtgggtgc tcgatctgga	11940
aattgagacc gccagtgaag aacatggcga tggggccacc gtatgtggag gacgtctcca	12000
cctgcgccg ataccagtca atttggtcag cctcaacgtc cttctcagat cgcttaacct	12060
tgcagttctt gtcggggatc cgttcggagc catcaaaatt gtgggacaaa atgaagaaga	12120
acgtcaagaa ggcaccagtt cgggtgggca ccagaggaat ggtgatcagc aatgaggtcc	12180
cagtgaggta ccagggcaag aatgcggtcc ggaagatgaa gaaagctcg atcacgaatg	12240
caagtgtgcg gtagcgccgc agaaagccat ttccccgat ggccgtgaca gactctggga	12300
tgggtgtcatt gtctgcacg cggaatatgt agagcgggtt gtacaccagc gataccccgt	12360
atgccccag cacaaggtac atgtaccaag cctggaagcg gtggaaccac tttcgggcgg	12420
tgtttgcgcg ggggtagttg tggaacacca ggaacggctc tgccgactcc gcatccaggt	12480
cgaggccgtg ctggttggtg taggtgtggt gccgcatgat gtgcgactgc agccaaatcc	12540
accgggacga tccgatgacg tcaatgccgt aggcgaagag ggcttgacc caagcttct	12600
tgctgatggc cccgtgacct gcatcatgct ggatggataa gccgatctgc atgtgatga	12660
atccagtcaa caggccccac agcaccatcc ccgtggtagc ccagtgccac tcgcaaaagg	12720
ccgtcacggc gatgatcca acggtgcgca gccagaagcc aggggttgcg taccagttcc	12780
tccctttcat cacctcgcgc attgtccgt taactcttg tgcgaaggga gaatcatacg	12840
tgtagacaat gtcttcgag gtgcggtcac gcttcctcg gacgaattgc tgcaggagct	12900
tggagttctc gggcttgacg tagggatgca tagagtagaa gaggggagag gcgtcggagg	12960

ctccagaggc cagaatgagg tctccaccac catgcacctt ggccaggccc tcgaggtcgt 13020

acaggatacc atcgatagcg acgagatcag gtcgctccag gagctgttcg gtagtgagag 13080

acagggccat ggccattgct gtagatatgt cttgtgtgta agggggttgg ggtggttgtt 13140

tgtgttcttg acttttgtgt tagcaaggga agacgggcaa aaaagtgagt gtggttggga 13200

gggagagacg agccttatat ataatgcttg tttgtgtttg tgcaagtgga cgccgaaacg 13260

ggcaggagcc aaactaaaca aggcagacaa tgcgagctta attggattgc ctgatgggca 13320

ggggttaggg ctcgatcaat gggggtgcga agtgacaaaa ttgggaatta ggttcgcaag 13380

caaggctgac aagactttgg cccaaacatt tgtacgcggt ggacaacagg agccacccat 13440

cgctctgcac gggctagccg gtcgtgcgtc ctgtcaggct ccacctaggc tccatgccac 13500

tccatacaat ccactagtgt taccgctagg ccgtttttag ctcccatcta agaccccccc 13560

aaaacctcca ctgtacagtg cactgtactg tgtggcgatc aagggaagg gaaaaaaggc 13620

gcaaacatgc acgcatggaa tgacgtaggt aaggcggtac tagactgaaa agtggcacat 13680

ttcggcgtgc caaagggtcc taggtgcgtt tcgagagctg ggccgagc caagccgctc 13740

caaaacgctt ctccgactcc ctccagcggc ctccatatcc ccatccctct ccacagcaat 13800

gttggttaagc ctgcaaacg aaaaaataga aaggctaata agcttccaat attgtggtgt 13860

acgtgcata acgcaacaat gagcgccaaa caacacacac acacagcaca cagcagcatt 13920

aacctat 13926

<210> 114

<211> 13926

<212> DNA

<213> Artificial Sequence

<220><223> Plasmid pZR5AU-555m

<400> 114

cgatccctaa ctgtatacta gcattctgaa ttaatatctt cctctttgat tgttccttag 60

gcatalcagc ggggatcccg tcaacagttt tatatatcgt agttacaacc atcaacactt 120

tttgtaagt gtaccattct atactccaac tggctctgaa ctgtacaagt agacatgtta 180

atggtagtta ataacatcta cagcagaacc tatggtaaag acattgcatt ttacaggaa 240

glatcgtcct acacgttgat aaatccaaag atgcggaact tcttccactt ttatcatcat 300

cccctactcg tacactcgta ctctttgttc gatcgcgatt catttctata aataatcttg 360

tatgtacatg cggccgctta gaactcctgg gcatcctgaa tgggttgagc aacacccacc 420

ttctgcatgt atcgaagagt ggacatgatg ttctcgagaa tggtaggata gtatgtgtaa 480

cggacgccat gtcgcttgca acatccccga acggtgtctt gaatgaaggg atagtgccaa 540
 gaggacattc ggggaaacaa gtgatgctcg atctgcaggt tgagaccgcc tgtaagcatc 600
 attcccacgc tgccaccgta ggtggcagca gtctcggcct gtgctcggtt ccagtcgact 660

 cgatcctctt ccacgtctct gacctccttg ccttccttag ctcgtttggc cttgcagttc 720
 ttctctgggg tccgttcggc tccctcgaag ttatgggaca gaacgaacac aaaggtgagg 780
 aagatgcctg tgaccgtggg caccagaaga atggtgagca gaggagaggc accagcaatg 840
 taccaggga gaaaggcgtt tcggaagatg aagagcagtc gaagggaac ggcgaaagct 900
 cgctgccgat tgaaaaagct gcctgggtctg agagaagtgg cctcggaat ggtgtcgttg 960
 tgctgcattg tgaagaggtt catgggattg tacaccatcg agacaccata cattcccaga 1020
 acgatgtaca tgiaccacgc ctgaaatcgg tggtagcact ttcgagaggc gttggtagca 1080

 ggataggagt ggaacaagat gaacggctct gccgaggcag catccaggtc gaggccatgc 1140
 tggttggtgt aagtgtgatg tcgcataatg tgggactgca gccagatcca acgagaggaa 1200
 ccaatggcgt cgatgccata agcgaaaaag gcgttgaccc agggcttttt gctgatggct 1260
 ccgtgagaac catcgtgttg aatggccaag ccaatctgcg agtggaaagta tccagtaaag 1320
 ataccccaca tcaactgtcc ggttgtgatc cagtaccact cgccaatggc agtacaagcg 1380
 atgatgagtg cggttctcag ccagaagcca ggagtggcgt accagtctcg acccttcattg 1440
 acctctcgaa cggactgctt gacgtcctga gcaaaggag agtcgaaggt gtacacggga 1500

 tccttggagc ttcgttcgtg cttgcctcgg gcaaatgtct gcagaagctt cgagttctct 1560
 ggcttgactc caggatgcat ggagtagaac agagcggttc cgtcggatgc accggcagct 1620
 tcgatgaggt tgctccggc atggaccttg gccagaccga agagatcgta caggactccg 1680
 tcgatggcaa cgagctcggg gtgttccagg agctgctcgg tagtcaggga gatggtggcc 1740
 atggttgtga attagggtgg tgagaatggt tggttgtagg gaagaatcaa aggccggtct 1800
 cgggatccgt gggatataat atatatatat atatatatcg tccttcgtta cctccctgtt 1860
 ctcaaaactg tggtttttcg ttttctgttt ttgtcttttt ttgatTTTT tagggccaac 1920

 taagcttcca gatttcgcta atcacctttg tactaattac aagaaaggaa gaagctgatt 1980
 agagttgggc tttttatgca actgtgctac tccttatctc tgatatgaaa gtgtagacct 2040
 aatcacatca tgtcatttag agttggtaat actgggagga tagataaggc acgaaaacga 2100
 gccatagcag acatgctggg tgtagccaag cagaagaaag tagatgggag ccaattgacg 2160
 agcgaggagg ctacgccaat ccgacatacg acacgtgag atcgtcttgg ccggggggta 2220
 cctacagatg tccaagggtt agtgcctgac tgtaattgta tgtctgagga caaatatgta 2280

gtcagccgta taaagtcata ccaggcacca gtgccatcat cgaaccacta actctctatg 2340

atacatgcct ccggtattat tgtaccatgc gtcgctttgt tacatacgtat tcttgcccttt 2400

ttctctcaga aactccagac ttgggtattt ggctcgagata agcccgacc atagttagtc 2460

tttcacactc tgtttaaaca agctaccacc aactcgttg ggtgcagtcg acgagtatct 2520

gtctgactcg tcattgccgc ctttggagta cgactccaac tatgagtgtg cttggatcac 2580

tttgacgata cattcttcgt tggaggctgt gggtctgaca gctgcgtttt cggcgcggtt 2640

ggccgacaac aatatcagct gcaacgtcat tgctggcttt catcatgac acatttttgt 2700

cggcaaaggc gacgcccaga gagccattga cgttctttct aatttgacc gatagccgta 2760

tagtccagtc taictataag ttcaactaac tcgtaactat taccataaca tatacttcac 2820

tgccccagat aaggttccga taaaaagttc tgcagactaa atttatttca gtctcctctt 2880

caccacaaaa atgcccctct acgaagctcg agctaacgtc cacaagtcg cctttgccgc 2940

tcgagtgtc aagctcgtgg cagccaagaa aaccaacctg tgtgcttctc tggatgttac 3000

caccaccaag gagctcattg agcttgccga taaggctcga ccttatgtgt gcatgatcaa 3060

aaccatata gacatcattg acgacttcac ctacgccgac actgtgtcc ccctcaagga 3120

acttgctctt aagcacggtt tcttctgtt cgaggacaga aagttcgag atattggcaa 3180

cactgtcaag caccagtacc ggtgtcaccg aatcgccgag tggtcgata tcaccaacgc 3240

ccacggtgta ccggaaccg gaatcattgc tggcctgca gctggtgccg aggaaactgt 3300

ctctgaacag aagaaggagg acgtctctga ctacgagaac tcccagtaca aggagtctct 3360

agtcacctct cccaacgaga agctggccag aggtctgtc atgtggccg agctgtcttg 3420

caagggtctt ctggccactg gcgagtactc caagcagacc attgagcttg cccgatccga 3480

ccccagttt gtggttggtt tcattgccca gaaccgacct aagggcgact ctgaggactg 3540

gcttattctg acccccgggg tgggtcttga cgacaaggga gacgtctcg gacagcagta 3600

ccgaactgtt gaggatgtca tgtctaccgg aacggatata ataattgtcg gccgaggtct 3660

gtacggccag aaccgagatc ctattgagga ggccaagcga taccagaagg ctggctggga 3720

ggcttaccag aagattaact gtttagaggt agactatgga tatgtaat t aactgtgtat 3780

atagagagcg tgcaagtatg gacgcttgt tcagcttgta tgatggtcag acgacctgtc 3840

tgatcgagta tgtatgata tgcacaacct gtgtatccgc atgatctgtc caatggggca 3900

tgttgttgtt ttctcagata cggagatgct gggtacagtg ctaatacgtt gaactactta 3960

tacttatatg aggtcgaag aaagctgact tgtgtatgac ttattctcaa ctacatcccc 4020

agtcacaata ccaccactgc actaccacta caccaaaacc atgatcaaac caccgatgga 4080
 cttcctggag gcagaagaac ttgttatgga aaagctcaag agagagaatt caggagagac 4140
 cgggttggcg gcgtatttgt gtcccaaaaa acagccccaa ttgccccgga gaagacggcc 4200
 aggccgccta gatgacaaat tcaacaactc acagctgact ttctgccatt gccactaggg 4260
 gggggccttt ttatatggcc aagccaagct ctccacgtcg gttgggctgc acccaacaat 4320
 aaatgggtag ggttgacca acaaagggat gggatggggg gtagaagata cgaggataac 4380
 ggggctcaat ggacaaaata agaacgaata ctgccattaa gactcgtgat ccagcgactg 4440

 acaccattgc atcatctaag ggctcaaaa ctacctgga actgctgcgc tgatctggac 4500
 accacagagg ttccgagcac tttaggttgc accaaatgtc ccaccagggtg caggcagaaa 4560
 acgttggac agcgtgtaca gtttgtctta acaaaaagtg agggcgctga ggtcgagcag 4620
 ggtggtgtga cttgttatag cctttagagc tgcgaaagcg cgtatggatt tggtcatca 4680
 ggccagattg aggtctcttg gacacatgtc atgttagtgt acttcaatcg cccctggat 4740
 atagccccga caataggccg tggcctcatt tttttgcctt ccgcacattt ccattgctcg 4800
 gtaccacac cttgtcttc ctcacttgc caaccttaat actggtttac attgaccaac 4860

 atcttacaag cggggggcctt gtctagggtat tatataaaca gtggctctcc caatcggttg 4920
 ccagtctctt ttttctttc ttccccaca gattcgaaat ctaaaactaca catcacacaa 4980
 tgcctgttac tgacgtcctt aagcgaaagt ccggtgtcat cgtcggcgac gatgtccgag 5040
 ccgtgagtat ccacgacaag atcagtgtcg agacgacgcg ttttgtgtaa tgacacaatc 5100
 cgaaagtgc tagcaacaca cactctctac acaaaactaac ccagtctctc atggctctct 5160
 cccttactac cgagcagctg ctcgagcgac ccgacctggt tgccatcgac ggcattctct 5220
 acgatctgga aggtcttgcc aaggtccatc ccgcatccga cttgatctc gcttctggtg 5280

 cctccgatgc ttctctctg ttctactcca tgcacctta cgtcaagccc gagaactcga 5340
 agctgcttca acagtctgtg cgaggcaagc acgaccgaac ctccaaggac attgtctaca 5400
 cctacgactc tccctttgca caggacgtca agcgaactat gcgagaggtc atgaaaggtc 5460
 ggaactggta tgccacacct ggattctggc tgcgaaccgt tggcatcatt gctgtcaccg 5520
 ccttttgca gtggcactgg gctactaccg gaatggtgct gtggggtctc ttgactggat 5580
 tcatgcatat gcagatcggc ctgtccattc agcacgatgg ttctcatggt gccatcagca 5640
 aaaagccctg ggtaacgct ctctttgcct acggcatcga cgtcattgga tcgtccagat 5700

 ggatctggct gcagtctcac atcatgcgac atcacaccta caccaatcag catggtctcg 5760
 acctggatgc cgagtccga gaaccattcc ttgtgttcca caactacct gctgccaaca 5820
 ctgctcgaat gtggtttcac cgattccagg cctggtacat gtacctctg cttggagcct 5880

acggcgtttc gctgggtgtac aacctctctt acatcttccg aatgcagcac aacgacacca 5940
 ttcccagatc tgtcacagcc atgcgagaga acggctttct gcgacggtac cgaacccttg 6000
 cattcgttat gcgagctttc ttcacttttc gaaccgcctt cttgccctgg tatctcactg 6060
 gaacctccct gctcatcacc attctcttgg tgcccactgc taccgggtgcc ttcctcacct 6120

tctttttcat cttgtctcac aacttcgatg gctcggagcg aatccccgac aagaactgca 6180
 aggtcaagag ctccgagaag gacgttgaag ccgatcagat cgactggtac agagctcagg 6240
 tggagacctc ttccacctac ggtggaccca ttgccatgtt ctttactggc ggtctcaact 6300
 tccagatcga gcatcacctc tttctctgaa tgtcgtcttg gcactatccc ttcgtgcagc 6360
 aagctgtccg agagtgttgc gaacgacacg gagttcggta cgtcttttac cctaccattg 6420
 tgggcaacat catttcacc ctcaagtaca tgcacaaagt cgggtgtggt cactgtgtca 6480
 aggacgtca ggattcctaa gcggccgcaa gtgtggatgg ggaagtgagt gcccggttct 6540

gtgtgcacaa ttggcaatcc aagatggatg gattcaacac agggatatag cgagctacgt 6600
 ggtgggtgca ggatatagca acggatattt atgtttgaca cttgagaatg tacgatacaa 6660
 gcactgtcca agtacaatac taaacatact gtacatactc atactcgtac ccgggcaacg 6720
 gtttcacttg agtgcagtgg ctagtgcctt tactcgtaca gtgtgcaata ctgcgtatca 6780
 tagtctttga tgtatatcgt attcattcat gttagtgtcg tacgacagac ggacaatgac 6840
 gagccaggtt cgcccaactg acggaatctc ctatagtctt ctcatcaagc gcataaatat 6900
 agtatataaa gaattacatg gtgggccaca gcttggcgat agtatgacgg tctgggctcc 6960

agatcacacc catgagggat gccgtgcgag tagatgtaga ggtgggtgctt ccagggtgtc 7020
 cccaacaaca atcgctactg gaaaataaccg tgaaagtacg tacaccaac tctgtgactc 7080
 gcgttgactt tcatctttca atttggtcgt gtcttacggt gcatgtgctg tattggacat 7140
 gaagaagggg ttggcaagct gaaccatgcc tgagggtctg gagtctgata ctgtctagca 7200
 gagggcgatc aaaagatcac catccacacg tttgtttcca tattcaatct ccaatttcca 7260
 acatctcctt ggctaatagc cactttgaga gatatttggg atcgcgtatt cacttgtctc 7320
 atatcccggtg tttctgtttc accaaggctt ggacgatatt atcgcgagaa aaaccgtatt 7380

gaggtattta ttagagtcca gagaggagga acaggtattc ggaagatgga aaagctgtac 7440
 agagcgagaa acgatcccat gtgcagctc ctgatggctt atactgagac atggaatcag 7500
 gatcttgccc ttcttcacag ctctacttgc agctacagaa ttgctttttc tcgtatccga 7560
 tcataattcc gctgctaiga taggcagcca ccgatacaat ataggagaga gtataggtat 7620
 tgactctata gacactgtat gatgtgaca aagctgactg agggattcga ctcagcttcg 7680
 ccgttgcttg atctccaatt aattatcgtg ggcgcgccag ctgcattaat gaatcgcca 7740

acgcgcgggg agaggcgggt tgcgtattgg gcgctcttcc gcttcctcgc tcaactgactc 7800

gctgcgctcg gtcgttcggc tgcggcgagc ggtatcagct cactcaaagg cggtaatacg 7860

gttatccaca gaatcagggg ataacgcagg aaagaacatg tgagcaaaag gccagcaaaa 7920

ggccaggaac cgtaaaaagg ccgcgttgct ggcggttttt cataggctcc gccccctga 7980

cgagcatcac aaaaatcgac gctcaagtca gaggtggcga aaccgcagag gactataaag 8040

ataccaggcg tttccccctg gaagctccct cgtgcgctct cctgttccga ccctgccgt 8100

taccggatac ctgtccgcct ttctcccttc gggaagcgtg gcgctttctc atagctcacg 8160

ctgtaggtag ctcaattcgg ttaggtcgt tcgctccaag ctgggctgtg tgcacgaacc 8220

ccccgttcag cccgaccgct gcgccttacc cggtaactat cgtcttgagt ccaaccgggt 8280

aagacacgac ttatcgccac tggcagcagc cactggtaac aggattagca gagcgaggta 8340

tgtaggcggg gctacagagt tcttgaagtg gtggcctaac tacggctaca ctagaagaac 8400

agtatttggg atctgcgctc tgcgaagcc agttaccttc ggaagagag ttggtagctc 8460

ttgatccggc aaacaaacca ccgctggtag cgggtggttt tttgtttgca agcagcagat 8520

tacgcgcaga aaaaaggat ctcaagaaga tctttgacg tttctacgg ggtctgacgc 8580

tcagtgggac gaaaactcac gtaagggat ttgggtcatg agattatcaa aaaggatctt 8640

cacctagatc cttttaaat aaaaatgaag ttttaaatca atctaaagta tatatgagta 8700

aacttgggtc gacagttacc aatgcttaat cagttaggca cctatctcag cgatctgtct 8760

atttcgttca tccatagttg cctgactccc cgtcgttag ataactacga tacgggaggg 8820

cttaccatct ggccccagtg ctgcaatgat accgcgagac ccacgctcac cggctccaga 8880

tttatcagca ataaaccagc cagccggaag ggccgagcgc agaagtggtc ctgcaacttt 8940

atccgcctcc atccagtcta ttaattgttg ccgggaagct agagtaagta gttgccagat 9000

taatagtttg cgcaacgttg ttgccattgc tacaggcatc gtggtgtcac gctcgtcgtt 9060

tggtatggct tcattcagct ccggttccca acgatcaagg cgagttacat gatccccat 9120

gttgtcaaaa aaagcggtta gctccttcgg tctccgacg gttgtcagaa gtaagttggc 9180

cgcagtgtaa tcaactatgg ttatggcagc actgcataat tctcttactg tcatgccatc 9240

cgtaagatgc ttttctgtga ctggtgagta ctcaaccaag tcattctgag aatagtgtat 9300

gcggcgaccg agttgctctt gcccgcgctc aatccgggat aatccgcgc cacatagcag 9360

aactttaaaa gtgctcatca ttggaaaacg ttcttcgggg cgaaaactct caaggatctt 9420

accgctgttg agatccagtt cgatgtaacc cactcgtgca cccaactgat cttcagcatc 9480

ttttactttt accagcgttt ctgggtgagc aaaaacagga aggcaaaatg cgcacaaaaa	9540
gggaataagg ggcacacgga aatgttgaat atcctacttc ttcctttttc aatattattg	9600
aagcatttat cagggttatt gtctcatgag cggatacata tttgaatgta tttagaaaaa	9660
taaacaataa ggggttccgc gcacatttcc ccgaaaagtg ccacctgatg cgggttgaaa	9720
taccgcacag atgcgtaagg agaaaatacc gcatcaggaa attgtaagcg ttaatatatt	9780
gttaaaattc gcgttaaat tttgttaaat cagctcattt tttaaccaat aggccgaaat	9840
cggcaaaatc cctataaat caaaagaata gaccgagata gggttgagtg ttgttccagt	9900
ttggaacaag agtccactat taaagaacgt ggactccaac gtcaaagggc gaaaaaccgt	9960
ctatcagggc gatggccac tacgtgaacc atcacctaa tcaagttttt tggggtcgag	10020
gtgccgtaaa gactaaatc ggaaccctaa agggagcccc cgatttagag ctgacgggg	10080
aaagccggcg aacgtggcga gaaaggaagg gaagaaagcg aaaggagcgg gcgctaggcg	10140
gtcggcaagt gtagcggta cgctcgcgt aaccaccaca cccgccgcgc ttaatgcgc	10200
gctacaggcg gcgtccatt gccattcagg ctgcgcaact gttgggaagg gcgatcggtg	10260
cgggcctctt cgctattacg ccagctggcg aaagggggat gtgctgcaag gcgattaagt	10320
tgggtaacgc cagggttttc ccagtcacga cgttgtaaaa cgacggccag tgaattgtaa	10380
tacgactcac tataggcgga attgggcccg acgtcgcag cattaccaac gttggcgcgc	10440
cacaacgtca gaaactctt tccatccgat aagccgtgtc taactatacg accctcaaac	10500
aaggtcagcc agaccaact ttagcactcg caaaacacac acccaaacc aggccgcgac	10560
gcagccggag cccgtctact tatactgttt cggccgagac gccaaagtac aaaaaagcg	10620
cacaaacct cccccagtg ggcggcgac atccacaaca aacgacccat aaacggcact	10680
cccacacgga acaaaaacgt ttgtcgggtt actcctcag ccaagacaca ctcacacaca	10740
caacctccca ctttcatca ccaactcaact acgggtacag tagtggtaga aaacacagag	10800
tccagagtc acacagacac cgcacagagt ctgtcacaca agcaccatgc caacaagcag	10860
accgttcatt tcgtgtgtg ttggaagctt ctccacgtcg tatcaacga acaccaccac	10920
acagacctcc gcatcgacag ccacgtcac caccggcccc acgtcgaga tgtacaagat	10980
gctcgacaac cactcgata ctgagcagcc ccagcagccg accctgactc agacacagcc	11040
agcggtcgag acggcgaga cggcgagac ggcgagcag gcaagtggga ctgcggacct	11100
cagcgaatcg ccacacgaca agctgtggat tggaggccga tcaaacgacg gacgggaaaa	11160
gttctaccgg ctggaaccgg ccaagcgaag ggcaccttc gatcgattt ctctggatca	11220
ggtgtctatt tgatgaggcg ccaaacagtg gccagaccaa gtctggatc tcaaccctc	11280
aaaattgata accaagtttc tgggtggagt caaaccaagg ggagctacgg agccgttcac	11340

aaggtgtcca cagaggttca aagcacattg caccaagacc gccatcatcc ttaattaaaa	11400
ggcgttgaaa cagaatgagc cagacagcaa ggacaaggtg gccaacagca aggagtccaa	11460
aaagccctct attgacgaga tccacgatgt tatgtctcat gaggtttccg agctcgatgc	11520
tgggaagaag aagtgatttg tatataagaa ataatgaga tatagtaaag gagtgcgaaga	11580
gaatggcaag gtggtcaaat tctatattac ttgcagtcac tggttcctcg ttgacatgaa	11640
tgaagttacc gttggcatag ctgatttaat atataactgt ccaactaact ctcacctaga	11700
tataacccat gtgtgtgttt ccaatcatca atgcggccgc ttaggaatcc tgtgctcct	11760
tcacgcagtg gacgacacc accttatgca tgtacttcag ggtggagatg atgttgccga	11820
cgatggtagg gtagaaaaca tatcgcactc catgtcgttc gcaacactcc cggaccgcct	11880
gctggacgaa ggggtagtgc caagacgaca tccggggaaa gaggtggtgc tcgatctgga	11940
aattgagacc gccagtgaag aacatggcga tggggccacc gtatgtggag gacgtctcca	12000
cctgcgcccc ataccagtca atttggtcag cctcaacgtc cttctcagag ctcttaacct	12060
tgcagttctt gtcggggatc cgttcggagc catcaaaatt gtgggacaaa atgaagaaga	12120
acgtcaagaa ggcaccagtt gcggtgggca ccagaggaat ggtgatcagc aatgaggtcc	12180
cagtgaggtg ccagggcaag aatgcggtcc ggaagatgaa gaaagctcgc atcacgaatg	12240
caagtgtgcg gtagcggcgc agaaagccat ttcccccat ggccgtgaca gactctggga	12300
tgggtgtcatt gtgtcgtatc cggaatatgt agagcgggtt gtacaccagc gataccccgt	12360
atgccccag cacaaggtac atgtaccaag cctggaagcg gtggaaccac tttcgggcgg	12420
tgtttcgggc ggggtagtgt tggaacacca ggaacggctc tgccgactcc gcatccaggt	12480
cgaggccgtg ctggttggtg taggtgtggt gccgatgat gtgcgactgc agccaaatcc	12540
accgggacga tccgatgacg tcaatgccgt aggcgaagag ggcgttgacc caaggttct	12600
tgctgatggc cccgtgacct gcatcatgct ggatggataa gccgatctgc atgtgcatga	12660
atccagtcaa caggccccac agcaccatcc ccgtggtagc ccagtggcac tcgcaaaagg	12720
ccgtcacggc gatgatccca acggtgcgca gccagaagcc aggggttgcg taccagttcc	12780
tccctttcat cacctcgcgc attgtccgt taactcttg tgcaaggga gaatcatacg	12840
tgtagacaat gtccttcgag gtgcggtcat gcttcctcg gacgaattgc tgcaggagct	12900
tggagtcttc gggcttgacg tagggatgca tagagtagaa gaggggagag gcgtcggagg	12960
ctccagaggc cagaatgagg tctccaccac catgcacatt ggccaggccc tcgaggtcgt	13020
acaggatacc atcgatagcg acgagatcag gtcgctccag gagctgttcg gtagtgagag	13080
acagggccat ggccattgct gtagatatgt cttgtgtgta aggggggttg ggtggttgtt	13140
tgtgttcttg acttttgtt tagcaaggga agacgggcaa aaaagtgagt gtggttggga	13200

gggagagacg agccttatat ataatgcttg tttgtgtttg tgcaagtga cgccgaaacg 13260

ggcaggagcc aaactaaaca aggcagacaa tgcgagctta attggattgc ctgatgggca 13320

ggggttaggg ctcgatcaat gggggtgcga agtgacaaaa ttgggaatta ggttcgcaag 13380

caaggctgac aagactttgg cccaacatt tgtacgcggt ggacaacagg agccacccat 13440

cgtctgtcac gggctagccg gtcgtgcgtc ctgtcaggct ccacctaggc tccatgccac 13500

tccatacaat ccactagtg taccgctagg ccgcttttag cteccatcta agaccccccc 13560

aaaacctcca ctgtacagtg cactgtactg tgtggcgatc aagggaagg gaaaaaaggc 13620

gcaaacatgc acgcatggaa tgacgtaggt aaggcggtac tagactgaaa agtggcacat 13680

ttcggcgtgc caaagggtcc taggtgcgtt tcgcgagctg ggcgccaggc caagccgctc 13740

caaaacgcct ctccgactcc ctccagcggc ctccatatcc ccatccctct ccacagcaat 13800

gttgtaagc cttgcaaac aaaaaataga aaggctaata agcttccaat attgtggtgt 13860

acgctgcata acgcaacaat gagcgccaaa caacacacac acacagcaca cagcagcatt 13920

aacctat 13926

<210> 115

<211> 7769

<212> DNA

<213> Artificial Sequence

<220><223> Plasmid pZUFmEgD9eS

<400> 115

ggccgcaagt gtggatgggg aagtgagtgc ccggttctgt gtgcacaatt ggcaatccaa 60

gatggatgga ttcaacacag ggatatagcg agctacgtgg tgggtcgagg atatagcaac 120

ggatatattat gtttgacact tgagaatgta cgatacaagc actgtccaag tacaatacta 180

aacatactgt acatactcat actcgtaccg gggcaacggg ttactttgag tgcagtggct 240

agtgtcttta ctcttacagt gtgcaatact gcgtatcata gtctttgatg tatatcgtat 300

tcattcatgt tagttgcgta cgagccggaa gcataaagtg taaagcctgg ggtgcctaata 360

gagttagcta actcacatta attgcgttgc gctcactgcc cgctttccag tcgggaaacc 420

tgtcgtgcc gctgcattaa tgaatcgcc aacgcgcggg gagaggcggt ttgcgtattg 480

ggcgctcttc cgcttcctcg ctactgact cgctgcgtc ggtcgttcgg ctgcggcgag 540

cggtatcagc tactcaaag gcgtaatac ggttatccac agaactcagg gataacgcag 600

gaaagaacat gtgagcaaaa ggccagcaaa aggccaggaa ccgtaaaaag gccgcgttgc 660

tggcggtttt ccataggctc cgccccctg acgagcatca caaaaatcga cgctcaagtc 720

agaggtggcg aaacccgaca ggactataaa gataccaggc gtttccccct ggaagctccc	780
tcgtgcgctc tctgtttccg accctgccgc ttaccggata cctgtccgcc ttctccctt	840
cgggaagcgt ggcgctttct catagctcac gctgtaggta tctcagttcg gtgtaggtcg	900
ttcgctccaa gctgggctgt gtgcacgaac cccccgttca gcccgaccgc tgcgccttat	960
ccggtaaacta tcgtcttgag tccaacccgg taagacacga cttatcgcca ctggcagcag	1020
ccactggtaa caggattagc agagcgaggt atgtaggcgg tgctacagag ttcttgaagt	1080
ggtggcctaa ctacggctac actagaagga cagtatttgg tatctgcgct ctgctgaagc	1140
cagttacctt cgaaaaaga gttggtagct ctgatccgg caaacaacc accgctggtta	1200
gcggtggttt ttttgtttgc aagcagcaga ttacgcgcag aaaaaagga tctcaagaag	1260
atcctttgat cttttctacg gggctctgacg ctcagtggaa cgaaaactca cgtaaggga	1320
ttttggtcat gagattatca aaaaggatct tcacctagat ctttttaaat taaaaatgaa	1380
gttttaaatc aatctaaagt atatatgagt aaacttggtc tgacagttac caatgcttaa	1440
tcagtgaggc acctatctca gcgatctgtc tatttcgttc atccatagtt gcctgactcc	1500
ccgtcgtgta gataactacg atacgggagg gcttaccatc tggccccagt gctgcaatga	1560
taccgcgaga cccacgctca ccggctccag atttatcagc aataaaccag ccagccggaa	1620
gggccgagcg cagaagtgtt cctgcaactt tatccgcctc catccagtct attaatgtt	1680
gccgggaagc tagagtaagt agttcgccag ttaatagttt gcgcaacgtt gttgccattg	1740
ctacaggcat cgtggtgtca cgctcgtcgt ttggtatggc ttcatcagc tccggttccc	1800
aacgatcaag gcgagttaca tgatccccca tgttgtgcaa aaaagcggtt agctccttcg	1860
gtcctccgat cgttgtcaga agtaagttag ccgcagtggt atcactcatg gttatggcag	1920
cactgcataa ttctcttact gtcatgcat ccgtaaatg cttttctgtg actggtgagt	1980
actcaaccaa gtcattctga gaatagtgtg tgcggcgacc gagttgctct tgccccgcgt	2040
caatacggga taataccgcg ccacatagca gaactttaaa agtgctcatc attggaaaac	2100
gttcttcggg gcgaaaactc tcaaggatct taccgctgtt gagatccagt tcgatgtaac	2160
ccactcgtgc acccaactga tcttcagcat cttttacttt caccagcgtt tctgggtgag	2220
caaaaacagg aaggcaaaat gccgcaaaaa agggaataag ggcgacacgg aaatgttgaa	2280
tactcatact cttccttttt caataattatt gaagcattta tcagggttat tgtctcatga	2340
gcggatacat atttgaatgt atttagaaaa ataaacaaat aggggttccg cgcacatttc	2400
cccgaagaat gccacctgac gcgccctgta gcggcgcat aagcgcgcg ggtgtggtgg	2460
ttacgcgcag cgtgaccgtt acacttgcca gcgccctagc gccgcctcct ttcgctttct	2520

tcccttcctt tctcgccacg ttccgcccgt ttccccgtca agctctaaat cgggggctcc	2580
ctttagggtt ccgatttagt gctttacggc acctcgaccc caaaaaactt gattagggtg	2640
atggttcacg tagtgggccca tcgccctgat agacgggttt tcgccctttg acgttggagt	2700
ccacgttctt taatagtgga ctcttgttcc aaactggaac aacactcaac cctatctcgg	2760
tctattcttt tgatttataa gggattttgc cgatttcggc ctattggtta aaaaatgagc	2820
tgatttaaca aaaatttaac gcgaatttta acaaaatatt aacgcttaca atttcattc	2880
gccattcagg ctgcgcaact gttgggaagg gcgatcgggtg cgggcctctt cgctattacg	2940
ccagctggcg aaagggggat gtgctgcaag gcgattaagt tgggtaacgc cagggttttc	3000
ccagtcacga cgttgtaaaa cgacggccag tgaattgtaa tacgactcac tataggcgca	3060
attgggtacc gggccccccc tcgaggtcga tgggtgctgat aagcttgata tcgaattcat	3120
gtcacacaaa ccgattctcg cctcaaggaa acctaatctt acatccgaga gactgccgag	3180
atccagtcta cactgattaa ttttcgggcc aataatttaa aaaaatcgtg ttatataata	3240
ttatatgtat tatatatata catcatgatg atactgacag tcatgtcca ttgctaaata	3300
gacagactcc atctgccgc tccaactgat gttctcaata ttttaagggt catctcgcat	3360
tgtttaataa taacagact ccatctaccg cctccaatg atgttctcaa aatatattgt	3420
atgaacttat ttttattact tagtattatt agacaactta cttgttttat gaaaaaact	3480
tcctatttag gaaacaattt ataatggcag ttcggttcatt taacaattta tgtagaataa	3540
atgttataaa tgcgtatggg aaatcttaaa tatggatagc ataaatgata tctgcattgc	3600
ctaattcgaa atcaacagca acgaaaaaaaa tcccttgtag aacataaata gtcacgaga	3660
aatatcaact atcaaagaac agctattcac acgttactat tgagattatt attggacgag	3720
aatcacacac tcaactgtct ttctctcttc tagaaataca ggtacaagta tgtactattc	3780
tcattgttca tacttctagt catttcatcc cacatatcc ttggatttct ctccaatgaa	3840
tgacattcta tcttgcaaat tcaacaatta taataagata taccaaagta gcggtatagt	3900
ggcaatcaaa aagcttctct ggtgtgcttc tcgtatttat ttttattcta atgatccatt	3960
aaaggatata atttatttct tggtatataa tccttttggt tattacatgg gctggataca	4020
taaaggatatt ttgatttaat tttttgctta aattcaatcc cccctcgctc agtgtcaact	4080
gtaatgtag gaaattacca tacttttgaa gaagcaaaaa aatgaaaga aaaaaaaaaat	4140
cgtatttcca ggtagacgt tccgcagaat ctagaatgcg gtatgcggta cattgttctt	4200
cgaacgtaaa agttgcgctc cctgagatat tgtacatttt tgcttttaca agtacaagta	4260

catcgtacaa ctatgtacta ctgttgatgc atccacaaca gttgttttg ttttttttg	4320
ttttttttt ttctaagat tcattaccgc tatgtatacc tacttgact ttagtaagc	4380
cgggttattg gcgttcaatt aatcatagac ttatgaatct gcacggtgtg cgctgcgagt	4440
tacttttagc ttatgcatgc tacttgggtg taatatggg atctgttcgg aaatcaacgg	4500
atgctcaatc gatttcgaca gtaattaatt aagtcataca caagtcagct ttcttcgagc	4560
ctcatataag tataagtagt tcaacgtatt agcactgtac ccagcatctc cgtatcgaga	4620
aacacaacaa catgccccat tggacagatc atgcggatac acaggttgtg cagtatcata	4680
catactcgat cagacaggtc gtctgacat catacaagct gaacaagcgc tccatacttg	4740
cacgctctct atatacacag ttaaattaca tatccatagt ctaacctcta acagttaatc	4800
ttctggtaag cctcccagcc agccttctgg tctcgcttg cctcctcaat aggatctcgg	4860
ttctggccgt acagacctg gccgacaatt atgatatccg ttccggtaga catgacatcc	4920
tcaacagttc ggtactgctg tccgagagcg tctcccttgt cgtcaagacc cccccgggg	4980
gtcagaataa gccagtcctc agagtcgcc ttaggtcggg tctgggcaat gaagccaacc	5040
acaaactcgg ggtcggatcg ggcaagctca atggtctgct tggagtactc gccagtggcc	5100
agagagccct tgcaagacag ctcgccagc atgagcagac ctctggccag cttctcgttg	5160
ggagagggga ctaggaactc cttgtactgg gagttctcgt agtcagagac gtctccttc	5220
ttctgttcag agacagtttc ctcgccacca gctcgaggc cagcaatgat tccggttccg	5280
ggtacaccgt gggcgttggg gatatcggac cactcggcga ttcggtgaca ccggtactgg	5340
tgcttgacag tgttgccaat atctgcgaac tttctgtcct cgaacaggaa gaaaccgtgc	5400
ttaagagcaa gttccttgag ggggagcaca gtgccggcgt aggtgaagtc gtcaatgatg	5460
tcgatatggg ttttgatcat gcacacataa ggtccgacct tatcggaag ctcaatgagc	5520
tccttggtgg tggtaacatc cagagaagca cacaggttgg ttttcttggc tgccacgagc	5580
ttgagcactc gagcggcaaa ggcggacttg tggacgttag ctcgagcttc gtaggagggc	5640
atcttggtgg tgaagaggag actgaaataa atttagctg cagaactttt tatcggaacc	5700
ttatctgggg cagtgaagta tatgttatgg taatagttac gagttagtig aacttataga	5760
tagactggac tatacggcta tcggtccaaa ttagaaagaa cgtcaatggc tctctgggcg	5820
tcgcttttgc cgacaaaaat gtgatcatga tgaaagccag caatgacgtt gcagctgata	5880
ttgttgtcgg ccaaccgcgc cgaaaacgca gctgtcagac ccacagcctc caacgaagaa	5940
tgtatcgtca aagtgatcca agcacactca tagttggagt cgtactcaa aggcggcaat	6000
gacgagtcag acagatactc gtcgacgttt aaacagtga cgcagatcta ctatagagga	6060
acatttaaat tgccccggag aagacggcca ggccgcctag atgacaaatt caacaactca	6120

cagctgactt tctgccattg ccactagggg ggggcctttt tataatggcca agccaagctc 6180
tccacgtcgg ttgggctgca cccaacaata aatgggtagg gttgcaccaa caaagggatg 6240
ggatgggggg tagaagatac gaggataacg gggctcaatg gcacaaataa gaacgaatac 6300
tgccattaag actcgtgata cagcgactga caccattgca tcatctaagg gcctcaaaac 6360

tacctcgga ctgctgcgt gatctggaca ccacagaggt tccgagcact ttaggttgca 6420
ccaaatgtcc caccaggtagc aggcagaaaa cgctggaaca gcgtgtacag ttgtcttaa 6480
caaaaagtga gggcgctgag gtcgagcagg gtggtgtgac ttgttatagc ctttagagct 6540
gcgaaagcgc gtatggattt ggctcatcag gccagattga gggctctgtg acacatgtca 6600
tgttagtgta ctcaatcgc cccctggata tagcccgac aataggccgt ggcctcattt 6660
ttttgccttc cgcacatttc cattgtctga taccacacc ttgcttctcc tgcattgcc 6720
aaccttaata ctggtttaca ttgaccaaca tcttacaagc ggggggcttg tctagggtat 6780

atataaacag tggctctccc aatcggttgc cagtctcttt tttcctttct tccccacag 6840
attcgaaatc taaactacac atcacagaat tccgagccgt gagtatccac gacaagatca 6900
gtgtcgagac gacgcgtttt gtgtaatgac acaatccgaa agtcgctagc aacacacact 6960
ctctacacaa actaaccagc ctctggtagc atggaggtcg tgaacgaaat cgtctccatt 7020
ggccaggagg ttcttcccaa ggtcgactat gctcagctct ggtctgatgc ctgcactgc 7080
gaggtgctgt acctctccat cgcttctgct atcctgaagt tcaccttgg tcctctcgga 7140
cccaagggtc agtctcgaat gaagtttgtg ttcaccaact acaacctgct catgtccatc 7200

tactcgtgg gctccttct ctctatggcc tacgcatgt acaccattgg tgtcatgtcc 7260
gacaactcgc agaaggcttt cgacaacaat gtcttccgaa tcaccactca gctgttctac 7320
ctcagcaagt tctcgagta cattgactcc ttctatctgc cctcatggg caagcctctg 7380
acctggttgc agttctttca ccatctcgga gctcctatgg acatgtggt gtcttacaac 7440
taccgaaacg aagccgtttg gatctttgtg ctgctcaacg gcttcattca ctggatcatg 7500
tacggctact attggaccgc actgatcaag ctcaagttcc ctatgcccaa gtccctgatt 7560
acttctatgc agatcattca gttaacgtt ggcttctaca tcgtctggaa gtaccggaac 7620

attcctgct accgacaaga tggaaatgaga atgtttggct ggtttttcaa ctacttctac 7680
gttggtactg tctgtgtct gtctctcaac ttctacgtgc agacctacat cgtccgaaag 7740
cacaaggag ccaaaaagat tcagtgcgc 7769

<210> 116

<211> 20

<212> DNA

<213> Artificial Sequence

<220><223> Primer pZUFm_6980_012208f

<400> 116

gctctggtac catggaggtc 20

<210> 117

<211> 20

<212> DNA

<213> Artificial Sequence

<220><223> Primer pZUFm_40_012208r

<400> 117

acagaaccgg gcactcatt 20

<210> 118

<211> 777

<212> DNA

<213> Euglena gracilis

<220><221> CDS

<222> (1)..(777)

<223> mutant delta-9 elongase "EgD9eS-L35F"

<400> 118

atg gag gtc gtg aac gaa atc gtc tcc att ggc cag gag gtt ctt ccc 48

Met Glu Val Val Asn Glu Ile Val Ser Ile Gly Gln Glu Val Leu Pro

1 5 10 15

aag gtc gac tat gct cag ctc tgg tct gat gcc tgc cac tgc gag gtg 96

Lys Val Asp Tyr Ala Gln Leu Trp Ser Asp Ala Ser His Cys Glu Val

20 25 30

ctg tac ttc tcc atc gcc ttc gtc atc ctg aag ttc acc ctt ggt cct 144

Leu Tyr Phe Ser Ile Ala Phe Val Ile Leu Lys Phe Thr Leu Gly Pro

35 40 45

ctc gga ccc aag ggt cag tct cga atg aag ttt gtg ttc acc aac tac 192

Leu Gly Pro Lys Gly Gln Ser Arg Met Lys Phe Val Phe Thr Asn Tyr

50 55 60

aac ctg ctc atg tcc atc tac tgc ctg ggc tcc ttc ctc tct atg gcc 240

Asn	Leu	Leu	Met	Ser	Ile	Tyr	Ser	Leu	Gly	Ser	Phe	Leu	Ser	Met	Ala	
65				70				75				80				
tac	gcc	atg	tac	acc	att	ggg	gtc	atg	tcc	gac	aac	tgc	gag	aag	gct	288
Tyr	Ala	Met	Tyr	Thr	Ile	Gly	Val	Met	Ser	Asp	Asn	Cys	Glu	Lys	Ala	
85				90				95								
ttc	gac	aac	aat	gtc	ttc	cga	atc	acc	act	cag	ctg	ttc	tac	ctc	agc	336
Phe	Asp	Asn	Asn	Val	Phe	Arg	Ile	Thr	Thr	Gln	Leu	Phe	Tyr	Leu	Ser	
100				105				110								
aag	ttc	ctc	gag	tac	att	gac	tcc	ttc	tat	ctg	ccc	ctc	atg	ggc	aag	384
Lys	Phe	Leu	Glu	Tyr	Ile	Asp	Ser	Phe	Tyr	Leu	Pro	Leu	Met	Gly	Lys	
115				120				125								
cct	ctg	acc	tgg	ttg	cag	ttc	ttt	cac	cat	ctc	gga	gct	cct	atg	gac	432
Pro	Leu	Thr	Trp	Leu	Gln	Phe	Phe	His	His	Leu	Gly	Ala	Pro	Met	Asp	
130				135				140								
atg	tgg	ctg	ttc	tac	aac	tac	cga	aac	gaa	gcc	gtt	tgg	atc	ttt	gtg	480
Met	Trp	Leu	Phe	Tyr	Asn	Tyr	Arg	Asn	Glu	Ala	Val	Trp	Ile	Phe	Val	
145				150				155				160				
ctg	ctc	aac	ggc	ttc	att	cac	tgg	atc	atg	tac	ggc	tac	tat	tgg	acc	528
Leu	Leu	Asn	Gly	Phe	Ile	His	Trp	Ile	Met	Tyr	Gly	Tyr	Tyr	Trp	Thr	
165				170				175								
cga	ctg	atc	aag	ctc	aag	ttc	cct	atg	ccc	aag	tcc	ctg	att	act	tct	576
Arg	Leu	Ile	Lys	Leu	Lys	Phe	Pro	Met	Pro	Lys	Ser	Leu	Ile	Thr	Ser	
180				185				190								
atg	cag	atc	att	cag	ttc	aac	gtt	ggc	ttc	tac	atc	gtc	tgg	aag	tac	624
Met	Gln	Ile	Ile	Gln	Phe	Asn	Val	Gly	Phe	Tyr	Ile	Val	Trp	Lys	Tyr	
195				200				205								
cgg	aac	att	ccc	tgc	tac	cga	caa	gat	ggg	atg	aga	atg	ttt	ggc	tgg	672
Arg	Asn	Ile	Pro	Cys	Tyr	Arg	Gln	Asp	Gly	Met	Arg	Met	Phe	Gly	Trp	
210				215				220								
ttt	ttc	aac	tac	ttc	tac	gtt	ggg	act	gtc	ctg	tgt	ctg	ttc	ctc	aac	720
Phe	Phe	Asn	Tyr	Phe	Tyr	Val	Gly	Thr	Val	Leu	Cys	Leu	Phe	Leu	Asn	

225 230 235 240
 ttc tac gtg cag acc tac atc gtc cga aag cac aag gga gcc aaa aag 768

Phe Tyr Val Gln Thr Tyr Ile Val Arg Lys His Lys Gly Ala Lys Lys
 245 250 255
 att cag tga 777
 Ile Gln

<210> 119

<211> 258

<212> PRT

<213> Euglena gracilis

<400> 119

Met Glu Val Val Asn Glu Ile Val Ser Ile Gly Gln Glu Val Leu Pro

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

65 70 75 80
 Tyr Ala Met Tyr Thr Ile Gly Val Met Ser Asp Asn Cys Glu Lys Ala
 85 90 95
 Phe Asp Asn Asn Val Phe Arg Ile Thr Thr Gln Leu Phe Tyr Leu Ser
 100 105 110
 Lys Phe Leu Glu Tyr Ile Asp Ser Phe Tyr Leu Pro Leu Met Gly Lys
 115 120 125
 Pro Leu Thr Trp Leu Gln Phe Phe His His Leu Gly Ala Pro Met Asp

130 135 140
 Met Trp Leu Phe Tyr Asn Tyr Arg Asn Glu Ala Val Trp Ile Phe Val

145 150 155 160
 Leu Leu Asn Gly Phe Ile His Trp Ile Met Tyr Gly Tyr Tyr Trp Thr
 165 170 175
 Arg Leu Ile Lys Leu Lys Phe Pro Met Pro Lys Ser Leu Ile Thr Ser
 180 185 190
 Met Gln Ile Ile Gln Phe Asn Val Gly Phe Tyr Ile Val Trp Lys Tyr

 195 200 205
 Arg Asn Ile Pro Cys Tyr Arg Gln Asp Gly Met Arg Met Phe Gly Trp
 210 215 220
 Phe Phe Asn Tyr Phe Tyr Val Gly Thr Val Leu Cys Leu Phe Leu Asn
 225 230 235 240
 Phe Tyr Val Gln Thr Tyr Ile Val Arg Lys His Lys Gly Ala Lys Lys
 245 250 255
 Ile Gln

<210> 120

<211> 7769

<212> DNA

<213> Artificial Sequence

<220><223> Plasmid pZUFmEgD9eS-L35F

<400> 120

ggccgcaagt gtggatgggg aagtgagtgc ccggttctgt gtgcacaatt ggcaatccaa 60
 gatggatgga ttcaacacag ggatatagcg agctacgtgg tgggtcgagg ataatagcaac 120
 ggatatattat gtttgacact tgagaatgta cgatacaagc actgtccaag tacaatacta 180
 aacataactgt acataactcat actcgtaccc gggcaacggt ttactttgag tgcagtggct 240
 agtgctctta ctctacagt gtgcaatact gcgtatcata gtctttgatg tatatcgtat 300
 tcattcatgt tagttgcgta cgagccggaa gcataaagtg taaagcctgg ggtgcctaata 360

 gagtgagcta actcacatta attgcgttgc gtcactgcc cgctttccag tcgggaaacc 420
 tgcgtgccca gctgcattaa tgaatcgcc aacgcgcggg gagaggcggt ttgcgtattg 480
 ggcgctcttc cgcttctctg ctactgact cgtgcgctc ggtcgttcgg ctgcgcgag 540
 cggtatcagc tactcaaag gcggttaatac ggttatccac agaatacagg gataacgcag 600
 gaaagaacat gtgagcaaaa ggccagcaaa aggccaggaa ccgtaaaaag gccgcgttgc 660

tgcggttttt ccataggtc cgccccctg acgagcatca caaaaatcga cgctcaagtc	720
agaggtggcg aaacccgaca ggactataaa gataccaggc gtttcccct ggaagctccc	780
tcgtgcgctc tectgttccg accctgccgc ttaccggata cctgtccgcc tttctccctt	840
cgggaagcgt ggcgctttct catagctcac gctgtaggta tctcagttcg gtgtaggtcg	900
ttcgctccaa gctgggctgt gtgcacgaac cccccgtca gcccgaccgc tgcgccttat	960
ccggtaaacta tegtcttgag tccaacccgg taagacacga cttatcgcca ctggcagcag	1020
ccactggtaa caggattagc agagcgaggt atgtaggcgg tgctacagag ttcttgaagt	1080
ggtggcctaa ctacggctac actagaagga cagtatttgg tatctgcgt ctgctgaagc	1140
cagttacctt cgaaaaaga gttggtagct cttgatccgg caaacaacc accgctggta	1200
gcggtggttt tttgtttgc aagcagcaga ttacgcgcag aaaaaagga tctcaagaag	1260
atcctttgat cttttctacg gggctgacg ctcagtggaa cgaaaactca cgttaaggga	1320
ttttggtcat gagattatca aaaaggatct tcacctagat ctttttaaat taaaaatgaa	1380
gttttaaatc aatctaaagt atatatgagt aaacttggc tgacagttac caatgcttaa	1440
tcagtgagge acctatctca gcgactgtc tatttcgttc atccatagtt gcctgactcc	1500
ccgtcgtgta gataactacg atacgggagg gcttaccatc tggccccagt gctgcaatga	1560
taccgcgaga cccacgtca ccggctccag atttatcagc aataaaccag ccagccggaa	1620
gggccgagcg cagaagtgtt cctgcaactt tatccgcctc catccagtct attaattgtt	1680
gccgggaagc tagagtaagt agttcgccag ttaatatgtt gcgcaacgtt gttgccattg	1740
ctacaggcat cgtgggtgca cgctcgtcgt ttggtatggc ttcatcagc tccggttccc	1800
aacgatcaag gcgagttaca tgatccccc tgttgtgcaa aaaagcgggt agctccttcg	1860
gtcctccgat cgttgtcaga agtaagttgg ccgcagtgtt atcactcatg gttatggcag	1920
cactgcataa ttctcttact gtcatgccaat ccgtaagatg cttttctgtg actggtgagt	1980
actcaaccaa gtcattctga gaatagtga tgcggcgacc gagttgctct tgccccgcgt	2040
caatacggga taataccgcg ccacatagca gaactttaaa agtgtcatc attggaaaac	2100
gttcttcggg gcgaaaactc tcaaggatct taccgtgtt gagatccagt tcgatgtaac	2160
ccactcgtgc acccaactga tcttcagcat cttttacttt caccagcgtt tctgggtgag	2220
caaaaacagg aaggcaaaa gccgcaaaaa agggaataag ggcgacacgg aatgttgaa	2280
tactcatact cticcttttt caatattatt gaagcattta tcagggttat tgttcatga	2340
gcggatacat atttgaatgt atttagaaaa ataaacaaat aggggttccg cgcacatttc	2400

cccgaaaagt gccacctgac gcgccttgta gcggcgcat t aagcgcgcg ggtgtggtgg 2460

 ttacgcgcag cgtgaccgct acacttgcca gcgccttagc gcccgctcct ttcgctttct 2520
 tcccttcctt tctcgccacg ttcccggtct tccccgtca agctctaaat cgggggctcc 2580
 ctttaggggt cggatttagt gctttacggc acctcgaccc caaaaaactt gattaggggtg 2640
 atggttcacg tagtgggcca tcgccctgat agacgggttt tcgccctttg acgttggagt 2700
 ccacgttctt taatagtga ctcttggtcc aaactggaac aacactcaac cctatctcgg 2760
 tctattcttt tgatttataa gggattttgc cgatttcggc ctattggtta aaaaatgagc 2820
 tgatttaaca aaaatttaac gcgaatttta acaaaatatt aacgcttaca atttcattc 2880

 gccattcagg ctgcgcaact gttgggaagg gcgatcggtg cgggcctctt cgctattacg 2940
 ccagctggcg aaagggggat gtgctgcaag gcgattaagt tgggtaacgc cagggttttc 3000
 ccagtcacga cgttgtaaaa cgacggccag tgaattgtaa tacgactcac tatagggcga 3060
 attgggtacc gggccccccc tcgaggtcga tgggtgctgat aagcttgata tcgaattcat 3120
 gtcacacaaa ccgactctcg cctcaaggaa acctaatctt acatccgaga gactgccgag 3180
 atccagtcta cactgattaa ttttcgggcc aataatttaa aaaaatcgtg ttatataata 3240
 ttatatgtat tatatatata catcatgatg atactgacag tcattgcca ttgctaaata 3300

 gacagactcc atctgccgc tccaactgat gttctcaata ttttaagggt catctcgcat 3360
 tgtttaataa taaacagact ccatctaccg cctccaaatg atgttctcaa aatatattgt 3420
 atgaacttat ttttattact tagtattatt agacaactta cttgctttat gaaaaacact 3480
 tcctatttag gaaacaattt ataatggcag ttcgttcatt taacaattta tgtagaataa 3540
 atgttataaa tgcgtatggg aaatcttaaa tatggatagc ataatgata tctgcattgc 3600
 ctaattcgaa atcaacagca acgaaaaaaa tcccttgtag aacataaata gtcactgaga 3660
 aatatcaact atcaaagaac agctattcac acgttactat tgagattatt attggacgag 3720

 aatcacacac tcaactgtct ttctctcttc tagaaatata ggtacaagta tgtactattc 3780
 tcattgttca tacttctagt catttcatcc cacatatcc ttggatttct ctccaatgaa 3840
 tgacattcta tcttgcaaat tcaacaatta taataagata taccaaagta gcggtatagt 3900
 ggcaatcaaa aagcttctct ggtgtgcttc tcgtatttat ttttattcta atgatccatt 3960
 aaaggatat atttatttct tggtatataa tcttttgtt tattacatgg gctggataca 4020
 taaaggattt ttgatttaat tttttgctta aattcaatcc cccctcgttc agtgtcaact 4080
 gtaatggtag gaaattacca tacttttgaa gaagcaaaaa aatgaaaga aaaaaaaaaat 4140

cgtattttcca ggtagacgt tccgcagaat ctagaatgcg gtatgcggtg cattgttctt	4200
cgaacgtaaa agttgcgctc cctgagatat tgtacatfff tgctttttaca agtacaagta	4260
catcgtacaa ctatgtacta ctgttgatgc atccacaaca gtttgttttg tttttttttg	4320
tttttttttt ttctaagtat tcattaccgc tatgtatacc tacttgtact ttagtaagc	4380
cgggttattg gcgttcaatt aatcatagac ttatgaatct gcacgggtg cgctgcgagt	4440
tacttttagc ttatgcagtc tacttgggtg taatatggg atctgttcgg aatcaacgg	4500
atgctcaatc gatttcgaca gtaattaatt aagtcataca caagtcagct ttcttcgagc	4560
ctcatataag tataagtagt tcaacgtatt agcactgtac ccagcatctc cgtatcgaga	4620
aacacaacaa catgccccat tggacagatc atgcggatac acaggttgtg cagtatcata	4680
catactcgat cagacaggtc gtctgacat catacaagct gaacaagcgc tccatacttg	4740
cacgtctct atatacacag ttaaattaca tatccatagt ctaacctcta acagttaatc	4800
ttctggttaag cctcccagcc agccttctgg tatcgcttgg cctcctcaat aggatctcgg	4860
ttctggccgt acagacctcg gccgacaatt atgatatccg ttccggtaga catgacatcc	4920
tcaacagttc ggtactgctg tccgagagcg tctcccttgt cgtcaagacc caccgccggg	4980
gtcagaataa gccagtcctc agagtcgccc ttaggtcggg tctgggcaat gaagccaacc	5040
acaaactcgg ggtcggatcg ggcaagctca atggtctgct tggagtactc gccagtggcc	5100
agagagccct tgcaagacag ctggccagc atgagcagac ctctggccag ctctcgttg	5160
ggagagggga ctaggaactc cttgtactgg gagttctcgt agtcagagac gtctccttc	5220
ttctgttcag agacagtttc ctggcacca gctcgcagc cagcaatgat tccggttccg	5280
ggtacaccgt gggcgttggg gatatcggac cactcggcga ttccgtgaca ccggtactgg	5340
tgcttgacag tgitgccaat atctgcgaac tttctgtcct cgaacaggaa gaaaccgtgc	5400
ttaagagcaa gttccttgag ggggagcaca gtgccggcgt aggtgaagtc gtcaatgatg	5460
tcgatatggg ttttgatcat gcacacataa ggtccgacct tatcggcaag ctcaatgagc	5520
tccttggtgg tggtaacatc cagagaagca cacaggttgg ttttcttggc tgccacgagc	5580
ttgagcactc gagcggcaaa ggcggacttg tggacgttag ctcgagcttc gtaggagggc	5640
atattggtgg tgaagaggag actgaaataa atttagtctg cagaactttt tatcggaacc	5700
ttatctgggg cagtgaagta tatgttatgg taatagttac gagttagtgt aacttataga	5760
tagactggac tatacggcta tcggtccaaa ttagaaagaa cgtcaatggc tctctgggcg	5820
tcgcctttgc cgacaaaaat gtgatcatga tgaaagccag caatgacgtt gcagctgata	5880
ttgttgtcgg ccaaccgcgc gaaaaacgca gctgtcagac ccacagcctc caacgaagaa	5940
tgtatcgtca aagtgatcca agcacactca tagttggagt cgtactcaa aggcggcaat	6000

gacgagtcag acagatactc gtcgacgttt aaacagtgtg cgcagatcta ctatagagga 6060
acatttaaat tgccccggag aagacggcca ggccgcctag atgacaaatt caacaactca 6120
cagctgactt tctgccattg ccactagggg ggggcctttt tataatggcca agccaagctc 6180
tccacgtcgg ttgggctgca cccaacaata aatgggtagg gttgcaccaa caaagggatg 6240

ggatgggggg tagaagatac gaggataacg gggtcctaatg gcacaaataa gaacgaatac 6300
tgccattaag actcgtgac cagcgactga caccattgca tcactaagg gcctcaaaac 6360
tacctcggaa ctgctgcgt gatctggaca ccacagaggt tccgagcact ttaggttgca 6420
ccaaatgtcc caccagtgcc aggcagaaaa cgctggaaca gcgtgtacag ttgtcttaa 6480
caaaaagtga gggcgctgag gtcgagcagg gtggtgtgac ttgttatagc ctttagagct 6540
gcgaaagcgc gtatggattt ggctcatcag gccagattga gggtctgtgg acacatgtca 6600
tgttagtgta cttcaatcgc ccctggata tagccccgac aataggccgt ggcctcattt 6660

ttttgccttc cgcacatttc cattgctcga taccacacc ttgcttctcc tgcacttgcc 6720
aaccttaata ctggtttaca ttgaccaaca tcttacaagc ggggggcttg tctagggtat 6780
atataaacag tggtctctcc aatcggttgc cagtctcttt tttctttct tccccacag 6840
attcgaaatc taaactacac atcacagaat tccgagccgt gagtatccac gacaagatca 6900
gtgtcgagac gacgcgtttt gtgtaatgac acaatccgaa agtcgctagc aacacacact 6960
ctctacaaa actaaccacg ctctgggtacc atggaggtcg tgaacgaaat cgtctccatt 7020
ggccaggagg ttcttcccaa ggtcgactat gctcagctct ggtctgatgc ctgcactgc 7080

gaggtgctgt acttctccat cgcttctgc atcctgaagt tcacccttgg tctctcggga 7140
cccaagggtc agtctcgaat gaagtttgtg ttaccaact acaacctgct catgtccatc 7200
tactcgtgg gctccttctt ctctatggcc tacgcatgt acaccattgg tgtcatgtcc 7260
gacaactcgc agaaggttt cgacaacaat gtcttccgaa tcaccactca gctgttctac 7320
ctcagcaagt tctcagagta cattgactcc ttctatctgc cctcatggg caagcctctg 7380
acctggttgc agttctttca ccatctcgga gctcctatgg acatgtggct gtcttacaac 7440
taccgaaacg aagccgtttg gatctttgtg ctgctcaacg gcttcattca ctggatcatg 7500

tacggctact attggacccg actgatcaag ctcaagttcc ctatgcccc gtcctgatt 7560
acttctatgc agatcattca gtccaactgt ggcttctaca tctctggaa gtaccggaac 7620
attccctgct accgacaaga tgggatgaga atgtttggct gggttttcaa ctacttctac 7680
gttgggtactg tctgtgtct gtctctcaac ttctacgtgc agacctacat cgtccgaaag 7740
cacaaggag ccaaaaagat tcagtgagc 7769

<210> 121

<211> 777

<212> DNA

<213> *Euglena gracilis*

<220><221> CDS

<222> (1)..(777)

<223> mutant delta-9 elongase "EgD9eS-K58R/I257T"

<400> 121

atg gag gtc gtg aac gaa atc gtc tcc att ggc cag gag gtt ctt ccc 48

Met Glu Val Val Asn Glu Ile Val Ser Ile Gly Gln Glu Val Leu Pro

1 5 10 15

aag gtc gac tat gct cag ctc tgg tct gat gcc tgc cac tgc gag gtg 96

Lys Val Asp Tyr Ala Gln Leu Trp Ser Asp Ala Ser His Cys Glu Val

20 25 30

ctg tac ctc tcc atc gcc ttc gtc atc ctg aag ttc acc ctt ggt cct 144

Leu Tyr Leu Ser Ile Ala Phe Val Ile Leu Lys Phe Thr Leu Gly Pro

35 40 45

ctc gga ccc aag ggt cag tct cga atg agg ttt gtg ttc acc aac tac 192

Leu Gly Pro Lys Gly Gln Ser Arg Met Arg Phe Val Phe Thr Asn Tyr

50 55 60

aac ctg ctc atg tcc atc tac tgc ctg ggc tcc ttc ctc tgc atg gcc 240

Asn Leu Leu Met Ser Ile Tyr Ser Leu Gly Ser Phe Leu Ser Met Ala

65 70 75 80

tac gcc atg tac acc att ggt gtc atg tcc gac aac tgc gag aag gct 288

Tyr Ala Met Tyr Thr Ile Gly Val Met Ser Asp Asn Cys Glu Lys Ala

85 90 95

ttc gac aac aat gtc ttc cga atc acc act cag ctg ttc tac ctc agc 336

Phe Asp Asn Asn Val Phe Arg Ile Thr Thr Gln Leu Phe Tyr Leu Ser

100 105 110

aag ttc ctc gag tac att gac tcc ttc tat ctg ccc ctc atg ggc aag 384

Lys Phe Leu Glu Tyr Ile Asp Ser Phe Tyr Leu Pro Leu Met Gly Lys

115 120 125

cct ctg acc tgg ttg caa ttc ttt cac cat ctc gga gct cct atg gac 432
Pro Leu Thr Trp Leu Gln Phe Phe His His Leu Gly Ala Pro Met Asp
130 135 140
atg tgg ctg ttc tac aac tac cga aac gaa gcc gtt tgg atc ttt gtg 480
Met Trp Leu Phe Tyr Asn Tyr Arg Asn Glu Ala Val Trp Ile Phe Val
145 150 155 160
ctg ctc aac ggc ttc att cac tgg atc atg tac ggc tac tat tgg acc 528
Leu Leu Asn Gly Phe Ile His Trp Ile Met Tyr Gly Tyr Tyr Trp Thr
165 170 175
cga ctg atc aag ctc aag ttc cct atg ccc aag tcc ctg att act tct 576
Arg Leu Ile Lys Leu Lys Phe Pro Met Pro Lys Ser Leu Ile Thr Ser
180 185 190
atg cag atc att cag ttc aac gtt ggc ttc tac atc gtc tgg aag tac 624
Met Gln Ile Ile Gln Phe Asn Val Gly Phe Tyr Ile Val Trp Lys Tyr
195 200 205
cgg aac att ccc tgc tac cga caa gat gga atg aga atg ttt ggc tgg 672
Arg Asn Ile Pro Cys Tyr Arg Gln Asp Gly Met Arg Met Phe Gly Trp
210 215 220
ttt ttc aac tac ttc tac gtt ggt act gtc ctg tgt ctg ttc ctc aac 720
Phe Phe Asn Tyr Phe Tyr Val Gly Thr Val Leu Cys Leu Phe Leu Asn
225 230 235 240
ttc tac gtg cag acc tac atc gtc cga aag cac aag gga gcc aaa aag 768
Phe Tyr Val Gln Thr Tyr Ile Val Arg Lys His Lys Gly Ala Lys Lys
245 250 255
act cag tga 777

Thr Gln

<210> 122

<211> 258

<212> PRT

<213> Euglena gracilis

<400> 122

Met Glu Val Val Asn Glu Ile Val Ser Ile Gly Gln Glu Val Leu Pro

1 5 10 15

Lys Val Asp Tyr Ala Gln Leu Trp Ser Asp Ala Ser His Cys Glu Val

20 25 30

Leu Tyr Leu Ser Ile Ala Phe Val Ile Leu Lys Phe Thr Leu Gly Pro

35 40 45

Leu Gly Pro Lys Gly Gln Ser Arg Met Arg Phe Val Phe Thr Asn Tyr

50 55 60

Asn Leu Leu Met Ser Ile Tyr Ser Leu Gly Ser Phe Leu Ser Met Ala

65 70 75 80

Tyr Ala Met Tyr Thr Ile Gly Val Met Ser Asp Asn Cys Glu Lys Ala

85 90 95

Phe Asp Asn Asn Val Phe Arg Ile Thr Thr Gln Leu Phe Tyr Leu Ser

100 105 110

Lys Phe Leu Glu Tyr Ile Asp Ser Phe Tyr Leu Pro Leu Met Gly Lys

115 120 125

Pro Leu Thr Trp Leu Gln Phe Phe His His Leu Gly Ala Pro Met Asp

130 135 140

Met Trp Leu Phe Tyr Asn Tyr Arg Asn Glu Ala Val Trp Ile Phe Val

145 150 155 160

Leu Leu Asn Gly Phe Ile His Trp Ile Met Tyr Gly Tyr Tyr Trp Thr

165 170 175

Arg Leu Ile Lys Leu Lys Phe Pro Met Pro Lys Ser Leu Ile Thr Ser

180 185 190

Met Gln Ile Ile Gln Phe Asn Val Gly Phe Tyr Ile Val Trp Lys Tyr

195 200 205

Arg Asn Ile Pro Cys Tyr Arg Gln Asp Gly Met Arg Met Phe Gly Trp

210 215 220

Phe Phe Asn Tyr Phe Tyr Val Gly Thr Val Leu Cys Leu Phe Leu Asn

225 230 235 240

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

245

250

255

Thr Gln

<210> 123

<211> 7769

<212> DNA

<213> Artificial Sequence

<220><223> Plasmid pZuFmEgD9eS-K58R/I257T

<400> 123

ggccgcaagt gtggatgggg aagtgagtgc cggtttctgt gtgcacaatt ggcaatccaa 60
gatggatgga ttcaacacag ggatatagcg agctacgtgg tgggtcgagg atatagcaac 120
ggatatattat gtttgacact tgagaatgta cgatacaagc actgtccaag tacaatacta 180

aacatactgt acatactcat actcgtaccc gggcaacggt ttactttgag tgcagtggct 240
agtgtcttta ctctacagt gtgcaatact gcgtatcata gtctttgatg tatatcgtat 300
tcattcatgt tagttgcgta cgagccggaa gcataaagtg taaagcctgg ggtgcctaata 360
gagttagcta actcacatta attgcgttgc gctcactgcc cgctttccag tcgggaaacc 420
tgtcgtgcc gctgcattaa tgaatcggcc aacgcgcggg gagaggcgggt ttgcgtattg 480
ggcgtcttct cgcttctctg ctactgact cgctgcgtc ggtcgttcgg ctgcgcgag 540
cggtatcagc tcaactcaag gcggttaatac ggttatccac agaatcaggg gataacgcag 600

gaaagaacat gtgagcaaaa ggccagcaaa aggccaggaa ccgtaaaaag gccgcgttgc 660
tggcgttttt ccataggtct cgccccctg acgagcatca caaaaatcga cgctcaagtc 720
agaggtaggc aaacccgaca ggactataaa gataccaggc gtttccccct ggaagctccc 780
tcgtgcgtc tctgtttccg acctgccgc ttaccggata cctgtccgcc tttctcctt 840
cgggaagcgt ggcgctttct catagctcac gctgtaggta tctcagttcg gtgtaggtcg 900
ttcgtccaa gctgggctgt gtgcacgaac cccccgtca gcccgaccgc tgcgccttat 960
ccggttaacta tcgtcttgag tccaacccgg taagacacga cttatcgcca ctggcagcag 1020

ccactggtaa caggattagc agagcgaggt atgtaggcgg tgctacagag ttcttgaagt 1080
ggtggcctaa ctacggctac actagaagga cagtatttgg tatctgcgt ctgctgaagc 1140
cagttacctt cgaaaaaaga gttgtagct cttgatccgg caaacaacc accgctggtta 1200
gcggtggttt ttttgtttgc aagcagcaga ttacgcgcag aaaaaagga tctcaagaag 1260
atcctttgat cttttctacg gggcttgacg ctactggaa cgaaaactca cgtaaggga 1320

ttttggtcat gagattatca aaaaggatct tcacctagat ccttttaaat taaaaatgaa	1380
gttttaaatc aatctaaagt atatatgagt aaacttggtc tgacagttac caatgcttaa	1440
tcagtgagge acctatctca gcgatctgtc tatttcgttc atccatagtt gcctgactcc	1500
ccgtcgtgta gataactacg atacgggagg gcttaccatc tggccccagt gctgcaatga	1560
taccgcgaga cccacgtca cgggtccag atttatcagc aataaaccag ccagccggaa	1620
gggccgagcg cagaagtggc cctgcaactt tatccgcctc catccagtct attaatgtt	1680
gccgggaagc tagagtaagt agttcgccag ttaatagttt gcgcaacgtt gttgccattg	1740
ctacaggcat cgtgggtgtc cgctcgtcgt ttggtatggc ttcattcagc tccggttccc	1800
aacgatcaag gcgagttaca tgatcccca tgttgtgcaa aaaagcgggt agctccttcg	1860
gtcctccgat cgttgtcaga agtaagttgg ccgcagtgtt atcactcatg gttatggcag	1920
cactgcataa ttctcttact gtcatgccat ccgtaagatg cttttctgtg actggtgagt	1980
actcaaccaa gtcatctga gaatagtga tgcggcgacc gagttgctct tccccggcgt	2040
caatacggga taataccgag ccacatagca gaactttaaa agtgctcatc attgaaaaac	2100
gttcttcggg gcgaaaactc tcaaggatct taccgctgtt gagatccagt tcgatgtaac	2160
ccactcgtgc acccaactga tcttcagcat cttttacttt caccagcgtt tctgggtgag	2220
caaaaacagg aaggcaaaat gccgcaaaa agggaataag ggcgacacgg aaatgttgaa	2280
tactcatact cticcttttt caatattatt gaagcattta tcagggttat tgtctcatga	2340
gcggatacat atttgaatgt atttagaaaa ataaacaaat aggggttccg cgcacatttc	2400
ccgaaaagt gccacctgac gcgcctgtga gcggcgcatc aagcgcgcg ggtgtggtgg	2460
ttacgcgag cgtgaccgt acacttgcca gcgccctagc gcccgctcct ttcgctttct	2520
tcccttctt tctcgccagc ttccgaggct ttccccgtca agctctaaat cgggggctcc	2580
ctttagggtt ccgatttagt gctttacggc acctcgaccc caaaaaactt gattagggtg	2640
atggttcacg tagtgggcca tcgccctgat agacggtttt tcgccctttg acgttggagt	2700
ccacgttctt taatagtga ctcttggttc aaactggaac aacctcaac cctatctcgg	2760
tctattcttt tgatttataa gggattttgc cgatttcggc ctattggtta aaaaatgagc	2820
tgatttaaca aaaatttaac gcgaatttta acaaaatatt aacgcttaca atttccattc	2880
gccattcagg ctgcgaact gttgggaagg gcgatcggtg cgggcctctt cgctattacg	2940
ccagctggcg aaagggggat gtgctgcaag gcgattaagt tgggtaacgc cagggttttc	3000
ccagtcacga cgttgtaaaa cgacggccag tgaattgtaa tacgactcac tataggcgca	3060

attgggtacc gggcccccc tcgaggtcga tgggtgcgat aagcttgata tcgaattcat	3120
gtcacacaaa ccgatcttcg cctcaaggaa acctaatctt acatccgaga gactgccgag	3180
atccagtcta cactgattaa ttttcgggcc aataatttaa aaaaatcgtg ttatataata	3240
ttatatgtat tatatatata catcatgatg atactgacag tcatgtccca ttgctaaata	3300
gacagactcc atctgccgcc tccaactgat gttctcaata ttttaaggggt catctcgcat	3360
tgtttaataa taaacagact ccatctaccg cctccaaatg atgttctcaa aatataattgt	3420
atgaacttat ttttattact tagtattatt agacaactta cttgctttat gaaaaaact	3480
tcctatttag gaaacaattt ataatggcag ttcgttcatt taacaattta tgtagaataa	3540
atgttataaa tgcgtatggg aaatcttaaa tatggatagc ataatgata tctgcattgc	3600
ctaattcgaa atcaacagca acgaaaaaaaa tcccttgtac aacataaata gtcacgaga	3660
aatatcaact atcaaagaac agctattcac acgttactat tgagattatt attggacgag	3720
aatcacacac tcaactgtct ttctctcttc tagaaatata ggtacaagta tgtactattc	3780
tcattgttca tacttctagt catttcatcc cacatatcc tttgatttct ctccaatgaa	3840
tgacattcta tcttgcaaat tcaacaatta taataagata taccaaagta gcggtatagt	3900
ggcaatcaaa aagcttctct ggtgtgcttc tcgtatttat ttttattcta atgatccatt	3960
aaaggtatat atttatttct tgttatataa tctttttgtt tattacatgg gctggataca	4020
taaaggtatt ttgattfaat tttttgctta aattcaatcc cccctcgctc agtgtcaact	4080
glaatggtag gaaattacca tacttttgaa gaagcaaaaa aaatgaaaga aaaaaaaaaat	4140
cgtatttcca ggtagacgt tccgcagaat ctagaatgcg gtatgcggta cattgttctt	4200
cgaacgtaaa agttgcgctc cctgagatat tglacatttt tgcttttaca agtacaagta	4260
catcgtacaa ctatgtacta ctgttgatgc atccacaaca gtttgttttg ttttttttg	4320
tttttttttt ttctaattgat tcattaccgc tatgtatacc tacttgtact tgtagtaagc	4380
cgggttattg gcgttcaatt aatcatagac ttatgaatct gcacggtgtg cgctgcgagt	4440
tacttttagc ttatgcatgc tacttgggtg taatatggg atctgttcgg aatcaacgg	4500
atgtcaatc gatttcgaca gtaattaatt aagtcataca caagtcagct ttcttcgagc	4560
ctcatataag tataagtagt tcaacgtatt agcactgtac ccagcatctc cgtatcgaga	4620
aacacaacaa catgccccat tggacagatc atgcggatac acaggttgtg cagtatcata	4680
catactcgat cagacaggtc gtctgacat catacaagct gaacaagcgc tccatacttg	4740
cacgtctct atatacacag ttaaattaca tatccatagt ctaacctcta acagttaatc	4800

ttctggtaag cctcccagcc agccttcttg tatcgcttgg cctcctcaat aggatctcgg	4860
ttctggccgt acagacctcg gccacaatt atgatatccg ttccggtaga catgacatcc	4920
tcaacagttc ggtactgctg tccgagagcg tctcccttgt cgtcaagacc caccgccggg	4980
gtcagaataa gccagtcttc agagtcgccc ttaggtcggg tctgggcaat gaagccaacc	5040
acaaactcgg ggtcggatcg ggcaagctca atggctctgt tggagtactc gccagtggcc	5100
agagagccct tgcaagacag ctcgccagc atgagcagac ctctggccag ctctcgttg	5160
ggagagggga ctaggaactc cttgtactgg gaggttctgt agtcagagac gtcctccttc	5220
ttctgttcag agacagtttc ctcgccacca gctcgcaggc cagcaatgat tccggttccg	5280
ggtagaccgt gggcgttggg gatatcggac cactcggcga ttccgtgaca ccggtactgg	5340
tgcttgacag tgttgccaat atctgcgaac ttctgtcct cgaacaggaa gaaaccgtgc	5400
ttaagagcaa gttccttagg ggggagcaca gtgccggcgt aggtgaagtc gtcaatgatg	5460
tcgatatggg ttttgatcat gcacacataa ggtccgacct tatcggaag ctcaatgagc	5520
tccttggtagg tggtaacatc cagagaagca cacaggttgg ttttcttggc tgccacgagc	5580
ttgagcactc gagcggcaaa ggccgacttg tggacgttag ctgagcttc gtaggagggc	5640
atcttggtagg tgaagaggag actgaataa atttagtctg cagaactttt tatcggaacc	5700
ttatctgggg cagtgaagta tatgttatgg taatagtac gaggtagttg aacttataga	5760
tagactggac tatacggcta tcggtccaaa ttgaaagaa cgtcaatggc tctctgggcg	5820
tcgcctttgc cgacaaaaat gtgatcatga tgaaagccag caatgacgtt gcagctgata	5880
ttgttgcagg ccaaccgcgc cgaaaacgca gctgtcagac ccacagcctc caacgaagaa	5940
tgtatcgtca aagtgatcca agcacactca tagttggagt cgtactccaa aggcggcaat	6000
gacgagtcag acagatactc gtcgacgttt aaacagtgtg cgcagatcta ctatagagga	6060
acatttaaat tgccccggag aagacggcca ggccgcctag atgacaaatt caacaactca	6120
cagctgactt tctgccattg ccactagggg ggggcctttt tataatggcca agccaagctc	6180
tccacgtcgg ttgggctgca cccaacaata aatgggtagg gttgcaccaa caaagggatg	6240
ggatgggggg tagaagatag gaggataacg gggctcaatg gcacaaataa gaacgaatac	6300
tgccattaag actcgtgatc cagcgactga caccattgca tcatctaagg gcctcaaac	6360
tacctcgaa ctgctgcgt gatctggaca ccacagaggt tccgagcact ttaggttgca	6420
ccaaatgtcc caccaggctc aggcagaaaa cgttgaaca gcgtgtacag tttgtcttaa	6480
caaaaagtga gggcgtgag gtcgagcagg gtggtgtgac ttgttatagc ctttagagct	6540
gcgaaagcgc gtatggattt ggctcatcag gccagattga gggctctgtg acacatgtca	6600
tgtagtgta ctcaatcgc cccctggata tagccccgac aataggccgt ggcctcattt	6660

ttttgccttc cgcacatttc cattgctcga taccacacc ttgcttctcc tgcacttgcc 6720
 aaccttaata ctggtttaca ttgaccaaca tcttacaagc ggggggcttg tctagggtat 6780
 atataaacag tggtctctcc aatcggttgc cagtctcttt tttcctttct tteccacag 6840
 attcgaaatc taaactacac atcacagaat tccgagccgt gagtatccac gacaagatca 6900

 gtgtcgagac gacgcgtttt gtgtaatgac acaatccgaa agtcgctagc aacacacact 6960
 ctctacaaa actaaccag ctctggtacc atggaggtcg tgaacgaaat cgtctccatt 7020
 ggccaggagg ttcttccaa ggtcgactat gctcagctct ggtctgatgc ctgcactgc 7080
 gaggtgctgt acctctccat cgccttcgtc atcctgaagt tcacccttgg tctctcga 7140
 cccaagggtc agtctcgaat gaggtttgtg ttcaccaact acaacctgct catgtccatc 7200
 tactcgttgg gctccttct ctcatggcc tacgcatgt acaccattgg tgtcatgtcc 7260
 gacaactgcg agaaggcttt cgacaacaat gtcttccgaa tcaccactca gctgttctac 7320

 ctacgaagt tctcgagta cattgactcc ttctatctgc cctcatggg caagcctctg 7380
 acctggttgc aattctttca ccatctcgga gctctatgg acatgtggct gttctacaac 7440
 tacgaaacg aagccgtttg gatctttgtg ctgctcaacg gcttcattca ctggatcatg 7500
 tacggctact attggaccg actgatcaag ctcaagttcc ctatgcccc gtcctgatt 7560
 acttctatgc agatcattca gttcaacgtt ggcttctaca tcgtctggaa gtaccggaac 7620
 attcctgct accgacaaga tggaatgaga atgtttggct ggtttttcaa ctacttctac 7680
 gttggtactg tctgtgtct gttctcaac ttctacgtgc agacctacat cgtccgaaag 7740

 cacaaggag ccaaaaagac tcagtgagc 7769

 <210> 124
 <211> 777
 <212> DNA
 <213> Euglena gracilis
 <220><221> CDS
 <222> (1)..(777)
 <223> mutant delta-9 elongase "EgD9eS-L130M/V243A1"
 <400> 124

 atg gag gtc gtg aac gaa atc gtc tcc att ggc cag gag gtt ctt ccc 48
 Met Glu Val Val Asn Glu Ile Val Ser Ile Gly Gln Glu Val Leu Pro
 1 5 10 15
 aag gtc gac tat gct cag ctc tgg tct gat gcc tcg cac tgc gag gtg 96

Lys Val Asp Tyr Ala Gln Leu Trp Ser Asp Ala Ser His Cys Glu Val	
20 25 30	
ctg tac ctc tcc atc gcc ttc gtc atc ctg aag ttc acc ctt ggt cct	144
Leu Tyr Leu Ser Ile Ala Phe Val Ile Leu Lys Phe Thr Leu Gly Pro	
35 40 45	
ctc gga ccc aag ggt cag tct cga atg aag ttt gtg ttc acc aac tac	192
Leu Gly Pro Lys Gly Gln Ser Arg Met Lys Phe Val Phe Thr Asn Tyr	
50 55 60	
aac ctg ctc atg tcc atc tac tcg ctg ggc tcc ttc ctc tct atg gcc	240
Asn Leu Leu Met Ser Ile Tyr Ser Leu Gly Ser Phe Leu Ser Met Ala	
65 70 75 80	
tac gcc atg tac acc att ggt gtc atg tcc gac aac tgc gag aag gct	288
Tyr Ala Met Tyr Thr Ile Gly Val Met Ser Asp Asn Cys Glu Lys Ala	
85 90 95	
ttc gac aac aat gtc ttc cga atc acc act cag ctg ttc tac ctc agc	336
Phe Asp Asn Asn Val Phe Arg Ile Thr Thr Gln Leu Phe Tyr Leu Ser	
100 105 110	
aag ttc ctc gag tac att gac tcc ttc tat ctg ccc ctc atg ggc aag	384
Lys Phe Leu Glu Tyr Ile Asp Ser Phe Tyr Leu Pro Leu Met Gly Lys	
115 120 125	
cct atg acc tgg ttg cag ttc ttt cac cat ctc gga gct cct atg gac	432
Pro Met Thr Trp Leu Gln Phe Phe His His Leu Gly Ala Pro Met Asp	
130 135 140	
atg tgg ctg ttc tac aat tac cga aac gaa gcc gtt tgg atc ttt gtg	480
Met Trp Leu Phe Tyr Asn Tyr Arg Asn Glu Ala Val Trp Ile Phe Val	
145 150 155 160	
ctg ctc aac ggc ttc att cac tgg atc atg tac ggc tac tat tgg acc	528
Leu Leu Asn Gly Phe Ile His Trp Ile Met Tyr Gly Tyr Tyr Trp Thr	
165 170 175	
cga ctg atc aag ctc aag ttc cct atg ccc aag tcc ctg att act tct	576
Arg Leu Ile Lys Leu Lys Phe Pro Met Pro Lys Ser Leu Ile Thr Ser	

180 185 190
atg cag atc att cag ttc aac gtt ggc ttc tac atc gtc tgg aag tac 624
Met Gln Ile Ile Gln Phe Asn Val Gly Phe Tyr Ile Val Trp Lys Tyr

195 200 205
cgg aac att ccc tgc tac cga caa gat gga atg aga atg ttt ggc tgg 672
Arg Asn Ile Pro Cys Tyr Arg Gln Asp Gly Met Arg Met Phe Gly Trp

210 215 220
ttt ttc aac tac ttc tac gtt ggt act gtc ctg tgt ctg ttc ctc aac 720
Phe Phe Asn Tyr Phe Tyr Val Gly Thr Val Leu Cys Leu Phe Leu Asn

225 230 235 240
ttc tac gcg cag acc tac atc gtc cga aag cac aag gga gcc aaa aag 768

Phe Tyr Ala Gln Thr Tyr Ile Val Arg Lys His Lys Gly Ala Lys Lys
245 250 255

att cag tga 777
Ile Gln

<210> 125

<211> 258

<212> PRT

<213> Euglena gracilis

<400> 125

Met Glu Val Val Asn Glu Ile Val Ser Ile Gly Gln Glu Val Leu Pro

1 5 10 15
Lys Val Asp Tyr Ala Gln Leu Trp Ser Asp Ala Ser His Cys Glu Val

20 25 30
Leu Tyr Leu Ser Ile Ala Phe Val Ile Leu Lys Phe Thr Leu Gly Pro

35 40 45
Leu Gly Pro Lys Gly Gln Ser Arg Met Lys Phe Val Phe Thr Asn Tyr

50 55 60
Asn Leu Leu Met Ser Ile Tyr Ser Leu Gly Ser Phe Leu Ser Met Ala

65 70 75 80

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

<210> 126

<211> 7769

<212> DNA

<213> Artificial Sequence

<220><223> Plasmid pZuFmEgD9eS-L130M/V243A1

<400> 126

ggccgcaagt gtggatgggg aagtgagtgc ccggttctgt gtcacaatt ggcaatccaa 60

gatggatgga ttcaacacag ggatatagcg agctacgtgg tgggtgcgagg atatagcaac 120

ggataatttat gtttgacact tgagaatgia cgatacaagc actgtccaag tacaatacta 180
 aacatactgt acatactcat actcgtaccc gggcaacggt ttcacttgag tgcagtggct 240
 agtgctctta ctctacaggt gtgcaatact gcgtatcata gtctttgatg tataatcgat 300
 tcattcatgt tagttgcgta cgagccggaa gcataaagtg taaagcctgg ggtgcctaata 360

 gagtgagcta actcacatta attgcgttgc gctcactgcc cgctttccag tcgggaaacc 420
 tgtcgtgcca gctgcattaa tgaatcggcc aacgcgcggg gagaggcggg ttgcgtattg 480
 ggcgctcttc cgcttcctcg ctactgact cgctgcgctc ggtcgttcgg ctgcggcgag 540
 cggatcagc tactcaaac gcggaatac ggttatccac agaatacagg gataacgcag 600
 gaaagaacat gtgagcaaaa ggccagcaaa aggccaggaa ccgtaaaaag gccgcgttgc 660
 tggcggtttt ccataggctc cgccccctg acgagcatca caaaaatcga cgctcaagtc 720
 agaggtggcg aaaccgcaca ggactataaa gataccaggc gtttccccct ggaagctccc 780

 tcgtgcgctc tctgttccg accctgccgc ttaccggata cctgtccgc tttctccctt 840
 cgggaagcgt ggcgctttct catagctcac gctgtaggta tctcagttcg gtgtaggtcg 900
 ttgcgtccaa gctgggctgt gtgcacgaac ccccggtca gcccgaccgc tgcgccttat 960
 ccggttaacta tcgtcttgag tccaaccggg taagacacga cttatcgcca ctggcagcag 1020
 ccactggtaa caggattagc agagcgaggt atgtaggcgg tgctacagag ttcttgaagt 1080
 ggtggcctaa ctacggctac actagaagga cagtatttgg tatctgcgt ctgctgaagc 1140
 cagttacctt cgaaaaaga gttggtagct cttgatccgg caaacaacc accgctggta 1200

 gcggtgggtt tttgtttgc aagcagcaga ttacgcgcag aaaaaagga tctcaagaag 1260
 atcctttgat cttttctac gggctctgacg ctacgtggaa cgaaaactca cgtaaggga 1320
 ttttggtcat gagattatca aaaaggatct tcacctagat ctttttaaat taaaaatgaa 1380
 gttttaaatc aatctaaagt atatatgagt aaacttggc tgacagttac caatgcttaa 1440
 tcagttaggc acctatctca gcgatctgtc ttttctgtc atccatagtt gcctgactcc 1500
 ccgtcgtgta gataactacg atacgggagg gcttaccatc tggccccagt gctgcaatga 1560
 taccgcgaga ccacgctca ccggctccag atttatcagc aataaaccag ccagccggaa 1620

 gggccgagcg cagaagtggc cctgcaactt tatccgctc catccagtct attaattgtt 1680
 gccgggaagc tagagtaagt agttcgccag ttaatatgtt gcgcaacgtt gttgccattg 1740
 ctacaggcat cgtgggtgca cgctcgtcgt ttggtatggc ttcatcagc tccggttccc 1800
 aacgatcaag gcgagttaca tgatcccca tgttgtgcaa aaaagcggtt agctccttcg 1860
 gtctccgat cgttgtcaga agtaagtgg ccgcagtgtt atactcatg gttatggcag 1920
 cactgcataa ttctcttact gtcatccat ccgtaagatg cttttctgtg actggtgagt 1980

actcaaccaa gtcattctga gaatagtgia tgcggcgacc gagttgctct tgcccggcgt 2040

caatacggga taataccgcg ccacatagca gaactttaaa agtgctcatc attgaaaaac 2100

gttcttcggg gcgaaaactc tcaaggatct taccgctgtt gagatccagt tcgatgtaac 2160

ccactcgtgc acccaactga tcttcagcat cttttacttt caccagcgtt tctgggtgag 2220

caaaaacagg aaggcaaaat gccgcaaaaa agggaataag ggcgacacgg aaatgttgaa 2280

tactcatact ctctcttttt caatattatt gaagcattta tcagggttat tgtctcatga 2340

gcggatacat atttgaatgt atttagaaaa ataaacaaat aggggttccg cgcacatttc 2400

cccgaagaat gccacctgac gcgccctgta gcggcgcatc aagcgcggcg ggtgtggtgg 2460

ttagcgcag cgtgaccgct acacttgcca gcgccctagc gcccgctcct ttcgctttct 2520

tcccttctct tctcgccacg ttcgccggct ttccccgtca agctctaaat cgggggctcc 2580

ctttagggtt ccgatttagt gctttacggc acctcgaccc caaaaaactt gattagggtg 2640

atggttcacg tagtgggcca tcgccctgat agacggtttt tcgccctttg acgttggagt 2700

ccacgttctt taatagtga ctcttgttcc aaactggaac aacactcaac cctatctcgg 2760

tctattcttt tgatttataa gggattttgc cgatttcggc ctattggtta aaaaatgagc 2820

tgatttaaca aaaatttaac gcgaatttta acaaaatatt aacgcttaca atttcattc 2880

gccattcagg ctgcgcaact gttgggaagg gcgatcgggtg cgggcctctt cgctattacg 2940

ccagctggcg aaagggggat gtgctgcaag gcgattaagt tgggtaacgc cagggttttc 3000

ccagtcacga cgttgtaaaa cgacggccag tgaattgtaa tacgactcac tataggcgca 3060

attgggtacc gggccccccc tcgaggtcga tgggtgcgat aagcttgata tcgaattcat 3120

gtcacacaaa ccgatcttcg cctcaaggaa acctaatctt acatccgaga gactgccgag 3180

atccagtcta cactgattaa ttttcgggcc aataatttaa aaaaatcgtg ttatataata 3240

ttatatgtat tatatatata catcatgatg atactgacag tcatgtccca ttgctaaata 3300

gacagactcc atctgccgc tccaactgat gttctcaata ttttaagggt catctcgcat 3360

tgtttaataa taaacagact ccatctaccg cctccaaatg atgttctcaa aatatattgt 3420

atgaacttat ttttattact tagtattatt agacaactta cttgctttat gaaaaacact 3480

tcctatttag gaaacaattt ataatggcag ttcgttcatt taacaattta tgtagaataa 3540

atgttataaa tgcgtatggg aaatcttaaa tatggatagc ataaatgata tctgcattgc 3600

ctaattcgaa atcaacagca acgaaaaaaa tccttgttac aacataaata gtcacgaga 3660

aatatcaact atcaaagaac agctattcac acgttactat tgagattatt attggacgag 3720

aatcacacac tcaactgtct ttctctcttc tagaaatata ggtacaagta tgtactattc	3780
tcattgttca tactttctagt catttcatcc cacatatcc ttggatttct ctccaatgaa	3840
tgacattcta tcttgcaaat tcaacaatta taataagata taccaaagta gcggtatagt	3900
ggcaatcaaa aagcttctct ggtgtgcttc tcgtatttat ttttattcta atgatccatt	3960
aaaggtatat atttatttct tgttatataa tccttttgtt tattacatgg gctggataca	4020
taaaggtatt ttgatttaatt tttttgctta aattcaatcc cccctcgttc agtgtcaact	4080
gtaatggtag gaaattacca tacttttgaa gaagcaaaaa aaatgaaaga aaaaaaaaaat	4140
cgattttcca ggttagacgt tccgcagaat ctagaatgcg gtatgcggtta cattgttctt	4200
cgaacgtaaa agttgcgctc cctgagatat tglacatttt tgctttttaca agtacaagta	4260
catcgtacaa ctatgtacta ctgttgatgc atccacaaca gtttgttttg ttttttttg	4320
tttttttttt ttctaattgat tcattaccgc tatgtatacc tacttgtact tgtagtaagc	4380
cgggttattg gcgttcaatt aatcatagac ttatgaatct gcacgggtgtg cgctgcgagt	4440
tacttttagc ttatgcatgc tacttgggtg taatattggg atctgttcgg aaatcaacgg	4500
atgctcaatc gatttcgaca gtaattaatt aagtcataca caagtcagct ttcttcgagc	4560
ctcatataag tataagtagt tcaacgtatt agcactgtac ccagcatctc cgtatcgaga	4620
aacacaacaa catgccccat tggacagatc atgcggatac acaggttgtg cagtatcata	4680
catactcgat cacagaggtc gtctgacat catacaagct gaacaagcgc tccatacttg	4740
cacgtctct atatacacag ttaaattaca tatccatagt ctaacctcta acagttaatc	4800
ttctggtaag cctcccagcc agccttctgg tatcgcttgg cctcctcaat aggatctcgg	4860
ttctggccgt acagacctcg gccgacaatt atgatatccg ttccggtaga catgacatcc	4920
tcaacagttc ggtactgctg tccgagagcg tctcccttgt cgtcaagacc caccgccggg	4980
gtcagaataa gccagtctc agagtgcgcc ttaggtcggt tctgggcaat gaagccaacc	5040
acaaactcgg ggtcggatcg ggcaagctca atggtctgct tggagtactc gccagtggcc	5100
agagagccct tgcaagacag ctcgccagc atgagcagac ctctggccag ctctctgttg	5160
ggagagggga ctaggaaactc cttgtactgg gagttctcgt agtcagagac gtccctcttc	5220
ttctgttcag agacagtttc ctcgccacca gctcgcaggc cagcaatgat tccggttccg	5280
ggtacaccgt gggcgttggg gatatcgac cactcggcga ttccgtgaca ccggtactgg	5340
tgcttgacag tgttgccaat atctgcgaac tttctgtcct cgaacaggaa gaaaccgtgc	5400
ttaagagcaa gttccttgag ggggagcaca gtgccgcgt aggtgaagtc gtcaatgatg	5460
tcgatatggg ttttgatcat gcacacataa ggtccgacct tatcggaag ctcaatgagc	5520
tccttgggtg tggtaacatc cagagaagca cacaggttgg ttttcttggc tgccacgagc	5580

ttgagcactc gagcggcaaa ggcggacttg tggacgttag ctcgagcttc gtaggagggc 5640
 attttggtgg tgaagaggag actgaaataa atttagtctg cagaactttt tatcggaacc 5700
 ttatctgggg cagtgaagta tatgttatgg taatagttaac gagttagttag aacttataga 5760
 tagactggac tatacggcta tcggtccaaa ttagaaagaa cgtcaatggc tctctgggcg 5820

 tcgcctttgc cgacaaaaat gtgatcatga tgaaagccag caatgacgtt gcagctgata 5880
 ttgttgctgg ccaaccgcgc cgaaaacgca gctgtcagac ccacagcctc caacgaagaa 5940
 tgtatcgtca aagtgatcca agcacactca tagttggagt cgtactccaa aggcggcaat 6000
 gagcagtcag acagatactc gtcgacgttt aaacagtgtg cgcagatcta ctatagagga 6060
 acatttaaat tgccccggag aagacggcca ggccgcctag atgacaaatt caacaactca 6120
 cagctgactt tctgccattg ccactagggg ggggcctttt tataatggcca agccaagctc 6180
 tccacgtcgg ttgggctgca cccaacaata aatgggtagg gttgcaccaa caaagggatg 6240

 ggatgggggg tagaagatac gaggataacg gggctcaatg gcacaaataa gaacgaatac 6300
 tgccattaag actcgtgac cagcgactga caccattgca tcatctaagg gcctcaaaac 6360
 tacctcggaa ctgctgcgt gatctggaca ccacagaggt tccgagcact ttaggttgca 6420
 ccaaatgtcc caccaggctc aggcagaaaa cgctggaaca gcgtgtacag tttgtcttaa 6480
 caaaaagtga gggcgtgag gtcgagcagg gtggtgtgac ttgttatagc ctttagagct 6540
 gcgaaagcgc gtatggattt ggctcatcag gccagattga gggctctgtg acacatgtca 6600
 tgttagtgta ctcaatcgc cccctggata tagccccgac aataggccgt ggcctcattt 6660

 ttttgccttc cgcacatttc cattgctcga taccacacc ttgcttctcc tgcacttgcc 6720
 aaccttaata ctggtttaca ttgaccaaca tcttacaagc ggggggcttg tctagggtat 6780
 atataaacag tggtctctcc aatcggttgc cagtctcttt tttctttct tccccacag 6840
 attcgaatc taaactacac atcacagaat tccgagccgt gagtatccac gacaagatca 6900
 gtgtcgagac gacgcgtttt gtgtaatgac acaatccgaa agtcgctagc aacacacact 6960
 ctctacaaa actaaccag ctctggtagc atggaggtcg tgaacgaaat cgtctccatt 7020
 ggccaggagg ttcttcccaa ggtcgactat gctcagctct ggtctgatgc ctgcactgc 7080

 gaggtgctgt acctctccat cgccttcgtc atcctgaagt tcacccttgg tcctctcgga 7140
 cccaagggtc agtcctgaat gaagtttgtg ttaccaact acaacctgct catgtccatc 7200
 tactcgctgg gctccttct ctctatggcc tacgcatgt acaccattgg tgtcatgtcc 7260
 gacaactcgc agaaggcttt cgacaacaat gtcttccgaa tcaccactca gctgttctac 7320
 ctgagcaagt tctcgagta cattgactcc ttctatctgc cctcatggg caagcctatg 7380
 acctggttgc agttctttca ccatctcgga gctcctatgg acatgtggct gttctacaat 7440

taccgaaacg aagccgtttg gatctttgtg ctgctcaacg gcttcattca ctggatcatg 7500

tacggctact attggaccgc actgatcaag ctcaagttcc ctatgcccaa gtccttgatt 7560

acttctatgc agatcattca gttcaacgtt ggcttctaca tcgtctggaa gtaccggaac 7620

attccctgct accgacaaga tggaatgaga atgtttggct gggttttcaa ctacttctac 7680

gttgggtactg tcctgtgtct gttcctcaac ttctacgcgc agacctacat cgtccgaaag 7740

cacaaggag ccaaaaagat tcagtgage 7769

<210> 127

<211> 777

<212> DNA

<213> *Euglena gracilis*

<220><221> CDS

<222> (1)..(777)

<223> mutant delta-9 elongase "EgD9eS-D98G"

<400> 127

atg gag gtc gtg aac gaa atc gtc tcc att ggc cag gag gtt ctt ccc 48

Met Glu Val Val Asn Glu Ile Val Ser Ile Gly Gln Glu Val Leu Pro

1 5 10 15

aag gtc gac tat gct cag ctc tgg tct gat gcc tcg cac tgc gag gtg 96

Lys Val Asp Tyr Ala Gln Leu Trp Ser Asp Ala Ser His Cys Glu Val

20 25 30

ctg tac ctc tcc atc gcc ttc gtc atc ctg aag ttc acc ctt ggt cct 144

Leu Tyr Leu Ser Ile Ala Phe Val Ile Leu Lys Phe Thr Leu Gly Pro

35 40 45

ctc gga ccc aag ggt cag tct cga atg aag ttt gtg ttc acc aac tac 192

Leu Gly Pro Lys Gly Gln Ser Arg Met Lys Phe Val Phe Thr Asn Tyr

50 55 60

aac ctg ctc atg tcc atc tac tcg ctg ggc tcc ttc ctc tct atg gcc 240

Asn Leu Leu Met Ser Ile Tyr Ser Leu Gly Ser Phe Leu Ser Met Ala

65 70 75 80

tac gcc atg tac acc att ggt gtc atg tcc gac aac tgc gag aag gct 288

Tyr Ala Met Tyr Thr Ile Gly Val Met Ser Asp Asn Cys Glu Lys Ala

85	90	95	
ttc ggc aac aat gtc ttc cga atc acc act cag ctg ttc tac ctc agc			336
Phe Gly Asn Asn Val Phe Arg Ile Thr Thr Gln Leu Phe Tyr Leu Ser			
100	105	110	
aag ttc ctc gag tac att gac tcc ttc tat ctg ccc ctc atg ggc aag			384
Lys Phe Leu Glu Tyr Ile Asp Ser Phe Tyr Leu Pro Leu Met Gly Lys			
115	120	125	
cct ctg acc tgg ttg cag ttc ttt cac cat ctc gga gct cct atg gac			432
Pro Leu Thr Trp Leu Gln Phe Phe His His Leu Gly Ala Pro Met Asp			
130	135	140	
atg tgg ctg ttc tac aac tac cga aac gaa gcc gtt tgg atc ttt gtg			480
Met Trp Leu Phe Tyr Asn Tyr Arg Asn Glu Ala Val Trp Ile Phe Val			
145	150	155	160
ctg ctc aac ggc ttc att cac tgg atc atg tac ggc tac tat tgg acc			528
Leu Leu Asn Gly Phe Ile His Trp Ile Met Tyr Gly Tyr Tyr Trp Thr			
165	170	175	
cga ctg atc aag ctc aag ttc cct atg ccc aag tcc ctg att act tct			576
Arg Leu Ile Lys Leu Lys Phe Pro Met Pro Lys Ser Leu Ile Thr Ser			
180	185	190	
atg cag atc att cag ttc aac gtt ggc ttc tac atc gtc tgg aag tac			624
Met Gln Ile Ile Gln Phe Asn Val Gly Phe Tyr Ile Val Trp Lys Tyr			
195	200	205	
cgg aac att ccc tgc tac cga caa gat gga atg aga atg ttt ggc tgg			672
Arg Asn Ile Pro Cys Tyr Arg Gln Asp Gly Met Arg Met Phe Gly Trp			
210	215	220	
ttt ttc aac tac ttc tac gtt ggt act gtc ctg tgt ctg ttc ctc aac			720
Phe Phe Asn Tyr Phe Tyr Val Gly Thr Val Leu Cys Leu Phe Leu Asn			
225	230	235	240
ttc tac gtg cag acc tac atc gtc cga aag cac aag gga gcc aaa aag			768
Phe Tyr Val Gln Thr Tyr Ile Val Arg Lys His Lys Gly Ala Lys Lys			
245	250	255	

att cag tga

777

Ile Gln

<210> 128

<211> 258

<212> PRT

<213> Euglena gracilis

<400> 128

Met Glu Val Val Asn Glu Ile Val Ser Ile Gly Gln Glu Val Leu Pro

1 5 10 15

Lys Val Asp Tyr Ala Gln Leu Trp Ser Asp Ala Ser His Cys Glu Val

20 25 30

Leu Tyr Leu Ser Ile Ala Phe Val Ile Leu Lys Phe Thr Leu Gly Pro

35 40 45

Leu Gly Pro Lys Gly Gln Ser Arg Met Lys Phe Val Phe Thr Asn Tyr

50 55 60

Asn Leu Leu Met Ser Ile Tyr Ser Leu Gly Ser Phe Leu Ser Met Ala

65 70 75 80

Tyr Ala Met Tyr Thr Ile Gly Val Met Ser Asp Asn Cys Glu Lys Ala

85 90 95

Phe Gly Asn Asn Val Phe Arg Ile Thr Thr Gln Leu Phe Tyr Leu Ser

100 105 110

Lys Phe Leu Glu Tyr Ile Asp Ser Phe Tyr Leu Pro Leu Met Gly Lys

115 120 125

Pro Leu Thr Trp Leu Gln Phe Phe His His Leu Gly Ala Pro Met Asp

130 135 140

Met Trp Leu Phe Tyr Asn Tyr Arg Asn Glu Ala Val Trp Ile Phe Val

145 150 155 160

Leu Leu Asn Gly Phe Ile His Trp Ile Met Tyr Gly Tyr Tyr Trp Thr

165 170 175

Arg Leu Ile Lys Leu Lys Phe Pro Met Pro Lys Ser Leu Ile Thr Ser

180 185 190
Met Gln Ile Ile Gln Phe Asn Val Gly Phe Tyr Ile Val Trp Lys Tyr
195 200 205
Arg Asn Ile Pro Cys Tyr Arg Gln Asp Gly Met Arg Met Phe Gly Trp
210 215 220
Phe Phe Asn Tyr Phe Tyr Val Gly Thr Val Leu Cys Leu Phe Leu Asn

225 230 235 240
Phe Tyr Val Gln Thr Tyr Ile Val Arg Lys His Lys Gly Ala Lys Lys
245 250 255
Ile Gln

<210> 129

<211> 7769

<212> DNA

<213> Artificial Sequence

<220><223> Plasmid pZuFmEgD9eS-D98G

<400> 129

ggccgcaagt gtggatgggg aagtgagtgc ccggttctgt gtgcacaatt ggcaatccaa 60
gatggatgga ttcaacacag ggatatagcg agctacgtgg tggatgcaggg atatacgcaac 120
ggatatattat gtttgacact tgagaatgta cgatacaagc actgtccaag tacaatacta 180

aacatactgt acatactcat actcgtaccc gggcaacggt ttcacttgag tgcagtggct 240
agtgtcttta ctctacagt gtgcaatact gcgtatcata gtctttgatg tatatcgtat 300
tcattcatgt tagttgcgta cgagccggaa gcataaagtg taaagcctgg ggtgcctaatt 360
gagtgcgcta actcacatta attgcgttgc gctcactgcc cgctttccag tcgggaaacc 420
tgtcgtgcca gctgcattaa tgaatcgccc aacgcgcggg gagaggcggt ttgcgtattg 480
ggcgcctcttc cgcttcctcg ctactgact cgctgcgctc ggtcgttcgg ctgcggcgag 540
cggatcagcgc tcaactcaag gcggtataac ggttatccac agaatacagg gataacgcag 600

gaaagaacat gtgagcaaaa ggccagcaaa aggccaggaa ccgtaaaaag gccgcgttgc 660
tggcggtttt ccataggctc cgccccctg acgagcatca caaaaatcga cgctcaagtc 720
agaggtggcg aaacccgaca ggactataaa gataccaggc gtttccccct ggaagctccc 780
tcgtgcgctc tctgtttccg accctgccgc ttaccggata cctgtccgcc tttctccctt 840
cgggaagcgt ggcgctttct catagctcac gctgtaggta tctcagttcg gtgtaggtcg 900

ttcgctccaa gctgggctgt gtgcacgaac cccccgttca gcccgaccgc tgcgccttat	960
ccggtaaacta tcgtcttgag tccaacccgg taagacacga cttatcgcca ctggcagcag	1020
ccactggtaa caggattagc agagcgaggt atgtaggcgg tgctacagag ttcttgaagt	1080
ggtggcctaa ctacggctac actagaagga cagtatttgg tatctgcgct ctgctgaagc	1140
cagttacctt cggaaaaaga gttggtagct cttgatccgg caaacaacc accgctggta	1200
gcggtggttt ttttgtttgc aagcagcaga ttacgcgcag aaaaaagga tctcaagaag	1260
atcctttgat cttttctacg gggctctgacg ctacagtggaa cgaaaactca cgtaaggga	1320
ttttggtcac gagattatca aaaaggatct tcacctagat ccttttaaat taaaaatgaa	1380
gttttaaatc aatctaaagt atatatgagt aaacttggtc tgacagttac caatgcttaa	1440
tcagtgaggc acctatctca gcgatctgtc tatttcgttc atccatagtt gcctgactcc	1500
ccgtcgtgta gataactacg atacgggagg gcttaccatc tggccccagt gctgcaatga	1560
taccgcgaga cccacgtca ccggctccag atttatcagc aataaaccag ccagccggaa	1620
gggccgagcg cagaagtgtt cctgcaactt tatccgcctc catccagtct attaatgtt	1680
gccgggaagc tagagtaagt agttcgccag ttaatatgtt gcgcaacgtt gttgccattg	1740
ctacaggcat cgtgggtgta cgctcgctgt ttggtatggc ttcatcagc tccggttccc	1800
aacgatcaag gcgagttaca tgatcccca tgttgtgcaa aaaagcggtt agctccttcg	1860
gtcctccgat cgttgtcaga agtaagttag ccgcagtggt atcactcatg gttatggcag	1920
cactgcataa ttctcttact gtcatgcat ccgtaagatg cttttctgtg actggtgagt	1980
actcaaccaa gtcattctga gaatagtgtg tgcggcgacc gagttgctct tgccccgct	2040
caatacggga taatacccg ccacatagca gaactttaaa agtgctcatc attggaaaac	2100
gttcttcggg gcgaaaactc tcaaggatct taccgtgtt gagatccagt tcgatgtaac	2160
ccactcgtgc acccaactga tcttcagcat cttttacttt caccagcgtt tctgggtgag	2220
caaaaacagg aaggcaaaat gccgcaaaa agggaataag ggcgacacgg aaatgttgaa	2280
tactcatact cttccttttt caatattatt gaagcattta tcagggttat tgtctcatga	2340
gcggatacat atttgaatgt atttagaaaa ataaacaaat aggggttccg cgcacatttc	2400
cccgaaaagt gccacctgac gcgccctgta gcggcgcat aagcgcggcg ggtgtggtgg	2460
ttacgcgcag cgtgaccgt acacttgcca gcgccctagc gcccgctctt ttcgctttct	2520
tcccttctt tctcgccacg ttcccggtt tccccgtca agctctaaat cgggggctcc	2580
ctttagggtt ccgatttagt gctttacggc acctcgacc caaaaaactt gattagggtg	2640

atggttcacg tagtgggcca tcgccctgat agacggtttt tcgccctttg acgttggagt	2700
ccacgttctt taatagtga cttttgttcc aaactggaac aacactcaac cctatctcgg	2760
tctattcttt tgatttataa gggattttgc cgatttcggc ctattggta aaaaatgagc	2820
tgatttaaca aaaatttaac gcgaatttta acaaaatatt aacgcttaca atttcattc	2880
gccattcagg ctgcgcaact gttgggaagg gcgacgggtg cgggcctctt cgctattacg	2940
ccagctggcg aaagggggat gtgctgcaag gcgattaagt tgggtaacgc cagggttttc	3000
ccagtcacga cgttgtaaaa cgacggccag tgaattgtaa tacgactcac tataggcgca	3060
attgggtacc gggccccccc tcgaggtcga tgggtgcgat aagcttgata tcgaattcat	3120
gtcacacaaa ccgatcttcg cctcaaggaa acctaattct acatccgaga gactgccgag	3180
atccagtcta cactgattaa ttttcgggcc aataatttaa aaaaatcgtg ttatataata	3240
ttatatgtat tatatatata catcatgatg atactgacag tcatgtccca ttgctaaata	3300
gacagactcc atctgccgcc tccaactgat gtcttcaata ttttaaggggt catctcgcat	3360
tgtttaataa taaacagact ccatctaccg cctccaaatg atgttctcaa aatatattgt	3420
atgaacttat ttttattact tagtattatt agacaactta cttgtcttat gaaaaacact	3480
tcctatttag gaaacaattt ataatggcag ttcgttcatt taacaattta tgtagaataa	3540
atgttataaa tgcgtatggg aaatcttaaa tatggatagc ataaatgata tctgcattgc	3600
ctaattcgaa atcaacagca acgaaaaaaaa tcctttgtac aacataaata gtcacgaga	3660
aatatcaact atcaaagaac agctattcac acgttactat tgagattatt attggacgag	3720
aatcacacac tcaactgtct ttctctcttc tagaaataca ggtacaagta tgtactattc	3780
tcattgttca tacttctagt catttcatcc cacatatcc ttggatttct ctccaatgaa	3840
tgacattcta tcttgcaaat tcaacaatta taataagata taccaaagta gcggtatagt	3900
ggcaatcaaa aagcttctct ggtgtgcttc tcgtatttat ttttattcta atgatccatt	3960
aaaggtatat atttatttct tgttatataa tccttttggt tattacatgg gctggataca	4020
taaaggtatt ttgatttaat tttttgctta aattcaatcc cccctcgttc agtgtcaact	4080
glaatggtag gaaattacca tacttttgaa gaagcaaaaa aaatgaaaga aaaaaaaaaat	4140
cgtatttcca ggtagacgt tccgcagaat ctagaatgcg gtatgcggtta cattgttctt	4200
cgaacgtaaa agttgcgtc cctgagatat tgfacatttt tgcttttaca agtacaagta	4260
catcgtacaa ctatgtacta ctgttgatgc atccacaaca gtttgttttg ttttttttg	4320
ttttttttt ttctaattgat tcattaccgc tatgtatacc tacttgtact tgtagtaagc	4380

cgggttattg gcgttcaatt aatcatagac ttatgaatct gcacggtgtg cgctgcgagt 4440
 tacttttagc ttatgcatgc tacttgggtg taatattggg atctgttcgg aaatcaacgg 4500
 atgctcaatc gatttcgaca gtaattaatt aagtcataca caagtcagct ttcttcgagc 4560
 ctcatataag tataagtagt tcaacgtatt agcactgtac ccagcatctc cgtatcgaga 4620
 aacacaacaa catgccccat tggacagatc atgcggatac acaggttgtg cagtatcata 4680
 catactcgat cagacaggtc gtctgacatc catacaagct gaacaagcgc tccatacttg 4740
 cacgctctct atatacacag ttaaattaca tatccatagt ctaacctcta acagttaatc 4800

 ttctggtaag cctcccagcc agccttctgg tatcgcttgg cctcctcaat aggatctcgg 4860
 ttctggccgt acagacctcg gccgacaatt atgatatccg ttccggtaga catgacatcc 4920
 tcaacagttc ggtactgctg tccgagagcg tctcccttgt cgtcaagacc caccgccggg 4980
 gtcagaataa gccagtcctc agagtcgccc ttaggtcggg tctgggcaat gaagccaacc 5040
 acaaactcgg ggtcggatcg ggcaagctca atggctctgt tggagtactc gccagtggcc 5100
 agagagccct tgcaagacag ctcgccagc atgagcagac ctctggccag ctctctgttg 5160
 ggagagggga ctaggaactc cttgtactgg gagtctctgt agtcagagac gtctccttc 5220

 ttctgttcag agacagtctc ctccgcacca gctcgcaggc cagcaatgat tccggttccg 5280
 ggtacaccgt gggcggttgg gatatcggac cactcggcga ttccgtgaca ccggtactgg 5340
 tgcttgacag tgttgccaat atctgcgaac ttctgtcct cgaacaggaa gaaaccgtgc 5400
 ttaagagcaa gtctcttagg ggggagcaca gtcccgccgt aggtgaagtc gtcaatgatg 5460
 tcgatatggg ttttgatcat gcacacataa ggtccgacct tatcggcaag ctcaatgagc 5520
 tccttggtag tggtaacatc cagagaagca cacaggtagg ttttcttggc tgccacgagc 5580
 ttgagcactc gagcggcaaa ggcggacttg tggacgttag ctcgagcttc gtaggagggc 5640

 attttggtag tgaagaggag actgaataa atttagtctg cagaactttt tatcggaacc 5700
 ttatctgggg cagtgaagta tatgttatgg taatagttac gagtagttg aacttataga 5760
 tagactggac tatacggcta tcggtccaaa ttagaaagaa cgtcaatggc tctctgggag 5820
 tcgcctttgc cgacaaaaat gtgatcatga tgaaagccag caatgacgtt gcagctgata 5880
 ttgttgtcgg ccaaccgcgc cgaaaacgca gctgtcagac ccacagctc caacgaagaa 5940
 tgtatcgtca aagtgatcca agcacactca tagttggagt cgtactcaa aggcggcaat 6000
 gacgagtcag acagatactc gtcgacgttt aaacagtgtg cgcagatcta ctatagagga 6060

 acatttaaat tgccccggag aagacggcca ggccgcctag atgacaaatt caacaactca 6120
 cagctgactt tctgceattg ccactagggg ggggcctttt tataatggca agccaagctc 6180
 tccacgtcgg ttgggctgca cccaacaata aatgggtagg gttgcacaa caaagggatg 6240

ggatgggggg tagaagatac gaggataacg gggctcaatg gcacaaataa gaacgaatac 6300
 tgccattaag actcgtgac cagcgactga caccattgca tcactaagg gcctcaaac 6360
 tacctcgaa ctgctgcgt gatctggaca ccacagaggt tccgagcact ttaggttgca 6420
 ccaaattgcc caccaggtgc aggcagaaaa cgctggaaca gcgtgtacag ttgtcttaa 6480

 caaaaagtga gggcgctgag gtcgagcagg gtggtgtgac ttgttatagc ctttagagct 6540
 gcgaaagcgc gtatggattt ggctcatcag gccagattga gggctctgtg acacatgtca 6600
 tgttagtgta ctcaatcgc cccctggata tagccccgac aataggccgt ggcctcattt 6660
 ttttgcttc cgcacatttc cattgctga taccacacc ttgcttctcc tgcacttgc 6720
 aaccttaata ctggtttaca ttgaccaaca tcttacaagc ggggggcttg tctagggtat 6780
 atataaacag tggctctccc aatcggttgc cagtctcttt tttccttct tccccacag 6840
 attcgaaatc taaactacac atcacagaat tccgagccgt gagtatccac gacaagatca 6900

 gtgtcgagac gacgcgtttt gtgtaatgac acaatccgaa agtcgctagc aacacacact 6960
 ctctacaaa actaaccag ctctgggtacc atggaggtcg tgaacgaaat cgtctccatt 7020
 ggccaggagg ttcttccaa ggtcgactat gtcagctct ggtctgatgc ctgcactgc 7080
 gaggtgctgt acctctccat cgccttcgtc atcctgaagt tcacccttgg tctctcga 7140
 cccaagggtc agtctgaat gaagtttgtg ttcaccaact acaacctgct catgtccatc 7200
 tactcgctgg gctccttct ctctatggcc tacgcatgt acaccattgg tgtcatgtcc 7260
 gacaactgcg agaaggcttt cggcaacaat gtcttccgaa tcaccactca gctgttctac 7320

 ctacgaagt tctcgagta cattgactcc ttctatctgc cctcatggg caagcctctg 7380
 acctggttgc agttcttca ccatctcgga gctcctatgg acatgtggct gttctacaac 7440
 taccgaaac aagccgtttg gatctttgtg ctgctcaacg gttcattca ctggatcatg 7500
 tacggetact attggaccgg actgatcaag ctcaagtcc ctatgccaa gtccctgatt 7560
 acttctatgc agatcattca gttcaacgtt ggcttctaca tcgtctggaa gtaccggaac 7620
 attcctgct accgacaaga tggaaatgaga atgtttggct ggtttttcaa ctacttctac 7680
 gtgggtactg tctgtgtct gttctcaac ttctacgtgc agacctacat cgtccgaaag 7740

 cacaaggag ccaaaaagat tcagtgagc 7769
 <210> 130
 <211> 777
 <212> DNA
 <213> *Euglena gracilis*
 <220><221> CDS

<222> (1)..(777)

<223> mutant delta-9 elongase "EgD9eS-L130M/V243A2"

<400> 130

atg gag gtc gtg aac gaa att gtc tcc att ggc cag gag gtt ctt ccc 48

Met Glu Val Val Asn Glu Ile Val Ser Ile Gly Gln Glu Val Leu Pro

1 5 10 15

aag gtc gac tat gct cag ctc tgg tct gat gcc tcg cac tgc gag gtg 96

Lys Val Asp Tyr Ala Gln Leu Trp Ser Asp Ala Ser His Cys Glu Val

20 25 30

ctg tac ctc tcc atc gcc ttc gtc atc ctg aag ttc acc ctt ggt cct 144

Leu Tyr Leu Ser Ile Ala Phe Val Ile Leu Lys Phe Thr Leu Gly Pro

35 40 45

ctc gga ccc aag ggt cag tct cga atg aag ttt gtg ttc acc aac tac 192

Leu Gly Pro Lys Gly Gln Ser Arg Met Lys Phe Val Phe Thr Asn Tyr

50 55 60

aac ctg ctc atg tcc atc tac tcg ctg ggc tcc ttc ctc tct atg gcc 240

Asn Leu Leu Met Ser Ile Tyr Ser Leu Gly Ser Phe Leu Ser Met Ala

65 70 75 80

tac gcc atg tac acc att ggt gtc atg tcc gac aac tgc gag aag gct 288

Tyr Ala Met Tyr Thr Ile Gly Val Met Ser Asp Asn Cys Glu Lys Ala

85 90 95

ttc gac aac aat gtc ttc cga atc acc act cag ctg ttc tac ctc agc 336

Phe Asp Asn Asn Val Phe Arg Ile Thr Thr Gln Leu Phe Tyr Leu Ser

100 105 110

aag ttc ctc gag tac att gac tcc ttc tat ctg ccc ctc atg ggc aag 384

Lys Phe Leu Glu Tyr Ile Asp Ser Phe Tyr Leu Pro Leu Met Gly Lys

115 120 125

cct atg acc tgg ttg cag ttc ttt cac cat ctc gga gct cct atg gac 432

Pro Met Thr Trp Leu Gln Phe Phe His His Leu Gly Ala Pro Met Asp

130 135 140

atg tgg ctg ttc tac aat tac cga aac gaa gcc gtt tgg atc ttt gtg 480

Met Trp Leu Phe Tyr Asn Tyr Arg Asn Glu Ala Val Trp Ile Phe Val
145 150 155 160
ctg ctc aac ggc ttc att cac tgg atc atg tac ggc tac tat tgg acc 528
Leu Leu Asn Gly Phe Ile His Trp Ile Met Tyr Gly Tyr Tyr Trp Thr
165 170 175
cga ctg atc aag ctc aag ttc cct atg ccc aag tcc ctg att act tct 576
Arg Leu Ile Lys Leu Lys Phe Pro Met Pro Lys Ser Leu Ile Thr Ser
180 185 190
atg cag atc att cag ttc aac gtt ggc ttc tac atc gtc tgg aag tac 624
Met Gln Ile Ile Gln Phe Asn Val Gly Phe Tyr Ile Val Trp Lys Tyr
195 200 205
cgg aac att ccc tgc tac cga caa gat gga atg aga atg ttt ggc tgg 672
Arg Asn Ile Pro Cys Tyr Arg Gln Asp Gly Met Arg Met Phe Gly Trp
210 215 220
ttt ttc aac tac ttc tac gtt ggt act gtc ctg tgt ctg ttc ctc aac 720
Phe Phe Asn Tyr Phe Tyr Val Gly Thr Val Leu Cys Leu Phe Leu Asn
225 230 235 240
ttc tac gcg cag acc tac atc gtc cga aag cac aag gga gcc aaa aag 768
Phe Tyr Ala Gln Thr Tyr Ile Val Arg Lys His Lys Gly Ala Lys Lys
245 250 255
att cag tga 777
Ile Gln
<210> 131
<211> 258
<212> PRT
<213> Euglena gracilis
<400> 131
Met Glu Val Val Asn Glu Ile Val Ser Ile Gly Gln Glu Val Leu Pro
1 5 10 15
Lys Val Asp Tyr Ala Gln Leu Trp Ser Asp Ala Ser His Cys Glu Val

20 25 30
 Leu Tyr Leu Ser Ile Ala Phe Val Ile Leu Lys Phe Thr Leu Gly Pro
 35 40 45
 Leu Gly Pro Lys Gly Gln Ser Arg Met Lys Phe Val Phe Thr Asn Tyr
 50 55 60
 Asn Leu Leu Met Ser Ile Tyr Ser Leu Gly Ser Phe Leu Ser Met Ala

 65 70 75 80
 Tyr Ala Met Tyr Thr Ile Gly Val Met Ser Asp Asn Cys Glu Lys Ala
 85 90 95
 Phe Asp Asn Asn Val Phe Arg Ile Thr Thr Gln Leu Phe Tyr Leu Ser
 100 105 110
 Lys Phe Leu Glu Tyr Ile Asp Ser Phe Tyr Leu Pro Leu Met Gly Lys
 115 120 125
 Pro Met Thr Trp Leu Gln Phe Phe His His Leu Gly Ala Pro Met Asp

 130 135 140
 Met Trp Leu Phe Tyr Asn Tyr Arg Asn Glu Ala Val Trp Ile Phe Val
 145 150 155 160
 Leu Leu Asn Gly Phe Ile His Trp Ile Met Tyr Gly Tyr Tyr Trp Thr
 165 170 175
 Arg Leu Ile Lys Leu Lys Phe Pro Met Pro Lys Ser Leu Ile Thr Ser
 180 185 190
 Met Gln Ile Ile Gln Phe Asn Val Gly Phe Tyr Ile Val Trp Lys Tyr

 195 200 205
 Arg Asn Ile Pro Cys Tyr Arg Gln Asp Gly Met Arg Met Phe Gly Trp
 210 215 220
 Phe Phe Asn Tyr Phe Tyr Val Gly Thr Val Leu Cys Leu Phe Leu Asn
 225 230 235 240
 Phe Tyr Ala Gln Thr Tyr Ile Val Arg Lys His Lys Gly Ala Lys Lys
 245 250 255
 Ile Gln

<210> 132

<211> 7769

<212> DNA

<213> Artificial Sequence

<220><223> Plasmid pZuFmEgD9eS-L130M/V243A2

<400> 132

```

ggccgcaagt gtggatgggg aagtgagtgc ccggttctgt gtgcacaatt ggcaatccaa      60
gatggatgga ttcaacacag ggatatagcg agctacgtgg tggtagcagg atatagcaac      120
ggatatattat gtttgacact tgagaatgta cgatacaagc actgtccaag tacaatacta      180
aacatactgt acatactcat actcgtaccc gggcaacggt ttcacttgag tgcagtggct      240
agtgtcttta ctctacagt gtgcaatact gcgtatcata gtctttgatg tatatcgtat      300
tcattcatgt tagttgcgta cgagccggaa gcataaagtg taaagcctgg ggtgcctaata      360

gagtgagcta atcacatta attgcgttgc gctcactgcc cgctttccag tcgggaaacc      420
tgtcgtgcca gctgcattaa tgaatcgcc aacgcgcggg gagaggcggg ttgcgtattg      480
ggcgctcttc cgcttccctg ctactgact cgctgcgctc ggtcgttcgg ctgcggcgag      540
cggatatcagc tactcaaaag gcggttaatac ggttatccac agaatacagg gataacgcag      600
gaaagaacat gtgagcaaaa ggccagcaaa aggccaggaa ccgtaaaaag gccgcgttgc      660
tggcggtttt ccataggttc cgccccctg acgagcatca caaaaatcga cgctcaagtc      720
agaggtggcg aaaccgcaca ggactataaa gataccaggc gtttccccct ggaagctccc      780

tcgtgcgctc tctgtttccg accctgccgc ttaccggata cctgtccgcc tttctccctt      840
cgggaagcgt ggcgctttct catagctcac gctgtaggta tctcagttcg gtgtaggtcg      900
ttcgtctcaa gctgggctgt gtgcacgaac cccccgttca gcccgaccgc tgcgccttat      960
ccggtaaacta tcgtcttgag tccaaccggt taagacacga cttatcgcca ctggcagcag      1020
ccactggtaa caggattagc agagcgaggt atgtaggcgg tgctacagag ttcttgaagt      1080
ggtggcctaa ctacggctac actagaagga cagtatttgg tatctgcgct ctgctgaagc      1140
cagttacctt cggaataaaga gttggtagct cttgatccgg caaacaacc accgctggta      1200

gcggtggttt ttttgtttgc aagcagcaga ttacgcgcag aaaaaagga tctcaagaag      1260
atcctttgat cttttctacg gggctcgacg ctcagtggaa cgaaaactca cgttaaggga      1320
ttttggtcat gagattatca aaaaggatct tcacctagat ctttttaaat taaaaatgaa      1380
gttttaaatc aatctaaagt atatatgagt aaacttggtc tgacagttac caatgcttaa      1440
tcagttaggc accatatctc gcgactgttc tatttcgttc atccatagtt gcctgactcc      1500
ccgtcgtgta gataactacg atacgggagg gcttaccatc tggccccagt gctgcaatga      1560

```

taccgcgaga cccacgctca ccggctccag atttatcagc aataaaccag ccagccggaa	1620
gggccgagcg cagaagtggc cctgcaactt tatccgcctc catccagtct attaatgtt	1680
gccgggaagc tagagtaagt agttcgccag ttaatatgtt gcgcaacgtt gttgccattg	1740
ctacaggcat cgtgggtgca cgctcgctgt ttggtatggc ttcatcagc tccggttccc	1800
aacgatcaag gcgagttaca tgatcccca tgttgtgcaa aaaagcgggt agtccttcg	1860
gtcctccgat cgttgtcaga agtaagtgg ccgcagtgtt atcactcatg gttatggcag	1920
cactgcataa ttctcttact gtcatgcat ccgtaagatg cttttctgtg actggtgagt	1980
actcaacca gtcattctga gaatagtgtg tgcggcgacc gattgtctt tgcggcggt	2040
caatacggga taataccgcg ccacatagca gaactttaaa agtgctcatc attggaaaac	2100
gttcttcggg gcgaaaactc tcaaggatct taccgtgtt gagatccagt tcgatgtaac	2160
ccactcgtgc acccaactga tcttcagcat cttttacttt caccagcgtt tctgggtgag	2220
caaaaacagg aaggcaaaat gccgcaaaa agggaataag ggcgacacgg aaatgttgaa	2280
tactcatact cttccttttt caatattatt gaagcattta tcagggttat tgtctcatga	2340
gcggatacat atttgaatgt atttagaaaa ataaacaaat aggggttcg gcacatttc	2400
ccgaaaagt gccacctgac gcgcctgtg gcggcgcat aagcgcgcg ggtgtggtgg	2460
ttacgcgcag cgtgaccgt acacttgcca gcgcctagc gccgcctctt ttcgctttct	2520
tcccttctt tctcgccacg ttcgccggt tccccgtca agctctaaat cgggggctcc	2580
ctttagggtt ccgatttagt gctttacggc acctcgacc caaaaaactt gattagggtg	2640
atggttcacg tagtgggcca tcgccctgat agacggttt tgcgccttg acgttgagat	2700
ccacgttctt taatagtga ctttgttcc aaactggaac aacactcaac cctatctcg	2760
tctattcttt tgatttataa gggattttgc cgatttcggc ctattggtta aaaaatgagc	2820
tgatttaaca aaaatttaac gcgaatttta acaaaatatt aacgttaca atttcattc	2880
gccattcagg ctgcgcaact gttgggaagg gcgatcgggt cgggcctctt cgctattacg	2940
ccagctggcg aaagggggat gtgctgcaag gcgattaagt tgggtaacgc cagggttttc	3000
ccagtcacga cgttgtaaaa cgacggccag tgaattgtaa tacgactcac tataggcgca	3060
attgggtacc gggcccccc tcgaggtcga tgggtgcgat aagcttgata tcgaattcat	3120
gtcacacaaa ccgattctcg cctcaaggaa acctaatct acatccgaga gactgccgag	3180
atccagtcta cactgattaa ttttcgggcc aataattta aaaaatcgtg ttatataata	3240
ttatatgtat tatatatata catcatgat atactgacag tcatgtcca ttgctaaata	3300

gacagactcc atctgccgcc tccaactgat gttctcaata ttttaagggt catctcgcat 3360
 tgtttaataa taaacagact ccatctaccg cctccaatg atgttctcaa aatatattgt 3420
 atgaacttat ttttattact tagtattatt agacaactta cttgctttat gaaaaaact 3480
 tcctatttag gaaacaattt ataatggcag ttcgttcatt taacaattta tgtagaataa 3540
 atgttataaa tgcgtatggg aaatcttaaa tatggatagc ataaatgata tctgcattgc 3600
 ctaattcgaa atcaacagca acgaaaaaaaa tcccttgtag aacataaata gtcacgaga 3660
 aatatcaact atcaaagaac agctattcac acgttactat tgagattatt attggacgag 3720

aatcacacac tcaactgtct ttctctcttc tagaaatata ggtacaagta tgtactattc 3780
 tcattgttca tacttctagt catttcatcc cacatatcc ttggatttct ctccaatgaa 3840
 tgacattcta tcttgcaaat tcaacaatta taataagata taccaaagta gcggtatagt 3900
 ggcaatcaaa aagcttctct ggtgtgcttc tcgtatttat ttttattcta atgatccatt 3960
 aaaggatat atttatttct tgttatataa tccttttggt tattacatgg gctggataca 4020
 taaaggatt ttgatttaat tttttgctta aattcaatcc cccctcgctc agtgtcaact 4080
 gtaatggtag gaaattacca tacttttgaa gaagcaaaaa aatgaaaga aaaaaaaaaat 4140

cgtatttcca ggtagacgt tccgcagaat ctagaatgcg gtatgcggtta cattgttctt 4200
 cgaacgtaaa agttgcgctc cctgagatat tgtacatttt tgcttttaca agtacaagta 4260
 catcgtacaa ctatgtacta ctgttgatgc atccacaaca gtttgtttg ttttttttg 4320
 ttttttttt ttctaagtat tcattaccgc tatgtatacc tacttgtact ttagtaagc 4380
 cgggttattg gcgttcaatt aatcatagac ttatgaatct gcacggtgtg cgtcgcgagt 4440
 tacttttagc ttatgcatgc tacttgggtg taatattggg atctgttcgg aaatcaacgg 4500
 atgctcaatc gatttcgaca gtaattaatt aagtcataca caagtcagct ttcttcgagc 4560

ctcatataag tataagtagt tcaacgtatt agcactgtac ccagcatctc cgtatcgaga 4620
 aacacaacaa catgccccat tggacagatc atgcggatac acaggttgtg cagtatcata 4680
 catactcgat cagacaggtc gtctgacat catacaagct gaacaagcgc tccatacttg 4740
 cagctctct atatacacag ttaaattaca tatccatagt ctaacctcta acagttaatc 4800
 ttctggtaag cctcccagcc agccttctgg tatcgcttg cctcctcaat aggatctcgg 4860
 ttctggccgt acagacctgc gccgacaatt atgatatccg ttccggtaga catgacatcc 4920
 tcaacagttc ggiactgctg tccgagagcg tctcccttgt cgtcaagacc caccgccggg 4980

gtcagaataa gccagtcctc agagtcgcc ttaggtcggg tctgggcaat gaagccaacc 5040
 acaaactcgg ggtcggatcg ggcaagctca atgggtctgct tggagtactc gccagtggcc 5100
 agagagccct tgcaagacag ctcgccagc atgagcagac ctctggccag cttctcgttg 5160

ggagagggga ctaggaactc cttgtactgg gatttctcgt agtcagagac gtcctccttc 5220
 ttctgttcag agacagtttc ctcggcacca gctcgcaggc cagcaatgat tccggttccg 5280
 ggtacaccgt gggcgttggg gatatacgac cactcggcga ttcggtgaca ccggtactgg 5340
 tgcttgacag tgttgccaat atctgcgaac tttctgtcct cgaacaggaa gaaaccgtgc 5400

 ttaagagcaa gttccttgag ggggagcaca gtgccggcgt aggtgaagtc gtcaatgatg 5460
 tcgatatggg ttttgatcat gcacacataa ggtccgacct tatcggaag ctcaatgagc 5520
 tccttgggtg tggtaacatc cagagaagca cacagggttg ttttcttggc tgccacgagc 5580
 ttgagcactc gagcggcaaa ggcggacttg tggacgttag ctcgagcttc gtaggagggc 5640
 attttgggtg tgaagaggag actgaaataa atttagtctg cagaactttt tatcggaacc 5700
 ttatctgggg cagtgaagta tatgttatgg taatagttac gagttagtgt aacttataga 5760
 tagactggac tatacggcta tcggtcctaaa ttagaaagaa cgtcaatggc tctctgggcg 5820

 tcgcctttgc cgacaaaaat gtgatcatga tgaaagccag caatgacgtt gcagctgata 5880
 ttgttgtcgg ccaaccgcgc cgaaaacgca gctgtcagac ccacagcctc caacgaagaa 5940
 tgtatcgtca aagtgatcca agcacactca tagttggagt cgtactccaa aggcggcaat 6000
 gacgagttag acagatactc gtcgacgttt aaacagtgtg cgcagatcta ctatagagga 6060
 acatttaaat tgccccggag aagacggcca ggccgcctag atgacaaatt caacaactca 6120
 cagctgactt tctgccattg ccactagggg ggggcctttt tataatggcca agccaagctc 6180
 tccacgtcgg ttgggctgca cccaacaata aatgggtagg gttgcaccaa caaagggatg 6240

 ggatgggggg tagaagatag gaggataacg gggctcaatg gcacaaataa gaacgaatac 6300
 tgccattaag actcgtgac cagcgactga caccattgca tcatctaagg gcctcaaaac 6360
 tacctcgga ctcgtgcgt gatctggaca ccacagaggt tccgagcact ttaggttgca 6420
 ccaaagtgc caccagggtc aggcagaaaa cgttggaaca gcgtgtacag tttgtcttaa 6480
 caaaaagtga gggcgtgag gtcgagcagg gtggtgtgac ttgttatagc ctttagagct 6540
 gcgaaagcgc gtatggattt ggctcatcag gccagattga gggctctgtg acacatgtca 6600
 tgtagtgta ctcaatcgc cccctggata tagccccgac aataggccgt ggcctcattt 6660

 ttttgcttc cgcacatttc cattgtctga taccacacc ttgcttctcc tgcattgcc 6720
 aaccttaata ctggtttaca ttgaccaaca tcttaacaag ggggggcttg tctagggtat 6780
 atataaacag tggctctccc aatcggttgc cagtctcttt tttctttct tccccacag 6840
 attcgaaatc taaactacac atcacagaat tccgagccgt gattatccac gacaagatca 6900
 gtgtcgagac gacgcgtttt gtgtaatgac acaatccgaa agtcgctagc aacacacact 6960
 ctctacacaa actaaccag ctctggtagc atggaggtcg tgaacgaaat tgtctccatt 7020

ggccaggagg ttcttcccaa ggtcgactat gctcagctct ggtctgatgc ctgcactgc 7080

gaggtgctgt acctctccat cgccttcgtc atcctgaagt tcacccttgg tcctctcgga 7140

cccaagggtc agtctcgaat gaagtttgtg ttcaccaact acaacctgct catgtccatc 7200

tactcgctgg gctccttctt ctctatggcc tacgcatgt acaccattgg tgtcatgtcc 7260

gacaactgcg agaaggtttt cgacaacaat gtcttccgaa tcaccactca gctgttctac 7320

ctcagcaagt tctcagagta cattgactcc ttctatctgc cctcatggg caagectatg 7380

acctggttgc agttctttca ccatctcgga gctcctatgg acatgtggct gttctacaat 7440

taccgaaacg aagccgtttg gatctttgtg ctgctcaacg gcttcattca ctggatcatg 7500

tacggctact attgaccgcg actgatcaag ctcaagttcc ctatgcccac gtccctgatt 7560

acttctatgc agatcattca gttcaacgtt ggcttctaca tcgtctggaa gtaccggaac 7620

attccctgct accgacaaga tggaatgaga atgtttggct ggtttttcaa ctacttctac 7680

gttggtactg tctgtgtct gttctcaac ttctacgcgc agacctacat cgtccgaaag 7740

cacaaggag ccaaaaagat tcagtgagc 7769

<210> 133

<211> 289

<212> PRT

<213> Ciona intestinalis

<400> 133

Met Asp Val Leu His Arg Phe Leu Gly Phe Tyr Glu Trp Thr Leu Thr

1 5 10 15
Phe Ala Asp Pro Arg Val Ala Lys Trp Pro Leu Ile Glu Asn Pro Leu

20 25 30
Pro Thr Ile Ala Ile Val Leu Leu Tyr Leu Ala Phe Val Leu Tyr Ile

35 40 45
Gly Pro Arg Phe Met Arg Lys Arg Ala Pro Val Asp Phe Gly Leu Phe

50 55 60
Leu Pro Gly Tyr Asn Phe Ala Leu Val Ala Leu Asn Tyr Tyr Ile Leu

65 70 75 80
Gln Glu Val Val Thr Gly Ser Tyr Gly Ala Gly Tyr Asp Leu Val Cys

85 90 95
Thr Pro Leu Arg Ser Asp Ser Tyr Asp Pro Asn Glu Met Lys Val Ala

100 105 110
 Asn Ala Val Trp Trp Tyr Tyr Val Ser Lys Ile Ile Glu Leu Phe Asp
 115 120 125
 Thr Val Leu Phe Thr Leu Arg Lys Arg Asp Arg Gln Val Thr Phe Leu

130 135 140
 His Val Tyr His His Ser Thr Met Pro Leu Leu Trp Trp Ile Gly Ala
 145 150 155 160
 Lys Trp Val Pro Gly Gly Gln Ser Phe Val Gly Ile Ile Leu Asn Ser
 165 170 175
 Ser Val His Val Ile Met Tyr Thr Tyr Tyr Gly Leu Ser Ala Leu Gly
 180 185 190
 Pro His Met Gln Lys Phe Leu Trp Trp Lys Lys Tyr Ile Thr Met Leu

195 200 205
 Gln Leu Val Gln Phe Val Leu Ala Ile Tyr His Thr Ala Arg Ser Leu
 210 215 220
 Tyr Val Lys Cys Pro Ser Pro Val Trp Met His Trp Ala Leu Ile Leu
 225 230 235 240
 Tyr Ala Phe Ser Phe Ile Leu Leu Phe Ser Asn Phe Tyr Met His Ala
 245 250 255
 Tyr Ile Lys Lys Ser Arg Lys Gly Lys Glu Asn Gly Ser Arg Gly Lys

260 265 270
 Gly Gly Val Ser Asn Gly Lys Glu Lys Leu His Ala Asn Gly Lys Thr
 275 280 285
 Asp

<210> 134

<211> 295

<212> PRT

<213> Oncorhynchus mykiss

<400> 134

Met Glu Thr Phe Asn Tyr Lys Leu Asn Met Tyr Ile Asp Ser Trp Met
 1 5 10 15

Gly Pro Arg Asp Glu Arg Val Gln Gly Trp Leu Leu Leu Asp Asn Tyr
20 25 30

Pro Pro Thr Phe Ala Leu Thr Val Met Tyr Leu Leu Ile Val Trp Met
35 40 45

Gly Pro Lys Tyr Met Arg His Arg Gln Pro Val Ser Cys Arg Gly Leu
50 55 60

Leu Leu Val Tyr Asn Leu Gly Leu Thr Ile Leu Ser Phe Tyr Met Phe
65 70 75 80

Tyr Glu Met Val Ser Ala Val Trp His Gly Asp Tyr Asn Phe Phe Cys
85 90 95

Gln Asp Thr His Ser Ala Gly Glu Thr Asp Thr Lys Ile Ile Asn Val
100 105 110

Leu Trp Trp Tyr Tyr Phe Ser Lys Leu Ile Glu Phe Met Asp Thr Phe
115 120 125

Phe Phe Ile Leu Arg Lys Asn Asn His Gln Ile Thr Phe Leu His Ile
130 135 140

Tyr His His Ala Ser Met Leu Asn Ile Trp Trp Phe Val Met Asn Trp
145 150 155 160

Val Pro Cys Gly His Ser Tyr Phe Gly Ala Ser Leu Asn Ser Phe Ile
165 170 175

His Val Leu Met Tyr Ser Tyr Tyr Gly Leu Ser Ala Val Pro Ala Leu
180 185 190

Arg Pro Tyr Leu Trp Trp Lys Lys Tyr Ile Thr Gln Val Gln Leu Ile
195 200 205

Gln Phe Phe Leu Thr Met Ser Gln Thr Ile Cys Ala Val Ile Trp Pro
210 215 220

Cys Asp Phe Pro Arg Gly Trp Leu Tyr Phe Gln Ile Phe Tyr Val Ile
225 230 235 240

Thr Leu Ile Ala Leu Phe Ser Asn Phe Tyr Ile Gln Thr Tyr Lys Lys
245 250 255

His Leu Val Ser Gln Lys Lys Glu Tyr His Gln Asn Gly Ser Val Ala

260 265 270
 Ser Leu Asn Gly His Val Asn Gly Val Thr Pro Thr Glu Thr Ile Thr
 275 280 285

 His Arg Lys Val Arg Gly Asp
 290 295
 <210> 135
 <211> 290
 <212> PRT
 <213> *Marchantia polymorpha*
 <400> 135
 Met Glu Ala Tyr Glu Met Val Asp Ser Phe Val Ser Lys Thr Val Phe
 1 5 10 15
 Glu Thr Leu Gln Arg Leu Arg Gly Gly Val Val Leu Thr Glu Ser Ala
 20 25 30
 Ile Thr Lys Gly Leu Pro Cys Val Asp Ser Pro Thr Pro Ile Val Leu
 35 40 45

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

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

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

Ile Lys Asn Asn Ser Pro Tyr Pro Gln Phe Leu Ile Gln Ile Leu Phe
 245 250 255
 Tyr Tyr Met Ile Ser Leu Leu Ala Leu Phe Gly Asn Phe Tyr Val His
 260 265 270
 Lys Tyr Val Ser Ala Pro Ala Lys Pro Ala Lys Ile Lys Ser Lys Lys
 275 280 285

Ala Glu

290

<210> 136

<211> 290

<212> PRT

<213> Physcomitrella patens

<400> 136

Met Glu Val Val Glu Arg Phe Tyr Gly Glu Leu Asp Gly Lys Val Ser

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

65 70 75 80
 Leu Gln Ala Leu Val Leu Val His Asn Leu Phe Cys Phe Ala Leu Ser
 85 90 95
 Leu Tyr Met Cys Val Gly Ile Ala Tyr Gln Ala Ile Thr Trp Arg Tyr
 100 105 110
 Ser Leu Trp Gly Asn Ala Tyr Asn Pro Lys His Lys Glu Met Ala Ile
 115 120 125
 Leu Val Tyr Leu Phe Tyr Met Ser Lys Tyr Val Glu Phe Met Asp Thr

 130 135 140
 Val Ile Met Ile Leu Lys Arg Ser Thr Arg Gln Ile Ser Phe Leu His
 145 150 155 160
 Val Tyr His His Ser Ser Ile Ser Leu Ile Trp Trp Ala Ile Ala His
 165 170 175
 His Ala Pro Gly Gly Glu Ala Tyr Trp Ser Ala Ala Leu Asn Ser Gly
 180 185 190
 Val His Val Leu Met Tyr Ala Tyr Tyr Phe Leu Ala Ala Cys Leu Arg

 195 200 205
 Ser Ser Pro Lys Leu Lys Asn Lys Tyr Leu Phe Trp Gly Arg Tyr Leu
 210 215 220
 Thr Gln Phe Gln Met Phe Gln Phe Met Leu Asn Leu Val Gln Ala Tyr
 225 230 235 240
 Tyr Asp Met Lys Thr Asn Ala Pro Tyr Pro Gln Trp Leu Ile Lys Ile
 245 250 255
 Leu Phe Tyr Tyr Met Ile Ser Leu Leu Phe Leu Phe Gly Asn Phe Tyr

 260 265 270
 Val Gln Lys Tyr Ile Lys Pro Ser Asp Gly Lys Gln Lys Gly Ala Lys
 275 280 285
 Thr Glu
 290
 <210> 137
 <211> 348
 <212> PRT

<213> *Marchantia polymorpha*

<400> 137

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

1 5 10 15

Ala Lys Met Lys Gln Ala Arg Ser Ser Pro Glu Gly Glu Ile Val Gly

20 25 30

Gly Asn Arg Met Gly Ser Gly Asn Gly Ala Glu Trp Thr Thr Ser Leu

35 40 45

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

50 55 60

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

65 70 75 80

Leu Glu Asp Val Val Gly Val Asp Ile Arg Lys Pro Asn Pro Val Thr

85 90 95

Lys Asp Leu Pro Met Val Glu Ser Pro Val Pro Val Leu Ala Cys Ile

100 105 110

Ser Leu Tyr Leu Leu Val Val Trp Leu Trp Ser Ser His Ile Lys Ala

115 120 125

Ser Gly Gln Lys Pro Arg Lys Glu Asp Pro Leu Ala Leu Arg Cys Leu

130 135 140

Val Ile Ala His Asn Leu Phe Leu Cys Cys Leu Ser Leu Phe Met Cys

145 150 155 160

Val Gly Leu Ile Ala Ala Ala Arg His Tyr Gly Tyr Ser Val Trp Gly

165 170 175

Asn Tyr Tyr Arg Glu Arg Glu Pro Ala Met Asn Leu Leu Ile Tyr Val

180 185 190

Phe Tyr Met Ser Lys Leu Tyr Glu Phe Met Asp Thr Ala Ile Met Leu

195 200 205

Phe Arg Arg Asn Leu Arg Gln Val Thr Tyr Leu His Val Tyr His His

210 215 220

Ala Ser Ile Ala Met Ile Trp Trp Ile Ile Cys Tyr Arg Phe Pro Gly

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

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

 <210> 138

 <
 211> 292
 <212> PRT
 <213> *Ostreococcus tauri*
 <400> 138

 Met Ser Gly Leu Arg Ala Pro Asn Phe Leu His Arg Phe Trp Thr Lys
 1 5 10 15
 Trp Asp Tyr Ala Ile Ser Lys Val Val Phe Thr Cys Ala Asp Ser Phe
 20 25 30
 Gln Trp Asp Ile Gly Pro Val Ser Ser Ser Thr Ala His Leu Pro Ala
 35 40 45
 Ile Glu Ser Pro Thr Pro Leu Val Thr Ser Leu Leu Phe Tyr Leu Val

 50 55 60
 Thr Val Phe Leu Trp Tyr Gly Arg Leu Thr Arg Ser Ser Asp Lys Lys
 65 70 75 80
 Ile Arg Glu Pro Thr Trp Leu Arg Arg Phe Ile Ile Cys His Asn Ala
 85 90 95

Phe Leu Ile Val Leu Ser Leu Tyr Met Cys Leu Gly Cys Val Ala Gln
 100 105 110
 Ala Tyr Gln Asn Gly Tyr Thr Leu Trp Gly Asn Glu Phe Lys Ala Thr
 115 120 125
 Glu Thr Gln Leu Ala Leu Tyr Ile Tyr Ile Phe Tyr Val Ser Lys Ile
 130 135 140
 Tyr Glu Phe Val Asp Thr Tyr Ile Met Leu Leu Lys Asn Asn Leu Arg
 145 150 155 160
 Gln Val Ser Phe Leu His Ile Tyr His His Ser Thr Ile Ser Phe Ile
 165 170 175
 Trp Trp Ile Ile Ala Arg Arg Ala Pro Gly Gly Asp Ala Tyr Phe Ser
 180 185 190
 Ala Ala Leu Asn Ser Trp Val His Val Cys Met Tyr Thr Tyr Tyr Leu
 195 200 205
 Leu Ser Thr Leu Ile Gly Lys Glu Asp Pro Lys Arg Ser Asn Tyr Leu
 210 215 220
 Trp Trp Gly Arg His Leu Thr Gln Met Gln Met Leu Gln Phe Phe Phe
 225 230 235 240
 Asn Val Leu Gln Ala Leu Tyr Cys Ala Ser Phe Ser Thr Tyr Pro Lys
 245 250 255
 Phe Leu Ser Lys Ile Leu Leu Val Tyr Met Met Ser Leu Leu Gly Leu
 260 265 270
 Phe Gly His Phe Tyr Tyr Ser Lys His Ile Ala Ala Ala Lys Leu Gln
 275 280 285
 Lys Lys Gln Gln
 290
 <210> 139
 <211> 277
 <212> PRT
 <213> Pavlova sp.
 <400> 139
 Met Met Leu Ala Ala Gly Tyr Leu Leu Val Leu Ser Ala Ala Arg Gln

1 5 10 15

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

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

 His Val Tyr His His Val Leu Leu Met Trp Ala Trp Phe Val Val Val
 145 150 155 160
 Lys Leu Gly Asn Gly Gly Asp Ala Tyr Phe Gly Gly Leu Met Asn Ser
 165 170 175
 Ile Ile His Val Met Met Tyr Ser Tyr Tyr Thr Met Ala Leu Leu Gly
 180 185 190
 Trp Ser Cys Pro Trp Lys Arg Tyr Leu Thr Gln Ala Gln Leu Val Gln
 195 200 205

 Phe Cys Ile Cys Leu Ala His Ser Thr Trp Ala Ala Val Thr Gly Ala
 210 215 220
 Tyr Pro Trp Arg Ile Cys Leu Val Glu Val Trp Val Met Val Ser Met
 225 230 235 240
 Leu Val Leu Phe Thr Arg Phe Tyr Arg Gln Ala Tyr Ala Lys Glu Ala
 245 250 255

Lys Ala Lys Glu Ala Lys Lys Leu Ala Gln Glu Ala Ser Gln Ala Lys
 260 265 270

Ala Val Lys Ala Glu
 275

<210> 140

<211> 302

<212> PRT

<213> Rebecca salina

<400> 140

Met Lys Ala Ala Ala Gly Lys Val Gln Gln Glu Ala Glu Arg Leu Thr
 1 5 10 15

Ala Gly Leu Trp Leu Pro Met Met Leu Ala Ala Gly Tyr Leu Leu Val
 20 25 30

Leu Ser Ala Asn Arg Ala Ser Phe Tyr Glu Asn Ile Asn Asn Glu Lys
 35 40 45

Gly Ala Tyr Ser Thr Ser Trp Phe Ser Leu Pro Cys Val Met Thr Ala
 50 55 60

Val Tyr Leu Gly Gly Val Phe Gly Leu Thr Lys Tyr Phe Glu Gly Arg
 65 70 75 80

Lys Pro Met Gln Gly Leu Lys Asp Tyr Met Phe Thr Tyr Asn Leu Tyr
 85 90 95

Gln Val Ile Ile Asn Val Trp Cys Ile Ala Ala Phe Val Val Glu Val
 100 105 110

Arg Arg Ala Gly Met Ser Ala Val Gly Asn Lys Val Asp Leu Gly Pro
 115 120 125

Asn Ser Phe Arg Leu Gly Phe Val Thr Trp Val His Tyr Asn Asn Lys
 130 135 140

Tyr Val Glu Leu Leu Asp Thr Leu Trp Met Val Leu Arg Lys Lys Thr
 145 150 155 160

Gln Gln Val Ser Phe Leu His Val Tyr His His Val Leu Leu Ile Trp
 165 170 175

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

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

<210> 141

<211> 300

<212> PRT

<213> *Ostreococcus tauri*

<400> 141

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

85 90 95
 Met Phe Ala Arg Glu Ile Ser Gly Leu Gly Gln Pro Val Trp Gly Ser
 100 105 110
 Thr Met Pro Trp Ser Asp Arg Lys Ser Phe Lys Ile Leu Leu Gly Val
 115 120 125
 Trp Leu His Tyr Asn Asn Lys Tyr Leu Glu Leu Leu Asp Thr Val Phe
 130 135 140
 Met Val Ala Arg Lys Lys Thr Lys Gln Leu Ser Phe Leu His Val Tyr
 145 150 155 160
 His His Ala Leu Leu Ile Trp Ala Trp Trp Leu Val Cys His Leu Met
 165 170 175
 Ala Thr Asn Asp Cys Ile Asp Ala Tyr Phe Gly Ala Ala Cys Asn Ser
 180 185 190
 Phe Ile His Ile Val Met Tyr Ser Tyr Tyr Leu Met Ser Ala Leu Gly
 195 200 205
 Ile Arg Cys Pro Trp Lys Arg Tyr Ile Thr Gln Ala Gln Met Leu Gln
 210 215 220
 Phe Val Ile Val Phe Ala His Ala Val Phe Val Leu Arg Gln Lys His
 225 230 235 240
 Cys Pro Val Thr Leu Pro Trp Ala Gln Met Phe Val Met Thr Asn Met
 245 250 255
 Leu Val Leu Phe Gly Asn Phe Tyr Leu Lys Ala Tyr Ser Asn Lys Ser
 260 265 270
 Arg Gly Asp Gly Ala Ser Ser Val Lys Pro Ala Glu Thr Thr Arg Ala
 275 280 285
 Pro Ser Val Arg Arg Thr Arg Ser Arg Lys Ile Asp
 290 295 300
 <210> 142
 <211> 358
 <212> PRT
 <213> Thalassiosira pseudonana

<400> 142

```

Met Cys Ser Ser Pro Pro Ser Gln Ser Lys Thr Thr Ser Leu Leu Ala
1           5           10           15
Arg Tyr Thr Thr Ala Ala Leu Leu Leu Leu Thr Leu Thr Thr Trp Cys
           20           25           30
His Phe Ala Phe Pro Ala Ala Thr Ala Thr Pro Gly Leu Thr Ala Glu
           35           40           45
Met His Ser Tyr Lys Val Pro Leu Gly Leu Thr Val Phe Tyr Leu Leu
           50           55           60

Ser Leu Pro Ser Leu Lys Tyr Val Thr Asp Asn Tyr Leu Ala Lys Lys
65           70           75           80
Tyr Asp Met Lys Ser Leu Leu Thr Glu Ser Met Val Leu Tyr Asn Val
           85           90           95
Ala Gln Val Leu Leu Asn Gly Trp Thr Val Tyr Ala Ile Val Asp Ala
           100          105          110
Val Met Asn Arg Asp His Pro Phe Ile Gly Ser Arg Ser Leu Val Gly
           115          120          125

Ala Ala Leu His Ser Gly Ser Ser Tyr Ala Val Trp Val His Tyr Cys
           130          135          140
Asp Lys Tyr Leu Glu Phe Phe Asp Thr Tyr Phe Met Val Leu Arg Gly
           145          150          155          160
Lys Met Asp Gln Val Ser Phe Leu His Ile Tyr His His Thr Thr Ile
           165          170          175
Ala Trp Ala Trp Trp Ile Ala Leu Arg Phe Ser Pro Gly Gly Asp Ile
           180          185          190

Tyr Phe Gly Ala Leu Leu Asn Ser Ile Ile His Val Leu Met Tyr Ser
           195          200          205
Tyr Tyr Ala Leu Ala Leu Leu Lys Val Ser Cys Pro Trp Lys Arg Tyr
           210          215          220
Leu Thr Gln Ala Gln Leu Leu Gln Phe Thr Ser Val Val Val Tyr Thr
           225          230          235          240

```

Gly Cys Thr Gly Tyr Thr His Tyr Tyr His Thr Lys His Gly Ala Asp
245 250 255

Glu Thr Gln Pro Ser Leu Gly Thr Tyr Tyr Phe Cys Cys Gly Val Gln
260 265 270

Val Phe Glu Met Val Ser Leu Phe Val Leu Phe Ser Ile Phe Tyr Lys
275 280 285

Arg Ser Tyr Ser Lys Lys Asn Lys Ser Gly Gly Lys Asp Ser Lys Lys
290 295 300

Asn Asp Asp Gly Asn Asn Glu Asp Gln Cys His Lys Ala Met Lys Asp
305 310 315 320

Ile Ser Glu Gly Ala Lys Glu Val Val Gly His Ala Ala Lys Asp Ala
325 330 335

Gly Lys Leu Val Ala Thr Ala Ser Lys Ala Val Lys Arg Lys Gly Thr
340 345 350

Arg Val Thr Gly Ala Met
355

<210> 143

<211> 272

<212> PRT

<213> Thalassiosira pseudonana

<400> 143

Met Asp Ala Tyr Asn Ala Ala Met Asp Lys Ile Gly Ala Ala Ile Ile
1 5 10 15

Asp Trp Ser Asp Pro Asp Gly Lys Phe Arg Ala Asp Arg Glu Asp Trp
20 25 30

Trp Leu Cys Asp Phe Arg Ser Ala Ile Thr Ile Ala Leu Ile Tyr Ile
35 40 45

Ala Phe Val Ile Leu Gly Ser Ala Val Met Gln Ser Leu Pro Ala Met
50 55 60

Asp Pro Tyr Pro Ile Lys Phe Leu Tyr Asn Val Ser Gln Ile Phe Leu
65 70 75 80

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

Leu His Val Tyr His His Thr Thr Ile Phe Leu Phe Tyr Trp Leu Asn
145 150 155 160
Ala Asn Val Leu Tyr Asp Gly Asp Ile Phe Leu Thr Ile Leu Leu Asn
165 170 175
Gly Phe Ile His Thr Val Met Tyr Thr Tyr Tyr Phe Ile Cys Met His
180 185 190
Thr Lys Asp Ser Lys Thr Gly Lys Ser Leu Pro Ile Trp Trp Lys Ser
195 200 205

Ser Leu Thr Ala Phe Gln Leu Leu Gln Phe Thr Ile Met Met Ser Gln
210 215 220
Ala Thr Tyr Leu Val Phe His Gly Cys Asp Lys Val Ser Leu Arg Ile
225 230 235 240
Thr Ile Val Tyr Phe Val Ser Leu Leu Ser Leu Phe Phe Leu Phe Ala
245 250 255
Gln Phe Phe Val Gln Ser Tyr Met Ala Pro Lys Lys Lys Lys Ser Ala
260 265 270

<210> 144

<211> 318

<212> PRT

<213> Mortierella alpina

<400> 144

Met Glu Ser Ile Ala Pro Phe Leu Pro Ser Lys Met Pro Gln Asp Leu
1 5 10 15
Phe Met Asp Leu Ala Thr Ala Ile Gly Val Arg Ala Ala Pro Tyr Val

20 25 30
 Asp Pro Leu Glu Ala Ala Leu Val Ala Gln Ala Glu Lys Tyr Ile Pro
 35 40 45
 Thr Ile Val His His Thr Arg Gly Phe Leu Val Ala Val Glu Ser Pro

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

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

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

 245 250 255
 Ser Ser Trp Asp Met Tyr Ala Met Lys Val Leu Gly Arg Pro Gly Tyr
 260 265 270

Pro Phe Phe Ile Thr Ala Leu Leu Trp Phe Tyr Met Trp Thr Met Leu
275 280 285

Gly Leu Phe Tyr Asn Phe Tyr Arg Lys Asn Ala Lys Leu Ala Lys Gln
290 295 300

Ala Lys Ala Asp Ala Ala Lys Glu Lys Ala Arg Lys Leu Gln

305 310 315

<210> 145

<211> 271

<212> PRT

<213> Thraustochytrium sp.

<400> 145

Met Met Glu Pro Leu Asp Arg Tyr Arg Ala Leu Ala Glu Leu Ala Ala
1 5 10 15

Arg Tyr Ala Ser Ser Ala Ala Phe Lys Trp Gln Val Thr Tyr Asp Ala
20 25 30

Lys Asp Ser Phe Val Gly Pro Leu Gly Ile Arg Glu Pro Leu Gly Leu
35 40 45

Leu Val Gly Ser Val Val Leu Tyr Leu Ser Leu Gln Ala Val Val Tyr
50 55 60

Ala Leu Arg Asn Tyr Leu Gly Gly Leu Met Ala Leu Arg Ser Val His
65 70 75 80

Asn Leu Gly Leu Cys Leu Phe Ser Gly Ala Val Trp Ile Tyr Thr Ser
85 90 95

Tyr Leu Met Ile Gln Asp Gly His Phe Arg Ser Leu Glu Ala Ala Thr
100 105 110

Cys Glu Pro Leu Lys His Pro His Phe Gln Leu Ile Ser Leu Leu Phe
115 120 125

Ala Leu Ser Lys Ile Trp Glu Trp Phe Asp Thr Val Leu Leu Ile Val
130 135 140

Lys Gly Asn Lys Leu Arg Phe Leu His Val Leu His His Ala Thr Thr
145 150 155 160

Phe Trp Leu Tyr Ala Ile Asp His Ile Phe Leu Ser Ser Ile Lys Tyr

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

Thr Pro Tyr Leu Phe Val Val Pro Phe Leu Ile Leu Phe Leu Asn Phe
 245 250 255
 Tyr Leu Gln Gln Tyr Val Leu Ala Pro Ala Lys Thr Lys Lys Ala
 260 265 270

<210> 146
 <211> 23
 <212> DNA
 <213> Artificial Sequence
 <220><223> Primer EgD9E_102_053008f
 <220><221> misc_feature
 <222> (10)..(11)
 <223> n is a, c, g, or t
 <400> 146
 gtgctgtacn nktccatcgc ctt

23

<210> 147
 <211> 24
 <212> DNA
 <213> Artificial Sequence
 <220><223> Primer EgD9E_760_053008r
 <400> 147
 tcccttgtgc tttcggacga tgta
 <210> 148
 <211> 7769

24

<212> DNA

<213> Artificial Sequence

<220><223> Plasmid pZuFmEgD9ES-L35G

<400> 148

ggccgcaagt gtggatgggg aagtgagtgc cggttctgt gtgcacaatt ggcaatccaa	60
gatggatgga ttcaacacag ggatatagcg agctacgtgg tggtagcagg atatatcaac	120
ggatatattat gtttgacact tgagaatgia cgatacaagc actgtccaag tacaatacta	180
aacatactgt acatactcat actcgtaccc gggcaacggt ttactttgag tgcagtggct	240
agtgccttta ctctacagt gtgcaatact gcgtatcata gtctttgatg tatatcgtat	300
tcattcatgt tagttgcgta cgagccggaa gcataaagtg taaagcctgg ggtgccta	360
gagttagcta actcacatta attgcgttgc gctcactgcc cgctttccag tcgggaaacc	420
tgtcgtgccg gcctcattaa tgaatcggcc aacgcgcggg gagaggcggg ttgcgtattg	480
ggcgcctctc cgcttcctcg ctactgact cgctgcgctc ggtcgttcgg ctgcggcgag	540
cggtatcagc tcaactcaaag gcggttaatac ggttatccac agaatacagg gataacgcag	600
gaaagaacat gtgagcaaaa ggccagcaaa aggccaggaa ccgtaaaaag gccgcgttgc	660
tggcggtttt ccataggctc cgccccctg acgagcatca caaaaatcga cgctcaagtc	720
agaggtggcg aaacccgaca ggactataaa gataccaggc gtttccccct ggaagctccc	780
tcgtgcgctc tctgtttccg accctgccgc ttaccggata cctgtccgcc tttctccctt	840
cgggaagcgt ggcgctttct catagctcac gctgtaggta tctcagttcg gtgtaggtcg	900
ttcgtccaa gcctgggtgt gtgcacgaac cccccgtca gcccgaccgc tgcgccttat	960
ccggttaacta tcgtcttgag tccaaccggg taagacacga cttatcgcca ctggcagcag	1020
ccactggtaa caggattagc agagcgaggt atgtaggcgg tgctacagag ttcttgaagt	1080
ggtggcctaa ctacggctac actagaagga cagtatttgg tatctgcgct ctgctgaagc	1140
cagttacctt cggaataaga gttggtagct cttgatccgg caaacaacc accgctggt	1200
gcggtgggtt ttttgtttgc aagcagcaga ttacgcgcag aaaaaagga tctcaagaag	1260
atcctttgat cttttctacg gggcttgacg ctactggaa cgaaaactca cgttaaggga	1320
ttttggtcat gagattatca aaaaggatct tcacctagat ccttttaaat taaaaatgaa	1380
gttttaaatc aatctaaagt atatatgagt aaacttggc tgacagttac caatgcttaa	1440
tcagtgagge acctatctca gcgatctgtc tatttcgttc atccatagtt gcctgactcc	1500
ccgtcgtgta gataactacg atacgggagg gcttaccatc tggccccagt gctgcaatga	1560
taccgcgaga cccacgtca ccggctccag atttatcagc aataaaccag ccagccggaa	1620

gggccgagcg cagaagtggc cctgcaactt tatccgcctc catccagtct attaatgtt 1680
 gccgggaagc tagagtaagt agttgccag ttaatatgtt gcgcaacgtt gttgccattg 1740
 ctacaggcat cgtgggtgca cgctcgtcgt ttggtatggc ttcattcagc tccggttccc 1800
 aacgatcaag gcgagttaca tgatccccc tgttgtgcaa aaaagcgggt agctccttcg 1860

 gtctccgat cgttgtcaga agtaagtggc ccgcagtgtt atcactcatg gttatggcag 1920
 cactgcataa ttctcttact gtcatgcat cgtgaagatg cttttctgtg actggtgagt 1980
 actcaaccaa gtcattctga gaatagtgtg tgcggcgacc gagttgctct tgcggcggt 2040
 caatacggga taataccgcg ccacatagca gaactttaaa agtgcctatc attggaanaa 2100
 gtcttcggg gcgaaaactc tcaaggatct taccgctgtt gagatccagt tcgatgtaac 2160
 ccactcgtgc acccaactga tcttcagcat cttttacttt caccagcgtt tctgggtgag 2220
 caaaaacagg aaggcaaaa gccgcaaaaa agggaataag ggcgacacgg aatgttgaa 2280

 tactcatact ctctctttt caatattatt gaagcattta tcagggttat tgtctcatga 2340
 gcggatacat atttgaatgt atttagaaaa ataaacaaat aggggttccg cgcacatttc 2400
 cccgaaaagt gccacctgac gcgcctgtga gcggcgcat aagcgcgcg ggtgtggtgg 2460
 ttacgcgcag cgtgaccgt acacttgcca gcgccctagc gcccgctcct ttcgctttct 2520
 tcccttctt tctcgccagc ttcgccggt tccccgtca agctctaaat cgggggctcc 2580
 ctttaggggt cggatttagt gctttacggc acctcgacce caaaaaactt gattaggggtg 2640
 atggttcacg tagtgggcca tcgccctgat agacgggttt tcgcccttg acgttgaggt 2700

 ccacgttctt taatagtga ctttgttcc aaactggaac aacactcaac cctatctcgg 2760
 tctattcttt tgatttataa gggattttgc cgatttcggc ctattggtta aaaaatgagc 2820
 tgatttaaca aaaaattaac gcgaatttta aaaaaatatt aacgcttaca atttcattc 2880
 gccattcagg ctgcgcaact gttgggaagg gcgacgggtg cgggcctctt cgctattacg 2940
 ccagctggcg aaagggggat gtgctgcaag gcgattaagt tgggtaacgc cagggttttc 3000
 ccagtcacga cgttgtaaaa cgacggccag tgaattgtaa tacgactcac tataggcgca 3060
 attgggtacc gggcccccc tcgaggtcga tgggtgcgat aagcttgata tcgaattcat 3120

 gtcacacaaa ccgatcttcg cctcaaggaa acctaatct acatccgaga gactgccgag 3180
 atccagtcta cactgattaa ttttcgggcc aataatttaa aaaaatcgtg ttatataata 3240
 ttatatgtat tatatatata catcatgat aactgacag tcatgtcca ttgctaaata 3300
 gacagactcc atcgcgcc tcctaactgat gttctcaata ttaagggtt catctcgcgt 3360
 tgtttaataa taaacagact ccatctaccg cctccaaatg atgttctcaa aatataatgt 3420
 atgaacttat ttttattact tagtattatt agacaactta cttgctttat gaaaaacact 3480

tcctattttag gaaacaatth ataattggcag ttcgttcatt taacaattta tgtagaataa	3540
atgtttataaa tgcgtatggg aaatcttaaa tatggatagc ataaatgata tctgcattgc	3600
ctaattcgaa atcaacagca acgaaaaaaaa tcccttgtac aacataaata gtcacgcaga	3660
aatatcaact atcaaagaac agctattcac acgttactat tgagattatt attggacgag	3720
aatcacacac tcaactgtct ttctctcttc tagaaatata ggtacaagta tgtactattc	3780
tcattgttca tacttctagt catttcatcc cacatatcc ttggatttct ctccaatgaa	3840
tgacattcta tcttgcaaat tcaacaatta taataagata taccaaagta gcggtatagt	3900
ggcaatcaaa aagcttctct ggtgtgcttc tcgtatttat ttttattcta atgatccatt	3960
aaaggtatat atttatttct tgttatataa tccttttggt tattacatgg gctggataca	4020
taaaggtatt ttgatttaat tttttgctta aattcaatcc cccctcgttc agtgtcaact	4080
gtaatggtag gaaattacca tacttttgaa gaagcaaaaa aaatgaaaga aaaaaaaaaat	4140
cgtatttcca ggtagacgt tccgcagaat ctagaatgcg gtatgcggtta cattgttctt	4200
cgaacgtaaa agttgcgtc cctgagatat tgtacatttt tgcttttaca agtacaagta	4260
catcgtacaa ctatgtacta ctgttgatgc atccacaaca gtttgttttg ttttttttg	4320
tttttttttt ttctaattgat tcattaccgc tatgtatacc tacttgtact tgtagtaagc	4380
cgggttattg gcgttcaatt aatcatagac ttatgaatct gcacgggtgt cgctgcgagt	4440
tacttttagc ttatgcatgc tacttgggtg taatatggg atctgttcgg aaatcaacgg	4500
atgtcaatc gatttcgaca gtaattaat aagtcataca caagtcagct ttcttcgagc	4560
ctcatataag tataagtagt tcaacgtatt agcactgtac ccagcatctc cgtatcgaga	4620
aacacaacaa catgccccat tggacagatc atgcggatac acaggttgtg cagtatcata	4680
catactcgat cagacaggtc gtctgacat catacaagct gaacaagcgc tccatacttg	4740
cacgtctct atatacacag ttaaatata tatccatagt ctaacctcta acagttaatc	4800
ttctggtaag cctcccagcc agccttctgg tatcgcttg cctcctcaat aggatctcgg	4860
ttctggccgt acagacctg gccgacaatt atgatatccg ttccggtaga catgacatcc	4920
tcaacagttc ggtactgtc tccgagagcg tctcccttgt cgtcaagacc caccctgggg	4980
gtcagaataa gccagtcctc agagtcgcc ttaggtcggg tctgggcaat gaagccaacc	5040
acaaactcgg ggtcggatcg ggcaagctca atggtctgct tggagtactc gccagtggcc	5100
agagagccct tgcaagacag ctgcggcagc atgagcagac ctctggccag cttctcgttg	5160
ggagagggga ctaggaactc cttgtactgg gagttctcgt agtcagagac gtctccttc	5220

ttctgttcag agacagtttc ctccggcacca gctcgcaggc cagcaatgat tccggttccg	5280
ggtacaccgt gggcgttggg gatatcggac cactcggcga ttcggtgaca ccggtactgg	5340
tgcttgacag tgttgccaat atctgcgaac tttctgtcct cgaacaggaa gaaaccgtgc	5400
ttaagagcaa gticctttag ggggagcaca gtgccggcgt aggtgaagtc gtcaatgatg	5460
tcgatatggg ttttgatcat gcacacataa ggtccgacct tatcggcaag ctcaatgagc	5520
tccttgggtg tggtaacatc cagagaagca cacaggttgg ttttcttggc tgccacgagc	5580
ttgagcactc gagcggcaaa ggcggacttg tggacgttag ctcgagcttc gtaggagggc	5640
attttgggtg tgaagaggag actgaaataa atttagtctg cagaactttt tatcggaacc	5700
ttatctgggg cagtgaagta tatgttatgg taatagttaac gagttagttag aacttataga	5760
tagactggac tatacggcta tcggtccaaa ttagaaagaa cgtcaatggc tctctgggcg	5820
tcgcctttgc cgacaaaaat gtgatcatga tgaagaccag caatgacgtt gcagctgata	5880
ttgttgcctg ccaaccgcgc cgaaaacgca gctgtcagac ccacagcctc caacgaagaa	5940
tgtatcgtca aagtgatcca agcacactca tagttggagt cgtactcaa aggcggcaat	6000
gacgagtcag acagatactc gtcgacgttt aaacagtgtg cgcagatcta ctatagagga	6060
acatttaaat tgccccggag aagacggcca ggccgccttag atgacaaatt caacaactca	6120
cagctgactt tctgccattg ccactagggg ggggcctttt tataatggcca agccaagctc	6180
tccacgtcgg ttgggctgca cccaacaata aatgggtagg gttgcaccaa caaagggatg	6240
ggatgggggg tagaagatag gaggataacg gggctcaatg gcacaaataa gaacgaatac	6300
tgccattaag actcgtgac cagcgactga caccattgca tcatctaagg gcctcaaaac	6360
tacctcgaa ctgctgcgct gatctggaca ccacagaggt tccgagcact ttaggttgca	6420
ccaaatgtcc caccaggctc aggcagaaaa cgctggaaca gcgtgtacag tttgtcttaa	6480
caaaaagtga gggcgtgag gtcgagcagg gtggtgtgac ttgttatagc ctttagagct	6540
gcgaaagcgc gtatggattt ggctcatcag gccagattga gggctctgtg acacatgtca	6600
tgtagtgta cttaatcgc cccctggata tagccccgac aataggccgt ggctcattt	6660
ttttgccttc cgcacatttc cattgctcga taccacacc ttgcttctcc tgcaattgcc	6720
aaccttaata ctggtttaca ttgaccaaca tcttacaagc ggggggcttg tctagggtat	6780
atataaacag tggtctctcc aatcggttgc cagtctcttt tttctttct tccccacag	6840
attcgaaatc taaactacac atcacagaat tccgagccgt gagtatccac gacaagatca	6900
gtgtcgagac gacgcgtttt gtgtaatgac acaatccgaa agtcgctagc aacacacact	6960
ctctacacaa actaaccag ctctggtagc atggaggtcg tgaacgaaat cgtctccatt	7020
ggccaggagg ttcttcccaa ggtcgactat gctcagctct ggtctgatgc ctgcactgc	7080

gaggtgctgt acgggtccat cgccttcgtc atcctgaagt tcacccttgg tcctctcgga 7140
 cccaagggtc agtctcgaat gaagtttgtg ttcaccaact acaacctgct catgtccatc 7200
 tactcgctgg gctccttcct ctctatggcc tacgcatgt acaccattgg tgtcatgtcc 7260
 gacaactgcg agaaggcttt cgacaacaat gtcttccgaa tcaccactca gctgttctac 7320

ctcagcaagt tcctcgagta cattgactcc ttctatctgc cctcatggg caagcctctg 7380
 acctggttgc agttctttca ccatctcgga gctcctatgg acatgtggct gtcttacaac 7440
 taccgaaacg aagccgtttg gatctttgtg ctgctcaacg gcttcattca ctggatcatg 7500
 tacggctact attggaccgc actgatcaag ctcaagttcc ctatgcccaa gtcctgatt 7560
 acttctatgc agatcattca gttcaacgtt ggcttctaca tcgtctggaa gtaccggaac 7620
 attcctgct accgacaaga tggaatgaga atgtttggct ggtttttcaa ctacttctac 7680
 gtggtactg tcctgtgtct gttcttcaac ttctacgtgc agacctacat cgtccgaaag 7740

cacaaggag ccaaaaagat tcagtgagc 7769

<210> 149

<211> 777

<212> DNA

<213> Euglena gracilis

<220><221> CDS

<222> (1)..(777)

<223> mutant delta-9 elongase "EgD9eS-L35M/Q107E"

<400> 149

atg gag gtc gtg aac gaa atc gtc tcc att ggc cag gag gtt ctt ccc 48
 Met Glu Val Val Asn Glu Ile Val Ser Ile Gly Gln Glu Val Leu Pro
 1 5 10 15
 aag gtc gac tat gct cag ctc tgg tct gat gcc tcg cac tgc gag gtg 96

Lys Val Asp Tyr Ala Gln Leu Trp Ser Asp Ala Ser His Cys Glu Val
 20 25 30
 ctg tac atg tcc atc gcc ttc gtc atc ctg aag ttc acc ctt ggt cct 144
 Leu Tyr Met Ser Ile Ala Phe Val Ile Leu Lys Phe Thr Leu Gly Pro
 35 40 45
 ctc gga ccc aag ggt cag tct cga atg aag ttt gtg ttc acc aac tac 192
 Leu Gly Pro Lys Gly Gln Ser Arg Met Lys Phe Val Phe Thr Asn Tyr

50	55	60	
aac ctg ctc atg tcc atc tac tgc ctg ggc tcc ttc ctc tct atg gcc			240
Asn Leu Leu Met Ser Ile Tyr Ser Leu Gly Ser Phe Leu Ser Met Ala			
65	70	75	80
tac gcc atg tac acc att ggt gtc atg tcc gac aac tgc gag aag gct			288
Tyr Ala Met Tyr Thr Ile Gly Val Met Ser Asp Asn Cys Glu Lys Ala			
	85	90	95
ttc gac aac aat gtc ttc cga atc acc act gag ctg ttc tac ctc agc			336
Phe Asp Asn Asn Val Phe Arg Ile Thr Thr Glu Leu Phe Tyr Leu Ser			
	100	105	110
aag ttc ctc gag tac att gac tcc ttc tat ctg ccc ctc atg ggc aag			384
Lys Phe Leu Glu Tyr Ile Asp Ser Phe Tyr Leu Pro Leu Met Gly Lys			
	115	120	125
cct ctg acc tgg ttg cag ttc ttt cac cat ctc gga gct cct atg gac			432
Pro Leu Thr Trp Leu Gln Phe Phe His His Leu Gly Ala Pro Met Asp			
	130	135	140
atg tgg ctg ttc tac aac tac cga aac gaa gcc gtt tgg atc ttt gtg			480
Met Trp Leu Phe Tyr Asn Tyr Arg Asn Glu Ala Val Trp Ile Phe Val			
	145	150	155
160			
ctg ctc aac ggc ttc att cac tgg atc atg tac ggc tac tat tgg acc			528
Leu Leu Asn Gly Phe Ile His Trp Ile Met Tyr Gly Tyr Tyr Trp Thr			
	165	170	175
cga ctg atc aag ctc aag ttc cct atg ccc aag tcc ctg att act tct			576
Arg Leu Ile Lys Leu Lys Phe Pro Met Pro Lys Ser Leu Ile Thr Ser			
	180	185	190
atg cag atc att cag ttc aac gtt ggc ttc tac atc gtc tgg aag tac			624
Met Gln Ile Ile Gln Phe Asn Val Gly Phe Tyr Ile Val Trp Lys Tyr			
	195	200	205
cgg aac att ccc tgc tac cga caa gat gga atg aga atg ttt ggc tgg			672
Arg Asn Ile Pro Cys Tyr Arg Gln Asp Gly Met Arg Met Phe Gly Trp			
	210	215	220

ttt ttc aac tac ttc tac gtt ggt act gtc ctg tgt ctg ttc ctc aac 720
 Phe Phe Asn Tyr Phe Tyr Val Gly Thr Val Leu Cys Leu Phe Leu Asn
 225 230 235 240
 ttc tac gtg cag acc tac atc gtc cga aag cac aag gga gcc aaa aag 768

 Phe Tyr Val Gln Thr Tyr Ile Val Arg Lys His Lys Gly Ala Lys Lys
 245 250 255
 att cag tga 777
 Ile Gln

 <210> 150
 <211> 258
 <212> PRT
 <213> Euglena gracilis
 <400> 150
 Met Glu Val Val Asn Glu Ile Val Ser Ile Gly Gln Glu Val Leu Pro

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

 65 70 75 80
 Tyr Ala Met Tyr Thr Ile Gly Val Met Ser Asp Asn Cys Glu Lys Ala
 85 90 95
 Phe Asp Asn Asn Val Phe Arg Ile Thr Thr Glu Leu Phe Tyr Leu Ser
 100 105 110
 Lys Phe Leu Glu Tyr Ile Asp Ser Phe Tyr Leu Pro Leu Met Gly Lys
 115 120 125
 Pro Leu Thr Trp Leu Gln Phe Phe His His Leu Gly Ala Pro Met Asp

130 135 140
Met Trp Leu Phe Tyr Asn Tyr Arg Asn Glu Ala Val Trp Ile Phe Val
145 150 155 160
Leu Leu Asn Gly Phe Ile His Trp Ile Met Tyr Gly Tyr Tyr Trp Thr
165 170 175
Arg Leu Ile Lys Leu Lys Phe Pro Met Pro Lys Ser Leu Ile Thr Ser
180 185 190
Met Gln Ile Ile Gln Phe Asn Val Gly Phe Tyr Ile Val Trp Lys Tyr

195 200 205
Arg Asn Ile Pro Cys Tyr Arg Gln Asp Gly Met Arg Met Phe Gly Trp
210 215 220
Phe Phe Asn Tyr Phe Tyr Val Gly Thr Val Leu Cys Leu Phe Leu Asn
225 230 235 240
Phe Tyr Val Gln Thr Tyr Ile Val Arg Lys His Lys Gly Ala Lys Lys
245 250 255
Ile Gln

<210> 151

<211> 7769

<212> DNA

<213> Artificial Sequence

<220><223> Plasmid pZuFmEgD9ES-L35M/Q107E

<400> 151

ggccgcaagt gtggatgggg aagtgagtgc ccggttctgt gtcacaatt ggcaatccaa	60
gatggatgga ttcaacacag ggatatagcg agctacgtgg tgggtcgagg atatagcaac	120
ggatatattat gtttgacact tgagaatgta cgatacaagc actgtccaag tacaatacta	180
aacatactgt acatactcat actcgtaccc gggcaacggt ttcacttgag tgcagtggct	240
agtgtcttta ctcgtacagt gtgcaatact gcgtatcata gtctttgatg tatatcgat	300
tcattcatgt tagttgcgta cgagccggaa gcataaagtg taaagcctgg ggtgccta	360
gagtgagcta actcacatta attgcgttgc gctcactgcc cgctttccag tcgggaaacc	420
tgctgtgcca gctgcattaa tgaatcggcc aacgcgcggg gagaggcggt ttgcgtattg	480
ggcgcctctc cgcttcctcg ctactgact cgctgcgctc ggtcgttcgg ctgcggcgag	540

cggtatcagc tcactcaaag gcggtaatat ggttatccac agaatacagg gataacgcag	600
gaaagaacat gtgagcaaaa ggccagcaaa aggccaggaa ccgtaaaaag gccgcgttgc	660
tggcggtttt ccataggctc cgccccctg acgagcatca caaaaatcga cgctcaagtc	720
agagggtggcg aaacccgaca ggactataaa gataccaggc gtttccccct ggaagctccc	780
tcgtgcgctc tctgtttccg accctgccgc ttaccggata cctgtccgcc ttttccctt	840
cgggaagcgt ggcgctttct catagctcac gctgtaggta tctcagttcg gtgtaggtcg	900
ttcgctccaa gctgggctgt gtgcacgaac cccccgttca gcccgaccgc tgcgccttat	960
ccggtaaacta tcgtcttgag tccaacccgg taagacacga cttatcgcca ctggcagcag	1020
ccactggtaa caggattagc agagcgaggt atgtaggcgg tgctacagag ttcttgaagt	1080
ggtggcctaa ctacggctac actagaagga cagtatttgg tatctgcgct ctgctgaagc	1140
cagttacctt cgaaaaaaga gttggtagct cttgatccgg caaacaacc accgctggta	1200
gcggtgggtt ttttgtttgc aagcagcaga ttacgcgcag aaaaaagga tctcaagaag	1260
atcctttgat cttttctacg gggctctgacg ctcagtggaa cgaaaactca cgtaaggga	1320
ttttggtcat gagattatca aaaaggatct tcacctagat ctttttaaat taaaaatgaa	1380
gttttaaatc aatctaaagt atatatgagt aaacttggtc tgacagttac caatgcttaa	1440
tcagtgaggc acctatctca gcgatctgtc tatttcgttc atccatagtt gcctgactcc	1500
ccgtcgtgta gataactacg atacgggagg gcttaccatc tggccccagt gctgcaatga	1560
taccgcgaga cccacgctca ccgggtccag atttatcagc aataaaccag ccagccggaa	1620
gggccgagcg cagaagtggc cctgcaactt tatccgcctc catccagtct attaatgtt	1680
gccgggaagc tagagtaagt agttcgccag ttaatatgtt gcgcaacgtt gttgccattg	1740
ctacaggcat cgtgggtgca cgctcgtcgt ttggtatggc ttcatcagc tccggttccc	1800
aacgatcaag gcgagttaca tgatccccca tgttgtgcaa aaaagcgggt agctccttcg	1860
gtcctccgat cgttgtcaga agtaagttag ccgcagtgtt atcactcatg gttatggcag	1920
cactgcataa ttctcttact gtcatgcat ccgtaaatg cttttctgtg actggtgagt	1980
actcaaccaa gtcattctga gaatagtga tgcggcgacc gagttgctct tgccccgct	2040
caatacggga taataccgcg ccacatagca gaactttaaa agtgctcatc attgaaaaac	2100
gttcttcggg gcgaaaactc tcaaggatct taccgtgtt gagatccagt tcgatgtaac	2160
ccactcgtgc acccaactga tcttcagcat cttttacttt caccagcgtt tctgggtgag	2220
caaaaacagg aaggcaaaat gccgcaaaa agggaataag ggcgacacgg aaatgttgaa	2280
tactcatact cttccttttt caatatatt gaagcattta tcagggttat tgttctcatga	2340
gcggatacat atttgaatgt atttagaaaa ataaacaaat aggggttccg cgcacatttc	2400

cccgaaaagt gccacctgac gcgccttgta gcggcgcat t aagcgcgcg ggtgtggtgg 2460

 ttacgcgcag cgtgaccgct acacttgcca gcgccttagc gcccgctcct ttcgctttct 2520
 tcccttctct tctcgccacg ttcccggtct tccccgtca agctctaaat cgggggctcc 2580
 ctttaggggt cggatttagt gctttacggc acctcgaccc caaaaaactt gattagggtg 2640
 atggttcacg tagtgggcca tcgccctgat agacgggttt tcgccctttg acgttggagt 2700
 ccacgttctt taatagtga ctcttggtcc aaactggaac aacactcaac cctatctcgg 2760
 tctattcttt tgatttataa gggattttgc cgatttcggc ctattggtta aaaaatgagc 2820
 tgatttaaca aaaatttaac gcgaatttta acaaaatatt aacgcttaca atttccattc 2880

 gccattcagg ctgcgcaact gttgggaagg gcgatcggtg cgggcctctt cgctattacg 2940
 ccagctggcg aaagggggat gtgctgcaag gcgattaagt tgggtaacgc cagggttttc 3000
 ccagtcacga cgttgtaaaa cgacggccag tgaattgtaa tacgactcac tataggcgca 3060
 attgggtacc gggccccccc tcgaggtcga tgggtgctgat aagcttgata tcgaattcat 3120
 gtcacacaaa ccgactctcg cctcaaggaa acctaatctt acatccgaga gactgccgag 3180
 atccagtcta cactgattaa ttttcgggcc aataatttaa aaaaatcgtg ttatataata 3240
 ttatatgtat tatatatata catcatgatg atactgacag tcatgtccca ttgctaaata 3300

 gacagactcc atctgccgc tccaactgat gttctcaata ttttaagggt catctcgcat 3360
 tgtttaataa taacagact ccatctaccg cctccaaatg atgttctcaa aatatattgt 3420
 atgaacttat ttttattact tagtattatt agacaactta cttgctttat gaaaaacact 3480
 tcctatttag gaaacaattt ataatggcag ttcgttcatt taacaattta tgtagaataa 3540
 atgttataaa tgcgtatggg aaatcttaaa tatggatagc ataaatgata tctgcattgc 3600
 ctaattcgaa atcaacagca acgaaaaaaa tcccttgtag aacataaata gtcacgaga 3660
 aatatcaact atcaaagaac agctattcac acgttactat tgagattatt attggacgag 3720

 aatcacacac tcaactgtct ttctctcttc tagaaataca ggtacaagta tgtactattc 3780
 tcattgttca tacttctagt catttcatcc cacatatcc ttggatttct ctccaatgaa 3840
 tgacattcta tcttgcaaat tcaacaatta taataagata taccaaagta gcggtatagt 3900
 ggcaatcaaa aagcttctct ggtgtgcttc tcgtatttat ttttattcta atgatccatt 3960
 aaaggatat atttatttct tggtatataa tcttttgtt tattacatgg gctggataca 4020
 taaaggattt ttgatttaat tttttgctta aattcaatcc cccctcgttc agtgtcaact 4080
 gtaatggtag gaaattacca tacttttgaa gaagcaaaaa aaatgaaaga aaaaaaaaaat 4140

cgtattttcca ggtagacgt tccgcagaat ctagaatgcg gtatgcggtg cattgttctt	4200
cgaacgtaaa agttgcgctc cctgagatat tgtacatfff tgctttttaca agtacaagta	4260
catcgtacaa ctatgtacta ctgttgatgc atccacaaca gtttgttttg tttttttttg	4320
tttttttttt ttctaagtat tcattaccgc tatgtatacc tacttgtact ttagtaagc	4380
cgggttattg gcgttcaatt aatcatagac ttatgaatct gcacgggtg cgctgcgagt	4440
tacttttagc ttatgcatgc tacttgggtg taatatggg atctgttcgg aaatcaacgg	4500
atgctcaatc gatttcgaca gtaattaatt aagtcataca caagtcagct ttcttcgagc	4560
ctcatataag tataagtagt tcaacgtatt agcactgtac ccagcatctc cgtatcgaga	4620
aacacaacaa catgccccat tggacagatc atgcggatac acaggttgtg cagtatcata	4680
catactcgat cagacaggtc gtctgacat catacaagct gaacaagcgc tccatacttg	4740
cacgtctct atatacacag ttaaattaca tatccatagt ctaacctcta acagttaatc	4800
ttctggttaag cctcccagcc agccttctgg tatcgcttgg cctcctcaat aggatctcgg	4860
ttctggccgt acagacctcg gccgacaatt atgatatccg ttccggtaga catgacatcc	4920
tcaacagttc ggtactgctg tccgagagcg tctcccttgt cgtcaagacc caccgccggg	4980
gtcagaataa gccagtcctc agagtcgccc ttaggtcggg tctgggcaat gaagccaacc	5040
acaaactcgg ggtcggatcg ggcaagctca atggtctgct tggagtactc gccagtggcc	5100
agagagccct tgcaagacag ctggccagc atgagcagac ctctggccag ctctcgttg	5160
ggagagggga ctaggaactc cttgtactgg gagttctcgt agtcagagac gtctccttc	5220
ttctgttcag agacagtttc ctggcacca gctcgcagc cagcaatgat tccggttccg	5280
ggtacaccgt gggcgttggg gatatcggac cactcggcga ttccgtgaca ccggtactgg	5340
tgcttgacag tgttgccaat atctgcgaac tttctgtcct cgaacaggaa gaaaccgtgc	5400
ttaagagcaa gttccttgag ggggagcaca gtgccggcgt aggtgaagtc gtcaatgatg	5460
tcgatatggg ttttgatcat gcacacataa ggtccgacct tatcggcaag ctcaatgagc	5520
tccttggtag tggtaacatc cagagaagca cacaggtagg ttttcttggc tgccacgagc	5580
ttgagcactc gagcggcaaa ggcggacttg tggacgttag ctcgagcttc gtaggagggc	5640
atattggtgg tgaagaggag actgaaataa atttagtctg cagaactttt tatcggaacc	5700
ttatctgggg cagtgaagta tatgttatgg taatagttac gagttagttg aacttataga	5760
tagactggac tatacggcta tcggtccaaa ttagaaagaa cgtcaatggc tctctggcgc	5820
tcgcctttgc cgacaaaaat gtgatcatga tgaaagccag caatgacgtt gcagctgata	5880
ttgttgtcgg ccaaccgcgc gaaaaacgca gctgtcagac ccacagctc caacgaagaa	5940
tgatcgtca aagtgatcca agcacactca tagttggagt cgtactcaa aggcggcaat	6000

gacgagtcag acagatactc gtcgacgttt aaacagtgtg cgcagatcta ctatagagga 6060
acatttaaat tgccccggag aagacggcca ggccgcctag atgacaaatt caacaactca 6120
cagctgactt tctgccattg ccactagggg ggggcctttt tataatggcca agccaagctc 6180
tccacgtcgg ttgggctgca cccaacaata aatgggtagg gttgcaccaa caaagggatg 6240

ggatgggggg tagaagatac gaggataacg gggtcctaatg gcacaaataa gaacgaatac 6300
tgccattaag actcgtgac cagcgactga caccattgca tcactaagg gcctcaaac 6360
tacctcggaa ctgctgcgt gatctggaca ccacagaggt tccgagcact ttaggttgca 6420
ccaaatgtcc caccagtgcc aggcagaaaa cgctggaaca gcgtgtacag ttgtcttaa 6480
caaaaagtga gggcgctgag gtcgagcagg gtggtgtgac ttgttatagc ctttagagct 6540
gcgaaagcgc gtatggattt ggctcatcag gccagattga gggctctgtg acacatgtca 6600
tgttagtgta cttcaatcgc ccctggata tagccccgac aataggccgt ggcctcattt 6660

ttttgccttc cgcacatttc cattgctcga taccacacc ttgcttctcc tgcacttgcc 6720
aaccttaata ctggtttaca ttgaccaaca tcttacaagc ggggggcttg tctagggtat 6780
atataaacag tggtctctcc aatcggttgc cagtctcttt tttctttct tccccacag 6840
attcgaaatc taaactacac atcacagaat tccgagccgt gagtatccac gacaagatca 6900
gtgtcgagac gacgcgtttt gtgtaatgac acaatccgaa agtcgctagc aacacacact 6960
ctctacaaa actaaccag ctctgggtacc atggaggtcg tgaacgaaat cgtctccatt 7020
ggccaggagg ttcttcccaa ggtcgactat gctcagctct ggtctgatgc ctgcactgc 7080

gaggtgctgt acatgtccat cgcttcgtc atcctgaagt tcacccttgg tctctcggga 7140
cccaagggtc agtctcgaat gaagtttgtg ttaccaact acaacctgct catgtccatc 7200
tactcgtgg gctccttct ctctatggcc tacgcatgt acaccattgg tgtcatgtcc 7260
gacaactcgc agaaggttt cgacaacaat gtcttcgaa tcaccactga gctgttctac 7320
ctcagcaagt tctcgagta cattgactcc ttctatctgc cctcatggg caagcctctg 7380
acctggttgc agttctttca ccatctcgga gctcctatgg acatgtggct gtcttacaac 7440
taccgaaacg aagccgtttg gatctttgtg ctgctcaacg gcttcattca ctggatcatg 7500

tacggctact attggaccgc actgatcaag ctcaagtcc ctatgcccc gtcctgatt 7560
acttctatgc agatcattca gttaacgtt ggcttctaca tcgtctggaa gtaccggaac 7620
attccctgct accgacaaga tggaatgaga atgtttggct gggttttcaa ctacttctac 7680
gttgggtactg tctgtgtct gtctctcaac ttctacgtgc agacctacat cgtccgaaag 7740
cacaaggag caaaaaagat tcagtgagc 7769

<210> 152

<211> 777

<212> DNA

<213> *Euglena gracilis*

<220><221> misc_feature

<222> (1)..(777)

<223> mutant delta-9 elongase "EgD9eS-L35G"

<400> 152

```
atggaggtcg tgaacgaaat cgtctccatt ggccaggagg ttcttccaa ggtcgactat      60
gtcagctctt ggtctgatgc ctgcactgc gaggtgctgt acggatccat cgccttcgtc      120
atcctgaagt tcaccttgg tcctctcgga cccaagggtc agtctcgaat gaagtttgtg      180
ttcaccaact acaacctgct catgtccatc tactcgctgg gtccttcct ctctatggcc      240
tacgcatgt acaccattgg tgtcatgtcc gacaactgcg agaaggcttt cgacaacaat      300
gtcttcgaa tcaccactca gctgttctac ctgagcaagt tcctcgagta cattgactcc      360
ttctatctgc cctcatggg caagcctctg acctgggtgc agttctttca ccatctcgga      420
```

```
gctcctatgg acatgtggct gttctacaac taccgaaacg aagccgtttg gatctttgtg      480
ctgctcaacg gcttcattca ctggatcatg tacggctact attggaccg actgatcaag      540
ctcaagttcc ctatgccaa gtccctgatt acttctatgc agatcattca gttcaacgtt      600
ggcttctaca tcgtctgaa gtaccggaac attccctgct accgacaaga tggaatgaga      660
atgtttggct ggtttttcaa ctacttctac gttgggtactg tcctgtgtct gttctcaac      720
ttctacgtgc agacctacat cgtccgaaag cacaaggag ccaaaaagat tcagtga      777
```

<210> 153

<211> 777

<212> DNA

<213> *Euglena gracilis*

<220><221> misc_feature

<222> (1)..(777)

<223> mutant delta-9 elongase "EgD9eS-L35G"

<400> 153

```
atggaggtcg tgaacgaaat cgtctccatt ggccaggagg ttcttccaa ggtcgactat      60
gtcagctctt ggtctgatgc ctgcactgc gaggtgctgt acggatccat cgccttcgtc      120
atcctgaagt tcaccttgg tcctctcgga cccaagggtc agtctcgaat gaagtttgtg      180
ttcaccaact acaacctgct catgtccatc tactcgctgg gtccttcct ctctatggcc      240
```


tacgccatgt acaccattgg tgtcatgtcc gacaactgcg agaaggcttt cgacaacaat 300
gtcttccgaa tcaccactca gctgttctac ctacagcaagt tcttcgagta cattgactcc 360

ttctatctgc cctcatggg caagcctctg acctgggtgc agttctttca ccatctcgga 420
gtcctatgg acatgtggct gttctacaac taccgaaacg aagccgtttg gatctttgtg 480
ctgctcaacg gcttcattca ctggatcatg tacggctact attggaccg actgatcaag 540
ctcaagtcc ctatgccaa gtccctgatt acttctatgc agatcattca gttcaacgtt 600
ggcttctaca tcgtctggaa gtaccggaac attccctgct accgacaaga tggaatgaga 660
atgtttggct ggtttttcaa ctacttctac gttggtactg tctgtgtct gttctcaac 720
ttctacgtgc agacctacat cgtccgaaag cacaaggag ccaaaaagat tcagtga 777

<210> 154

<211> 777

<212> DNA

<213> Euglena gracilis

<220><221> misc_feature

<222> (1)..(777)

<223> mutant delta-9 elongase "EgD9eS-L35G"

<400> 154

atggaggtcg tgaacgaaat cgtctccatt ggccaggagg ttcttccaa ggtcgactat 60
gtcagctct ggtctgatgc ctgcactgc gaggtgctgt acggttccat cgccttcgtc 120
atcctgaagt tcaccttgg tctctcgga cccaagggtc agtctcgaat gaagtttgtg 180
ttaccaact acaacctgct catgtccatc tactcgctgg gctccttct ctctatggcc 240
tacgccatgt acaccattgg tgtcatgtcc gacaactgcg agaaggcttt cgacaacaat 300

gtcttccgaa tcaccactca gctgttctac ctacagcaagt tcttcgagta cattgactcc 360
ttctatctgc cctcatggg caagcctctg acctgggtgc agttctttca ccatctcgga 420
gtcctatgg acatgtggct gttctacaac taccgaaacg aagccgtttg gatctttgtg 480
ctgctcaacg gcttcattca ctggatcatg tacggctact attggaccg actgatcaag 540
ctcaagtcc ctatgccaa gtccctgatt acttctatgc agatcattca gttcaacgtt 600
ggcttctaca tcgtctggaa gtaccggaac attccctgct accgacaaga tggaatgaga 660
atgtttggct ggtttttcaa ctacttctac gttggtactg tctgtgtct gttctcaac 720

ttctacgtgc agacctacat cgtccgaaag cacaaggag ccaaaaagat tcagtga 777

<210> 155

<211> 22
 <212> DNA
 <213> Artificial Sequence
 <220><223> Primer Eg_9D_122709f
 <400> 155
 gaaatcgtcg acattggcca gg 22
 <210> 156
 <211> 27
 <212> DNA
 <213> Artificial Sequence
 <220><223> Primer Eg_84C_122709r
 <400> 156
 gacaccaatg gtacacatgg cgtaggc 27
 <210> 157
 <211> 27
 <212> DNA
 <213> Artificial Sequence

 <220><223> Primer Eg_84C_122709f
 <400> 157
 gcctacgcc a tgtgtacat tggtgtc 27
 <210> 158
 <211> 22
 <212> DNA
 <213> Artificial Sequence
 <220><223> Primer Eg_9D_122709r
 <400> 158
 cctggccaat gtcgacgatt tc 22
 <210> 159
 <211> 25
 <212> DNA
 <213> Artificial Sequence
 <220><223> Primer Eg_161T_122709f
 <400> 159
 gatctttgtg accctcaacg gcttc 25

<

210> 160

<211> 25

<212> DNA

<213> Artificial Sequence

<220><223> Primer Eg_179R_122709r

<400> 160

gaacttgagc ttctgcagtc gggtc 25

<210> 161

<211> 25

<212> DNA

<213> Artificial Sequence

<220><223> Primer Eg_179R_122709f

<400> 161

gacccgactg cgaaagctca agttc 25

<210> 162

<211> 25

<212> DNA

<213> Artificial Sequence

<220><223> Primer Eg_161T_122709r

<400> 162

gaagccgttg agggtcacaa agatc 25

<210> 163

<211> 22

<212> DNA

<213> Artificial Sequence

<220><223> Primer Eg_244N_122709f

<400> 163

cttctacgtg aacacctaca tc 22

<210> 164

<211> 23

<212> DNA

<213> Artificial Sequence

<220><223> Primer Eg_21V_010710r

<400> 164
gaccagagct gcacatagtc gac 23

<210> 165
<211> 23
<212> DNA
<213> Artificial Sequence
<220><223> Primer Eg_21V_010710f
<400> 165
gtcgactatg tgcagctctg gtc 23

<210> 166
<211> 22
<212> DNA
<213> Artificial Sequence
<220><223> Primer Eg_244N_122709r
<400> 166
gatgtaggtg ttcacgtaga ag 22

<210> 167
<211> 22
<212> DNA
<213> Artificial Sequence
<220><223> Primer Eg_32F_010710f
<400> 167
gcactgcgag tttctgtacc tc 22

<210> 168
<211> 22
<212> DNA
<213> Artificial Sequence
<220><223> Primer Eg_108G_010710r
<400> 168
gaggtagaaa ccctgagtgg tg 22

<210> 169
<211> 22
<212> DNA

<213> Artificial Sequence

<220><223> Primer Eg_108G_010710f

<400> 169

caccactcag ggtttctacc tc

22

<210> 170

<211> 22

<212> DNA

<213> Artificial Sequence

<220><223> Primer Eg_32F_010710r

<400> 170

gaggtacaga aactcgcagt gc

22

<210> 171

<211> 22

<212> DNA

<213> Artificial Sequence

<220><223> Primer Eg_132T_010710f

<400> 171

gcctctgacc accttgacgt tc

22

<210> 172

<211> 22

<212> DNA

<213> Artificial Sequence

<220><223> Primer Eg_143N_010710r

<400> 172

gccacatgtc gtaggagct cc

22

<210> 173

<211> 22

<212> DNA

<213> Artificial Sequence

<220><223> Primer Eg_143N_010710f

<400> 173

ggagctccta acgacatgtg gc

22

<210> 174

<211> 22

<212> DNA
 <213> Artificial Sequence
 <220><223> Primer Eg_132T_010710r
 <400> 174
 gaactgcaag gtggtcagag gc 22

<210> 175
 <211> 23
 <212> DNA
 <213> Artificial Sequence
 <220><223> Primer Eg_236N_010710f
 <400> 175
 gtactgtcct gaacctgttc etc 23

<210> 176
 <211> 23
 <212> DNA
 <213> Artificial Sequence
 <220><223> Primer Eg_236N_010710r
 <400> 176
 gaggaacagg ttcaggacag tac 23

<210> 177
 <211> 777
 <212> DNA
 <213> *Euglena gracilis*
 <220><221> CDS
 <222> (1)..(777)
 <223> mutant delta-9 elongase "EgD9eS-A21V/L35G/L108G/I179R"

<400> 177
 atg gag gtc gtg aac gaa atc gtc tcc att ggc cag gag gtt ctt ccc 48
 Met Glu Val Val Asn Glu Ile Val Ser Ile Gly Gln Glu Val Leu Pro
 1 5 10 15
 aag gtc gac tat gtg cag ctc tgg tct gat gcc tgc cac tgc gag gtg 96
 Lys Val Asp Tyr Val Gln Leu Trp Ser Asp Ala Ser His Cys Glu Val
 20 25 30

ctg tac ggg tcc atc gcc ttc gtc atc ctg aag ttc acc ctt ggt cct	144
Leu Tyr Gly Ser Ile Ala Phe Val Ile Leu Lys Phe Thr Leu Gly Pro	
35 40 45	
ctc gga ccc aag ggt cag tct cga atg aag ttt gtg ttc acc aac tac	192
Leu Gly Pro Lys Gly Gln Ser Arg Met Lys Phe Val Phe Thr Asn Tyr	
50 55 60	
aac ctg ctc atg tcc atc tac tgc ctg ggc tcc ttc ctc tct atg gcc	240
Asn Leu Leu Met Ser Ile Tyr Ser Leu Gly Ser Phe Leu Ser Met Ala	
65 70 75 80	
tac gcc atg tac acc att ggt gtc atg tcc gac aac tgc gag aag gct	288
Tyr Ala Met Tyr Thr Ile Gly Val Met Ser Asp Asn Cys Glu Lys Ala	
85 90 95	
ttc gac aac aat gtc ttc cga atc acc act cag ggt ttc tac ctc agc	336
Phe Asp Asn Asn Val Phe Arg Ile Thr Thr Gln Gly Phe Tyr Leu Ser	
100 105 110	
aag ttc ctc gag tac att gac tcc ttc tat ctg ccc ctc atg ggc aag	384
Lys Phe Leu Glu Tyr Ile Asp Ser Phe Tyr Leu Pro Leu Met Gly Lys	
115 120 125	
cct ctg acc tgg ttg cag ttc ttt cac cat ctc gga gct cct atg gac	432
Pro Leu Thr Trp Leu Gln Phe Phe His His Leu Gly Ala Pro Met Asp	
130 135 140	
atg tgg ctg ttc tac aac tac cga aac gaa gcc gtt tgg atc ttt gtg	480
Met Trp Leu Phe Tyr Asn Tyr Arg Asn Glu Ala Val Trp Ile Phe Val	
145 150 155 160	
ctg ctc aac ggc ttc att cac tgg atc atg tac ggc tac tat tgg acc	528
Leu Leu Asn Gly Phe Ile His Trp Ile Met Tyr Gly Tyr Tyr Trp Thr	
165 170 175	
cga ctg cga aag ctc aag ttc cct atg ccc aag tcc ctg att act tct	576
Arg Leu Arg Lys Leu Lys Phe Pro Met Pro Lys Ser Leu Ile Thr Ser	

180	185	190	
atg cag atc att cag ttc aac gtt ggc ttc tac atc gtc tgg aag tac			624
Met Gln Ile Ile Gln Phe Asn Val Gly Phe Tyr Ile Val Trp Lys Tyr			
195	200	205	
cgg aac att ccc tgc tac cga caa gat gga atg aga atg ttt ggc tgg			672
Arg Asn Ile Pro Cys Tyr Arg Gln Asp Gly Met Arg Met Phe Gly Trp			
210	215	220	
ttt ttc aac tac ttc tac gtt ggt act gtc ctg tgt ctg ttc ctc aac			720
Phe Phe Asn Tyr Phe Tyr Val Gly Thr Val Leu Cys Leu Phe Leu Asn			
225	230	235	240
ttc tac gtg cag acc tac atc gtc cga aag cac aag gga gcc aaa aag			768
Phe Tyr Val Gln Thr Tyr Ile Val Arg Lys His Lys Gly Ala Lys Lys			
245	250	255	
att cag tga			777

Ile Gln

<210> 178

<211> 258

<212> PRT

<213> Euglena gracilis

<400> 178

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

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

100 105 110
Lys Phe Leu Glu Tyr Ile Asp Ser Phe Tyr Leu Pro Leu Met Gly Lys

115 120 125
Pro Leu Thr Trp Leu Gln Phe Phe His His Leu Gly Ala Pro Met Asp

130 135 140
Met Trp Leu Phe Tyr Asn Tyr Arg Asn Glu Ala Val Trp Ile Phe Val

145 150 155 160
Leu Leu Asn Gly Phe Ile His Trp Ile Met Tyr Gly Tyr Tyr Trp Thr

165 170 175
Arg Leu Arg Lys Leu Lys Phe Pro Met Pro Lys Ser Leu Ile Thr Ser

180 185 190
Met Gln Ile Ile Gln Phe Asn Val Gly Phe Tyr Ile Val Trp Lys Tyr

195 200 205
Arg Asn Ile Pro Cys Tyr Arg Gln Asp Gly Met Arg Met Phe Gly Trp

210 215 220
Phe Phe Asn Tyr Phe Tyr Val Gly Thr Val Leu Cys Leu Phe Leu Asn

225 230 235 240
Phe Tyr Val Gln Thr Tyr Ile Val Arg Lys His Lys Gly Ala Lys Lys

245 250 255
Ile Gln

<210> 179

<211> 7769

<212> DNA

<213> Artificial Sequence

<220><223> Plasmid pZuFmEgD9ES-A21V/L35G/L108G/I179R

<400> 179

ggccgcaagt gtggatgggg aagtgagtgc ccggttctgt gtgcacaatt ggcaatccaa 60

gatggatgga ttcaacacag ggatatagcg agctacgtgg tgggtcggagg atatatgcaac 120

ggatatttat gtttgacact tgagaatgia cgatacaagc actgtccaag tacaatacta 180

aacatactgt acatactcat actcgtaccc gggcaacggt ttcacttgag tgcagtggct 240

agtgtcttta ctctacagt gtgcaatact gcgtatcata gtctttgatg tatatcgat 300

tcattcatgt tagttgcgta cgagccggaa gcataaagtg taaagcctgg ggtgcctaata 360

gagttagcta actcacatta attgcgttgc gctcactgcc cgctttccag tcgggaaacc 420

tgtcgtgcca gctgcattaa tgaatcggcc aacgcgcggg gagaggcggg ttgcgtattg 480

ggcgtctctc cgcttcctcg ctactgact cgctgcgctc ggtcgttcgg ctgcggcgag 540

cggtatcagc tcaactcaaag gcggaatac ggttatccac agaatacagg gataacgcag 600

gaaagaacat gtgagcaaaa ggccagcaaa aggccaggaa ccgtaaaaag gccgcgttgc 660

tggcggtttt ccataggctc cgccccctg acgagcatca caaaaatcga cgctcaagtc 720

agaggtagcg aaacccgaca ggactataaa gataccaggc gtttccccct ggaagctccc 780

tcgtgcgctc tctgtttccg accctgccgc ttaccggata cctgtccgcc tttctccctt 840

cgggaagcgt ggcgctttct catagctcac gctgtaggta tctcagttcg gtgtaggtcg 900

ttcgtccaa gctgggctgt gtgcacgaac ccccggtca gcccgaccgc tgcgccttat 960

ccggtacta tcgtcttgag tccaacccgg taagacacga cttatcgcca ctggcagcag 1020

ccactggtaa caggattagc agagcgaggt atgtaggcgg tgctacagag ttcttgaagt 1080

ggtggcctaa ctacggctac actagaagga cagtatttgg tatctgcgt ctgctgaagc 1140

cagttacctt cgaaaaaga gttgtagct ctgatccgg caaacaacc accgctggtta 1200

gcggtgggtt tttgtttgc aagcagcaga ttacgcgag aaaaaagga tctcaagaag 1260

atcctttgat cttttctac gggctctgacg ctacgtggaa cgaaaactca cgtaaggga 1320

ttttggtcat gagattatca aaaaggatct tcacctagat ctttttaaat taaaaatgaa 1380

gttttaaatc aatctaaagt atatatgagt aaacttggtc tgacagttac caatgcttaa 1440

tcagtaggc acctatctca gcgatctgtc tatttcgttc atccatagtt gcctgactcc 1500

ccgtcgtgta gataactacg atacgggagg gcttaccatc tggccccagt gctgcaatga 1560

taccgcgaga cccacgtca ccggctccag atttatcagc aataaaccag ccagccggaa 1620

gggccgagcg cagaagtggc cctgcaactt tatccgctc catccagtct attaattgtt 1680

gccgggaagc tagagtaagt agttcgccag ttaatatgtt gcgcaacgtt gttgccattg 1740

ctacaggcat cgtgggtgca cgctcgtcgt ttggtatggc ttcatcagc tccggttccc 1800

aacgatcaag gcgagttaca tgatcccca tgtgtgcaa aaaagcggtt agtccttcg 1860

gtcctccgat cgttgtcaga agtaagttgg ccgcagtggt atcactcatg gttatggcag 1920
 cactgcataa ttctcttact gtcatgcat ccgtaagatg cttttctgtg actggtgagt 1980
 actcaaccaa gtcattctga gaatagtgtg tgcggcgacc gagttgctct tgccccggcgt 2040
 caatacggga taataaccgc ccacatagca gaactttaaa agtgctcatc attggaaaac 2100
 gttcttcggg gcgaaaactc tcaaggatct taccgtgtt gagatccagt tcgatgtaac 2160
 ccaactcgtgc acccaactga tcttcagcat cttttacttt caccagcgtt tctgggtgag 2220
 caaaaacagg aaggcaaaat gccgcaaaaa agggaataag ggcgacacgg aaatgttgaa 2280

tactcatact cttccttttt caatattatt gaagcattta tcagggttat tgtctcatga 2340
 gcggatacat atttgaatgt atttagaaaa ataaacaaat aggggttccg cgcacatttc 2400
 cccgaaaagt gccacctgac gcgccctgtg gcggcgcat aagcgcggcg ggtgtggtgg 2460
 ttacgcgcag cgtgaccgt acacttgcca gcgccctagc gcccgctcct ttcgctttct 2520
 tcccttctt tctcgccacg ttccgccggt tccccgtca agctctaaat cgggggctcc 2580
 ctttaggggt cggatttagt gctttacggc acctcgacc caaaaaactt gattaggggtg 2640
 atgggtcacg tagtgggcca tcgccctgat agacggtttt tcgccctttg acgttggtg 2700

ccacgttctt taatagtga ctcttgttcc aaactggaac aacactcaac cctatctcgg 2760
 tctattcttt tgatttataa gggattttgc cgatttcggc ctattggtta aaaaatgagc 2820
 tgatttaaca aaaatttaac gcgaatttta acaaaatatt aacgcttaca atttcattc 2880
 gccattcagg ctgcgcaact gttgggaagg gcgatcgggt cgggcctctt cgctattacg 2940
 ccagctggcg aaagggggat gtgctgcaag gcgattaagt tgggtaacgc cagggttttc 3000
 ccagtcacga cgttgtaaaa cgacggccag tgaattgtaa tacgactcac tataggcgca 3060
 attgggtacc gggccccccc tcgaggtcga tgggtgtgat aagcttgata tcgaattcat 3120

gtcacacaaa ccgactctcg cctcaaggaa acctaatct acatccgaga gactgccgag 3180
 atccagtcta cactgattaa ttttcgggcc aataatttaa aaaaatcgtg ttataataa 3240
 ttatatgtat tatatatata catcatgatg atactgacag tcatgtccca ttgctaaata 3300
 gacagactcc atctgccgc tccaactgat gttctcaata ttttaagggt catctcgcat 3360
 tgtttaataa taaacagact ccatctaccg cctccaaatg atgttctcaa aatataattgt 3420
 atgaacttat ttttattact tagtattatt agacaactta cttgctttat gaaaaacact 3480
 tcctatttag gaaacaattt ataattggcag ttcgttcatt taacaattta tgtagaataa 3540

atgttataaa tgcgtatggg aaactctaaa tatggatagc ataaatgata tctgcattgc 3600
 ctaattcgaa atcaacagca acgaaaaaaa tcccttgtac aacataaata gtcacgaga 3660
 aatatcaact atcaaagaac agctattcac acgttactat tgagattatt attggacgag 3720

aatcacacac tcaactgtct ttctctcttc tagaaatata ggtacaagta tgtactattc	3780
tcattgttca tacttttagt catttcatcc cacatatcc ttggatttct ctccaatgaa	3840
tgacattcta tcttgcaaat tcaacaatta taataagata taccaaagta gcggtatagt	3900
ggcaatcaaa aagcttctct ggtgtgcttc tcgtatttat ttttattcta atgatccatt	3960
aaaggtatat atttatttct tgttatataa tccttttgtt tattacatgg gctggataca	4020
taaaggtatt ttgatttaat tttttgctta aattcaatcc cccctcgctc agtgtcaact	4080
glaatggtag gaaattacca tacttttgaa gaagcaaaaa aaatgaaaga aaaaaaaaaat	4140
cgtatttcca ggtagacgt tccgcagaat ctagaatgcg gtatgcggta cattgttctt	4200
cgaacgtaaa agttgcgctc cctgagatat tglacatttt tgcttttaca agtacaagta	4260
catcgtaaa ctatgtacta ctgttgatgc atccacaaca gtttgttttg ttttttttg	4320
tttttttttt ttctaattgat tcattaccgc tatgtatacc tacttgtact tgtagtaagc	4380
cgggttattg gcgttcaatt aatcatagac ttatgaatct gcacgggtg cgctgcgagt	4440
tacttttagc ttatgcatgc tacttgggtg taatatggg atctgttcgg aaatcaacgg	4500
atgtcaate gatttcgaca gtaattaatt aagtcataca caagtcagct ttcttcgagc	4560
ctcatataag tataagtagt tcaacgtatt agcactgtac ccagcatctc cgtatcgaga	4620
aacacaacaa catgccccat tggacagatc atgcggatac acaggtttgt cagtatcata	4680
catactcatg cagacaggct gtctgacat catacaagct gaacaagcgc tccatacttg	4740
cacgtctct atatacacag ttaaattaca tatccatagt ctaacctcta acagttaatc	4800
ttctggttaag cctcccagcc agccttctgg tatcgcttgg cctcctcaat aggatctcgg	4860
ttctggccgt acagacctcg gccgacaatt atgatatccg ttccggtaga catgacatcc	4920
tcaacagttc ggtactgctg tccgagagcg tctcccttgt cgtcaagacc caccctgggg	4980
gtcagaataa gccagtcctc agagtcgcc ttaggtcggt tctgggcaat gaagccaacc	5040
acaaactcgg ggtcggatcg ggcaagctca atggtctgct tggagtactc gccagtggcc	5100
agagagccct tgcaagacag ctcgccagc atgagcagac ctctggccag ctctctgttg	5160
ggagagggga ctaggaactc cttgtactgg gagttctcgt agtcagagac gtctccttc	5220
ttctgttcag agacagtttc ctcgccacca gctcgaggc cagcaatgat tccggttccg	5280
ggtacaccgt gggtgttgt gatatcggac cactcgcgga ttcggtgaca ccggtactgg	5340
tgcttgacag tgttgccaat atctgcgaac ttctgtcct cgaacaggaa gaaaccgtgc	5400
ttaagagcaa gttccttgag ggggagcaca gtgccgcgt aggtgaagtc gtcaatgatg	5460
tcgatatggg ttttgatcat gcacacataa ggtccgacct tatcggaag ctcaatgagc	5520
tccttgggtg tggtaacatc cagagaagca cacaggttg tttcttggc tgccacgagc	5580

ttgagcactc gagcggcaaa ggcggacttg tggacgttag ctcgagcttc gtaggagggc	5640
atTTTggtgg tgaagaggag actgaaataa atttagtctg cagaactttt tATcggaacc	5700
ttatctgggg cagtgaagta tatgttatgg taatagttag gagttagtTg aacttataga	5760
tagactggac tatacggcta tcggtccaaa ttgaaagaa cgtcaatggc tctctgggcg	5820
tcgcctttgc cgacaaaaat gtgatcatga tgaaagccag caatgacgtt gcagctgata	5880
ttgttgtcgg ccaaccgcgc cgaaaacgca gctgtcagac ccacagcctc caacgaagaa	5940
tgtatcgtca aagtgatcca agcacactca tagttggagt cgtactccaa aggcggcaat	6000
gacgagtcag acagatactc gtcgacgttt aaacagtgtg cgcagatcta ctatagagga	6060
acatttAAAT tgccccggag aagacggcca ggccgcctag atgacaaatt caacaactca	6120
cagctgactt tctgccattg ccactagggg ggggcctttt tataTggcca agccaagctc	6180
tccacgtcgg ttgggctgca cccaacaata aatgggtagg gttgcaccaa caaagggatg	6240
ggatgggggg tagaagatag gaggataacg gggctcaatg gcacaaataa gaacgaatac	6300
tgccattaag actcgtgac cagcgactga caccattgca tcatctaagg gcctcaaaac	6360
tacctcggaA ctgtcgcgt gatctggaca ccacagaggt tccgagcact ttaggttgca	6420
ccaaatgtcc caccaggtgc aggcagaaaa cgctggaaca gcgtgtacag tttgtcttaa	6480
caaaaagtga gggcgctgag gtcgagcagg gtggtgtgac ttgttatagc ctttagagct	6540
gcgaaagcgc gtatggattt ggctcatcag gccagattga gggTctgtgg acacatgtca	6600
tgttagtgta cttaacTgc cccctggata tagccccgac aataggccgt ggcctcattt	6660
ttttgccttc cgcacatttc cattgtctga taccacacc ttgtctctcc tgcacttgcc	6720
aaccttaata ctggtttaca ttgaccaaca tcttacaagc ggggggcttg tctagggtat	6780
atataaacag tggtctctcc aatcggttgc cagtctcttt tttctttct tccccacag	6840
attcgaaatc taaactacac atcacagaat tccgagccgt gagtatccac gacaagatca	6900
gtgtcgagac gacgcgtttt gtgtaatgac acaatccgaa agtcgctagc aacacacact	6960
ctctacacaa actaaccag ctctggtagc atggaggtcg tgaacgaaat cgtctccatt	7020
ggccaggagg ttcttcccaa ggtcgactat gtgcagctct ggtctgatgc ctgcactgc	7080
gaggtgctgt acgggtccat cgccttcgtc atcctgaagt tcacccttgg tcctctcgga	7140
cccaagggtc agtctcgaat gaagtttgtg ttaccaact acaacctgct catgtccatc	7200
tactcgtgg gctcttctct ctctatggcc tacgcatgt acaccattgg tgtcatgtcc	7260
gacaactcgc agaaggcttt cgacaacaat gtcttccgaa tcaccactca gggtttctac	7320

ctcagcaagt tctcagagta cattgactcc ttctatctgc cctcatggg caagcctctg 7380
 acctggttgc agttctttca ccatactcgga gctcctatgg acatgtggct gttctacaac 7440
 taccgaaacg aagccgtttg gatctttgtg ctgctcaacg gcttcattca ctggatcatg 7500
 tacggctact attggacccg actgcgaaaag ctcaagtcc ctatgcccaa gtccctgatt 7560
 acttctatgc agatcattca gttcaacgtt ggcttctaca tcgtctggaa gtaccggaac 7620
 attccctgct accgacaaga tggaatgaga atgtttggct ggtttttcaa ctacttctac 7680
 gttggtactg tctgtgtct gttcctcaac ttctacgtgc agacctacat cgtccgaaag 7740

cacaagggag ccaaaaagat tcagtgagc 7769

<210> 180

<211> 4

<212> PRT

<213> Artificial Sequence

<220><223> HXGX motif

<220><221> misc_feature

<222> (2)..(2)

<223> Xaa can be any naturally occurring amino acid

<220><221> misc_feature

<222> (4)..(4)

<223> Xaa can be any naturally occurring amino acid

<400> 180

His Xaa Gly Xaa

1

<210> 181

<211> 4

<212> PRT

<213> Artificial Sequence

<220><223> HPGG motif

<400>

> 181

His Pro Gly Gly

1

<210> 182

<211> 5

<212> PRT

<213> Artificial Sequence

<220><223> HXXXH motif

<220><221> misc_feature

<222> (2)..(4)

<223> Xaa can be any naturally occurring amino acid

<400> 182

His Xaa Xaa Xaa His

1 5

<210> 183

<211> 5

<212> PRT

<213> Artificial Sequence

<220><223> HDASH motif

<400> 183

His Asp Ala Ser His

1 5

<210> 184

<211> 1350

<212> DNA

<213> Euglena gracilis

<

220><221> CDS

<222> (1)..(1350)

<223> delta-5 desaturase

<300><302> DELTA-5 DESATURASE AND ITS USE IN MAKING POLYUNSATURATED FATTY

ACIDS

<310> WO 2007/136671

<311> 2007-05-16

<312> 2007-11-29

<313> (1)..(1350)

<300><302> DELTA-5 DESATURASE AND ITS USE IN MAKING POLYUNSATURATED FATTY

ACIDS

<310> US 7,678,560

<311> 2007-05-15

<312> 2010-03-16

<313> (1)..(1350)

<400> 184

atg gct ctc agt ctt acc aca gaa cag ctg tta gaa cgc cct gat ttg 48

Met Ala Leu Ser Leu Thr Thr Glu Gln Leu Leu Glu Arg Pro Asp Leu

1 5 10 15
gtt gcg att gat ggc atc ctc tac gac ctt gaa ggg ctt gcc aaa gtt 96

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

20 25 30
cat cca gga gga gat ttg att ctc gct tct ggt gcc tct gat gcc tcc 144

His Pro Gly Gly Asp Leu Ile Leu Ala Ser Gly Ala Ser Asp Ala Ser

35 40 45

cct ctc ttttat tca atg cat cca tac gtc aaa ccg gag aat tcc aaa 192

Pro Leu Phe Tyr Ser Met His Pro Tyr Val Lys Pro Glu Asn Ser Lys

50 55 60
ttg ctt caa cag ttc gtc cga ggg aag cat gac cgc acc tcg aag gac 240

Leu Leu Gln Gln Phe Val Arg Gly Lys His Asp Arg Thr Ser Lys Asp

65 70 75 80
att gtc tac acg tat gat tct ccc ttc gca caa gac gtt aag cgg aca 288

Ile Val Tyr Thr Tyr Asp Ser Pro Phe Ala Gln Asp Val Lys Arg Thr

85 90 95
atg cgc gag gtg atg aaa ggg agg aac tgg tac gca acc cct ggc ttc 336

Met Arg Glu Val Met Lys Gly Arg Asn Trp Tyr Ala Thr Pro Gly Phe

100 105 110
tgg ctg cgc acc gtt ggg atc atc gcc gtg acg gcc ttt tgc gag tgg 384

Trp Leu Arg Thr Val Gly Ile Ile Ala Val Thr Ala Phe Cys Glu Trp

115 120 125
cac tgg gct acc acg ggg atg gtg ctg tgg ggc ctg ttg act gga ttc 432

His Trp Ala Thr Thr Gly Met Val Leu Trp Gly Leu Leu Thr Gly Phe

130 135 140
atg cac atg cag atc ggc tta tcc atc cag cat gat gcg tcc cac ggg 480

Met His Met Gln Ile Gly Leu Ser Ile Gln His Asp Ala Ser His Gly

145	150	155	160	
gcc atc agc aag aag cct tgg gtc aac gcc ctc ttc gcc tac ggc att				528
Ala Ile Ser Lys Lys Pro Trp Val Asn Ala Leu Phe Ala Tyr Gly Ile				
	165	170	175	
gac gtc atc gga tcg tcc cgg tgg att tgg ctg cag tcg cac atc atg				576
Asp Val Ile Gly Ser Ser Arg Trp Ile Trp Leu Gln Ser His Ile Met				
	180	185	190	
cgg cac cac acc tac acc aac cag cac ggc ctc gac ctg gat gcg gag				624
Arg His His Thr Tyr Thr Asn Gln His Gly Leu Asp Leu Asp Ala Glu				
	195	200	205	
tcg gca gag ccg ttc ctg gtg ttc cac aac tac ccc gcc gca aac acc				672
Ser Ala Glu Pro Phe Leu Val Phe His Asn Tyr Pro Ala Ala Asn Thr				
	210	215	220	
gcc cga aag tgg ttc cac cgc ttc caa gct tgg tac atg tac ctt gtg				720
Ala Arg Lys Trp Phe His Arg Phe Gln Ala Trp Tyr Met Tyr Leu Val				
225	230	235	240	
ctg ggg gca tac ggg gta tcg ctg gtg tac aac ccg ctc tac att ttc				768
Leu Gly Ala Tyr Gly Val Ser Leu Val Tyr Asn Pro Leu Tyr Ile Phe				
	245	250	255	
cgg atg cag cac aat gac acc atc cca gag tct gtc acg gcc atg cgg				816
Arg Met Gln His Asn Asp Thr Ile Pro Glu Ser Val Thr Ala Met Arg				
	260	265	270	
gag aat ggc ttt ctg cgg cgc tac cgc aca ctt gca ttc gtg atg cga				864
Glu Asn Gly Phe Leu Arg Arg Tyr Arg Thr Leu Ala Phe Val Met Arg				
	275	280	285	
gct ttc ttc atc ttc cgg acc gca ttc ttg ccc tgg tac ctc act ggg				912
Ala Phe Phe Ile Phe Arg Thr Ala Phe Leu Pro Trp Tyr Leu Thr Gly				
	290	295	300	
acc tca ttg ctg atc acc att cct ctg gtg ccc act gca act ggt gcc				960

Thr Ser Leu Leu Ile Thr Ile Pro Leu Val Pro Thr Ala Thr Gly Ala
305 310 315 320
ttc ttg acg ttc ttc ttc att ttg tcc cac aat ttt gat ggc tcc gaa 1008
Phe Leu Thr Phe Phe Phe Ile Leu Ser His Asn Phe Asp Gly Ser Glu
325 330 335
cgg atc ccc gac aag aac tgc aag gtt aag agc tct gag aag gac gtt 1056
Arg Ile Pro Asp Lys Asn Cys Lys Val Lys Ser Ser Glu Lys Asp Val
340 345 350
gag gct gac caa att gac tgg tat cgg gcg cag gtg gag acg tcc tcc 1104
Glu Ala Asp Gln Ile Asp Trp Tyr Arg Ala Gln Val Glu Thr Ser Ser
355 360 365
aca tac ggt ggc ccc atc gcc atg ttc ttc act ggc ggt ctc aat ttc 1152
Thr Tyr Gly Gly Pro Ile Ala Met Phe Phe Thr Gly Gly Leu Asn Phe
370 375 380
cag atc gag cac cac ctc ttt ccc cgg atg tcg tct tgg cac tac ccc 1200
Gln Ile Glu His His Leu Phe Pro Arg Met Ser Ser Trp His Tyr Pro
385 390 395 400
ttc gtc cag cag gcg gtc cgg gag tgt tgc gaa cgc cat gga gtg cga 1248
Phe Val Gln Gln Ala Val Arg Glu Cys Cys Glu Arg His Gly Val Arg
405 410 415
tat gtt ttc tac cct acc atc gtc ggc aac atc atc tcc acc ctg aag 1296
Tyr Val Phe Tyr Pro Thr Ile Val Gly Asn Ile Ile Ser Thr Leu Lys
420 425 430
tac atg cat aag gtg ggt gtc gtc cac tgc gtg aag gac gca cag gat 1344
Tyr Met His Lys Val Gly Val Val His Cys Val Lys Asp Ala Gln Asp
435 440 445
tcc tga 1350
Ser

<210> 185

<211> 449

<212> PRT

<213> *Euglena gracilis*

<400> 185

Met Ala Leu Ser Leu Thr Thr Glu Gln Leu Leu Glu Arg Pro Asp Leu

1 5 10 15

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

20 25 30

His Pro Gly Gly Asp Leu Ile Leu Ala Ser Gly Ala Ser Asp Ala Ser

35 40 45

Pro Leu Phe Tyr Ser Met His Pro Tyr Val Lys Pro Glu Asn Ser Lys

50 55 60

Leu Leu Gln Gln Phe Val Arg Gly Lys His Asp Arg Thr Ser Lys Asp

65 70 75 80

Ile Val Tyr Thr Tyr Asp Ser Pro Phe Ala Gln Asp Val Lys Arg Thr

85 90 95

Met Arg Glu Val Met Lys Gly Arg Asn Trp Tyr Ala Thr Pro Gly Phe

100 105 110

Trp Leu Arg Thr Val Gly Ile Ile Ala Val Thr Ala Phe Cys Glu Trp

115 120 125

His Trp Ala Thr Thr Gly Met Val Leu Trp Gly Leu Leu Thr Gly Phe

130 135 140

Met His Met Gln Ile Gly Leu Ser Ile Gln His Asp Ala Ser His Gly

145 150 155 160

Ala Ile Ser Lys Lys Pro Trp Val Asn Ala Leu Phe Ala Tyr Gly Ile

165 170 175

Asp Val Ile Gly Ser Ser Arg Trp Ile Trp Leu Gln Ser His Ile Met

180 185 190

Arg His His Thr Tyr Thr Asn Gln His Gly Leu Asp Leu Asp Ala Glu

195 200 205

Ser Ala Glu Pro Phe Leu Val Phe His Asn Tyr Pro Ala Ala Asn Thr

210 215 220

Ala Arg Lys Trp Phe His Arg Phe Gln Ala Trp Tyr Met Tyr Leu Val

225	230							235							240						
Leu	Gly	Ala	Tyr	Gly	Val	Ser	Leu	Val	Tyr	Asn	Pro	Leu	Tyr	Ile	Phe						
				245							250							255			
Arg	Met	Gln	His	Asn	Asp	Thr	Ile	Pro	Glu	Ser	Val	Thr	Ala	Met	Arg						
				260							265							270			
Glu	Asn	Gly	Phe	Leu	Arg	Arg	Tyr	Arg	Thr	Leu	Ala	Phe	Val	Met	Arg						
				275							280							285			
Ala	Phe	Phe	Ile	Phe	Arg	Thr	Ala	Phe	Leu	Pro	Trp	Tyr	Leu	Thr	Gly						
				290							295							300			
Thr	Ser	Leu	Leu	Ile	Thr	Ile	Pro	Leu	Val	Pro	Thr	Ala	Thr	Gly	Ala						
305							310							315			320				
Phe	Leu	Thr	Phe	Phe	Phe	Ile	Leu	Ser	His	Asn	Phe	Asp	Gly	Ser	Glu						
				325							330							335			
Arg	Ile	Pro	Asp	Lys	Asn	Cys	Lys	Val	Lys	Ser	Ser	Glu	Lys	Asp	Val						
				340							345							350			
Glu	Ala	Asp	Gln	Ile	Asp	Trp	Tyr	Arg	Ala	Gln	Val	Glu	Thr	Ser	Ser						
				355							360							365			
Thr	Tyr	Gly	Gly	Pro	Ile	Ala	Met	Phe	Phe	Thr	Gly	Gly	Leu	Asn	Phe						
370							375							380							
Gln	Ile	Glu	His	His	Leu	Phe	Pro	Arg	Met	Ser	Ser	Trp	His	Tyr	Pro						
385							390							395			400				
Phe	Val	Gln	Gln	Ala	Val	Arg	Glu	Cys	Cys	Glu	Arg	His	Gly	Val	Arg						
				405							410							415			
Tyr	Val	Phe	Tyr	Pro	Thr	Ile	Val	Gly	Asn	Ile	Ile	Ser	Thr	Leu	Lys						
				420							425							430			
Tyr	Met	His	Lys	Val	Gly	Val	Val	His	Cys	Val	Lys	Asp	Ala	Gln	Asp						
435							440							445							
Ser																					

<210> 186

<211> 5

<212> PRT
 <213> Artificial Sequence
 <220><223> His-rich motif
 <220><221> misc_feature
 <222> (2)..(4)
 <223> Xaa can be any naturally occurring amino acid
 <400> 186
 His Xaa Xaa Xaa His
 1 5
 <210> 187
 <211> 6
 <212> PRT
 <213> Artificial Sequence
 <220><223> His-rich motif
 <220><221> misc_feature

 <222> (2)..(5)
 <223> Xaa can be any naturally occurring amino acid
 <400> 187
 His Xaa Xaa Xaa Xaa His
 1 5
 <210> 188
 <211> 5
 <212> PRT
 <213> Artificial Sequence
 <220><223> His-rich motif
 <220><221> misc_feature
 <222> (2)..(3)
 <223> Xaa can be any naturally occurring amino acid
 <400> 188
 His Xaa Xaa His His
 1 5
 <210> 189
 <211> 6
 <212> PRT

<213> Artificial Sequence
 <220><223> His-rich motif
 <220><221> misc_feature

 <222> (2)..(4)
 <223> Xaa can be any naturally occurring amino acid
 <400> 189
 His Xaa Xaa Xaa His His
 1 5
 <210> 190
 <211> 5
 <212> PRT
 <213> Artificial Sequence
 <220><223> His-rich motif
 <220><221> MISC_FEATURE
 <222> (1)..(1)
 <223> Xaa = Gln [Q] or His [H]
 <220><221> misc_feature
 <222> (2)..(3)
 <223> Xaa can be any naturally occurring amino acid
 <400> 190
 Xaa Xaa Xaa His His
 1 5
 <210> 191
 <211> 6
 <212> PRT
 <213>
 Artificial Sequence
 <220><223> His-rich motif
 <220><221> MISC_FEATURE
 <222> (1)..(1)
 <223> Xaa = Gln [Q] or His [H]
 <220><221> misc_feature
 <222> (2)..(4)
 <223> Xaa can be any naturally occurring amino acid

<400> 191

Xaa Xaa Xaa Xaa His His

1 5

<210> 192

<211> 1350

<212> DNA

<213> *Euglena gracilis*

<220><221> CDS

<222> (1)..(1350)

<223> delta-5 desaturase ("EgD5R")

<400> 192

atg gct ctc agt ctt acc aca gaa cag ctg tta gaa cgc cct gat ttg 48

Met Ala Leu Ser Leu Thr Thr Glu Gln Leu Leu Glu Arg Pro Asp Leu

1 5 10 15

gtt gcg att gat ggc atc ctc tac gac ctt gaa ggg ctt gcc aaa gtt 96

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

20 25 30

cat cca gga gga gat ttg att ctc gct tct ggt gcc tct gat gcc tcc 144

His Pro Gly Gly Asp Leu Ile Leu Ala Ser Gly Ala Ser Asp Ala Ser

35 40 45

cct ctc ttt tat tca atg cat cca tac gtc aaa ccg gag aat tcc aaa 192

Pro Leu Phe Tyr Ser Met His Pro Tyr Val Lys Pro Glu Asn Ser Lys

50 55 60

ttg ctt caa cag ttc gtc cga ggg aag cat gac cgc acc tcg aag gac 240

Leu Leu Gln Gln Phe Val Arg Gly Lys His Asp Arg Thr Ser Lys Asp

65 70 75 80

att gtc tac acg tat gat tct ccc ttc gca caa gac gtt aag cgg aca 288

Ile Val Tyr Thr Tyr Asp Ser Pro Phe Ala Gln Asp Val Lys Arg Thr

85 90 95

atg cgc gag gtg atg aaa ggg agg aac tgg tac gca acc cct ggc ttc 336

Met Arg Glu Val Met Lys Gly Arg Asn Trp Tyr Ala Thr Pro Gly Phe

100 105 110

tgg ctg cgc acc gtt ggg atc atc gcc gtg acg gcc ttt tgc gag tgg	384
Trp Leu Arg Thr Val Gly Ile Ile Ala Val Thr Ala Phe Cys Glu Trp	
115 120 125	
cac tgg gct acc acg ggg atg gtg ctg tgg ggc ctg ttg act gga ttc	432
His Trp Ala Thr Thr Gly Met Val Leu Trp Gly Leu Leu Thr Gly Phe	
130 135 140	
atg cac atg cag atc ggc tta tcc atc cag cat gat gcg tcc cac ggg	480
Met His Met Gln Ile Gly Leu Ser Ile Gln His Asp Ala Ser His Gly	
145 150 155 160	
gcc atc agc aag aag cct tgg gtc aac gcc ctc ttc gcc tac ggc att	528
Ala Ile Ser Lys Lys Pro Trp Val Asn Ala Leu Phe Ala Tyr Gly Ile	
165 170 175	
gac gtc atc gga tcg tcc cgg tgg att tgg ctg cag tcg cac atc atg	576
Asp Val Ile Gly Ser Ser Arg Trp Ile Trp Leu Gln Ser His Ile Met	
180 185 190	
cgg cac cac acc tac acc aac cag cac ggc ctc gac ctg gat gcg gag	624
Arg His His Thr Tyr Thr Asn Gln His Gly Leu Asp Leu Asp Ala Glu	
195 200 205	
tcg gca gag ccg ttc ctg gtg ttc cac aac tac ccc gcc gca aac acc	672
Ser Ala Glu Pro Phe Leu Val Phe His Asn Tyr Pro Ala Ala Asn Thr	
210 215 220	
gcc cga aag tgg ttc cac cgc ttc caa gct tgg tac atg tac ctt gtg	720
Ala Arg Lys Trp Phe His Arg Phe Gln Ala Trp Tyr Met Tyr Leu Val	
225 230 235 240	
ctg ggg gca tac ggg gta tcg ctg gtg tac aac ccg ctc tac att ttc	768
Leu Gly Ala Tyr Gly Val Ser Leu Val Tyr Asn Pro Leu Tyr Ile Phe	
245 250 255	
cgg atg cag cac aat gac acc atc cca gag tct gtc acg gcc atg cgg	816
Arg Met Gln His Asn Asp Thr Ile Pro Glu Ser Val Thr Ala Met Arg	

260	265	270	
gaa aat ggc ttt ctg cgg cgc tac cgc aca ctt gca ttc gtg atg cga			864
Glu Asn Gly Phe Leu Arg Arg Tyr Arg Thr Leu Ala Phe Val Met Arg			
275	280	285	
gct ttc ttc atc ttc cgg acc gca ttc ttg ccc tgg tac ctc act ggg			912
Ala Phe Phe Ile Phe Arg Thr Ala Phe Leu Pro Trp Tyr Leu Thr Gly			
290	295	300	
acc tca ttg ctg atc acc att cct ctg gtg ccc acc gca act ggt gcc			960
Thr Ser Leu Leu Ile Thr Ile Pro Leu Val Pro Thr Ala Thr Gly Ala			
305	310	315	320
ttc ttg acg ttc ttc ttc att ttg tcc cac aat ttt gat ggc tcc gaa			1008
Phe Leu Thr Phe Phe Phe Ile Leu Ser His Asn Phe Asp Gly Ser Glu			
325	330	335	
cgg atc ccc gac aag aac tgc aag gtt aag aga tct gag aag gac gtt			1056
Arg Ile Pro Asp Lys Asn Cys Lys Val Lys Arg Ser Glu Lys Asp Val			
340	345	350	
gag gct gac caa att gac tgg tat cgg gcg cag gtg gag acg tcc tcc			1104
Glu Ala Asp Gln Ile Asp Trp Tyr Arg Ala Gln Val Glu Thr Ser Ser			
355	360	365	
aca tac ggt ggc ccc atc gcc atg ttc ttc act ggc ggt ctc aat ttc			1152
Thr Tyr Gly Gly Pro Ile Ala Met Phe Phe Thr Gly Gly Leu Asn Phe			
370	375	380	
cag atc gag cac cac ctc ttt ccc cgg atg tcg tct tgg cac tac ccc			1200
Gln Ile Glu His His Leu Phe Pro Arg Met Ser Ser Trp His Tyr Pro			
385	390	395	400
ttc gtc cag cag gcg gtc cgg gag tgt tgc gaa cgc cat gga gtg cga			1248
Phe Val Gln Gln Ala Val Arg Glu Cys Cys Glu Arg His Gly Val Arg			
405	410	415	
tat gtt ttc tac cct acc atc gtc ggc aac atc atc tcc acc ctg aag			1296
Tyr Val Phe Tyr Pro Thr Ile Val Gly Asn Ile Ile Ser Thr Leu Lys			
420	425	430	

tac atg cat aag gtg ggt gtc gtc cac tgc gtg aag gac gca cag gat 1344
Tyr Met His Lys Val Gly Val Val His Cys Val Lys Asp Ala Gln Asp
435 440 445
tcc taa 1350

Ser

<210> 193

<211> 449

<212> PRT

<213> Euglena gracilis

<400> 193

Met Ala Leu Ser Leu Thr Thr Glu Gln Leu Leu Glu Arg Pro Asp Leu
1 5 10 15
Val Ala Ile Asp Gly Ile Leu Tyr Asp Leu Glu Gly Leu Ala Lys Val
20 25 30
His Pro Gly Gly Asp Leu Ile Leu Ala Ser Gly Ala Ser Asp Ala Ser
35 40 45
Pro Leu Phe Tyr Ser Met His Pro Tyr Val Lys Pro Glu Asn Ser Lys
50 55 60
Leu Leu Gln Gln Phe Val Arg Gly Lys His Asp Arg Thr Ser Lys Asp
65 70 75 80
Ile Val Tyr Thr Tyr Asp Ser Pro Phe Ala Gln Asp Val Lys Arg Thr
85 90 95
Met Arg Glu Val Met Lys Gly Arg Asn Trp Tyr Ala Thr Pro Gly Phe
100 105 110
Trp Leu Arg Thr Val Gly Ile Ile Ala Val Thr Ala Phe Cys Glu Trp
115 120 125
His Trp Ala Thr Thr Gly Met Val Leu Trp Gly Leu Leu Thr Gly Phe
130 135 140
Met His Met Gln Ile Gly Leu Ser Ile Gln His Asp Ala Ser His Gly
145 150 155 160

Ala Ile Ser Lys Lys Pro Trp Val Asn Ala Leu Phe Ala Tyr Gly Ile

165 170 175

Asp Val Ile Gly Ser Ser Arg Trp Ile Trp Leu Gln Ser His Ile Met

180 185 190

Arg His His Thr Tyr Thr Asn Gln His Gly Leu Asp Leu Asp Ala Glu

195 200 205

Ser Ala Glu Pro Phe Leu Val Phe His Asn Tyr Pro Ala Ala Asn Thr

210 215 220

Ala Arg Lys Trp Phe His Arg Phe Gln Ala Trp Tyr Met Tyr Leu Val

225 230 235 240

Leu Gly Ala Tyr Gly Val Ser Leu Val Tyr Asn Pro Leu Tyr Ile Phe

245 250 255

Arg Met Gln His Asn Asp Thr Ile Pro Glu Ser Val Thr Ala Met Arg

260 265 270

Glu Asn Gly Phe Leu Arg Arg Tyr Arg Thr Leu Ala Phe Val Met Arg

275 280 285

Ala Phe Phe Ile Phe Arg Thr Ala Phe Leu Pro Trp Tyr Leu Thr Gly

290 295 300

Thr Ser Leu Leu Ile Thr Ile Pro Leu Val Pro Thr Ala Thr Gly Ala

305 310 315 320

Phe Leu Thr Phe Phe Phe Ile Leu Ser His Asn Phe Asp Gly Ser Glu

325 330 335

Arg Ile Pro Asp Lys Asn Cys Lys Val Lys Arg Ser Glu Lys Asp Val

340 345 350

Glu Ala Asp Gln Ile Asp Trp Tyr Arg Ala Gln Val Glu Thr Ser Ser

355 360 365

Thr Tyr Gly Gly Pro Ile Ala Met Phe Phe Thr Gly Gly Leu Asn Phe

370 375 380

Gln Ile Glu His His Leu Phe Pro Arg Met Ser Ser Trp His Tyr Pro

385 390 395 400

Phe Val Gln Gln Ala Val Arg Glu Cys Cys Glu Arg His Gly Val Arg

405 410 415
Tyr Val Phe Tyr Pro Thr Ile Val Gly Asn Ile Ile Ser Thr Leu Lys

420 425 430
Tyr Met His Lys Val Gly Val Val His Cys Val Lys Asp Ala Gln Asp

435 440 445
Ser

<210> 194

<211> 8438

<212> DNA

<213> Artificial Sequence

<220><223> Plasmid pDMW367

<400> 194

catggctctc agtcttacc aagaacagct gttagaacgc cctgatttgg ttgcgattga 60
tggcatcctc tacgaccttg aagggtctgc caaagttcat ccaggaggag atttgattct 120
cgcttctggt gcctctgatg cctccctctt cttttattca atgcatccat acgtcaaacc 180

ggagaattcc aaattgcttc aacagttcgt ccgaggaag catgaccgca cctcgaagga 240
cattgtctac acgtatgatt ctcccttcgc acaagacgtt aagcggacaa tgcgcgaggt 300
gatgaaaggg aggaactggt acgcaacccc tggtctctgg ctgcgcaccg ttgggatcat 360
cgccgtgacg gccttttgcg agtggcactg ggctaccacg gggatggtgc tgtggggcct 420
gttgactgga ttcatgcaca tgcagatcgg cttatccatc cagcatgatg cgtcccacgg 480
ggccatcagc aagaagcctt gggatcaacgc cctcttcgcc tacggcattg acgtcatcgg 540
atcgtcccgg tggatttggc tgcagtcgca catcatgcgg caccacacct acaccaacca 600

gcacggcctc gacctggatg cggagtcggc agagccgttc ctggtgttcc acaactacc 660
cgccgcaaac accgcccga agtggttcca ccgcttccaa gcttgggtaca tgtaccttgt 720
gctgggggca tacgggggtat cgctggtgta caaccgctc tacattttcc ggatgcagca 780
caatgacacc atcccagagt ctgtcacggc catgcgggaa aatggcttcc tgcggcgcta 840
ccgcacactt gcattcgtga tgcgagcttt cttcatcttc cggaccgcat tcttgccttg 900
gtacctcact gggacctcat tgctgatcac cattcctctg gtgcccaccg caactgggtgc 960
cttcttgacg ttcttttcca ttttgtccca caattttgat ggctccgaac ggatccccga 1020

caagaactgc aagggttaaga gatctgagaa ggacgttgag gctgacaaa ttgactggta 1080

tcgggcgcag gtggagacgt cctccacata cggtaggcccc atcgccatgt tcttcactgg	1140
cgggtctcaat ttccagatcg agcaccacct ctttccccgg atgtcgtctt ggcactaccc	1200
cttcgtccag caggcgggtcc gggagtgttg cgaacgccat ggagtgcgat atgttttcta	1260
ccctaccatc gtcggcaaca tcatctccac cctgaagtac atgcataagg tgggtgtcgt	1320
ccactgcgtg aaggacgcac aggattccta agcggccgca agtgtggatg gggaagtgag	1380
tgcccggttc tgtgtgcaca attggcaatc caagatggat ggattcaaca cagggatata	1440
gcgagctacg tggtaggtgcg aggatatagc aacggatatt tatgtttgac acttgagaat	1500
gtacgataca agcactgtcc aagtacaata ctaaacatac tgtacatact catactcgta	1560
ccccggcaac ggtttcactt gagtgcagtg gctagtgtctc ttactcgtac agtgtgcaat	1620
actgcgtatc atagtctttg atgtatatcg tattcattca tgttagttgc gtacgagccg	1680
gaagcataaa gtgtaaagcc tggggtgcct aatgagtgag ctaactcaca ttaattgcgt	1740
tgcgtcact gcccgtttc cagtcgggaa acctgtcgtg ccagctgcat taatgaatcg	1800
gccaacgcgc ggggagaggc ggtttgcgta ttgggcgctc ttccgcttcc tcgctcactg	1860
actcgtcgcg ctcggtcgtt cggtcgcggc gagcggatc agtcactca aaggcggtaa	1920
tacggttatc cacagaatca ggggataacg caggaaagaa catgtgagca aaaggccagc	1980
aaaaggccag gaaccgtaaa aaggccgcgt tgcgtggcgtt tttccatagg ctccgcccc	2040
ctgacgagca tcacaaaaat cgacgtctca gtcagaggtg gcgaaaccg acaggactat	2100
aaagatacca ggcgtttccc cctggaagct ccctcgtcgc ctctcctgtt ccgacctgc	2160
cgtttaccgg atacctgtcc gccctttctc cttcggaag cgtggcgtt tctcatagct	2220
cacgtgtag gtatctcagt tcggtgtagg tcgttcgctc caagctgggc tgtgtgcacg	2280
aacccccgt tcagcccgac cgctgcgcct tatccgtaa ctatcgtctt gagtccaacc	2340
cggtaagaca cgacttatcg ccaactggcag cagccactgg taacaggatt agcagagcga	2400
ggtatgtagg cggtagctca gagtcttga agtggtaggc taactacggc tacactagaa	2460
ggacagtatt tggtagctgc gctctgctga agccagttac cttcgaaaa agagttggt	2520
gctcttgatc cggcaaaaa accaccgtg gtagcggtagg ttttttgtt tgcaagcagc	2580
agattacgcg cagaaaaaa ggatctcaag aagatcctt gatctttct acggggtctg	2640
acgtcagtg gaacgaaac tcacgttaag ggattttggt catgagatta tcaaaaagga	2700
tcttcaccta gatcctttta aattaaaaat gaagtttta atcaatctaa agtatatatg	2760
agtaaacttg gtctgacagt taccaatgct taatcagtga ggcacctatc tcagcgatct	2820
gtctatttcg ttcatccata gttgcctgac tccccgtcgt gtagataact acgatacggg	2880
agggcttacc atctggcccc agtgcgtcaa tgataccgcg agaccacgc tcaccggctc	2940

cagattttatc agcaataaac cagccagccg gaagggccga gcgcagaagt ggtcctgcaa	3000
ctttatccgc ctccatccag tctattaatt gttgccggga agctagagta agtagttcgc	3060
cagttaatag tttgcgcaac gttgttgcca ttgctacagg catcgtgggtg tcacgctcgt	3120
cgtttgggtat ggcttcattc agtcccggtt cccaacgata aaggcgagtt acatgatccc	3180
ccatgttgtg caaaaaagcg gttagctcct tcggctcctcc gatcgttgtc agaagtaagt	3240
tggccgcagt gttatcactc atggttatgg cagcactgca taattctctt actgtcatgc	3300
catccgtaag atgcttttct gtgactgggtg agtactcaac caagtcattc tgagaatagt	3360
gtatgcggcg accgagttgc tcttgcccgg cgtcaatacg ggataatacc gcgccacata	3420
gcagaacttt aaaagtgtc atcattggaa aacgtttctt cggggcgaata ctctcaagga	3480
tcttaccgct gttgagatcc agttcgatgt aaccactcg tgcaccaac tgatcttcag	3540
catctttttac tttcaccagc gtttctgggt gagcaaaaac aggaaggcaa aatgccgcaa	3600
aaaagggaat aaggcgca cggaaatgtt gaatactcat actcttcctt tttcaatatt	3660
attgaagcat ttatcagggt tattgtctca tgagcggata catatttgaa tgtatttaga	3720
aaaataaaca aatagggtt ccgcgcacat tccccgaaa agtgccacct gacgcgccct	3780
glagcggcgc attaaagcgc gcgggtgttg tggttacgcg cagcgtgacc gctacacttg	3840
ccagcgccct agcgcccgt cctttcgtt tcttccttc ctttctcgcc acgttcgccg	3900
gctttccccg tcaagctcta aatcgggggc tccctttagg gttccgattt agtgccttac	3960
ggcacctcga cccccaaaaa cttgattagg gtgatgttc acgtagtggg ccatcgccct	4020
gataagcgtt ttttcgccct ttgacgttgg agtccacgtt ctttaatagt ggactcttgt	4080
tccaaactgg aacaacactc aacctatct cggctctatc ttttgattta taagggattt	4140
tgccgatttc ggcctattgg ttaaaaaatg agctgattta acaaaaattt aacgcgaatt	4200
ttaacaaaat attaacgtt acaatttcca ttgcctatc aggetgcga actgttggga	4260
agggcgatcg gtgcgggcct cttcgtatc acgccagctg gcgaaagggg gatgtgctgc	4320
aaggcgatta agttgggtta cgccagggtt ttcccagtc cgacgttgta aaacgacggc	4380
cagtgaattg taatacgact cactataggg cgaattgggt accgggcccc cctcagagt	4440
cgatggtgtc gataagcttg atatcgaatt catgtcacac aaaccgatct tcgcctcaag	4500
gaaacctaat tctacatccg agagactgcc gagatccagt ctacactgat taattttcgg	4560
gccataaatt taaaaaatc gtgttatata atatttatg tattatatat atacatcatg	4620
atgatactga cagtcattgc ccattgctaa atagacagac tccatctgcc gcctccaact	4680
gatgttctca atatttaagg ggtcatctcg catgttttaa taataaacag actccatcta	4740

ccgcctccaa atgatgttct caaaatatat tgtatgaact tatttttatt acttagtatt	4800
attagacaac ttacttgctt tatgaaaaac acttcctatt taggaaacaa ttataaatgg	4860
cagttcgttc atttaacaat ttatgtagaa taaatgttat aaatgcgtat gggaaatctt	4920
aaatatggat agcataaatg atacttgcat tgcctaattc gaaatcaaca gcaacgaaaa	4980
aaatcccttg tacaacataa atagtcacg agaaatatca actatcaaag aacagctatt	5040
cacacgttac tattgagatt attattggac gagaatcaca cactcaactg tctttctctc	5100
ttctagaaat acaggtacaa gtatgtacta ttctcattgt tcatacttct agtcatttca	5160
tcccacatat tccttggatt tctctccaat gaatgacatt ctatcttgca aattcaacaa	5220
ttataataag atataccaaa gtagcgggat agtggcaatc aaaaagcttc tctgggtgac	5280
ttctcgtatt tatttttatt ctaatgatcc attaaaggta tatatttatt tcttggtata	5340
taatcctttt gtttattaca tgggctggat acataaaggat attttgattt aattttttgc	5400
ttaaattcaa tccccctcg ttcagtgtca actgtaatgg taggaaatta ccatactttt	5460
gaagaagcaa aaaaaatgaa agaaaaaaaa aatcgtattt ccaggttaga cgttcgcgag	5520
aatctagaat gcggtatgag gtacattgtt ctctgaacgt aaaagttagc ctccctgaga	5580
tattgtacat ttttgctttt acaagtacaa gtacatcgta caactatgta ctactgttga	5640
tgcattcaca acagtttgtt ttgttttttt ttgttttttt tttttctaata gattcattac	5700
cgctatgtat acctacttgt actttagtaga agccgggtta ttggcgttca attaatcata	5760
gacttatgaa tctgcacggt gtgcgctgag agttactttt agcttatgca tgctacttgg	5820
gtgtaatat ttggatctgtt cggaaatcaa cggatgctca atcgatttcg acagtaatta	5880
attaagtcat acacaagtca gctttcttcg agcctcatat aagtataagt agttcaacgt	5940
attagcactg taccagcat ctccgtatcg agaaacacaa caaatgccc cattggacag	6000
atcatcgga tacacaggtt gtgcagtatc atacatactc gatcagacag gtcgtctgac	6060
catcatacaa gctgaacaag cgctccatac ttgcacgctc tctatataca cagttaaatt	6120
acatatccat agtctaacct ctaacagtta atcttctggg aagcctccca gccagccttc	6180
tggtatcgct tggcctcttc aataggatct cggttctggc cgtacagacc tcggccgaca	6240
attatgatat ccgttccggt agacatgaca tctcaacag ttcggtagtg ctgtccgaga	6300
gcgtctccct tgtcgtcaag acccaccctg ggggtcagaa taagccagtc ctgagagtcg	6360
cccttaggtc ggttctgggc aatgaagcca accacaaact cggggtcgga tcgggcaagc	6420
tcaatggtct gcttggagta ctgccagtg gccagagagc ccttgcaaga cagctcggcc	6480

agcatgagca gacctctggc cagcttctcg ttgggagagg ggactaggaa ctctttgtac	6540
tgggagttct cgtagtcaga gacgtcctcc ttctttctgtt cagagacagt ttcttcggca	6600
ccagctcgca ggccagcaat gattccggtt ccgggtacac cgtgggcgtt ggtgatatcg	6660
gaccactcgg cgattcggtg acaccggtagc ttgtgcttga cagtgttgcc aatatctgcg	6720
aactttctgt cctcgaacag gaagaaaccg tgcttaagag caagttcctt gagggggagc	6780
acagtgccgg cgtaggtgaa gtctgcaatg atgtcgatat gggttttgat catgcacaca	6840
taaggtcgca ccttatcggc aagctcaatg agctccttgg ttgtggtaac atccagagaa	6900
gcacacaggt ttgttttctt ggctgccacg agcttgagca ctcgagcggc aaaggcggac	6960
ttgtggacgt tagctcgagc ttcttaggag ggcatTTTgg ttgtgaagag gagactgaaa	7020
taaatttagt ctgcagaact ttttatcgga accttatctg gggcagtgaa gtatatgtta	7080
tggtaatagt tacgagttag ttgaacttat agatagactg gactatacgg ctatcggtcc	7140
aaattagaaa gaacgtcaat ggctctctgg gcgtcgctt tgccgacaaa aatgtgatca	7200
tgatgaaagc cagcaatgac gttgcagctg atattgttgt cggccaaccg cgccgaaac	7260
gcagctgtca gaccacagc ctccaacgaa gaatgtatcg tcaaagtgat ccaagcacac	7320
tcatagttgg agtcgtactc caaaggcggc aatgacgagt cagacagata ctctcgact	7380
caggcgacga cggaattcct gcagcccatc tgcagaattc aggagagacc gggttggcgg	7440
cgtatttgtg tcccaaaaaa cagcccaat tgccccggag aagacggcca ggccgcctag	7500
atgacaaatt caacaactca cagctgactt tctgccattg ccactagggg ggggcctttt	7560
tatatggcca agccaagctc tccacgtcgg ttgggctgca cccaacaata aatgggtagg	7620
gttgaccaa caaagggatg ggatgggggg tagaagatac gaggataacg gggctcaatg	7680
gcacaaataa gaacgaatac tgccattaag actcgtgatc cagcgactga caccattgca	7740
tcactaagg gcctcaaac tacctcgga cgtctgcgt gatctggaca ccacagaggt	7800
tccgagcact ttaggttgca ccaaatgtcc caccaggtgc aggcagaaaa cgctggaaca	7860
gcgtgtacag ttgtcttaa caaaaagtga gggcgtgag gtcgagcagg gtggtgtgac	7920
ttgttatagc ctttagagct gcgaaagcgc gtatggattt ggctcatcag gccagattga	7980
gggtctgtgg acacatgtca tgttagtgta cttcaatcg cccctggata tagccccgac	8040
aataggccgt ggctcattt tttgccttc cgcacattc cattgtcgg tacccacacc	8100
ttgttctcc tgcacttgcc aaccttaata ctggtttaca ttgaccaaca tttacaagc	8160
ggggggcctt tctagggtat atataaacag tggtctctcc aatcggttgc cagtctcttt	8220
tttctttct ttccccacag attcgaaatc taaactacac atcacacaat gcctgttact	8280
gacgtcctta agcgaaagtc cgggtgtcatc gtcggcgacg atgtccgagc cgtgagtatc	8340

cacgacaaga tcagtgtcga gacgacgcgt ttgtgtaat gacacaatcc gaaagtcgct 8400

agcaacacac actctctaca caaactaacc cagctctc 8438

<210> 195

<211> 1350

<212> DNA

<213> *Euglena gracilis*

<220><221> CDS

<222> (1)..(1350)

<223> synthetic delta-5 desaturase (codon-optimized for *Yarrowia*

lipolytica)

<300><302> DELTA-5 DESATURASE AND ITS USE IN MAKING POLYUNSATURATED FATTY

ACIDS

<310> WO 2007/136671

<311> 2007-05-16

<312> 2007-11-29

<313> (1)..(1350)

<300><302> DELTA-5 DESATURASE AND ITS USE IN MAKING POLYUNSATURATED FATTY

ACIDS

<310> US 2007-0292924-A1

<311> 2007-05-15

<312> 2007-12-20

<313> (1)..(1350)

<400> 195

atg gct ctc tcc ctt act acc gag cag ctg ctc gag cga ccc gac ctg 48

Met Ala Leu Ser Leu Thr Thr Glu Gln Leu Leu Glu Arg Pro Asp Leu

1 5 10 15

gtt gcc atc gac ggc att ctc tac gat ctg gaa ggt ctt gcc aag gtc 96

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

20 25 30

cat ccc gga ggc gac ttg atc ctc gct tct ggt gcc tcc gat gct tct 144

His Pro Gly Gly Asp Leu Ile Leu Ala Ser Gly Ala Ser Asp Ala Ser

35 40 45

cct ctg ttc tac tcc atg cac cct tac gtc aag ccc gag aac tcg aag	192
Pro Leu Phe Tyr Ser Met His Pro Tyr Val Lys Pro Glu Asn Ser Lys	
50 55 60	
ctg ctt caa cag ttc gtg cga ggc aag cac gac cga acc tcc aag gac	240
Leu Leu Gln Gln Phe Val Arg Gly Lys His Asp Arg Thr Ser Lys Asp	
65 70 75 80	
att gtc tac acc tac gac tct ccc ttt gca cag gac gtc aag cga act	288
Ile Val Tyr Thr Tyr Asp Ser Pro Phe Ala Gln Asp Val Lys Arg Thr	
85 90 95	
atg cga gag gtc atg aaa ggt cgg aac tgg tat gcc aca cct gga ttc	336
Met Arg Glu Val Met Lys Gly Arg Asn Trp Tyr Ala Thr Pro Gly Phe	
100 105 110	
tgg ctg cga acc gtt ggc atc att gct gtc acc gcc ttt tgc gag tgg	384
Trp Leu Arg Thr Val Gly Ile Ile Ala Val Thr Ala Phe Cys Glu Trp	
115 120 125	
cac tgg gct act acc gga atg gtg ctg tgg ggt ctc ttg act gga ttc	432
His Trp Ala Thr Thr Gly Met Val Leu Trp Gly Leu Leu Thr Gly Phe	
130 135 140	
atg cac atg cag atc ggc ctg tcc att cag cac gat gcc tct cat ggt	480
Met His Met Gln Ile Gly Leu Ser Ile Gln His Asp Ala Ser His Gly	
145 150 155 160	
gcc atc agc aaa aag ccc tgg gtc aac gct ctc ttt gcc tac ggc atc	528
Ala Ile Ser Lys Lys Pro Trp Val Asn Ala Leu Phe Ala Tyr Gly Ile	
165 170 175	
gac gtc att gga tcg tcc aga tgg atc tgg ctg cag tct cac atc atg	576
Asp Val Ile Gly Ser Ser Arg Trp Ile Trp Leu Gln Ser His Ile Met	
180 185 190	
cga cat cac acc tac acc aat cag cat ggt ctc gac ctg gat gcc gag	624
Arg His His Thr Tyr Thr Asn Gln His Gly Leu Asp Leu Asp Ala Glu	
195 200 205	
tcc gca gaa cca ttc ctt gtg ttc cac aac tac cct gct gcc aac act	672

Ser Ala Glu Pro Phe Leu Val Phe His Asn Tyr Pro Ala Ala Asn Thr
210 215 220
gct cga aag tgg ttt cac cga ttc cag gcc tgg tac atg tac ctc gtg 720
Ala Arg Lys Trp Phe His Arg Phe Gln Ala Trp Tyr Met Tyr Leu Val
225 230 235 240
ctt gga gcc tac ggc gtt tcg ctg gtg tac aac cct ctc tac atc ttc 768
Leu Gly Ala Tyr Gly Val Ser Leu Val Tyr Asn Pro Leu Tyr Ile Phe
245 250 255
cga atg cag cac aac gac acc att ccc gag tct gtc aca gcc atg cga 816
Arg Met Gln His Asn Asp Thr Ile Pro Glu Ser Val Thr Ala Met Arg
260 265 270
gag aac ggc ttt ctg cga cgg tac cga acc ctt gca ttc gtt atg cga 864
Glu Asn Gly Phe Leu Arg Arg Tyr Arg Thr Leu Ala Phe Val Met Arg
275 280 285
gct ttc ttc atc ttt cga acc gcc ttc ttg ccc tgg tat ctc act gga 912
Ala Phe Phe Ile Phe Arg Thr Ala Phe Leu Pro Trp Tyr Leu Thr Gly
290 295 300
acc tcc ctg ctc atc acc att cct ctg gtg ccc act gct acc ggt gcc 960
Thr Ser Leu Leu Ile Thr Ile Pro Leu Val Pro Thr Ala Thr Gly Ala
305 310 315 320
ttc ctc acc ttc ttt ttc atc ttg tct cac aac ttc gat ggc tcg gag 1008
Phe Leu Thr Phe Phe Phe Ile Leu Ser His Asn Phe Asp Gly Ser Glu
325 330 335
cga atc ccc gac aag aac tgc aag gtc aag agc tcc gag aag gac gtt 1056
Arg Ile Pro Asp Lys Asn Cys Lys Val Lys Ser Ser Glu Lys Asp Val
340 345 350
gaa gcc gat cag atc gac tgg tac aga gct cag gtg gag acc tct tcc 1104
Glu Ala Asp Gln Ile Asp Trp Tyr Arg Ala Gln Val Glu Thr Ser Ser
355 360 365
acc tac ggt gga ccc att gcc atg ttc ttt act ggc ggt ctc aac ttc 1152
Thr Tyr Gly Gly Pro Ile Ala Met Phe Phe Thr Gly Gly Leu Asn Phe

370 375 380

cag atc gag cat cac ctc ttt cct cga atg tcg tct tgg cac tat ccc 1200
 Gln Ile Glu His His Leu Phe Pro Arg Met Ser Ser Trp His Tyr Pro
 385 390 395 400
 ttc gtg cag caa gct gtc cga gag tgt tgc gaa cga cac gga gtt cgg 1248
 Phe Val Gln Gln Ala Val Arg Glu Cys Cys Glu Arg His Gly Val Arg
 405 410 415
 tac gtc ttc tac cct acc att gtg ggc aac atc att tcc acc ctc aag 1296

Tyr Val Phe Tyr Pro Thr Ile Val Gly Asn Ile Ile Ser Thr Leu Lys
 420 425 430
 tac atg cac aaa gtc ggt gtg gtt cac tgt gtc aag gac gct cag gat 1344
 Tyr Met His Lys Val Gly Val Val His Cys Val Lys Asp Ala Gln Asp
 435 440 445
 tcc taa 1350
 Ser

<210> 196

<211> 449

<212> PRT

<213> Euglena gracilis

<400> 196

Met Ala Leu Ser Leu Thr Thr Glu Gln Leu Leu Glu Arg Pro Asp Leu
 1 5 10 15
 Val Ala Ile Asp Gly Ile Leu Tyr Asp Leu Glu Gly Leu Ala Lys Val
 20 25 30
 His Pro Gly Gly Asp Leu Ile Leu Ala Ser Gly Ala Ser Asp Ala Ser
 35 40 45

Pro Leu Phe Tyr Ser Met His Pro Tyr Val Lys Pro Glu Asn Ser Lys
 50 55 60
 Leu Leu Gln Gln Phe Val Arg Gly Lys His Asp Arg Thr Ser Lys Asp
 65 70 75 80

Ile Val Tyr Thr Tyr Asp Ser Pro Phe Ala Gln Asp Val Lys Arg Thr
85 90 95

Met Arg Glu Val Met Lys Gly Arg Asn Trp Tyr Ala Thr Pro Gly Phe
100 105 110

Trp Leu Arg Thr Val Gly Ile Ile Ala Val Thr Ala Phe Cys Glu Trp
115 120 125

His Trp Ala Thr Thr Gly Met Val Leu Trp Gly Leu Leu Thr Gly Phe
130 135 140

Met His Met Gln Ile Gly Leu Ser Ile Gln His Asp Ala Ser His Gly
145 150 155 160

Ala Ile Ser Lys Lys Pro Trp Val Asn Ala Leu Phe Ala Tyr Gly Ile
165 170 175

Asp Val Ile Gly Ser Ser Arg Trp Ile Trp Leu Gln Ser His Ile Met
180 185 190

Arg His His Thr Tyr Thr Asn Gln His Gly Leu Asp Leu Asp Ala Glu
195 200 205

Ser Ala Glu Pro Phe Leu Val Phe His Asn Tyr Pro Ala Ala Asn Thr
210 215 220

Ala Arg Lys Trp Phe His Arg Phe Gln Ala Trp Tyr Met Tyr Leu Val
225 230 235 240

Leu Gly Ala Tyr Gly Val Ser Leu Val Tyr Asn Pro Leu Tyr Ile Phe
245 250 255

Arg Met Gln His Asn Asp Thr Ile Pro Glu Ser Val Thr Ala Met Arg
260 265 270

Glu Asn Gly Phe Leu Arg Arg Tyr Arg Thr Leu Ala Phe Val Met Arg
275 280 285

Ala Phe Phe Ile Phe Arg Thr Ala Phe Leu Pro Trp Tyr Leu Thr Gly
290 295 300

Thr Ser Leu Leu Ile Thr Ile Pro Leu Val Pro Thr Ala Thr Gly Ala
305 310 315 320

Phe Leu Thr Phe Phe Phe Ile Leu Ser His Asn Phe Asp Gly Ser Glu

325 330 335
 Arg Ile Pro Asp Lys Asn Cys Lys Val Lys Ser Ser Glu Lys Asp Val
 340 345 350
 Glu Ala Asp Gln Ile Asp Trp Tyr Arg Ala Gln Val Glu Thr Ser Ser
 355 360 365

Thr Tyr Gly Gly Pro Ile Ala Met Phe Phe Thr Gly Gly Leu Asn Phe
 370 375 380
 Gln Ile Glu His His Leu Phe Pro Arg Met Ser Ser Trp His Tyr Pro
 385 390 395 400
 Phe Val Gln Gln Ala Val Arg Glu Cys Cys Glu Arg His Gly Val Arg
 405 410 415
 Tyr Val Phe Tyr Pro Thr Ile Val Gly Asn Ile Ile Ser Thr Leu Lys
 420 425 430

Tyr Met His Lys Val Gly Val Val His Cys Val Lys Asp Ala Gln Asp
 435 440 445
 Ser

<210> 197

<211> 1350

<212> DNA

<213> Euglena gracilis

<220><221> CDS

<222> (1)..(1350)

<223> delta-5 desaturase ("EgD5R*")

<400> 197

atg gct ctc agt ctt acc aca gaa cag ctg tta gaa cgc cct gat ttg 48
 Met Ala Leu Ser Leu Thr Thr Glu Gln Leu Leu Glu Arg Pro Asp Leu
 1 5 10 15

gtt gcg att gat ggc atc ctc tac gac ctt gaa ggg ctt gcc aaa gtt 96
 Val Ala Ile Asp Gly Ile Leu Tyr Asp Leu Glu Gly Leu Ala Lys Val
 20 25 30

cat cca gga gga gat ttg att ctc gct tct ggt gcc tct gat gcc tcc 144

His Pro Gly Gly Asp Leu Ile Leu Ala Ser Gly Ala Ser Asp Ala Ser	
35 40 45	
cct ctc ttt tat tca atg cat cca tac gtc aaa ccg gag aac tcc aaa	192
Pro Leu Phe Tyr Ser Met His Pro Tyr Val Lys Pro Glu Asn Ser Lys	
50 55 60	
ttg ctt caa cag ttc gtc cga ggg aag cat gac cgc acc tcg aag gac	240
Leu Leu Gln Gln Phe Val Arg Gly Lys His Asp Arg Thr Ser Lys Asp	
65 70 75 80	
att gtc tac acg tat gat tct ccc ttc gca caa gac gtt aag cgg aca	288
Ile Val Tyr Thr Tyr Asp Ser Pro Phe Ala Gln Asp Val Lys Arg Thr	
85 90 95	
atg cgc gag gtg atg aaa ggg agg aac tgg tac gca acc cct ggc ttc	336
Met Arg Glu Val Met Lys Gly Arg Asn Trp Tyr Ala Thr Pro Gly Phe	
100 105 110	
tgg ctg cgc acc gtt ggg atc atc gcc gtg acg gcc ttt tgc gag tgg	384
Trp Leu Arg Thr Val Gly Ile Ile Ala Val Thr Ala Phe Cys Glu Trp	
115 120 125	
cac tgg gct acc acg ggg atg gtg ctg tgg ggc ctg ttg act gga ttc	432
His Trp Ala Thr Thr Gly Met Val Leu Trp Gly Leu Leu Thr Gly Phe	
130 135 140	
atg cac atg cag atc ggc tta tcc atc cag cat gat gcg tcc cac ggg	480
Met His Met Gln Ile Gly Leu Ser Ile Gln His Asp Ala Ser His Gly	
145 150 155 160	
gcc atc agc aag aag cct tgg gtc aac gcc ctc ttc gcc tac ggc att	528
Ala Ile Ser Lys Lys Pro Trp Val Asn Ala Leu Phe Ala Tyr Gly Ile	
165 170 175	
gac gtc atc gga tcg tcc cgg tgg att tgg ctg cag tcg cac atc atg	576
Asp Val Ile Gly Ser Ser Arg Trp Ile Trp Leu Gln Ser His Ile Met	
180 185 190	
cgg cac cac acc tac acc aac cag cac ggc ctc gac ctg gat gcg gag	624

Arg His His Thr Tyr Thr Asn Gln His Gly Leu Asp Leu Asp Ala Glu

195	200	205	
tcg gca gag ccg ttc ctg gtg ttc cac aac tac ccc gcc gca aac acc			672
Ser Ala Glu Pro Phe Leu Val Phe His Asn Tyr Pro Ala Ala Asn Thr			
210	215	220	
gcc cga aag tgg ttc cac cgc ttc cag gct tgg tac atg tac ctt gtg			720
Ala Arg Lys Trp Phe His Arg Phe Gln Ala Trp Tyr Met Tyr Leu Val			
225	230	235	240
ctg ggg gca tac ggg gta tcg ctg gtg tac aac ccg ctc tac att ttc			768
Leu Gly Ala Tyr Gly Val Ser Leu Val Tyr Asn Pro Leu Tyr Ile Phe			
245	250	255	
cgg atg cag cac aat gac acc atc cca gag tct gtc acg gcc atg cgg			816
Arg Met Gln His Asn Asp Thr Ile Pro Glu Ser Val Thr Ala Met Arg			
260	265	270	
gaa aat ggc ttt ctg cgg cgc tac cgc aca ctt gca ttc gtg atg cga			864
Glu Asn Gly Phe Leu Arg Arg Tyr Arg Thr Leu Ala Phe Val Met Arg			
275	280	285	
gct ttc ttc atc ttc cgg acc gca ttc ttg ccc tgg tac ctc act ggg			912
Ala Phe Phe Ile Phe Arg Thr Ala Phe Leu Pro Trp Tyr Leu Thr Gly			
290	295	300	
acc tca ttg ctg atc acc att cct ctg gtg ccc acc gca act ggt gcc			960
Thr Ser Leu Leu Ile Thr Ile Pro Leu Val Pro Thr Ala Thr Gly Ala			
305	310	315	320
ttc ttg acg ttc ttc ttc att ttg tcc cac aat ttt gat ggc tcc gaa			1008
Phe Leu Thr Phe Phe Phe Ile Leu Ser His Asn Phe Asp Gly Ser Glu			
325	330	335	
cgg atc ccc gac aag aac tgc aag gtt aag cga tct gag aag gac gtt			1056
Arg Ile Pro Asp Lys Asn Cys Lys Val Lys Arg Ser Glu Lys Asp Val			
340	345	350	

gag gct gac caa att gac tgg tat cgg gcg cag gtg gag acg tcc tcc 1104
 Glu Ala Asp Gln Ile Asp Trp Tyr Arg Ala Gln Val Glu Thr Ser Ser
 355 360 365

aca tac ggt ggc ccc atc gcc atg ttc ttc act ggc ggt ctc aat ttc 1152
 Thr Tyr Gly Gly Pro Ile Ala Met Phe Phe Thr Gly Gly Leu Asn Phe
 370 375 380

cag atc gag cac cac ctc ttt ccc cgg atg tcg tct tgg cac tac ccc 1200

Gln Ile Glu His His Leu Phe Pro Arg Met Ser Ser Trp His Tyr Pro
 385 390 395 400

ttc gtc cag cag gcg gtc cgg gag tgt tgc gaa cga cat gga gtg cga 1248
 Phe Val Gln Gln Ala Val Arg Glu Cys Cys Glu Arg His Gly Val Arg
 405 410 415

tat gtt ttc tac cct acc atc gtc ggc aac atc atc tcc acc ctg aag 1296
 Tyr Val Phe Tyr Pro Thr Ile Val Gly Asn Ile Ile Ser Thr Leu Lys
 420 425 430

tac atg cat aag gtg ggt gtc gtc cac tgc gtg aag gac gca cag gat 1344
 Tyr Met His Lys Val Gly Val Val His Cys Val Lys Asp Ala Gln Asp
 435 440 445

tcc taa 1350
 Ser

<210> 198

<211> 449

<212> PRT

<213> Euglena gracilis

<400> 198

Met Ala Leu Ser Leu Thr Thr Glu Gln Leu Leu Glu Arg Pro Asp Leu
 1 5 10 15

Val Ala Ile Asp Gly Ile Leu Tyr Asp Leu Glu Gly Leu Ala Lys Val
 20 25 30

His Pro Gly Gly Asp Leu Ile Leu Ala Ser Gly Ala Ser Asp Ala Ser
 35 40 45

Pro Leu Phe Tyr Ser Met His Pro Tyr Val Lys Pro Glu Asn Ser Lys

50 55 60

Leu Leu Gln Gln Phe Val Arg Gly Lys His Asp Arg Thr Ser Lys Asp

65 70 75 80

Ile Val Tyr Thr Tyr Asp Ser Pro Phe Ala Gln Asp Val Lys Arg Thr

85 90 95

Met Arg Glu Val Met Lys Gly Arg Asn Trp Tyr Ala Thr Pro Gly Phe

100 105 110

Trp Leu Arg Thr Val Gly Ile Ile Ala Val Thr Ala Phe Cys Glu Trp

115 120 125

His Trp Ala Thr Thr Gly Met Val Leu Trp Gly Leu Leu Thr Gly Phe

130 135 140

Met His Met Gln Ile Gly Leu Ser Ile Gln His Asp Ala Ser His Gly

145 150 155 160

Ala Ile Ser Lys Lys Pro Trp Val Asn Ala Leu Phe Ala Tyr Gly Ile

165 170 175

Asp Val Ile Gly Ser Ser Arg Trp Ile Trp Leu Gln Ser His Ile Met

180 185 190

Arg His His Thr Tyr Thr Asn Gln His Gly Leu Asp Leu Asp Ala Glu

195 200 205

Ser Ala Glu Pro Phe Leu Val Phe His Asn Tyr Pro Ala Ala Asn Thr

210 215 220

Ala Arg Lys Trp Phe His Arg Phe Gln Ala Trp Tyr Met Tyr Leu Val

225 230 235 240

Leu Gly Ala Tyr Gly Val Ser Leu Val Tyr Asn Pro Leu Tyr Ile Phe

245 250 255

Arg Met Gln His Asn Asp Thr Ile Pro Glu Ser Val Thr Ala Met Arg

260 265 270

Glu Asn Gly Phe Leu Arg Arg Tyr Arg Thr Leu Ala Phe Val Met Arg

275 280 285

Ala Phe Phe Ile Phe Arg Thr Ala Phe Leu Pro Trp Tyr Leu Thr Gly

290 295 300
Thr Ser Leu Leu Ile Thr Ile Pro Leu Val Pro Thr Ala Thr Gly Ala

305 310 315 320
Phe Leu Thr Phe Phe Phe Ile Leu Ser His Asn Phe Asp Gly Ser Glu

325 330 335
Arg Ile Pro Asp Lys Asn Cys Lys Val Lys Arg Ser Glu Lys Asp Val

340 345 350
Glu Ala Asp Gln Ile Asp Trp Tyr Arg Ala Gln Val Glu Thr Ser Ser

355 360 365
Thr Tyr Gly Gly Pro Ile Ala Met Phe Phe Thr Gly Gly Leu Asn Phe

370 375 380
Gln Ile Glu His His Leu Phe Pro Arg Met Ser Ser Trp His Tyr Pro

385 390 395 400
Phe Val Gln Gln Ala Val Arg Glu Cys Cys Glu Arg His Gly Val Arg

405 410 415
Tyr Val Phe Tyr Pro Thr Ile Val Gly Asn Ile Ile Ser Thr Leu Lys

420 425 430
Tyr Met His Lys Val Gly Val Val His Cys Val Lys Asp Ala Gln Asp

435 440 445
Ser

<210> 199

<211> 8438

<212> DNA

<213> Artificial Sequence

<220><223> Plasmid pDMW367-M4

<400> 199

catggctctc agtcttacca cagaacagct gtagaacgc cctgatttgg ttgcgattga	60
tggcatcctc tacgacctg aagggttgc caaagttcat ccaggaggag atttgattct	120
cgcttctggt gcctctgatg cctccctct cttttattca atgcatccat acgtcaaacc	180
ggagaactcc aaattgcttc aacagttcgt ccgaggaag catgaccgca cctcgaagga	240

cattgtctac acgtatgatt ctcccttcgc acaagacgtt aagcggacaa tgcgcgaggt 300

gatgaaaggg aggaactggg acgcaacccc tggtttctgg ctgcgcaccg ttgggatcat 360

cgccgtgacg gccttttgcg agtggcactg ggtaccacg gggatggtgc tgtggggcct 420

gttgactgga ttcatgcaca tgcagatcgg ctatccatc cagcatgatg cgtcccacgg 480

ggccatcagc aagaagcctt gggtaacgc cctcttcgcc tacggcattg acgtcatcgg 540

atcgtcccgg tggatttggc tgcagtcgca catcatgcgg caccacacct acaccaacca 600

gcacggcctc gacctggatg cggagtcggc agagccgttc ctggtgttcc acaactaccc 660

cgccgcaaac accgcccga aagtgttcca ccgcttcag gcttgggtaca tgtaccttgt 720

gctgggggca tacgggggat cgctggtgta caacccgctc tacatcttcc ggatgcagca 780

caatgacacc atcccagagt ctgtcacggc catgcgggaa aatggctttc tgcggcgcta 840

ccgcacactt gcattcgtga tgcgagcttt ctcatcttc cggaccgcat tcttgccctg 900

gtacctcact gggacctcat tgcgtatcac cattcctctg gtgcccaccg caactggtgc 960

cttcttgacg ttcttcttca ttttgtcca caattttgat ggctccgaac ggatccccga 1020

caagaactgc aaggtaagc gatctgagaa ggacgttgag gctgacaaa ttgactggta 1080

tcgggcgag gtggagacgt cctccacata cgggtggccc atcgccatgt tcttcactgg 1140

cggtctcaat ttccagatcg agcaccacct ctctccccgg atgtcgtctt ggcaactacc 1200

cttcgtccag caggcggctc gggagtgttg cgaacgacat ggagtgcgat atgttttcta 1260

ccctaccatc gtcggcaaca tcatctccac cctgaagtac atgcataagg tgggtgtcgt 1320

ccactgcgtg aaggacgcac aggattccta agcggccgca agtgtggatg gggaagtgag 1380

tgcccggttc tgtgtgcaca attggcaatc caagatggat ggattcaaca cagggatata 1440

gcgagctacg tgggtgtgcg aggatatagc aacggatatt tatgtttgac acttgagaat 1500

gtacgataca agcactgtcc aagtacaata ctaaacatac tgtacatact catactcgta 1560

cccgggcaac ggtttcactt gactgcagtg gctagtgtc ttactcgtac agtgtgcaat 1620

actgcgtatc atagtctttg atgtatatcg taticattca tgttagttgc gtacgagccg 1680

gaagcataaa gtgtaaagcc tgggtgtcct aatgagttag ctaactcaca ttaattgcgt 1740

tgcgtcact gcccgtttc cagtcgggaa acctgtcgtg ccagctgcat taatgaatcg 1800

gccaacgcgc ggggagaggg ggtttgcgta ttgggcgctc ttccgcttcc tcgctcactg 1860

actcgtcgcg ctcggtcgtt cggctgcggc gagcggatc agctcactca aaggcggtaa 1920

tacggttacc cacagaatca ggggataacg caggaaagaa catgtgagca aaaggccagc 1980

aaaaggccag gaaccgtaaa aaggccgcgt tgctggcggt tttccatagg ctccgcccc	2040
ctgacgagca tcacaaaaat cgacgtcaa gtcagagggt gcgaaccgc acaggactat	2100
aaagatacca ggcgtttccc cctggaagct ccctcgtgcg ctctcctgtt ccgaccctgc	2160
cgcttaccgg atacctgtcc gcctttctcc ctccgggaag cgtggcgctt tctcatagct	2220
cacgctgtag gtatctcagt tcggtgtagg tcgttcgctc caagctgggc tgtgtgcacg	2280
aacccccgt tcagcccgac cgctgcgctt tatccgtaa ctatcgtctt gagtccaacc	2340
cggtaagaca cgacttatcg ccactggcag cagccactgg taacaggatt agcagagcga	2400
ggtatgtagg cgggtctaca gagttcttga agtgggtggc taactacggc tacactagaa	2460
ggacagtatt tggatctgc gctctgctga agccagttac ctccggaaaa agagttggta	2520
gctcttgatc cggcaaaaa accaccgtg gtagcgggtg ttttttgtt tgcaagcagc	2580
agattacgcg cagaaaaaaa ggatctcaag aagatccttt gatctttctt acggggtctg	2640
acgctcagtg gaacgaaaa tcacgttaag ggatttttgt catgagatta tcaaaaagga	2700
tcttcaccta gatcctttta aattaaaaat gaagtttta atcaatctaa agtatatatg	2760
agtaaacttg gtctgacagt taccaatgct taatcagtga ggacacctc tcagcgatct	2820
gtctatttcg ttcattcata gttgcctgac tccccgtcgt gtagataact acgatacggg	2880
agggttacc atctggcccc agtgcgtcaa tgataccgcg agaccacgc tcaccggctc	2940
cagatttacc agcaataaac cagccagccg gaaggggcga gcgcagaagt ggtcctgcaa	3000
ctttatccgc ctccatccag tctattaatt gttgccggga agctagagta agtagttcgc	3060
cagttaatag ttgcgcaac gttgttgcca ttgctacagg catcgtggtg tcacgctcgt	3120
cgtttggtat ggttcattc agctccggtt cccaacgac aaggcgagtt acatgatccc	3180
ccatgttgtg caaaaaagcg gttagctcct tcggctcctc gatcgttgtc agaagtaagt	3240
tggccgcagt gttatcactc atggttatgg cagcactgca taattctctt actgtcatgc	3300
catccgtaag atgcttttct gtgactggtg agtactcaac caagtcattc tgagaatagt	3360
gtatgcggcg accgagttgc tcttgcccgg cgtcaatacg ggataatacc gcgccacata	3420
gcagaacttt aaaagtgtc atcattggaa aacgtttctt ggggcgaaaa ctctcaagga	3480
tcttaccgct gttgagatcc agttcgatgt aaccactcg tgcaccaac tgatcttcag	3540
catcttttac tttcaccagc gtttctgggt gagcaaaaac aggaaggcaa aatgccgcaa	3600
aaaagggaat aaggcgaca cggaatgtt gaatactcat actcttctt tttcaatatt	3660
attgaagcat ttatcagggt tattgtctca tgagcggata catatttgaa tgtatttaga	3720
aaaataaaca aatagggtt ccgcgcacat tccccgaaa agtgccacct gacgcgcctt	3780
gtagcggcgc attaagcgcg gcgggtgtgg tggttacgcg cagcgtgacc gctacacttg	3840

ccagcgccct agcgcccgt cctttcgctt tcttcccttc ctttctcgcc acgttcgccg	3900
gctttccccc tcaagctcta aatcgggggc tccctttagg gttccgattt agtgctttac	3960
ggcacctcga ccccaaaaaa cttgattagg gtgatgttc acgtagtggg ccatcgccct	4020
gatagacggt ttttcgccct ttgacgttgg agtccacgtt ctttaatagt ggactcttgt	4080
tccaaactgg aacaacactc aacctatct cggtctattc ttttgattta taagggattt	4140
tgccgatttc ggctatttgg ttaaaaaatg agctgattta acaaaaattt aacgcgaatt	4200
ttaacaaaat attaacgctt acaatttcca ttccgcatte aggctgcgca actgttggga	4260
agggcgatcg gtgcgggcct cttcgtatt acgccagctg gcgaaagggg gatgtgctgc	4320
aaggcgatta agttgggtaa cgccagggtt ttcccagtca cgacgttcta aaacgacggc	4380
cagtgaattg taatacgact cactataggg cgaattgggt accgggcccc cctcagagt	4440
cgatggtgtc gataagcttg atatcgaatt catgtcacac aaaccgatct tcgcctcaag	4500
gaaacctaat tctacatccg agagactgcc gagatccagt ctacactgat taattttcgg	4560
gccaataatt taaaaaaatc gtgttatata atattatatg tattatatat atacatcatg	4620
atgatactga cagtcattgc ccattgctaa atagacagac tccatctgcc gcctccaact	4680
gatgttctca atatttaagg ggtcatctcg cattgtttta taataaacag actccatcta	4740
ccgcctccaa atgatgttct caaaatatat tgtatgaact tatttttatt acttagtatt	4800
attagacaac ttacttgcct tatgaaaaac acttcctatt taggaaacaa ttataatgg	4860
cagttcgttc atttaacaat ttatgtagaa taaatgttat aaatgcgtat gggaaatctt	4920
aaatatggat agcataaatg atatctgcat tgcctaattc gaaatcaaca gcaacgaaaa	4980
aatcccttg tacaacataa atagtcacg agaaatatca actatcaaag aacagctatt	5040
cacacgttac tattgagatt attattggac gagaatcaca cactcaactg tctttctctc	5100
ttctagaaat acaggtacaa gtatgtacta ttctcattgt tcatacttct agtcatttca	5160
tccacatat tcttggatt tcttccaat gaatgacatt ctatcttgca aattcaacaa	5220
ttataataag atataccaaa gtagcggtat agtggcaatc aaaaagcttc tctggtgtgc	5280
ttctcgtatt tatttttatt ctaatgatcc attaaaggta tataattatt tcttgttata	5340
taatcctttt gtttattaca tgggctggat acataaaggt attttgattt aattttttgc	5400
ttaaattcaa tccccctcg ttcagtgtca actgtaatgg taggaaatta ccatactttt	5460
gaagaagcaa aaaaaatgaa agaaaaaaa aatcgatttt ccaggttaga cgttccgcag	5520
aatctagaat gcggtatgcg gtacattgtt cttcgaacgt aaaagttgcg ctccctgaga	5580
tattgtacat ttttgccttt acaagtacaa gtacatcgta caactatgta ctactgttga	5640
tgcatccaca acagtttgtt ttgttttttt ttgttttttt tttttctaata gattcattac	5700

cgctatgtat acctacttgt actttagtia agccgggtta ttggcgttca attaatacata 5760

gacttatgaa tctgcacggt gtgcgtgctg agttactttt agcttatgca tgctacttgg 5820

gtgtaataatt gggatctgtt cggaaatcaa cggatgctca atcgatttcg acagtaatta 5880

attaagtcac acacaagica gctttcttcg agcctcatat aagtataagt agttcaacgt 5940

attagcactg taccagcat ctccgtatcg agaacacaaa caaatgccc cattggacag 6000

atcatgcgga tacacaggtt gtgcagtatc atacatactc gatcagacag gtcgtctgac 6060

catcatacaa gctgaacaag cgctccatac ttgcacgctc tctatataca cagttaaatt 6120

acatatccat agtctaacct ctaacagtta atcttctggt aagcctccca gccagccttc 6180

tggtatcgct tggcctctc aataggatct cggttctggc cgtacagacc tcggccgaca 6240

attatgatat ccgttccggt agacatgaca tctcaacag ttcggtactg ctgtccgaga 6300

gcgtctccct tgcgtcaag accaccccg ggggtcagaa taagccagtc ctgagagtcg 6360

cccttaggtc ggttctgggc aatgaagcca accacaaact cggggtcgga tcgggcaagc 6420

tcaatggctc gcttgagta ctgccagtg gccagagagc ccttgcaaga cagctcgcc 6480

agcatgagca gacctctggc cagcttctcg ttgggagagg ggactaggaa ctcttctac 6540

tgggagttct cgtagtcaga gacgtcctc ttcttctgtt cagagacagt ttctcgga 6600

ccagctcgca ggccagcaat gattccggtt ccgggtacac cgtgggcgtt ggtgatatcg 6660

gaccactcgg cgattcgggtg acaccggtac tggtgcttga cagtgttgcc aatatctgcg 6720

aactttctgt cctgaacag gaagaaaccg tgcttaagag caagttcctt gagggggagc 6780

acagtcccg cgtaggtgaa gtctcaatg atgtcgatat gggttttgat catgcacaca 6840

taaggtcga cctatcggc aagctcaatg agctccttgg tggtggtaac atccagagaa 6900

gcacacaggt tggttttctt ggctgccacg agcttgagca ctgagcggc aaaggcggac 6960

ttgtggacgt tagctcgagc ttcttaggag ggcattttgg tggatgaagag gagactgaaa 7020

taaatttagt ctgcagaact tttatcgga acctatctg gggcagtga gatatgtta 7080

tggtaatagt tacgagttag ttgaacttat agatagactg gactatacgg ctatcggtcc 7140

aaattagaaa gaacgtcaat ggctctctgg gcgtgcctt tgccgacaaa aatgtgatca 7200

tgatgaaagc cagcaatgac gttgcagctg atattgttgt cggccaaccg cgccgaaaac 7260

gcagctgtca gaccacagc ctccaacgaa gaatgtatcg tcaaagtgat ccaagcacac 7320

tcatagttgg agtcgtactc caaaggcggc aatgacgagt cagacagata ctgctcgact 7380

caggcgacga cggaattcct gcagcccatc tgcagaattc aggagagacc gggttggcgg 7440

cgtatttgtg tcccaaaaaa cagccccaat tgccccggag aagacggcca ggccgcctag 7500
atgacaaatt caacaactca cagctgactt tctgccattg ccactagggg ggggcctttt 7560
tatatggcca agccaagctc tccacgtcgg ttgggctgca cccaacaata aatgggtagg 7620
gttgcaccaa caaagggatg ggatgggggg tagaagatac gaggataacg gggctcaatg 7680
gcacaataa gaacgaatac tgccattaag actcgtgatc cagcgactga caccattgca 7740
tcatctaagg gcctcaaaac tacctcggaa ctgctgcgct gatctggaca ccacagaggt 7800
tccgagcact ttaggttgca ccaaattgcc caccaggtgc aggcagaaaa cgctggaaca 7860

gcgtgtacag tttgtcttaa caaaaagtga gggcgctgag gtcgagcagg gtggtgtgac 7920
ttgttatagc ctttagagct gcgaaagcgc gtatggattt ggctcatcag gccagattga 7980
gggtctgtgg acacatgtca tgtagtgta cttcaatcgc cccctggata tagccccgac 8040
aataggccgt ggcctcatTT ttttgccttc cgcacatttc cattgtcgg taccacacc 8100
ttgtcttcc tgcattgcc aaccttaata ctggtttaca ttgaccaaca tcttacaagc 8160
ggggggcttg tctagggtat atataaacag tggtctctcc aatcggttgc cagtctcttt 8220
tttcttttct ttccccacag attcgaaatc taaactacac atcacacaat gcctgttact 8280

gacgtcctta agcgaagtc cgggtgtcatc gtcggcgacg atgtccgagc cgtgagatc 8340
cacgacaaga tcagtgtcga gacgacgcgt tttgtgtaat gacacaatcc gaaagtcgct 8400
agcaacacac actctctaca caaactaacc cagctctc 8438

<210> 200

<211> 31

<212> DNA

<213> Artificial Sequence

<220><223> Primer YL 813

<400> 200

cgtaaaccg gagaactcca aattgcttca a 31

<210> 201

<211> 31

<212> DNA

<213> Artificial Sequence

<220><223> Primer YL 814

<400>

> 201

ttgaagcaat ttggagtctt ccggtttgac g 31

<210> 202

<211> 31	
<212> DNA	
<213> Artificial Sequence	
<220><223> Primer YL 815	
<400> 202	
aactgcaagg ttaagcgatc tgagaaggac g	31
<210> 203	
<211> 31	
<212> DNA	
<213> Artificial Sequence	
<220><223> Primer YL 816	
<400> 203	
cgtccttctc agatcgctta accttgcagt t	31
<210> 204	
<211> 8438	
<212> DNA	
<213> Artificial Sequence	
<220><223> Plasmid pDMW367-M2	
<400> 204	
catggetctc agtcttacca cagaacagct gtagaacgc cctgatttgg ttgcgattga	60
tggcatcctc tacgaccttg aagggttgc caaagttcat ccaggaggag atttgattct	120
cgtttctggt gcctctgatg cctccctct cttttattca atgcatccat acgtcaaacc	180
ggagaactcc aaattgcttc aacagtctgt ccgagggaag catgaccgca cctcgaagga	240
cattgtctac acgtatgatt ctcccttcgc acaagacgtt aagcggacaa tgcgcgaggt	300
gatgaaaggg aggaactggt acgcaacccc tggttcttgg ctgcgcaccg ttgggatcat	360
cgccgtgacg gccttttgcg agtggcactg ggctaccacg gggatggtgc tgtggggcct	420
gttgactgga ttcatgcaca tgcagatcgg cttatccatc cagcatgatg cgtcccacgg	480
ggccatcagc aagaagcctt gggtaacgc cctcttcgcc tacggcattg acgtcatcgg	540
atcgtcccgg tggatttggc tgcagtcgca catcatcgcg caccacacct acaccaacca	600
gcacggcctc gacctggatg cggagtcggc agagccgttc ctggtgttcc acaactacce	660
cgccgcaaac accgcccga agtggttcca ccgcttccaa gcttgggtaca tgtaccttgt	720
gctgggggca tacgggggtat cgctggtgta caaccgctc tacattttcc ggatgcagca	780

caatgacacc atcccagagt ctgtcacggc catgcgggaa aatggctttc tgcggcgcta	840
ccgcacactt gcattcgtga tgcgagcttt ctccatcttc cggaccgcat tcttgccctg	900
gtacctcact gggacctcat tgcctgacac cttcctcttg gtgcccacg caactgggtg	960
cttcttgacg ttcttcttca ttttgtccca caattttgat ggctccgaac ggatccccga	1020
caagaactgc aaggtttaagc gatctgagaa ggacgttgag gctgacaaa ttgactggta	1080
tccggcgagc gtggagacgt cctccacata cggtagggccc atcgccatgt tcttactgg	1140
cggctcctaat ttccagatcg agcaccacct ctttccccgg atgtcgtctt ggcactaccc	1200
cttcgtccag caggcgggcc gggagtgttg cgaacgccat ggagtgcgat atgttttcta	1260
ccctaccatc gtcggcaaca tcatctccac cctgaagtac atgcataagg tgggtgtcgt	1320
ccactgcgtg aaggacgcac aggattccta agcggccgca agtgtggatg gggaagttag	1380
tgcccggttc tgtgtgcaca attggcaatc caagatggat ggattcaaca cagggatata	1440
gcgagctacg tggtaggtgcg aggatatagc aacggatatt tatgtttgac acttgagaat	1500
gtacgataca agcactgtcc aagtacaata ctaaacatac tgtacatact catactcgta	1560
cccgggcaac ggtttcactt gagtgcagtg gctagtgttc ttactcgtac agtgtgcaat	1620
actgcgtatc atagtctttg atgtatatcg tattcattca tgttagttgc gtacgagccg	1680
gaagcataaa gtgtaaagcc tggggtgcct aatgagttag ctaactcaca ttaattgcgt	1740
tgcgtcactt gcccgctttc cagtgcgggaa acctgtcgtg ccagctgcat taatgaatcg	1800
gccaacgcgc ggggagagcg ggtttgcgta ttgggcgttc ttccgcttcc tcgtcactg	1860
actcgtcgcg ctccgtcgtt cggctgcggc gagcggatc agtcactca aaggcggtaa	1920
tacggttatc cacagaatca ggggataacg caggaaagaa catgtgagca aaaggccagc	1980
aaaaggccag gaaccgtaaa aaggccgcgt tgcgtggcgtt tttccatagg ctccgcccc	2040
ctgacgagca tcacaaaaat cgacgtcaa gtcagagggtg gcgaaaccg acaggactat	2100
aaagatacca ggcgtttccc cctggaagct ccctcgtcgc ctctcctgtt ccgaccctgc	2160
cgtttaccgg atacctgtcc gccctttctc cttcggaag cgtggcgtt tctcatagct	2220
cacgtgttag gtatctcagt tcggtgtagg tcgttcgtc caagctgggc tgtgtgcacg	2280
aacccccgt tcagcccgac cgctgcgcct tatccgtaa ctatcgtctt gagtccaacc	2340
cggtaagaca cgacttatcg ccaactggcag cagccactgg taacaggatt agcagagcga	2400
ggtatgtagg cgggtgtaca gagttcttga agtgggtggc taactacggc tacactagaa	2460
ggacagtatt tggatctgc gctctgtga agccagttac cttcgaaaa agagttggta	2520

gctcttgatc cggcaaacaa accaccgctg gtagcgggtg tttttttgtt tgcaagcagc 2580
 agattacgcg cagaaaaaaa ggatctcaag aagatccttt gatcttttct acggggtctg 2640
 acgctcagtg gaacgaaaaa tcacgttaag ggatttttgt catgagatta tcaaaaagga 2700
 tcttcaccta gatcctttta aattaaaaat gaagttttta atcaatctaa agtatatatg 2760
 agtaaaactg gtctgacagt taccaatgct taatcagtga ggcacctatc tcagcgatct 2820
 gtctatttcg ttcattccata gttgcctgac tccccgtcgt gtagataact acgatacggg 2880
 agggcttacc atctggcccc agtgctgcaa tgataccgcg agaccacgc tcaccggctc 2940

 cagatttatt agcaataaac cagccagccg gaagggccga gcgcagaagt ggtcctgcaa 3000
 ctttatccgc ctccatccag tctattaatt gttgccggga agctagagta agtagttcgc 3060
 cagttaatat ttgcgcaac gttgttgcca ttgctacagg catcgtggtg tcacgctcgt 3120
 cgtttggtat ggcttcattc agctccggtt cccaacgac aaggcgagtt acatgatccc 3180
 ccatgttgtg caaaaaagcg gttagctcct tcggctctcc gatcgtttgc agaagtaagt 3240
 tggccgcagt gttatcactc atggttatgg cagcactgca taattctctt actgtcatgc 3300
 catccgtaag atgcttttct gtgactggtg agtactcaac caagtcattc tgagaatagt 3360

 gtagcggcg accgagttgc tcttgccccg cgtcaatacg ggataatacc gcgccacata 3420
 gcagaacttt aaaagtgtc atcattggaa aacgtttctt ggggcgaaaa ctctcaagga 3480
 tcttaccgct gttgagatcc agttcgatgt aaccactcg tgcaccaac tgatcttcag 3540
 catcttttac tttcaccagc gtttctgggt gagcaaaaac aggaaggcaa aatgccgcaa 3600
 aaaagggaat aaggcgaca cggaaatgtt gaatactcat actcttctt tttcaatatt 3660
 attgaagcat ttatcagggt tattgtctca tgagcgata catatttgaa tgtatttaga 3720
 aaaataaaca aatagggggt cgcgcacat tccccgaaa agtgccacct gacgcgcct 3780

 gtagcggcg attaagcgcg gcgggtgttg tggttacgcg cagcgtgacc gctacacttg 3840
 ccagcgccct agcgcccgct ctttcgctt tcttccttc ctttctgcc acgttcgccg 3900
 gctttccccg tcaagctcta aatcgggggc tcccttagg gttccgattt agtgctttac 3960
 ggcacctga ccccaaaaaa cttgattagg gtgatggtt acgtagtggg ccatcgccct 4020
 gatagacggt ttttcgccct ttgacgttg agtccacgtt ctttaatagt ggactcttgt 4080
 tccaaactgg aacaacactc aacctatct cgtctattc ttttgattta taagggattt 4140
 tgccgatttc ggctatttg ttaaaaaatg agctgattta acaaaaattt aacgcgaatt 4200

 ttaacaaaat attaacgctt acaatttcca ttgccattc aggtcgcga actgttgga 4260
 agggcgatcg gtgcgggcct cttcgtatt acgccagctg gcgaaaggg gatgtgctgc 4320
 aaggcgatta agttgggtta cgccagggtt ttccagtcg cgacgttgta aaacgacggc 4380

cagtgaattg taatacgact cactataggg cgaattgggt accgggcccc ccctcgaggt 4440
cgatggtgtc gataagcttg atatcgaatt catgtcacac aaaccgatct tcgcctcaag 4500
gaaacctaatt tctacatccg agagactgcc gagatccagt ctacactgat taattttcgg 4560
gccaataatt taaaaaaatc gtgttatata atattatatg tattatata atacatcatg 4620

atgatactga cagtcacgtc ccattgctaa atagacagac tccatctgcc gcctccaact 4680
gatgttctca atatttaagg ggtcatctcg cattgtttaa taataaacag actccatcta 4740
ccgcctccaa atgatgttct caaaatata tgtatgaact ttttttatt acttagtatt 4800
attagacaac ttacttgctt tatgaaaaac acttctatt taggaaacaa ttataatgg 4860
cagttcgttc atttaacaat ttatgtagaa taaatgttat aaatgcgtat gggaaatctt 4920
aaatatggat agcataaatg atatctgcat tgcctaattc gaaatcaaca gcaacgaaaa 4980
aaatcccttg tacaacataa atagtcacgt agaaatatca actatcaaag aacagctatt 5040

cacacgttac tattgagatt attattggac gagaatcaca cactcaactg tctttctctc 5100
ttctagaaat acaggtacaa gtatgtacta ttctcattgt tcatacttct agtcatttca 5160
tcccacatat tecttggaat tctctccaat gaatgacatt ctatcttgca aattcaacaa 5220
ttataataag atataccaaa gtagcgggtat agtggcaatc aaaaagcttc tctgggtgtc 5280
ttctcgtatt ttttttatt ctaatgatcc attaaaggta tatatttatt tcttgttata 5340
taatcctttt gtttattaca tgggctggat acataaagggt attttgattt aattttttgc 5400
ttaaattcaa tccccctcg ttcagtgtca actgtaatgg taggaaatta ccatactttt 5460

gaagaagcaa aaaaaatgaa agaaaaaaa aatcgatatt ccaggttaga cgttcgcgag 5520
aatctagaat gcggtatgcg gtacattgtt ctcgaacgt aaaagttgcg ctccctgaga 5580
tattgtacat ttttgccttt acaagtacaa gtacatcgta caactatgta ctactgttga 5640
tgcatccaca acagtttgtt ttgttttttt ttgttttttt ttttctaat gattcattac 5700
cgctatgtat acctacttgt actttagta agccgggtta ttggcgttca attaatcata 5760
gacttatgaa tctgcacggt gtgcgctgcg agttactttt agcttatgca tgctacttgg 5820
gtgtaatatt gggatctgtt cggaatatca cggatgctca atcgatttcg acagtaatta 5880

attaagtcac acacaagtca gctttcttcg agcctcatat aagtataagt agttcaacgt 5940
attagcactg taccagcat ctccgtatcg agaacacaa caaatgccc cattggacag 6000
atcatgcgga tacacaggtt gtgcagtatc atacatactc gatcagacag gtcgtctgac 6060
catcatacaa gctgaacaag cgctccatac ttgcacgtc tctatataca cagttaaatt 6120
acatatccat agtctaacct ctaacagtta atcttctggg aagcctccca gccagccttc 6180
tggtatcgct tggcctctc aataggatct cggttctggc cgtacagacc tcggccgaca 6240

attatgatat ccgttccggt agacatgaca tcctcaacag ttcggtactg ctgtccgaga 6300

gcgtctccct tgtcgtcaag acccaccg ggggtcagaa taagccagtc ctgagagtcg 6360

cccttaggtc ggttctgggc aatgaagcca accacaaact cggggtcgga tcgggcaagc 6420

tcaatggtct gcttggagta ctgccagtg gccagagagc ccttgcaaga cagctcggcc 6480

agcatgagca gacctctggc cagcttctcg ttgggagagg ggactaggaa ctcttgtac 6540

tgggagtctt cgtagttaga gacgtcctcc ttcttctgtt cagagacagt ttcctcggca 6600

ccagctcgca ggccagcaat gattccggtt ccgggtacac cgtgggcgtt ggtgatatcg 6660

gaccactcgg cgattcgggtg acaccggtac tggtgcttga cagtgttgcc aatatctgcg 6720

aactttctgt cctcgaacag gaagaaaccg tgcttaagag caagttcctt gagggggagc 6780

acagtgccgg cgtaggtaga gtcgtcaatg atgtcgatat gggttttgat catgcacaca 6840

taagggtcga cttatcggc aagctcaatg agctccttgg tggtagtaac atccagagaa 6900

gcacacaggt tggttttctt ggctgccacg agcttgagca ctcgagcggc aaaggcggac 6960

ttgtggacgt tagctcgagc ttcttaggag ggcatcttgg tggtagagag gagactgaaa 7020

taaatttagt ctgcagaact ttttatcgga acctatctg gggcagtgaa gtatatgtta 7080

tggtaatagt tacagtttag ttgaacttat agatagactg gactatacgg ctatcggtcc 7140

aaattagaaa gaacgtcaat ggctctctgg gcgtcgctt tgccgacaaa aatgtgatca 7200

tgatgaaagc cagcaatgac gttgcagctg atattgttgt cggccaaccg cgccgaaaac 7260

gcagctgtca gaccacagc ctccaacgaa gaatgtatcg tcaaagtgat ccaagcacac 7320

tcatagttag agtcgtactc caaaggcggc aatgacgagt cagacagata ctgctcgact 7380

caggcgacga cggaattcct gcagcccatc tgcagaattc aggagagacc gggttggcgg 7440

cgtatttgtg tcccaaaaaa cagccccaat tgccccggag aagacggcca ggccgcctag 7500

atgacaaatt caacaactca cagctgactt tctgccattg ccaactagggg ggggcctttt 7560

tatatggcca agccaagtc tccacgtcgg ttgggctgca cccaacaata aatgggtagg 7620

gttgacacaa caaagggatg ggatgggggg tagaagatac gaggataacg gggctcaatg 7680

gcacaaataa gaacgaatac tgccattaag actcgtgatc cagcgactga caccattgca 7740

tcatctaagg gcctcaaaac tacctcgga ctgctgcgt gatctggaca ccacagaggt 7800

tccgagcact ttaggttgca ccaaatgtcc caccaggtgc aggcagaaaa cgctggaaca 7860

gcgtgtacag ttgtcttaa caaaaagtga gggcgctgag gtcgagcagg gtggtgtgac 7920

ttgttatagc ctttagagct gcgaaagcgc gtatggattt ggctcatcag gccagattga 7980

gggtctgtgg acacatgtca tgtagtgta cttcaatcgc cccctggata tagccccgac 8040
aataggccgt ggcctcattt ttttgccttc cgcacatttc cattgctcgg taccacacacc 8100
ttgtctctcc tgcacttgcc aaccttaata ctggtttaca ttgaccaaca tcttacaagc 8160
ggggggcctt tctagggtat atataaacag tggctctccc aatcggttgc cagtctcttt 8220
tttcttttct ttccccacag attcgaaatc taaactacac atcacacaat gcctgttact 8280
gacgtcctta agcgaaagtc cgggtgtcatc gtcggcgacg atgtccgagc cgtgagtatc 8340
cacgacaaga tcagtgtcga gacgacgcgt tttgtgtaat gacacaatcc gaaagtcgct 8400

agcaacacac actctctaca caaactaacc cagctctc 8438

<210> 205

<211> 31

<212> DNA

<213> Artificial Sequence

<220><223> Primer YL 829

<400> 205

gttccaccgc ttccaggctt ggtacatgta c 31

<210> 206

<211> 31

<212> DNA

<213> Artificial Sequence

<220><223> Primer YL 830

<400> 206

gtacatgtac caagcctgga agcgggtggaa c 31

<210> 207

<211> 32

<212> DNA

<213> Artificial Sequence

<220><223> Primer YL 831

<400> 207

ggagtgttgc gaacgacatg gagtgcgata tg 32

<210> 208

<211> 32

<212> DNA

<213> Artificial Sequence

<220><223> Primer YL 832

<400> 208

catatcgcac tccatgtcgt tcgcaacact cc 32

<210> 209

<211> 39

<212> DNA

<213> Artificial Sequence

<220><223> Primer

<400> 209

ttatccatcc agcatgattg ctcccacggg gccatcagc 39

<210> 210

<211> 39

<212> DNA

<213> Artificial Sequence

<220><223> Primer

<400> 210

gtgatggcc ccgtgggagc aatcatgctg gatggataa 39

<210> 211

<211> 37

<212> DNA

<213> Artificial Sequence

<220><223> Primer

<400> 211

tatccatcca gcatgatgac tcccacgggg ccatcag 37

<210> 212

<211> 37

<212> DNA

<213> Artificial Sequence

<220><223> Primer

<400> 212

ctgatggccc cgtgggagtc atcatgctgg atggata 37

<210> 213

<211>

> 37

<212> DNA
 <213> Artificial Sequence
 <220><223> Primer
 <400> 213
 tatccatcca gcatgatgag tcccacgggg ccatcag 37
 <210> 214
 <211> 37
 <212> DNA
 <213> Artificial Sequence
 <220><223> Primer
 <400> 214
 ctgatggccc cgtgggactc atcatgctgg atggata 37
 <210> 215
 <211> 39
 <212> DNA
 <213> Artificial Sequence
 <220><223> Primer
 <400> 215
 ttatccatcc agcatgatatt ctcccacggg gccatcagc 39
 <210> 216
 <211> 39
 <212> DNA
 <213> Artificial Sequence
 <220><223> Primer
 <400> 216
 gctgatggcc ccgtgggaga aatcatgctg gatggataa 39
 <210> 217
 <211> 39
 <212> DNA
 <213> Artificial Sequence
 <220><223> Primer
 <400> 217
 ttatccatcc agcatgatgg ctcccacggg gccatcagc 39
 <210> 218

<211> 39
 <212> DNA
 <213> Artificial Sequence
 <220><223> Primer
 <400> 218
 gctgatggcc ccgtgggagc catcatgctg gatggataa 39

<210> 219
 <211> 37
 <212> DNA
 <213> Artificial Sequence
 <220><223> Primer
 <400> 219
 tatccatcca gcatgatcac tcccacgggg ccatcag 37

<210> 220
 <211> 37
 <212> DNA
 <213> Artificial Sequence
 <220><223> Primer
 <400> 220
 ctgatggccc cgtgggagtg atcatgctgg atggata 37

<210> 221
 <211> 37
 <212> DNA
 <213> Artificial Sequence
 <220><223> Primer
 <400> 221
 tatccatcca gcatgatatc tcccacgggg ccatcag 37

<210> 222
 <211> 37
 <212> DNA
 <213> Artificial Sequence
 <220><223> Primer
 <400> 222

ctgatggccc cgtgggagat atcatgctgg atggata 37

<210> 223

<211> 37

<212> DNA

<213> Artificial Sequence

<220><223> Primer

<400> 223

tatccatcca gcatgataag tcccacgggg ccatcag 37

<210> 224

<211> 37

<212> DNA

<213> Artificial Sequence

<220><223> Primer

<400> 224

ctgatggccc cgtgggactt atcatgctgg atggata 37

<210> 225

<211> 37

<212> DNA

<213> Artificial Sequence

<220><223> Primer

<400> 225

tatccatcca gcatgatctg tcccacgggg ccatcag 37

<210> 226

<211> 37

<212> DNA

<213> Artificial Sequence

<220><223> Primer

<400> 226

ctgatggccc cgtgggacag atcatgctgg atggata 37

<210> 227

<211> 37

<212> DNA

<213> Artificial Sequence

<220><223> Primer

<400> 227
tatccatcca gcatgatatg tcccacgggg ccatcag 37

<210> 228
<211> 37
<212> DNA
<213> Artificial Sequence
<220><223> Primer
<400> 228
ctgatggccc cgtgggacat atcatgctgg atggata 37

<210> 229
<211> 37
<212> DNA
<213> Artificial Sequence
<220><223> Primer
<400> 229
tatccatcca gcatgataac tcccacgggg ccatcag 37

<210> 230
<211> 37
<212> DNA
<213> Artificial Sequence
<220><223> Primer
<400> 230
ctgatggccc cgtgggagtt atcatgctgg atggata 37

<210> 231
<211> 37
<212> DNA
<213> Artificial Sequence
<220><223> Primer
<400> 231
tatccatcca gcatgatccc tcccacgggg ccatcag 37

<210> 232
<211> 37
<212> DNA

<213> Artificial Sequence

<220><223> Primer

<400> 232

ctgatggccc cgtgggaggg atcatgctgg atggata 37

<210> 233

<211> 37

<212> DNA

<213> Artificial Sequence

<220><223> Primer

<400> 233

tatccatcca gcatgatcaa tcccacgggg ccatcag 37

<210> 234

<211> 37

<212> DNA

<213> Artificial Sequence

<220><223> Primer

<400> 234

ctgatggccc cgtgggattg atcatgctgg atggata 37

<210> 235

<211> 37

<212> DNA

<213> Artificial Sequence

<220><223> Primer

<400> 235

tatccatcca gcatgatcga tcccacgggg ccatcag 37

<210> 236

<211> 37

<212> DNA

<213> Artificial Sequence

<220><223> Primer

<400> 236

ctgatggccc cgtgggatcg atcatgctgg atggata 37

<210> 237

<211> 39
 <212> DNA
 <213> Artificial Sequence
 <220><223> Primer
 <400> 237
 ttatccatcc agcatgattc ctcccacggg gccatcagc 39
 <210> 238
 <211> 39
 <212> DNA
 <213> Artificial Sequence
 <220><223> Primer
 <400> 238
 gctgatggcc ccgtgggagg aatcatgctg gatggataa 39
 <210> 239
 <211> 37
 <212> DNA
 <213> Artificial Sequence
 <220><223> Primer
 <400> 239
 tatccatcca gcatgatacc tcccacgggg ccatcag 37

 <210> 240
 <211> 37
 <212> DNA
 <213> Artificial Sequence
 <220><223> Primer
 <400> 240
 ctgatggccc cggtgggaggt atcatgctgg atggata 37
 <210> 241
 <211> 37
 <212> DNA
 <213> Artificial Sequence
 <220><223> Primer
 <400> 241
 tatccatcca gcatgatgtc tcccacgggg ccatcag 37

<210> 242
 <211> 37
 <212> DNA
 <213> Artificial Sequence
 <220><223> Primer
 <400> 242
 ctgatggccc cgtgggagac atcatgctgg atggata 37

<210> 243
 <211> 37
 <212> DNA
 <213> Artificial Sequence
 <220><223> Primer
 <400> 243
 tatccatcca gcatgattgg tcccacgggg ccatcag 37

<210> 244
 <211> 37
 <212> DNA
 <213> Artificial Sequence
 <220><223> Primer
 <400> 244
 ctgatggccc cgtgggacca atcatgctgg atggata 37

<210> 245
 <211> 39
 <212> DNA
 <213> Artificial Sequence
 <220><223> Primer
 <400> 245
 ttatccatcc agcatgatta ctcccacggg gccatcagc 39

<210> 246
 <211> 39
 <212> DNA
 <213> Artificial Sequence
 <220><223> Primer

<400> 246

gctgatggcc ccgtgggagt aatcatgctg gatggataa

39

<210> 247

<211> 449

<212> PRT

<213> Euglena gracilis

<400> 247

Met Ala Leu Ser Leu Thr Thr Glu Gln Leu Leu Glu Arg Pro Asp Leu

1 5 10 15

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

20 25 30

His Pro Gly Gly Asp Leu Ile Leu Ala Ser Gly Ala Ser Asp Ala Ser

35 40 45

Pro Leu Phe Tyr Ser Met His Pro Tyr Val Lys Pro Glu Asn Ser Lys

50 55 60

Leu Leu Gln Gln Phe Val Arg Gly Lys His Asp Arg Thr Ser Lys Asp

65 70 75 80

Ile Val Tyr Thr Tyr Asp Ser Pro Phe Ala Gln Asp Val Lys Arg Thr

85 90 95

Met Arg Glu Val Met Lys Gly Arg Asn Trp Tyr Ala Thr Pro Gly Phe

100 105 110

Trp Leu Arg Thr Val Gly Ile Ile Ala Val Thr Ala Phe Cys Glu Trp

115 120 125

His Trp Ala Thr Thr Gly Met Val Leu Trp Gly Leu Leu Thr Gly Phe

130 135 140

Met His Met Gln Ile Gly Leu Ser Ile Gln His Asp Gly Ser His Gly

145 150 155 160

Ala Ile Ser Lys Lys Pro Trp Val Asn Ala Leu Phe Ala Tyr Gly Ile

165 170 175

Asp Val Ile Gly Ser Ser Arg Trp Ile Trp Leu Gln Ser His Ile Met

180 185 190

Arg His His Thr Tyr Thr Asn Gln His Gly Leu Asp Leu Asp Ala Glu

195	200	205	
Ser Ala Glu Pro Phe Leu Val Phe His Asn Tyr Pro Ala Ala Asn Thr			
210	215	220	
Ala Arg Lys Trp Phe His Arg Phe Gln Ala Trp Tyr Met Tyr Leu Val			
225	230	235	240
Leu Gly Ala Tyr Gly Val Ser Leu Val Tyr Asn Pro Leu Tyr Ile Phe			
245	250	255	
Arg Met Gln His Asn Asp Thr Ile Pro Glu Ser Val Thr Ala Met Arg			
260	265	270	
Glu Asn Gly Phe Leu Arg Arg Tyr Arg Thr Leu Ala Phe Val Met Arg			
275	280	285	
Ala Phe Phe Ile Phe Arg Thr Ala Phe Leu Pro Trp Tyr Leu Thr Gly			
290	295	300	
Thr Ser Leu Leu Ile Thr Ile Pro Leu Val Pro Thr Ala Thr Gly Ala			
305	310	315	320
Phe Leu Thr Phe Phe Phe Ile Leu Ser His Asn Phe Asp Gly Ser Glu			
325	330	335	
Arg Ile Pro Asp Lys Asn Cys Lys Val Lys Arg Ser Glu Lys Asp Val			
340	345	350	
Glu Ala Asp Gln Ile Asp Trp Tyr Arg Ala Gln Val Glu Thr Ser Ser			
355	360	365	
Thr Tyr Gly Gly Pro Ile Ala Met Phe Phe Thr Gly Gly Leu Asn Phe			
370	375	380	
Gln Ile Glu His His Leu Phe Pro Arg Met Ser Ser Trp His Tyr Pro			
385	390	395	400
Phe Val Gln Gln Ala Val Arg Glu Cys Cys Glu Arg His Gly Val Arg			
405	410	415	
Tyr Val Phe Tyr Pro Thr Ile Val Gly Asn Ile Ile Ser Thr Leu Lys			
420	425	430	
Tyr Met His Lys Val Gly Val Val His Cys Val Lys Asp Ala Gln Asp			
435	440	445	

Ser

<210> 248

<211> 449

<212> PRT

<213> *Euglena gracilis*

<400> 248

Met Ala Leu Ser Leu Thr Thr Glu Gln Leu Leu Glu Arg Pro Asp Leu

1 5 10 15

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

20 25 30

His Pro Gly Gly Asp Leu Ile Leu Ala Ser Gly Ala Ser Asp Ala Ser

35 40 45

Pro Leu Phe Tyr Ser Met His Pro Tyr Val Lys Pro Glu Asn Ser Lys

50 55 60

Leu Leu Gln Gln Phe Val Arg Gly Lys His Asp Arg Thr Ser Lys Asp

65 70 75 80

Ile Val Tyr Thr Tyr Asp Ser Pro Phe Ala Gln Asp Val Lys Arg Thr

85 90 95

Met Arg Glu Val Met Lys Gly Arg Asn Trp Tyr Ala Thr Pro Gly Phe

100 105 110

Trp Leu Arg Thr Val Gly Ile Ile Ala Val Thr Ala Phe Cys Glu Trp

115 120 125

His Trp Ala Thr Thr Gly Met Val Leu Trp Gly Leu Leu Thr Gly Phe

130 135 140

Met His Met Gln Ile Gly Leu Ser Ile Gln His Asp Ser Ser His Gly

145 150 155 160

Ala Ile Ser Lys Lys Pro Trp Val Asn Ala Leu Phe Ala Tyr Gly Ile

165 170 175

Asp Val Ile Gly Ser Ser Arg Trp Ile Trp Leu Gln Ser His Ile Met

180 185 190

Arg His His Thr Tyr Thr Asn Gln His Gly Leu Asp Leu Asp Ala Glu

195 200 205
Ser Ala Glu Pro Phe Leu Val Phe His Asn Tyr Pro Ala Ala Asn Thr

210 215 220
Ala Arg Lys Trp Phe His Arg Phe Gln Ala Trp Tyr Met Tyr Leu Val
225 230 235 240
Leu Gly Ala Tyr Gly Val Ser Leu Val Tyr Asn Pro Leu Tyr Ile Phe
245 250 255
Arg Met Gln His Asn Asp Thr Ile Pro Glu Ser Val Thr Ala Met Arg
260 265 270
Glu Asn Gly Phe Leu Arg Arg Tyr Arg Thr Leu Ala Phe Val Met Arg

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

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

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

Ser

<210> 249

<211> 39

<212> DNA

<213> Artificial Sequence

<220><223> Primer

<400> 249

tccatccagc atgatgcggc tcacggggcc atcagcaag 39

<210> 250

<211> 39

<212> DNA

<213> Artificial Sequence

<220><223> Primer

<400> 250

cttgctgatg gccccgtgag ccgcatcatg ctggatgga 39

<210> 251

<211> 39

<212> DNA

<213> Artificial Sequence

<220><223> Primer

<400> 251

tccatccagc atgatgcgtg ccacggggcc atcagcaag 39

<210> 252

<211> 39

<212> DNA

<213> Artificial Sequence

<220><223> Primer

<400> 252

cttgctgatg gccccgtggc acgcatcatg ctggatgga 39

<210> 253

<211> 39

<212> DNA

<213> Artificial Sequence

<220><223> Primer

<400> 253

tccatccagc atgatgcgga ccacggggcc atcagcaag 39

<210> 254

<211> 39

<212> DNA

<213> Artificial Sequence

<220><223> Primer

<400> 254

cttgctgatg gccccgtggt ccgcatcatg ctggatgga 39

<210> 255

<211> 39

<212> DNA

<213> Artificial Sequence

<220><223> Primer

<400> 255

tccatccagc atgatgcgga gcacggggcc atcagcaag 39

<210> 256

<211> 39

<212> DNA

<213> Artificial Sequence

<220><223> Primer

<400> 256

cttgctgatg gccccgtgct ccgcatcatg ctggatgga 39

<210> 257

<211> 39

<212> DNA

<213> Artificial Sequence

<220><223> Primer

<400> 257

tccatccagc atgatgcgtt ccacggggcc atcagcaag 39

<210> 258

<211> 39

<212> DNA
 <213> Artificial Sequence
 <220><223> Primer
 <400> 258
 cttgctgatg gccccgtgga acgcatcatg ctggatgga 39

<210> 259
 <211> 39
 <212> DNA
 <213> Artificial Sequence
 <220><223> Primer
 <400> 259
 tccatccagc atgatgcggg tcacggggcc atcagcaag 39

<210> 260
 <211> 39
 <212> DNA
 <213> Artificial Sequence
 <220><223> Primer
 <400> 260
 cttgctgatg gccccgtgac ccgcatcatg ctggatgga 39

<210> 261
 <211> 39
 <212> DNA
 <213> Artificial Sequence
 <220><223> Primer
 <400> 261
 tccatccagc atgatgcgca ccacggggcc atcagcaag 39

<210> 262
 <211> 39
 <212> DNA
 <213> Artificial Sequence
 <220><223> Primer
 <400> 262
 cttgctgatg gccccgtggt gcgcatcatg ctggatgga 39

<210> 263
 <211> 39
 <212> DNA
 <213> Artificial Sequence
 <220><223> Primer
 <400> 263
 tccatccagc atgatgcgat ccacggggcc atcagcaag 39
 <210> 264
 <211> 39
 <212> DNA
 <213> Artificial Sequence
 <220><223> Primer
 <400> 264
 cttgctgatg gccccgtgga tcgcatcatg ctggatgga 39

 <210> 265
 <211> 39
 <212> DNA
 <213> Artificial Sequence
 <220><223> Primer
 <400> 265
 tccatccagc atgatgcgaa gcacggggcc atcagcaag 39
 <210> 266
 <211> 39
 <212> DNA
 <213> Artificial Sequence
 <220><223> Primer
 <400> 266
 cttgctgatg gccccgtgct tcgcatcatg ctggatgga 39
 <210> 267
 <211> 39
 <212> DNA
 <213> Artificial Sequence
 <220><223> Primer
 <400> 267

tccatccagc atgatgcgct gcacggggcc atcagcaag 39

<210> 268

<211> 39

<212> DNA

<213> Artificial Sequence

<220><223> Primer

<400> 268

cttgctgatg gcccctgca gcgcatcatg ctggatgga 39

<210> 269

<211> 39

<212> DNA

<213> Artificial Sequence

<220><223> Primer

<400> 269

tccatccagc atgatgcgat gcacggggcc atcagcaag 39

<210> 270

<211> 39

<212> DNA

<213> Artificial Sequence

<220><223> Primer

<400> 270

cttgctgatg gcccctgca tcgcatcatg ctggatgga 39

<210> 271

<211> 39

<212> DNA

<213> Artificial Sequence

<220><223> Primer

<400> 271

tccatccagc atgatgcgaa ccacggggcc atcagcaag 39

<210> 272

<211> 39

<212> DNA

<213> Artificial Sequence

<220><223> Primer
 <400> 272
 cttgctgatg gccccgtggt tcgcatcatg ctggatgga 39
 <210> 273
 <211> 39
 <212> DNA
 <213> Artificial Sequence
 <220><223> Primer
 <400> 273
 tccatccagc atgatgcgcc tcacggggcc atcagcaag 39
 <210> 274
 <211> 39
 <212> DNA
 <213> Artificial Sequence
 <220><223> Primer
 <400> 274
 cttgctgatg gccccgtgag gcgcatcatg ctggatgga 39
 <210> 275
 <211> 39
 <212> DNA
 <213> Artificial Sequence
 <220><223> Primer
 <400> 275
 tccatccagc atgatgcgca gcacggggcc atcagcaag 39
 <210> 276
 <211> 39
 <212> DNA
 <213> Artificial Sequence
 <220><223> Primer
 <400> 276
 cttgctgatg gccccgtgct gcgcatcatg ctggatgga 39
 <210> 277
 <211> 39

<212> DNA
 <213> Artificial Sequence
 <220><223> Primer
 <400> 277
 tccatccagc atgatgcgcg acacggggcc atcagcaag 39
 <210> 278
 <211> 39
 <212> DNA
 <213> Artificial Sequence
 <220><223> Primer
 <400> 278
 cttgctgatg gccccgtgtc gcgcatcatg ctggatgga 39
 <210> 279
 <211> 39
 <212> DNA
 <213> Artificial Sequence
 <220><223> Primer
 <400> 279
 tccatccagc atgatgcgac ccacggggcc atcagcaag 39
 <210> 280
 <211> 39
 <212> DNA
 <213> Artificial Sequence
 <220><223> Primer
 <400> 280
 cttgctgatg gccccgtggg tcgcatcatg ctggatgga 39
 <210> 281
 <211> 39
 <212> DNA
 <213> Artificial Sequence
 <220><223> Primer
 <400> 281
 tccatccagc atgatgcggt ccacggggcc atcagcaag 39
 <210> 282

<211> 39
 <212> DNA
 <213> Artificial Sequence
 <220><223> Primer
 <400> 282
 cttgctgatg gccccgtgga cgcacatcatg ctggatgga 39

<210> 283
 <211> 39
 <212> DNA
 <213> Artificial Sequence
 <220><223> Primer
 <400> 283
 tccatccagc atgatgcgtg gcacggggcc atcagcaag 39

<210> 284
 <211> 39
 <212> DNA
 <213> Artificial Sequence
 <220><223> Primer
 <400> 284
 cttgctgatg gccccgtgcc acgcacatcatg ctggatgga 39

<210> 285
 <211> 39
 <212> DNA
 <213> Artificial Sequence
 <220><223> Primer
 <400> 285
 tccatccagc atgatgcgta ccacggggcc atcagcaag 39

<210> 286
 <211> 39
 <212> DNA
 <213> Artificial Sequence
 <220><223> Primer
 <400> 286

cttgctgatg gccccgtggg acgcatcatg ctggatgga

39

<210> 287

<211> 449

<212> PRT

<213> Euglena gracilis

<400> 287

Met Ala Leu Ser Leu Thr Thr Glu Gln Leu Leu Glu Arg Pro Asp Leu

1 5 10 15

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

20 25 30

His Pro Gly Gly Asp Leu Ile Leu Ala Ser Gly Ala Ser Asp Ala Ser

35 40 45

Pro Leu Phe Tyr Ser Met His Pro Tyr Val Lys Pro Glu Asn Ser Lys

50 55 60

Leu Leu Gln Gln Phe Val Arg Gly Lys His Asp Arg Thr Ser Lys Asp

65 70 75 80

Ile Val Tyr Thr Tyr Asp Ser Pro Phe Ala Gln Asp Val Lys Arg Thr

85 90 95

Met Arg Glu Val Met Lys Gly Arg Asn Trp Tyr Ala Thr Pro Gly Phe

100 105 110

Trp Leu Arg Thr Val Gly Ile Ile Ala Val Thr Ala Phe Cys Glu Trp

115 120 125

His Trp Ala Thr Thr Gly Met Val Leu Trp Gly Leu Leu Thr Gly Phe

130 135 140

Met His Met Gln Ile Gly Leu Ser Ile Gln His Asp Ala Ala His Gly

145 150 155 160

Ala Ile Ser Lys Lys Pro Trp Val Asn Ala Leu Phe Ala Tyr Gly Ile

165 170 175

Asp Val Ile Gly Ser Ser Arg Trp Ile Trp Leu Gln Ser His Ile Met

180 185 190

Arg His His Thr Tyr Thr Asn Gln His Gly Leu Asp Leu Asp Ala Glu

195 200 205

Ser Ala Glu Pro Phe Leu Val Phe His Asn Tyr Pro Ala Ala Asn Thr
210 215 220

Ala Arg Lys Trp Phe His Arg Phe Gln Ala Trp Tyr Met Tyr Leu Val
225 230 235 240

Leu Gly Ala Tyr Gly Val Ser Leu Val Tyr Asn Pro Leu Tyr Ile Phe
245 250 255

Arg Met Gln His Asn Asp Thr Ile Pro Glu Ser Val Thr Ala Met Arg
260 265 270

Glu Asn Gly Phe Leu Arg Arg Tyr Arg Thr Leu Ala Phe Val Met Arg
275 280 285

Ala Phe Phe Ile Phe Arg Thr Ala Phe Leu Pro Trp Tyr Leu Thr Gly
290 295 300

Thr Ser Leu Leu Ile Thr Ile Pro Leu Val Pro Thr Ala Thr Gly Ala
305 310 315 320

Phe Leu Thr Phe Phe Phe Ile Leu Ser His Asn Phe Asp Gly Ser Glu
325 330 335

Arg Ile Pro Asp Lys Asn Cys Lys Val Lys Arg Ser Glu Lys Asp Val
340 345 350

Glu Ala Asp Gln Ile Asp Trp Tyr Arg Ala Gln Val Glu Thr Ser Ser
355 360 365

Thr Tyr Gly Gly Pro Ile Ala Met Phe Phe Thr Gly Gly Leu Asn Phe
370 375 380

Gln Ile Glu His His Leu Phe Pro Arg Met Ser Ser Trp His Tyr Pro
385 390 395 400

Phe Val Gln Gln Ala Val Arg Glu Cys Cys Glu Arg His Gly Val Arg
405 410 415

Tyr Val Phe Tyr Pro Thr Ile Val Gly Asn Ile Ile Ser Thr Leu Lys
420 425 430

Tyr Met His Lys Val Gly Val Val His Cys Val Lys Asp Ala Gln Asp
435 440 445

Ser

<210> 288

<211> 449

<212> PRT

<213> *Euglena gracilis*

<400> 288

Met Ala Leu Ser Leu Thr Thr Glu Gln Leu Leu Glu Arg Pro Asp Leu

1 5 10 15

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

20 25 30

His Pro Gly Gly Asp Leu Ile Leu Ala Ser Gly Ala Ser Asp Ala Ser

35 40 45

Pro Leu Phe Tyr Ser Met His Pro Tyr Val Lys Pro Glu Asn Ser Lys

50 55 60

Leu Leu Gln Gln Phe Val Arg Gly Lys His Asp Arg Thr Ser Lys Asp

65 70 75 80

Ile Val Tyr Thr Tyr Asp Ser Pro Phe Ala Gln Asp Val Lys Arg Thr

85 90 95

Met Arg Glu Val Met Lys Gly Arg Asn Trp Tyr Ala Thr Pro Gly Phe

100 105 110

Trp Leu Arg Thr Val Gly Ile Ile Ala Val Thr Ala Phe Cys Glu Trp

115 120 125

His Trp Ala Thr Thr Gly Met Val Leu Trp Gly Leu Leu Thr Gly Phe

130 135 140

Met His Met Gln Ile Gly Leu Ser Ile Gln His Asp Ala Gly His Gly

145 150 155 160

Ala Ile Ser Lys Lys Pro Trp Val Asn Ala Leu Phe Ala Tyr Gly Ile

165 170 175

Asp Val Ile Gly Ser Ser Arg Trp Ile Trp Leu Gln Ser His Ile Met

180 185 190

Arg His His Thr Tyr Thr Asn Gln His Gly Leu Asp Leu Asp Ala Glu

195 200 205

Ser Ala Glu Pro Phe Leu Val Phe His Asn Tyr Pro Ala Ala Asn Thr

210 215 220

Ala Arg Lys Trp Phe His Arg Phe Gln Ala Trp Tyr Met Tyr Leu Val

225 230 235 240

Leu Gly Ala Tyr Gly Val Ser Leu Val Tyr Asn Pro Leu Tyr Ile Phe

245 250 255

Arg Met Gln His Asn Asp Thr Ile Pro Glu Ser Val Thr Ala Met Arg

260 265 270

Glu Asn Gly Phe Leu Arg Arg Tyr Arg Thr Leu Ala Phe Val Met Arg

275 280 285

Ala Phe Phe Ile Phe Arg Thr Ala Phe Leu Pro Trp Tyr Leu Thr Gly

290 295 300

Thr Ser Leu Leu Ile Thr Ile Pro Leu Val Pro Thr Ala Thr Gly Ala

305 310 315 320

Phe Leu Thr Phe Phe Phe Ile Leu Ser His Asn Phe Asp Gly Ser Glu

325 330 335

Arg Ile Pro Asp Lys Asn Cys Lys Val Lys Arg Ser Glu Lys Asp Val

340 345 350

Glu Ala Asp Gln Ile Asp Trp Tyr Arg Ala Gln Val Glu Thr Ser Ser

355 360 365

Thr Tyr Gly Gly Pro Ile Ala Met Phe Phe Thr Gly Gly Leu Asn Phe

370 375 380

Gln Ile Glu His His Leu Phe Pro Arg Met Ser Ser Trp His Tyr Pro

385 390 395 400

Phe Val Gln Gln Ala Val Arg Glu Cys Cys Glu Arg His Gly Val Arg

405 410 415

Tyr Val Phe Tyr Pro Thr Ile Val Gly Asn Ile Ile Ser Thr Leu Lys

420 425 430

Tyr Met His Lys Val Gly Val Val His Cys Val Lys Asp Ala Gln Asp

435 440 445

Ser

<210> 289

<211> 8438

<212> DNA

<213> Artificial Sequence

<220><223> Plasmid pDMW367M4-157g

<400> 289

catggctctc agtcttacca cagaacagct gttagaacgc cctgatttgg ttgcgattga 60

tggtatcctc tacgaccttg aagggttgc caaagttcat ccaggaggag atttgattct 120

cgcttctggt gcctctgatg cctccctctt cttttattca atgcatccat acgtcaaacc 180

ggagaactcc aaattgttc aacagttcgt ccgagggaag catgaccgca cctcgaagga 240

cattgtctac acgtatgatt ctcccttcgc acaagacgtt aagcggacaa tgcgcgaggt 300

gatgaaaggg aggaactggt acgcaacccc ttgcttctgg ctgcgcaccg ttgggatcat 360

cgccgtgacg gccttttgcg agtggcactg ggctaccacg gggatggtgc tgtggggcct 420

gttgactgga ttcatgcaca tgcagatcgg cttatccatc cagcatgatg gctcccacgg 480

ggccatcagc aagaagcctt gggtaacgc cctcttcgcc tacggcattg acgtcatcgg 540

atcgtcccgg tggatttggc tgcagtcgca catcatgcgg caccacacct acaccaacca 600

gcacggcctc gacctggatg cggagtcggc agagccgttc ctggtgttcc acaactaccc 660

cgccgcaaac accgcccga agtggttcca ccgcttcag gcttgggtaca tgtaccttgt 720

gttgggggca tacgggggat cgctgggtga caaccgctc tacattttcc ggatgcagca 780

caatgacacc atcccagagt ctgtcacggc catgcgggaa aatggcttcc tgcggcgcta 840

ccgcacactt gcattcgtga tgcgagcttt ctcatcttc cggaccgcat tcttgccctg 900

gtacctcact gggacctcat tgctgatcac cattcctctg gtgcccaccg caactggtgc 960

cttcttgacg ttctttctca ttttgtccca caattttgat ggctccgaac ggatccccga 1020

caagaactgc aagggttaagc gatctgagaa ggacgttgag gctgacaaa ttgactggta 1080

tcgggagcag gtggagacgt cctccacata cgggtggccc atcgccatgt tcttactgg 1140

cggctcctat ttccagatcg agcaccacct ctttccccgg atgtcgtctt ggcaactacc 1200

cttcgtccag caggcgggtc gggagtgttg cgaacgacat ggagtgcgat atgttttcta 1260

ccctaccatc gtcggcaaca tcatctccac cctgaagtac atgcataagg tgggtgtcgt 1320

ccactgcgtg aaggacgcac aggattccta agcggccgca agtgtggatg gggaagtgag 1380

tgccccgttc tgtgtgcaca attggcaatc caagatggat ggattcaaca cagggatata 1440

gcgagctacg tggtagtgcg aggatatagc aacggatatt tatgtttgac acttgagaat 1500
gtacgataca agcactgtcc aagtacaata ctaaacatac tgtacatact catactcgta 1560
cccgggcaac ggtttcactt gactgcagtg gctagtgtct ttactcgtac agtgtgcaat 1620
actgcgtatc atagtctttg atgtatatcg tattcattca tgttagttgc gtacgagccg 1680
gaagcataaa gtgtaaagcc tggggtgcct aatgagttag ctaactcaca ttaattgcgt 1740

tgcgctcact gcccgttttc cagtgcggaa acctgtcgtg ccagctgcat taatgaatcg 1800
gccaacgcgc ggggagaggg ggtttgcgta ttgggcgtct ttcgcttcc tcgctcactg 1860
actcgtcgcg ctcggtcgtt cggctgcggc gagcggatc agctcactca aaggcggtaa 1920
tacggttata cacagaatca ggggataacg caggaaagaa catgtgagca aaaggccagc 1980
aaaaggccag gaaccgtaaa aaggccgcgt tgctggcggt tttccatagg ctccgcccc 2040
ctgacgagca tcacaaaaat cgacgtcaa gttagagggt gcgaaaccg acaggactat 2100
aaagatacca ggcgtttccc cctggaagct cctcgtgcg ctctcctgtt ccgaccctgc 2160

cgcttaccgg atacctgtcc gcctttctcc ctccgggaag cgtggcgctt tctcatagct 2220
cacgtgtag gtatctcagt tcggttagg tcgttcgtc caagctgggc tgtgtgcacg 2280
aaccctccgt tcagcccgac cgctgcgcct tatccggtaa ctatcgtctt gactccaacc 2340
cggtaaagca cgacttatcg ccaactggcag cagccactgg taacaggatt agcagagcga 2400
ggtatgtagg cggtgtcaca gatttcttga agtggtggcc taactacggc tacactagaa 2460
ggacagtatt tggtagctgc gctctgtga agccagttac ctccgaaaa agagtggta 2520
gctcttgatc cggcaaaaca accaccgtg gtagcggtgg ttttttgtt tgcaagcagc 2580

agattacgcg cagaaaaaaa ggatctcaag aagatccttt gatcttttct acggggtctg 2640
acgtcagtga gaacgaaac tcacgttaag ggatttttgt catgagatta tcaaaaagga 2700
tcttcaccta gatcctttta aattaaaaat gaagttttta atcaatctaa agtatatatg 2760
agtaaacctg gtctgacagt taccaatgct taatcagtga ggacacctat tcagcgatct 2820
gtctatttcg ttcattcata gttgcctgac tccccgtgt gtagataact acgatacggg 2880
agggcttacc atctggcccc agtgcgtcaa tgataccgag agaccacgc tcaccggctc 2940
cagatttatc agcaataaac cagccagccg gaagggccga gcgcagaagt ggtcctgcaa 3000

ctttatccgc ctccatccag tctattaatt gttgccggga agctagagta agtagttcgc 3060
cagttaatag tttagcgaac gttgttgcca ttgctacagg catcgtgggt tcacgtcgt 3120
cgtttggtat ggcttcattc agctccggtt cccaacgac aaggcgagtt acatgatccc 3180
ccatgttgtg caaaaaagcg gttagctcct tcggctctcc gatcgttgtc agaagtaagt 3240
tgcccgaggt gttatcactc atggttatgg cagcactgca taattctctt actgtcatgc 3300

catccgtaag atgcttttct gtgactgggtg agtactcaac caagtcattc tgagaatagt	3360
glatgcggcg accgagttgc tcttgcccgg cgtcaatacg ggataatacc gcgccacata	3420
gcagaacttt aaaagtgtc atcattggaa aacgtttctt ggggcgaaaa ctctcaagga	3480
tcttaccgct gttgagatcc agttcgatgt aaccactcg tgcaccaac tgatcttcag	3540
catcttttac tttcaccagc gtttctgggt gagcaaaaac aggaaggcaa aatgccgcaa	3600
aaaagggaaat aaggcgaca cggaaatgtt gaatactcat actcttcctt tttcaatatt	3660
attgaagcat ttatcagggt tattgtctca tgagcggata catatttgaa tgtatttaga	3720
aaaataaaca aataggggtt ccgcgacat tccccgaaa agtgccacct gacgcgcct	3780
gtagcggcg attaaagcgc gcgggtgttg tggttacgcg cagcgtgacc gctacacttg	3840
ccagcgcct agcgcctgt cctttcgtt tcttccctt ctttctgcc acgttcgccg	3900
gttttccccg tcaagctcta aatcgggggc tcccttagg gttccgattt agtgcttac	3960
ggcacctga ccccaaaaaa cttgattagg gtgatgttc acgtagtggg ccatcgccct	4020
gatagacggt ttttcgcct ttgacgttg agtcacgtt ctttaatagt ggactcttgt	4080
tccaaactgg aacaacactc aacctatct cggctctt ttttgattta taagggattt	4140
tgccgatttc ggcctattgg ttaaaaaatg agctgattta acaaaaattt aacgcgaatt	4200
ttaacaaaat attaacgtt acaatttcca ttgccattc aggtcgcga actgttgga	4260
aggcgatcg gtgcgggcct cttcgtatt acgccagtg gcgaaaggg gatgtgctgc	4320
aaggcgatta agttgggtta gccagggtt tcccagtc cgacgttgta aaacgacgc	4380
cagtgaattg taatacgact cactatagg cgaattgggt accgggcccc cctcagagt	4440
cgatggtgtc gataagcttg atatgaatt catgtcacac aaaccgatct tcgcctcaag	4500
gaaacctaat tctacatccg agagactgcc gagatccagt ctacactgat taattttcgg	4560
gccataaatt taaaaaaatc gtgttatata atattatag tattatatat atacatcatg	4620
atgatactga cagtcagtgc ccattgctaa atagacagac tccatctgcc gcctccaact	4680
gatgttctca atatttaagg ggtcatctcg catgttttaa taataaacag actccatcta	4740
ccgcctccaa atgatgttct caaaatatat tgtatgaact tatttttatt acttagtatt	4800
attagacaac ttacttgctt tatgaaaaac acttcctatt taggaaacaa tttataatgg	4860
cagttcgttc atttaacaat ttatgtagaa taaatgttat aaatgcgtat gggaaatctt	4920
aaatatggat agcataaatg atatctgcat tgcctaattc gaaatcaaca gcaacgaaaa	4980
aaatcccttg tacaacataa atagtcatcg agaaatatca actatcaaag aacagctatt	5040

cacacgttac tattgagatt attattggac gagaatcaca cactcaactg tctttctctc	5100
ttctagaaat acaggtacaa gtatgtacta ttctcattgt tcatacttct agtcatttca	5160
tcccacatat tecttggatt tctctccaat gaatgacatt ctatcttgca aattcaacaa	5220
ttataataag atataccaaa gtagcgggat agtggcaatc aaaaagcttc tctgggtgtc	5280
ttctcgtatt tatttttatt ctaatgatcc attaaaggta tatatttatt tcttggtata	5340
taatcctttt gtttattaca tgggctggat acataaaggat attttgattt aattttttgc	5400
ttaaattcaa tccccctcg ttcagtgtca actgtaatgg taggaaatta ccatactttt	5460
gaagaagcaa aaaaaatgaa agaaaaaaaa aatcgtattt ccaggttaga cgttccgcag	5520
aatctagaat gcggtatgcg gtacattgtt ctccgaacgt aaaagttgcg ctccctgaga	5580
tattgtacat ttttgctttt acaagtacaa gtacatcgta caactatgta ctactgttga	5640
tgcattcaca acagtttgtt ttgttttttt ttgttttttt tttttctaata gattcattac	5700
cgctatgtat acctacttgt actttagta agccgggtta ttggcgttca attaatcata	5760
gacttatgaa tctgcacggt gtgcgctgcg agttactttt agcttatgca tgctacttgg	5820
gtgtaataatt gggatctgtt cggaatacaa cggatgctca atcgatttcg acagtaatta	5880
attaagtcac acacaagtca gctttcttcg agcctcatat aagtataagt agttcaacgt	5940
attagcactg taccagcat ctccgtatcg agaaacacaa caacatgccc cattggacag	6000
atcatgcgga tacacaggtt gtgcagtatc atacatactc gatcagacag gtcgtctgac	6060
catcatacaa gctgaacaag cgctccatac ttgcacgctc tctatataca cagttaaatt	6120
acataccat agtctaacct ctaacagtta atcttctggt aagcctccca gccagccttc	6180
tggatcgcct tggcctcctc aataggatct cggttctggc cgtacagacc tcggccgaca	6240
attatgatat cgttccggt agacatgaca tctcaacag ttcggtactg ctgtccgaga	6300
gcgtctccct tgtcgtcaag acccaccctg ggggtcagaa taagccagtc ctacagatcg	6360
cccttaggtc ggttctgggc aatgaagcca accacaaact cggggtcgga tcgggcaagc	6420
tcaatggtct gcttggagta ctgccagtg gccagagagc ccttgcaaga cagctcggcc	6480
agcatgagca gacctctggc cagcttctcg ttgggagagg ggactaggaa ctcttgtac	6540
tgggagttct cgtagtcaga gacgtcctcc ttcttctgtt cagagacagt ttctcggca	6600
ccagctcgca ggccagcaat gattccggtt ccgggtacac cgtgggcgtt ggtgatatcg	6660
gaccactcgg cgattcgggtg acaccggtac tgggtgcttga cagtgttgc aatatctgcg	6720
aactttctgt cctcgaacag gaagaaaccg tgcttaagag caagttcctt gagggggagc	6780

acagtgccgg cgtagggtgaa gtcgtcaatg atgtcgatat gggttttgat catgcacaca 6840
 taaggctccga ctttatcggc aagctcaatg agtccttgg tggtagtaac atccagagaa 6900
 gcacacaggt tggttttctt ggctgccacg agcttgagca ctcgagcggc aaaggcggac 6960
 ttgtggacgt tagctcgagc ttcgtaggag ggcatTTTgg tggtagaagag gagactgaaa 7020
 taaatttagt ctgcagaact ttttatcgga accttatctg gggcagttaa gtatatgta 7080
 tggtaatagt tacgagttag ttgaacttat agatagactg gactatacgg ctatcggtcc 7140
 aaattagaaa gaacgtcaat ggctctctgg gcgtcgctt tggcgacaaa aatgtgatca 7200

 tgatgaaagc cagcaatgac gttgcagctg atattgtgt cgccaaccg cgccgaaaac 7260
 gcagctgtca gaccacagc ctccaacgaa gaatgtatcg tcaaagtga ccaagcacac 7320
 tcatagttgg agtcgtactc caaaggcggc aatgacgagt cagacagata ctctcgact 7380
 caggcgacga cggaattcct gcagcccatc tgcagaattc aggagagacc gggttggcgg 7440
 cgtatttgtg tcccaaaaaa cagccccaat tgccccggag aagacggcca ggccgcttag 7500
 atgacaaatt caacaactca cagctgactt tctgccattg ccactagggg ggggcctttt 7560
 tatatggcca agccaagctc tccacgtcgg ttgggctgca cccaacaata aatgggtagg 7620

 gttgcaccaa caaagggatg ggatgggggg tagaagatac gaggataacg gggtcaatg 7680
 gcacaaataa gaacgaatac tgccattaag actcgtgatc cagcgactga caccattgca 7740
 tcatctaagg gctcaaaaac tacctcgga ctgctgcgt gatctggaca ccacagaggt 7800
 tccgagcact ttaggttgca ccaaatgtcc caccaggtgc aggagaaaa cgctggaaca 7860
 gcgtgtacag tttgtcttaa caaaaagtga gggcgctgag gtcgagcagg gtggtgtgac 7920
 ttgttatagc ctttagagct gcgaaagcgc gtatggattt ggctcatcag gccagattga 7980
 gggctctgtg acacatgtca tgttagtgta cttcaatgc cccctggata tagccccgac 8040

 aataggccgt ggcctcattt ttttgccttc cgcacatttc cattgctcgg tacccacacc 8100
 ttgcttctcc tgcacttgcc aaccttaata ctggtttaca ttgaccaaca tcttacaagc 8160
 ggggggcttg tctagggtat atataaacag tggtctctcc aatcggttgc cagtctcttt 8220
 tttcctttct tccccacag attcgaaatc taaactacac atcacacaat gcctgttact 8280
 gacgtcctta agcgaaagtc cgggtgtcatc gtcggcgacg atgtccgagc cgtgagtatc 8340
 cagcacaaga tcagtgtcga gacgacgct tttgtgtaat gacacaatcc gaaagtcgct 8400
 agcaacacac actctctaca caaactaacc cagctctc 8438

<210> 290

<211> 8438

<212> DNA

<213> Artificial Sequence

<220><223> Plasmid pDMW367M4-158a

<400> 290

catggctctc agtcttacca cagaacagct gttagaacgc cctgatttgg ttgcgattga	60
tggcatcctc tacgaccttg aagggttgc caaagtcat ccaggaggag atttgattct	120
cgcttctggt gcctctgatg cctccctct cttttattca atgcatccat acgtcaaacc	180
ggagaactcc aaattgcttc aacagttcgt ccgagggaag catgaccgca cctcgaagga	240
cattgtctac acgtatgatt ctcccttgc acaagacgtt aagcggacaa tgcgcgaggt	300
gatgaaaggg aggaactggt acgcaacccc tggcttctgg ctgcgcaccg ttgggatcat	360
cgccgtgacg gccttttgcg agtggcactg ggctaccacg gggatggtgc tgtgggcct	420
gttgactgga ttcatgcaca tgcagatcgg ctatccatc cagcatgatg cggtcacgg	480
ggccatcagc aagaagcctt gggtaacgc cctcttcgcc tacggcattg acgtcatcgg	540
atcgtcccgg tggatttggc tgcagtcgca catcatcggg caccacacct acaccaacca	600
gcacggcctc gacctggatg cggagtcggc agagccgttc ctggtgttcc acaactaccc	660
cgccgaaac accgcccga agtggttcca ccgcttcag gcttgggtaca tgtacctgt	720
gctgggggca tacggggtat cgctgggtga caaccgctc tacattttcc ggatgcagca	780
caatgacacc atcccagagt ctgtcacggc catgcgggaa aatggcttcc tgcggcgcta	840
ccgcacactt gcattctgta tgcgagcttt ctcatcttc cggaccgcat tcttgcctg	900
gtacctact gggacctcat tgcgtatcac cttcctctg gtgcccaccg caactggtgc	960
cttcttgacg ttcttttca ttttgtccca caattttgat ggctccgaac ggatccccga	1020
caagaactgc aaggtaagc gatctgagaa ggacgttgag gctgacaaa ttgactggta	1080
tcgggcgcag gtggagacgt cctccacata cgggtggccc atcgccatgt tcttactgg	1140
cggcttcaat ttccagatcg agcaccacct ctttccccgg atgtcgtctt ggcactaccc	1200
cttcgtccag caggcggctc gggagtgttg cgaacgacat ggagtgcgat atgttttcta	1260
ccctaccatc gtcggcaaca tcatctccac cctgaagtac atgcataagg tgggtgtcgt	1320
ccactgcgtg aaggacgcac aggattccta agcggccgca agtgtggatg gggaagtgag	1380
tgcccgttc tgtgtgcaca attggcaatc caagatggat ggattcaaca cagggatata	1440
gcgagctacg tgggtgtgcg aggatatagc aacggatatt tatgtttgac acttgagaat	1500
gtacgataca agcactgtcc aagtacaata ctaaacatac tgtacatact catactcgta	1560
cccgggcaac ggtttcactt gactgcagtg gctagtgtc ttactcgtac agtgtgcaat	1620
actgcgtatc atagtctttg atgtatatcg tattcattca tgttagttgc gtacgagccg	1680

gaagcataaa gtgtaaagcc tggggtgcct aatgagttag ctaactcaca ttaattgcgt	1740
tgcgtcact gcccgcttcc cagtcgggaa acctgtctg ccagctgcat taatgaatcg	1800
gccaacgcgc ggggagaggg ggtttgcgta ttgggcgctc ttccgcttcc tcgtcactg	1860
actcgtcgc ctcggtcgtt cggctgcggc gagcggatc agtcactca aaggcggtaa	1920
tacggttatc cacagaatca ggggataacg caggaaagaa catgtgagca aaaggccagc	1980
aaaaggccag gaaccgtaaa aaggccgctg tgcctggcgtt tttccatagg ctccgcccc	2040
ctgacgagca tcacaaaaat cgacgtcaa gtcagaggtg gcgaaacccg acaggactat	2100
aaagatacca ggcgcttccc cctggaagct ccctcgtgcg ctctcctgtt ccgacctgc	2160
cgcttaccgg atacctgtcc gcctttctcc cttcggaag cgtggcgctt tctcatagct	2220
cacgtgttag gtatctcagt tcggttagg tcgttcgctc caagctgggc tgtgtgcacg	2280
aacccccctg tcagcccgac cgctgcgctt tatccgtaaa ctatcgtctt gagtccaacc	2340
cggtaagaca cgacttatcg ccactggcag cagccactgg taacaggatt agcagagcga	2400
ggtatgtagg cgggtgtaca gatttcttga agtgggtggc taactacggc tacactagaa	2460
ggacagtatt tggatatcgc gctctgctga agccagttac cttcgaaaaa agagttggta	2520
gctcttgatc cggcaaaaaa accaccgtg gtagcggtag tttttttgtt tgcaagcagc	2580
agattacgcg cagaaaaaaa ggatctcaag aagatccttt gatcttttct acggggtctg	2640
acgtcagtg gaacgaaaaa tcacgttaag ggatttttgt catgagatta tcaaaaagga	2700
tcttcaccta gatcctttta aattaaaaat gaagttttta atcaatctaa agtatatatg	2760
agtaaaactg gtctgacagt taccaatgct taatcagtga ggacacctat tcagcgatct	2820
gtctatttcg ttcattcata gttgcctgac tccccgtcgt gtagataact acgatacggg	2880
agggttacc atctggcccc agtgcgtcaa tgataccgcg agaccacgc tcaccggctc	2940
cagatttatc agcaataaac cagccagccg gaagggccga gcgcagaagt ggtcctgcaa	3000
ctttatccgc ctccatccag tctattaatt gttgccggga agctagagta agtagttcgc	3060
cagttaatag tttgcgcaac gttgttgcca ttgctacagg catcgtggtg tcacgtcgt	3120
cgtttggat ggcttcattc agctccggtt cccaacgac aaggcgagtt acatgatccc	3180
ccatgttgtg caaaaaagcg gttagctcct tcggtcctcc gatcgttgtc agaagtaagt	3240
tggccgcagt gttatcactc atggttatgg cagcactgca taattctctt actgtcatgc	3300
catccgtaag atgcttttct gtgactggig agtactcaac caagtcattc tgagaatagt	3360
glatcgggcg accgagttgc tcttgcccgg cgtcaatacg ggataatacc gcgccacata	3420
gcagaacttt aaaagtgtc atcattggaa aacgtttctt ggggcgaaaa ctctcaagga	3480
tcttaccgct gttgagatcc agttcgatgt aaccactcg tgcaccaac tgatcttcag	3540

catcttttac tttcaccagc gtttctgggt gagcaaaaac aggaaggcaa aatgccgcaa	3600
aaaagggaat aaggcgaca cggaatgtt gaatactcat actcttcctt tttcaatatt	3660
attgaagcat ttatcagggt tattgtctca tgagcggata catatttgaa tgtatttaga	3720
aaaataaaca aataggggtt ccgcgcacat ttccccgaaa agtgccacct gacgcgcctt	3780
gtagcggcgc attaagcgcg gcgggtgttg tggttacgcg cagcgtgacc gctacacttg	3840
ccagcgcctt agcgcgcgtt ctttctgctt tcttcccttc ctttctcgcc acgttcgcgc	3900
gctttccccg tcaagctcta aatcgggggc tccttttagg gttccgattt agtgctttac	3960
ggcacctcga cccccaaaaa cttgattagg gtgatgttc acgtagtggg ccatcgccct	4020
gatagacggt ttttcgcctt ttgacgttgg agtccacgtt ctttaaatag ggactcttgt	4080
tccaaactgg aacaacactc aacctatct cggctctatc ttttgattta taagggattt	4140
tgccgatttc ggctatttg ttaaaaaatg agctgattta acaaaaattt aacgcgaatt	4200
ttaacaaaat attaacgtt acaatttcca ttcgccattc aggcctgcga actgttggga	4260
agggcgatcg gtgcgggcct cttcgtatt acgccagctg gcgaaaggg gatgtgctgc	4320
aaggcgatta agttgggtaa cgccagggtt ttcccagtca cgacgttgta aaacgacggc	4380
cagtgaattg taatacact cactataggg cgaattgggt accgggcccc cctcagagt	4440
cgatggtgtc gataagcttg atatcgaatt catgtcacac aaaccgatct tcgcctcaag	4500
gaaacctaat tctacatccg agagactgcc gagatccagt ctacactgat taattttcgg	4560
gccaataatt taaaaaaatc gtgttatata atatttatg tattatatat atacatcatg	4620
atgatactga cagtcattgc ccattgctaa atagacagac tccatctgcc gcctccaact	4680
gatgttctca atatttaagg ggtcattctc cattgtttta taataaacag actccatcta	4740
ccgcctccaa atgatgttct caaaatatat tgtatgaact tatttttatt acttagtatt	4800
attagacaac ttacttgcct tatgaaaaac acttccctatt taggaaacaa tttataatgg	4860
cagttcgttc atttaacaat ttatgtagaa taaatgttat aaatgcgtat gggaaatctt	4920
aaatatggat agcataaatg atatctgcat tgcctaattc gaaatcaaca gcaacgaaaa	4980
aaatcccttg tacaacataa atagtcacgc agaaatatca actatcaaag aacagctatt	5040
cacacgttac tattgagatt attattggac gagaatcaca cactcaactg tctttctctc	5100
ttctagaaat acaggtacaa gtatgtacta ttctcattgt tcatacttct agtcatttca	5160
tcccacatat tccttggatt tctctccaat gaatgacatt ctatcttgca aattcaacaa	5220
ttataataag atataccaaa gtagcggat agtggcaatc aaaaagcttc tctgggtgtgc	5280
ttctcgtatt tatttttatt ctaatgatcc attaaaggta tataatttatt tcttgttata	5340

taatcctttt gtttattaca tgggctggat acataaagg tttttgattt aattttttgc	5400
ttaaattcaa tccccctcg ttcagtgtca actgtaatgg taggaaatta ccatactttt	5460
gaagaagcaa aaaaaatgaa agaaaaaaaa aatcgatatt ccaggttaga cgttccgcag	5520
aatctagaat gcggtatgcg gtacattgtt cttcgaacgt aaaagttagc ctccttgaga	5580
tattgtacat ttttgccttt acaagtacaa gtacatcgta caactatgta ctactgttga	5640
tgcatccaca acagtttgtt ttgttttttt ttgttttttt tttttcta gattcattac	5700
cgctatgtat acctacttgt actttagta agccgggtta ttggcgttca attaatacata	5760
gacttatgaa tctgcacggt gtgcgctgcg agttactttt agcttatgca tgctacttgg	5820
gtgtaatat gggatctgtt cggaatcaa cggatgctca atcgatttcg acagtaatta	5880
attaagtcac acacaagtc gctttcttcg agcctcatat aagtataagt agttcaacgt	5940
attagcactg taccagcat ctcgtatcg agaacacaa caacatgccc cattggacag	6000
atcatgcgga tacacaggtt gtgcagtatc atacatactc gatcagacag gtcgtctgac	6060
catcatacaa gctgaacaag cgctccatac ttgcacgctc tctatataca cagttaaatt	6120
acataccat agtctaacct ctaacagtta atcttctggg aagcctccca gccagccttc	6180
tggtatcgct tggcctctc aataggatct cggttctggc cgtacagacc tcggccgaca	6240
attatgatat ccgttccggt agacatgaca tctcaacag ttcggtactg ctgtccgaga	6300
gcgtctccct tgcgtcaag acccaccctg ggggtcagaa taagccagtc ctgagatcg	6360
cccttaggtc ggttctgggc aatgaagcca accacaaact cggggtcgga tcgggcaagc	6420
tcaatggctt gcttgagta ctgccagtg gccagagagc ccttgcaaga cagctcgcc	6480
agcatgagca gacctctggc cagcttctcg ttgggagagg ggactaggaa ctcttgtac	6540
tgggagttct cgtagtca gacgtcctc ttcttctgtt cagagacagt ttctcggca	6600
ccagctcgca ggccagcaat gattccggtt ccgggtacac cgtgggcgtt ggtgatatcg	6660
gaccactcgg cgattcgggt acaccggtac tggtgcttga cagtgttgcc aatatctgcg	6720
aactttctgt cctgaacag gaagaaaccg tgcttaagag caagttcctt gagggggagc	6780
acagtcccg cgtaggtaga gtcgtcaatg atgtcgatat gggttttgat catgcacaca	6840
taaggctcca ccttatcggc aagctcaatg agctccttgg tggtagtaac atccagagaa	6900
gcacacaggt tggttttctt ggctgccacg agcttgagca ctgagcggc aaaggcggac	6960
ttgtggacgt tagctcgagc ttcgtaggag ggcatatttg tggtagag gagactgaaa	7020
taaatttagt ctgcagaact ttttatcgga accttatctg gggcagtga gtatatgta	7080

tggtaatagt tacgagttag ttgaacttat agatagactg gactatacgg ctatcgggcc 7140
 aaattagaaa gaacgtcaat ggctctctgg gcgtcgccctt tgccgacaaa aatgtgatca 7200
 tgatgaaagc cagcaatgac gttgcagctg atattgttgt cggccaaccg cgccgaaaac 7260
 gcagctgtca gaccacagc ctccaacgaa gaatgtatcg tcaaagtgat ccaagcacac 7320
 tcatagttag agtcgtactc caaaggcggc aatgacgagt cagacagata ctctcgact 7380
 caggcgacga cggaattcct gcagcccatc tgcagaattc aggagagacc gggttggcgg 7440
 cgtatttgtg tcccaaaaaa cagccccaat tgccccggag aagacggcca ggccgcctag 7500

atgacaaatt caacaactca cagctgactt tctgccattg ccactagggg ggggcctttt 7560
 tatatggcca agccaagctc tccacgtcgg ttgggctgca cccaacaata aatgggtagg 7620
 gttgcaccaa caaagggatg ggatgggggg tagaagatac gaggataacg gggtcaatg 7680
 gcacaaataa gaacgaatac tgccattaag actcgtgac cagcgactga caccattgca 7740
 tcatctaagg gcctcaaaac tacctcgga ctgctgcgt gatctggaca ccacagaggt 7800
 tccgagcact ttaggttgca ccaaatgtcc caccaggtgc aggcagaaaa cgctggaaca 7860
 gcgtgtacag tttgtcttaa caaaaagtga gggcgctgag gtcgagcagg gtggtgtgac 7920

ttgttatagc ctttagagct gcgaaagcgc gtatggattt ggctcatcag gccagattga 7980
 gggctctgtg acacatgtca tgtagtgta cttcaatcgc cccctggata tagccccgac 8040
 aataggccgt ggctcatctt tttgccttc cgcacatttc cattgctcgg tacccacacc 8100
 ttgcttctcc tgcacttgcc aaccttaata ctggtttaca ttgaccaaca tcttacaagc 8160
 ggggggcttg tctagggtat atataaacag tggtctctcc aatcggttgc cagtctcttt 8220
 ttctctttct tccccacag attcgaaatc taaactacac atcacacaat gcctgttact 8280
 gacgtcctta agcgaagtc cgggtgtcatc gtcggcgacg atgtccgagc cgtgagtatc 8340

cacgacaaga tcagtgtcga gacgacgct tttgtgtaat gacacaatcc gaaagtcgct 8400
 agcaacacac actctctaca caaactaacc cagctctc 8438

<210> 291

<211> 8438

<212> DNA

<213> Artificial Sequence

<220><223> Plasmid pDMW367M4-158g

<400> 291

catggctctc agtcttacca cagaacagct gttagaacgc cctgatttgg ttgcgattga 60
 tggcatcctc tacgaccttg aagggttgc caaagttcat ccaggaggag atttgattct 120
 cgcttctggt gcctctgatg cctccctct cttttattca atgcatccat acgtcaaacc 180

ggagaactcc aaattgcttc aacagttcgt ccgaggggaag catgaccgca cctcgaagga 240

cattgtctac acgtatgatt ctcccttcgc acaagacgtt aagcggacaa tgcgcgaggt 300

gatgaaaggg aggaactggg acgcaacccc tggctttctgg ctgcgcaccg ttgggatcat 360

cgccgtgacg gccttttgcg agtggcacitg ggctaccacg gggatggtgc tgtggggcct 420

gttgactgga ttcatgcaca tgcagatcgg ctatccatc cagcatgatg cgggtcacgg 480

ggccatcagc aagaagcctt gggatcaacgc cctcttcgcc tacggcattg acgtcatcgg 540

atcgtcccgg tggatttggc tgcagtcgca catcatgcgg caccacacct acaccaacca 600

gcacggcctc gacctggatg cggagtcggc agagccgttc ctggtgttcc acaactaccc 660

cgccgcaaac accgcccgaag agtgggttcca ccgcttccag gcttgggtaca tgtaccttgt 720

gctgggggca tacgggggtat cgctgggtgta caaccgctc tacattttcc ggatgcagca 780

caatgacacc atcccagagt ctgtcacggc catgcgggaa aatggcttcc tgcggcgcta 840

ccgcacactt gcattcgtga tgcgagcttt ctcatcttc cggaccgcat tcttgccctg 900

gtacctcact gggacctcat tgcgtatcac cattcctctg gtgcccaccg caactgggtgc 960

cttcttgacg ttctttctca ttttgtccca caattttgat ggctccgaac ggateccccga 1020

caagaactgc aagggttaagc gatctgagaa ggacgttgag gctgacaaa ttgactggta 1080

tccggcgagc gtggagacgt cctccacata cgggtggccc atcgccatgt tcttcaactgg 1140

cgggtctcaat ttccagatcg agcaccacct ctctccccgg atgtcgtctt ggcactaccc 1200

cttcgtccag caggcggctc gggagtgttg cgaacacat ggagtgcgat atgttttcta 1260

ccctaccatc gtcggcaaca tcatctccac cctgaagtac atgcataagg tgggtgtcgt 1320

ccactgcgtg aaggacgcac aggattccta agcggccgca agtgtggatg gggaagtgag 1380

tgcccggctc tgtgtgcaca attgcaatc caagatggat ggattcaaca cagggatata 1440

gcgagctacg tgggtgtgag aggatatagc aacggatatt tatgtttgac acttgagaat 1500

gtacgataca agcactgtcc aagtacaata ctaaacatac tgtacatact catactcgta 1560

cccgggcaac ggtttcactt gagtgcagtg gctagtgtc ttactcgtac agtgtgcaat 1620

actgcgtatc atagtctttg atgtatatcg tattcattca tgttagttgc gtacgagccg 1680

gaagcataaa gtgtaaagcc tggggtgcct aatgagtgag ctaactcaca ttaattgcgt 1740

tgcgtcact gcccgtttc cagtgcggaa acctgtcgtg ccagctgcat taatgaatcg 1800

gccaacgcgc ggggagaggg ggtttgcgta ttgggcgtc ttccgcttcc tcgctcactg 1860

actcgtcgc ctcggctgtt cggctgcggc gagcggatc agctcactca aaggcggtaa 1920

tacggttatc cacagaatca ggggataacg caggaaagaa catgtgagca aaaggccagc	1980
aaaaggccag gaaccgtaaa aaggccgcgt tgctggcggt tttccatagg ctccgcccc	2040
ctgacgagca tcacaaaaat cgacgtcaa gtacagagggt gcgaaaccg acaggactat	2100
aaagatacca ggcgtttccc cctggaagct cctcgtgcg ctctcctgtt ccgacctgc	2160
cgcttaccgg atacctgtcc gcctttctcc ctccgggaag cgtggcgctt tctcatagct	2220
cacgtgtag gtatctcagt tcgggtgtagg tcgttcgctc caagctgggc tgtgtgcacg	2280
aacccccgt tcagcccgac cgctgcgcct tatccggtaa ctatcgtctt gagtccaacc	2340
cggtaaagaca cgacttatcg ccactggcag cagccactgg taacaggatt agcagagcga	2400
ggtatgtagg cgggtgtaca gagttcttga agtgggtggc taactacggc tacactagaa	2460
ggacagtatt tggatctgc gctctgctga agccagttac ctccgaaaa agagttggta	2520
gctcttgatc cggcaaaaa accaccgtg gtagcggtgg ttttttgtt tgcaagcagc	2580
agattacgcg cagaaaaaaa ggatctcaag aagatccttt gatctttctt acggggtctg	2640
acgctcagt gaacgaaac tcacgttaag ggattttgt catgagatta tcaaaaagga	2700
tcttcaccta gatcctttta aattaaaaat gaagtttta atcaatctaa agtatatatg	2760
agtaaaacttg gtctgacagt taccaatgct taatcagtga ggcacctatc tcagcgatct	2820
gtctatttcg ttcattcata gttgcctgac tccccgtcgt gtagataact acgatacggg	2880
agggcttacc atctggcccc agtctgtcaa tgataccgag agaccacgc tcaccggctc	2940
cagatttacc agcaataaac cagccagccg gaagggccga gcgcagaagt ggtcctgcaa	3000
ctttatccgc ctccatccag tctattaatt gttgccggga agctagagta agtagttcgc	3060
cagttaatag ttgcgcaac gttgttgcca ttgctacagg catcgtgggtg tcacgctcgt	3120
cgtttggtat ggcttcattc agctccggtt cccaacgac aaggcgagtt acatgatccc	3180
ccatgttgtg caaaaaagcg gttagctcct tcggctctcc gatcgttgtc agaagtaagt	3240
tggccgcagt gttatcactc atggttatgg cagcactgca taattctctt actgtcatgc	3300
catccgtaag atgcttttct gtgactggtg agtactcaac caagtcattc tgagaatagt	3360
glatgcggcg accgagttgc tcttgcccg cgtaataac ggataatacc gcgccacata	3420
gcagaacttt aaaagtgtc atcattggaa aacgttcttc gggcgcaaaa ctctcaagga	3480
tcttaccgct gttgagatcc agttcgatgt aacccactcg tgcacccaac tgatcttcag	3540
catcttttac tttcaccagc gtttctgggt gagcaaaaac aggaaggcaa aatgccgcaa	3600
aaaagggaat aaggcgaca cggaatgtt gaatactcat actcttctt tttcaatatt	3660
attgaagcat ttatcagggt tattgtctca tgagcgata catatttgaa tgtatttaga	3720
aaaataaaca aatagggtt ccgcgcacat tccccgaaa agtgccacct gacgcgcct	3780

gtagcgcgc attaagcgcg gcgggtgtgg tggttacgcg cagcgtgacc gctacacttg 3840
 ccagcgccct agcgcccgct ccttctcgct tcttcccttc ctttctcgcc acgttcgccc 3900
 gctttccccg tcaagctcta aatcgggggc tccctttagg gttccgattt agtgctttac 3960
 ggcacctcga ccccaaaaaa cttgattagg gtgatggttc acgtagtggg ccatcgccct 4020

 gatagacggt ttttcgccct ttgacgttgg agtccacgtt ctttaatagt ggactcttgt 4080
 tccaaactgg aacaacactc aacctatct cggtctatc ttttgattta taagggattt 4140
 tgccgatttc ggcctattgg ttaaaaaatg agctgattta acaaaaattt aacgcgaatt 4200
 ttaacaaaat attaacgctt acaatttcca ttcgccattc aggctgcgca actgttggga 4260
 agggcgatcg gtgcgggcct cttcgtatc acgccagctg gcgaaagggg gatgtgctgc 4320
 aaggcgatta agttgggtaa cgccagggtt tccccagtca cgacgttgta aaacgacggc 4380
 cagtgaattg taatacgact cactataggg cgaattgggt accgggcccc ccctcgaggt 4440

 cgatggtgtc gataagcttg atatgaatt catgtcacac aaaccgatct tcgcctcaag 4500
 gaaacctaat tctacatccg agagactgcc gagatccagt ctacactgat taattttcgg 4560
 gccataaatt taaaaaaatc gtgttatata atattatatg tattatatat atacatcatg 4620
 atgatactga cagtcattgc ccattgctaa atagacagac tccatctgcc gcctccaact 4680
 gatgttctca atatttaagg ggtcatctcg cattgtttta taataaacag actccatcta 4740
 ccgcctcaa atgatgttct caaaatatat tgtatgaact tatttttatt acttagtatt 4800
 attagacaac ttacttgctt tatgaaaaac acttcctatt taggaaacaa ttataatgg 4860

 cagttcgttc atttaacaat ttatgtagaa taaatgttat aaatgcgtat gggaaatctt 4920
 aaatatggat agcataaatg atatctgcat tgcctaattc gaaatcaaca gcaacgaaaa 4980
 aaatcccttg tacaacataa atagtcacg agaaatatca actatcaaag aacagctatt 5040
 cacacgttac tattgagatt attattggac gagaatcaca cactcaactg tctttctctc 5100
 ttctagaaat acaggtagaa gtatgtacta ttctcattgt tcatacttct agtcatttca 5160
 tcccacatat tcttggatt tctctccaat gaatgacatt ctatcttgca aattcaacaa 5220
 ttataataag atataccaaa gtagcggtat agtggcaatc aaaaagcttc tctggtgtgc 5280

 ttctcgtatt tatttttatt ctaatgatcc attaaaggta tatatttatt tcttgttata 5340
 taatcccttt gtttattaca tgggctggat acataaagggt attttgattt aattttttgc 5400
 ttaaattcaa tccccctcg ttcagtgtca actgtaatgg taggaaatta ccatactttt 5460
 gaagaagcaa aaaaaatgaa agaaaaaaa aatcgatatt ccaggttaga cgttccgcag 5520
 aatctagaat gcggtatgcg gtacattgtt cttcgaacgt aaaagtgcg ctccctgaga 5580
 tattgtacat ttttgctttt acaagtacaa gtacatcgta caactatgta ctactgttga 5640

tgcatccaca acagtttgtt ttgttttttt ttgttttttt tttttctaata gattcattac	5700
cgctatgtat acctacttgt actttagta agccgggtta ttggcgttca attaatcata	5760
gacttatgaa tctgcacggt gtgcgctgcg agttactttt agcttatgca tgctacttgg	5820
gtgtaataatt gggatctgtt cggaatacaa cggatgctca atcgatttcg acagtaatta	5880
attaagtcac acacaagtc gctttcttcg agcctcatat aagtataagt agttcaacgt	5940
attagcactg taccagcat ctccgtatcg agaacacaa caacatgccc cattggacag	6000
atcatgcgga tacacaggtt gtgcagtatc atacatactc gatcagacag gtcgtctgac	6060
catcatacaa gctgaacaag cgctccatac ttgcacgctc tctatataca cagttaaatt	6120
acatatccat agictaacct ctaacagtta atcttctggt aagcctccca gccagccttc	6180
tggtatcgct tggcctcctc aataggatct cggttctggc cgtacagacc tcggccgaca	6240
attatgatat ccgttccggt agacatgaca tctcaacag ttcggtactg ctgtccgaga	6300
gcgtctccct tgcgtcaag acccaccctg ggggtcagaa taagccagtc ctgagagtcg	6360
cccttaggtc ggttctgggc aatgaagcca accacaaact cggggtcgga tcgggcaagc	6420
tcaatggtct gcttggagta ctgccagtg gccagagagc ccttgcaaga cagctcggcc	6480
agcatgagca gacctctggc cagcttctcg ttgggagagg ggactaggaa ctcttgtac	6540
ttgggagtct cgtagtcaaa gacgtcctcc ttcttctgtt cagagacagt ttctcggca	6600
ccagctcgca ggccagcaat gattccggtt ccgggtacac cgtgggcgtt ggtgatatcg	6660
gaccactcgg cgattcgggtg acaccggtac ttgtgcttga cagtgttgc aatatctgcg	6720
aactttctgt cctgaacag gaagaaaccg tgcctaagag caagtctctt gagggggagc	6780
acagtgccgg cgtaggtaga gtcgtcaatg atgtcgatat gggttttgat catgcacaca	6840
taaggtcga ccttatcgcc aagctcaatg agctccttgg ttgtggtaac atccagagaa	6900
gcacacaggt ttgttttctt ggctgccacg agcttgagca ctcgagcggc aaaggcggac	6960
ttgtggacgt tagctcgagc ttcttaggag ggcatcttgg ttgtgaagag gagactgaaa	7020
taaatttagt ctgcagaact ttttatcgga accttatctg gggcagtgaa gtatatgtta	7080
ttgtaatagt tacgagttag ttgaacttat agatagactg gactatacgg ctatcggctc	7140
aaattagaaa gaacgtcaat ggctctctgg gcgtcgctt tgccgacaaa aatgtgatca	7200
tgatgaaagc cagcaatgac gttgcagctg atattgttgt cgcccaaccg cgccgaaaac	7260
gcagctgtca gaccacagc ctccaacgaa gaatgtatcg tcaaagtgat ccaagcacac	7320
tcatagttgg agtcgtactc caaaggcggc aatgacgagt cagacagata ctgctcgact	7380

caggcgacga cggaattcct gcagcccatc tgcagaattc aggagagacc gggttggcgg 7440

cgtatttgtg tcccaaaaaa cagccccaat tgccccggag aagacggcca ggccgcctag 7500

atgacaaatt caacaactca cagctgactt tctgccattg ccactagggg ggggcctttt 7560

tatatggcca agccaagctc tccacgtcgg ttgggctgca cccaacaata aatgggtagg 7620

gttgaccaa caaagggatg ggatgggggg tagaagatac gaggataacg gggctcaatg 7680

gcacaaataa gaacgaatac tgccattaag actcgtgac cagcgactga caccattgca 7740

tcatctaagg gcctcaaac tacctcgga ctgctgcgct gatctggaca ccacagaggt 7800

tccgagcact ttaggttgca ccaaatgtcc caccaggtgc aggcagaaaa cgctggaaca 7860

gcgtgtacag tttgtcttaa caaaaagtga gggcgtgag gtcgagcagg gtggtgtgac 7920

ttgttatagc ctttagagct gcgaaagcgc gtatggattt ggctcatcag gccagattga 7980

gggtctgtgg acacatgtca tgttagtgta cttcaatgc cccctggata tagccccgac 8040

aataggccgt ggctcatctt tttgccttc cgcacatttc cattgctcgg taccacacc 8100

ttgcttctcc tgcacttgcc aaccttaata ctggtttaca ttgaccaaca tcttacaagc 8160

ggggggcttg tctagggtat atataaacag tggctctccc aatcggttgc cagtctcttt 8220

tttctttct tccccacag attcgaaatc taaactacac atcacacaat gcctgttact 8280

gagctcctta agcgaaagtc cgggtgtcatc gtcggcgacg atgtccgagc cgtgagtatc 8340

cacgacaaga tcagtgtcga gacgacgcgt tttgtgtaat gacacaatcc gaaagtcgct 8400

agcaacacac actctctaca caaactaacc cagctctc 8438

<210> 292

<211> 38

<212> DNA

<213> Artificial Sequence

<220><223> Primer

<400> 292

ggcttgccaa agttcatggc ggaggagatt tgattctc 38

<210> 293

<211> 38

<212>

> DNA

<213> Artificial Sequence

<220><223> Primer

<400> 293

gagaatcaaa tctcctccgc catgaacttt ggcaagcc 38

<210> 294
 <211> 38
 <212> DNA
 <213> Artificial Sequence
 <220><223> Primer
 <400> 294
 ggcttgccaa agttcatcac ggaggagatt tgattctc 38
 <210> 295
 <211> 38
 <212> DNA
 <213> Artificial Sequence
 <220><223> Primer
 <400> 295
 gagaatcaaa tctcctccgt gatgaacttt ggcaagcc 38
 <210> 296
 <211> 38
 <212> DNA
 <213> Artificial Sequence
 <220><223> Primer
 <400> 296
 ccaaagttca tccaggatcc gatttgattc tcgcttct 38
 <210> 297
 <211> 38
 <212> DNA
 <213> Artificial Sequence
 <220><223> Primer
 <400> 297
 agaagcgaga atcaaatcgg atcctggatg aactttgg 38
 <210> 298
 <211> 1350
 <212> DNA
 <213> *Euglena gracilis*
 <220><221> CDS
 <222> (1)..(1350)

<400> 298

atg gct ctc agt ctt acc aca gaa cag ctg tta gaa cgc cct gat ttg 48

Met Ala Leu Ser Leu Thr Thr Glu Gln Leu Leu Glu Arg Pro Asp Leu

1 5 10 15

gtt gcg att gat ggc atc ctc tac gac ctt gaa ggg ctt gcc aaa gtt 96

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

20 25 30

cat ggt gga gga gat ttg att ctc gct tct ggt gcc tct gat gcc tcc 144

His Gly Gly Gly Asp Leu Ile Leu Ala Ser Gly Ala Ser Asp Ala Ser

35 40 45

cct ctc ttt tat tca atg cat cca tac gtc aaa ccg gag aac tcc aaa 192

Pro Leu Phe Tyr Ser Met His Pro Tyr Val Lys Pro Glu Asn Ser Lys

50 55 60

ttg ctt caa cag ttc gtc cga ggg aag cat gac cgc acc tcg aag gac 240

Leu Leu Gln Gln Phe Val Arg Gly Lys His Asp Arg Thr Ser Lys Asp

65 70 75 80

att gtc tac acg tat gat tct ccc ttc gca caa gac gtt aag cgg aca 288

Ile Val Tyr Thr Tyr Asp Ser Pro Phe Ala Gln Asp Val Lys Arg Thr

85 90 95

atg cgc gag gtg atg aaa ggg agg aac tgg tac gca acc cct ggc ttc 336

Met Arg Glu Val Met Lys Gly Arg Asn Trp Tyr Ala Thr Pro Gly Phe

100 105 110

tgg ctg cgc acc gtt ggg atc atc gcc gtg acg gcc ttt tgc gag tgg 384

Trp Leu Arg Thr Val Gly Ile Ile Ala Val Thr Ala Phe Cys Glu Trp

115 120 125

cac tgg gct acc acg ggg atg gtg ctg tgg ggc ctg ttg act gga ttc 432

His Trp Ala Thr Thr Gly Met Val Leu Trp Gly Leu Leu Thr Gly Phe

130 135 140

atg cac atg cag atc ggc tta tcc atc cag cat gat ggc tcc cac ggg 480

Met His Met Gln Ile Gly Leu Ser Ile Gln His Asp Gly Ser His Gly

145 150 155 160

gcc atc agc aag aag cct tgg gtc aac gcc ctc ttc gcc tac ggc att 528

Ala Ile Ser Lys Lys Pro Trp Val Asn Ala Leu Phe Ala Tyr Gly Ile

 165 170 175

gac gtc atc gga tgc tcc cgg tgg att tgg ctg cag tgc cac atc atg 576

Asp Val Ile Gly Ser Ser Arg Trp Ile Trp Leu Gln Ser His Ile Met

 180 185 190

cgg cac cac acc tac acc aac cag cac ggc ctc gac ctg gat gcg gag 624

Arg His His Thr Tyr Thr Asn Gln His Gly Leu Asp Leu Asp Ala Glu

 195 200 205

tcg gca gag ccg ttc ctg gtg ttc cac aac tac ccc gcc gca aac acc 672

Ser Ala Glu Pro Phe Leu Val Phe His Asn Tyr Pro Ala Ala Asn Thr

 210 215 220

gcc cga aag tgg ttc cac cgc ttc cag gct tgg tac atg tac ctt gtg 720

Ala Arg Lys Trp Phe His Arg Phe Gln Ala Trp Tyr Met Tyr Leu Val

225 230 235 240

ctg ggg gca tac ggg gta tgc ctg gtg tac aac ccg ctc tac att ttc 768

Leu Gly Ala Tyr Gly Val Ser Leu Val Tyr Asn Pro Leu Tyr Ile Phe

 245 250 255

cgg atg cag cac aat gac acc atc cca gag tct gtc acg gcc atg cgg 816

Arg Met Gln His Asn Asp Thr Ile Pro Glu Ser Val Thr Ala Met Arg

 260 265 270

gaa aat ggc ttt ctg cgg cgc tac cgc aca ctt gca ttc gtg atg cga 864

Glu Asn Gly Phe Leu Arg Arg Tyr Arg Thr Leu Ala Phe Val Met Arg

 275 280 285

gct ttc ttc atc ttc cgg acc gca ttc ttg ccc tgg tac ctc act ggg 912

Ala Phe Phe Ile Phe Arg Thr Ala Phe Leu Pro Trp Tyr Leu Thr Gly

 290 295 300

acc tca ttg ctg atc acc att cct ctg gtg ccc acc gca act ggt gcc 960
 Thr Ser Leu Leu Ile Thr Ile Pro Leu Val Pro Thr Ala Thr Gly Ala
 305 310 315 320
 ttc ttg acg ttc ttc ttc att ttg tcc cac aat ttt gat ggc tcc gaa 1008
 Phe Leu Thr Phe Phe Phe Ile Leu Ser His Asn Phe Asp Gly Ser Glu
 325 330 335
 cgg atc ccc gac aag aac tgc aag gtt aag cga tct gag aag gac gtt 1056
 Arg Ile Pro Asp Lys Asn Cys Lys Val Lys Arg Ser Glu Lys Asp Val
 340 345 350
 gag gct gac caa att gac tgg tat cgg gcg cag gtg gag acg tcc tcc 1104
 Glu Ala Asp Gln Ile Asp Trp Tyr Arg Ala Gln Val Glu Thr Ser Ser
 355 360 365
 aca tac ggt ggc ccc atc gcc atg ttc ttc act ggc ggt ctc aat ttc 1152
 Thr Tyr Gly Gly Pro Ile Ala Met Phe Phe Thr Gly Gly Leu Asn Phe
 370 375 380
 cag atc gag cac cac ctc ttt ccc cgg atg tcg tct tgg cac tac ccc 1200
 Gln Ile Glu His His Leu Phe Pro Arg Met Ser Ser Trp His Tyr Pro
 385 390 395 400
 ttc gtc cag cag gcg gtc cgg gag tgt tgc gaa cga cat gga gtg cga 1248
 Phe Val Gln Gln Ala Val Arg Glu Cys Cys Glu Arg His Gly Val Arg
 405 410 415
 tat gtt ttc tac cct acc atc gtc ggc aac atc atc tcc acc ctg aag 1296
 Tyr Val Phe Tyr Pro Thr Ile Val Gly Asn Ile Ile Ser Thr Leu Lys
 420 425 430
 tac atg cat aag gtg ggt gtc gtc cac tgc gtg aag gac gca cag gat 1344
 Tyr Met His Lys Val Gly Val Val His Cys Val Lys Asp Ala Gln Asp
 435 440 445
 tcc taa 1350

Ser

<210> 299

<211> 449

<212> PRT

<213> *Euglena gracilis*

<400> 299

Met Ala Leu Ser Leu Thr Thr Glu Gln Leu Leu Glu Arg Pro Asp Leu

1 5 10 15

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

20 25 30

His Gly Gly Gly Asp Leu Ile Leu Ala Ser Gly Ala Ser Asp Ala Ser

35 40 45

Pro Leu Phe Tyr Ser Met His Pro Tyr Val Lys Pro Glu Asn Ser Lys

50 55 60

Leu Leu Gln Gln Phe Val Arg Gly Lys His Asp Arg Thr Ser Lys Asp

65 70 75 80

Ile Val Tyr Thr Tyr Asp Ser Pro Phe Ala Gln Asp Val Lys Arg Thr

85 90 95

Met Arg Glu Val Met Lys Gly Arg Asn Trp Tyr Ala Thr Pro Gly Phe

100 105 110

Trp Leu Arg Thr Val Gly Ile Ile Ala Val Thr Ala Phe Cys Glu Trp

115 120 125

His Trp Ala Thr Thr Gly Met Val Leu Trp Gly Leu Leu Thr Gly Phe

130 135 140

Met His Met Gln Ile Gly Leu Ser Ile Gln His Asp Gly Ser His Gly

145 150 155 160

Ala Ile Ser Lys Lys Pro Trp Val Asn Ala Leu Phe Ala Tyr Gly Ile

165 170 175

Asp Val Ile Gly Ser Ser Arg Trp Ile Trp Leu Gln Ser His Ile Met

180 185 190

Arg His His Thr Tyr Thr Asn Gln His Gly Leu Asp Leu Asp Ala Glu

195 200 205

Ser Ala Glu Pro Phe Leu Val Phe His Asn Tyr Pro Ala Ala Asn Thr

210 215 220

Ala Arg Lys Trp Phe His Arg Phe Gln Ala Trp Tyr Met Tyr Leu Val

225 230 235 240

Leu Gly Ala Tyr Gly Val Ser Leu Val Tyr Asn Pro Leu Tyr Ile Phe

245 250 255

Arg Met Gln His Asn Asp Thr Ile Pro Glu Ser Val Thr Ala Met Arg

260 265 270

Glu Asn Gly Phe Leu Arg Arg Tyr Arg Thr Leu Ala Phe Val Met Arg

275 280 285

Ala Phe Phe Ile Phe Arg Thr Ala Phe Leu Pro Trp Tyr Leu Thr Gly

290 295 300

Thr Ser Leu Leu Ile Thr Ile Pro Leu Val Pro Thr Ala Thr Gly Ala

305 310 315 320

Phe Leu Thr Phe Phe Phe Ile Leu Ser His Asn Phe Asp Gly Ser Glu

325 330 335

Arg Ile Pro Asp Lys Asn Cys Lys Val Lys Arg Ser Glu Lys Asp Val

340 345 350

Glu Ala Asp Gln Ile Asp Trp Tyr Arg Ala Gln Val Glu Thr Ser Ser

355 360 365

Thr Tyr Gly Gly Pro Ile Ala Met Phe Phe Thr Gly Gly Leu Asn Phe

370 375 380

Gln Ile Glu His His Leu Phe Pro Arg Met Ser Ser Trp His Tyr Pro

385 390 395 400

Phe Val Gln Gln Ala Val Arg Glu Cys Cys Glu Arg His Gly Val Arg

405 410 415

Tyr Val Phe Tyr Pro Thr Ile Val Gly Asn Ile Ile Ser Thr Leu Lys

420 425 430

Tyr Met His Lys Val Gly Val Val His Cys Val Lys Asp Ala Gln Asp

435 440 445

Ser

<210> 300

<211> 1350

<212> DNA

<213> *Euglena gracilis*

<220><221> CDS

<222> (1)..(1350)

<400> 300

atg gct ctc agt ctt acc aca gaa cag ctg tta gaa cgc cct gat ttg 48

Met Ala Leu Ser Leu Thr Thr Glu Gln Leu Leu Glu Arg Pro Asp Leu

1 5 10 15

gtt gcg att gat ggc atc ctc tac gac ctt gaa ggg ctt gcc aaa gtt 96

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

20 25 30

cat ggt gga gga gat ttg att ctc gct tct ggt gcc tct gat gcc tcc 144

His Gly Gly Gly Asp Leu Ile Leu Ala Ser Gly Ala Ser Asp Ala Ser

35 40 45

cct ctc ttt tat tca atg cat cca tac gtc aaa ccg gag aac tcc aaa 192

Pro Leu Phe Tyr Ser Met His Pro Tyr Val Lys Pro Glu Asn Ser Lys

50 55 60

ttg ctt caa cag ttc gtc cga ggg aag cat gac cgc acc tcg aag gac 240

Leu Leu Gln Gln Phe Val Arg Gly Lys His Asp Arg Thr Ser Lys Asp

65 70 75 80

att gtc tac acg tat gat tct ccc ttc gca caa gac gtt aag cgg aca 288

Ile Val Tyr Thr Tyr Asp Ser Pro Phe Ala Gln Asp Val Lys Arg Thr

85 90 95

atg cgc gag gtg atg aaa ggg agg aac tgg tac gca acc cct ggc ttc 336

Met Arg Glu Val Met Lys Gly Arg Asn Trp Tyr Ala Thr Pro Gly Phe

100 105 110

tgg ctg cgc acc gtt ggg atc atc gcc gtg acg gcc ttt tgc gag tgg 384

Trp Leu Arg Thr Val Gly Ile Ile Ala Val Thr Ala Phe Cys Glu Trp

115 120 125

cac tgg gct acc acg ggg atg gtg ctg tgg ggc ctg ttg act gga ttc 432

His Trp Ala Thr Thr Gly Met Val Leu Trp Gly Leu Leu Thr Gly Phe
130 135 140
atg cac atg cag atc ggc tta tcc atc cag cat gat gcg gct cac ggg 480
Met His Met Gln Ile Gly Leu Ser Ile Gln His Asp Ala Ala His Gly
145 150 155 160
gcc atc agc aag aag cct tgg gtc aac gcc ctc ttc gcc tac ggc att 528
Ala Ile Ser Lys Lys Pro Trp Val Asn Ala Leu Phe Ala Tyr Gly Ile
165 170 175
gac gtc atc gga tcg tcc cgg tgg att tgg ctg cag tcg cac atc atg 576
Asp Val Ile Gly Ser Ser Arg Trp Ile Trp Leu Gln Ser His Ile Met
180 185 190
cgg cac cac acc tac acc aac cag cac ggc ctc gac ctg gat gcg gag 624
Arg His His Thr Tyr Thr Asn Gln His Gly Leu Asp Leu Asp Ala Glu
195 200 205
tcg gca gag ccg ttc ctg gtg ttc cac aac tac ccc gcc gca aac acc 672
Ser Ala Glu Pro Phe Leu Val Phe His Asn Tyr Pro Ala Ala Asn Thr
210 215 220
gcc cga aag tgg ttc cac cgc ttc cag gct tgg tac atg tac ctt gtg 720
Ala Arg Lys Trp Phe His Arg Phe Gln Ala Trp Tyr Met Tyr Leu Val
225 230 235 240
ctg ggg gca tac ggg gta tcg ctg gtg tac aac ccg ctc tac att ttc 768
Leu Gly Ala Tyr Gly Val Ser Leu Val Tyr Asn Pro Leu Tyr Ile Phe
245 250 255
cgg atg cag cac aat gac acc atc cca gag tct gtc acg gcc atg cgg 816
Arg Met Gln His Asn Asp Thr Ile Pro Glu Ser Val Thr Ala Met Arg
260 265 270
gaa aat ggc ttt ctg cgg cgc tac cgc aca ctt gca ttc gtg atg cga 864
Glu Asn Gly Phe Leu Arg Arg Tyr Arg Thr Leu Ala Phe Val Met Arg
275 280 285
gct ttc ttc atc ttc cgg acc gca ttc ttg ccc tgg tac ctc act ggg 912
Ala Phe Phe Ile Phe Arg Thr Ala Phe Leu Pro Trp Tyr Leu Thr Gly

290	295	300	
acc tca ttg ctg atc acc att cct ctg gtg ccc acc gca act ggt gcc			960
Thr Ser Leu Leu Ile Thr Ile Pro Leu Val Pro Thr Ala Thr Gly Ala			
305	310	315	320
ttc ttg acg ttc ttc ttc att ttg tcc cac aat ttt gat ggc tcc gaa			1008
Phe Leu Thr Phe Phe Phe Ile Leu Ser His Asn Phe Asp Gly Ser Glu			
	325	330	335
cgg atc ccc gac aag aac tgc aag gtt aag cga tct gag aag gac gtt			1056
Arg Ile Pro Asp Lys Asn Cys Lys Val Lys Arg Ser Glu Lys Asp Val			
	340	345	350
gag get gac caa att gac tgg tat cgg gcg cag gtg gag acg tcc tcc			1104
Glu Ala Asp Gln Ile Asp Trp Tyr Arg Ala Gln Val Glu Thr Ser Ser			
	355	360	365
aca tac ggt ggc ccc atc gcc atg ttc ttc act ggc ggt ctc aat ttc			1152
Thr Tyr Gly Gly Pro Ile Ala Met Phe Phe Thr Gly Gly Leu Asn Phe			
	370	375	380
cag atc gag cac cac ctc ttt ccc cgg atg tcg tct tgg cac tac ccc			1200
Gln Ile Glu His His Leu Phe Pro Arg Met Ser Ser Trp His Tyr Pro			
385	390	395	400
ttc gtc cag cag gcg gtc cgg gag tgt tgc gaa cga cat gga gtg cga			1248
Phe Val Gln Gln Ala Val Arg Glu Cys Cys Glu Arg His Gly Val Arg			
	405	410	415
tat gtt ttc tac cct acc atc gtc ggc aac atc atc tcc acc ctg aag			1296
Tyr Val Phe Tyr Pro Thr Ile Val Gly Asn Ile Ile Ser Thr Leu Lys			
	420	425	430
tac atg cat aag gtg ggt gtc gtc cac tgc gtg aag gac gca cag gat			1344
Tyr Met His Lys Val Gly Val Val His Cys Val Lys Asp Ala Gln Asp			
	435	440	445
tcc taa			1350
Ser			

<210> 301

<211> 449

<212> PRT

<213> *Euglena gracilis*

<400> 301

Met Ala Leu Ser Leu Thr Thr Glu Gln Leu Leu Glu Arg Pro Asp Leu

1 5 10 15

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

20 25 30

His Gly Gly Gly Asp Leu Ile Leu Ala Ser Gly Ala Ser Asp Ala Ser

35 40 45

Pro Leu Phe Tyr Ser Met His Pro Tyr Val Lys Pro Glu Asn Ser Lys

50 55 60

Leu Leu Gln Gln Phe Val Arg Gly Lys His Asp Arg Thr Ser Lys Asp

65 70 75 80

Ile Val Tyr Thr Tyr Asp Ser Pro Phe Ala Gln Asp Val Lys Arg Thr

85 90 95

Met Arg Glu Val Met Lys Gly Arg Asn Trp Tyr Ala Thr Pro Gly Phe

100 105 110

Trp Leu Arg Thr Val Gly Ile Ile Ala Val Thr Ala Phe Cys Glu Trp

115 120 125

His Trp Ala Thr Thr Gly Met Val Leu Trp Gly Leu Leu Thr Gly Phe

130 135 140

Met His Met Gln Ile Gly Leu Ser Ile Gln His Asp Ala Ala His Gly

145 150 155 160

Ala Ile Ser Lys Lys Pro Trp Val Asn Ala Leu Phe Ala Tyr Gly Ile

165 170 175

Asp Val Ile Gly Ser Ser Arg Trp Ile Trp Leu Gln Ser His Ile Met

180 185 190

Arg His His Thr Tyr Thr Asn Gln His Gly Leu Asp Leu Asp Ala Glu

195 200 205
 Ser Ala Glu Pro Phe Leu Val Phe His Asn Tyr Pro Ala Ala Asn Thr
 210 215 220
 Ala Arg Lys Trp Phe His Arg Phe Gln Ala Trp Tyr Met Tyr Leu Val
 225 230 235 240
 Leu Gly Ala Tyr Gly Val Ser Leu Val Tyr Asn Pro Leu Tyr Ile Phe

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

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

 370 375 380
 Gln Ile Glu His His Leu Phe Pro Arg Met Ser Ser Trp His Tyr Pro
 385 390 395 400
 Phe Val Gln Gln Ala Val Arg Glu Cys Cys Glu Arg His Gly Val Arg
 405 410 415
 Tyr Val Phe Tyr Pro Thr Ile Val Gly Asn Ile Ile Ser Thr Leu Lys
 420 425 430
 Tyr Met His Lys Val Gly Val Val His Cys Val Lys Asp Ala Gln Asp

 435 440 445

Ser

<210> 302

<211> 1350

<212> DNA

<213> *Euglena gracilis*

<220><221> CDS

<222> (1)..(1350)

<400> 302

atg gct ctc agt ctt acc aca gaa cag ctg tta gaa cgc cct gat ttg 48

Met Ala Leu Ser Leu Thr Thr Glu Gln Leu Leu Glu Arg Pro Asp Leu

1 5 10 15

gtt gcg att gat ggc atc ctc tac gac ctt gaa ggg ctt gcc aaa gtt 96

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

20 25 30

cat ggt gga gga gat ttg att ctc gct tct ggt gcc tct gat gcc tcc 144

His Gly Gly Gly Asp Leu Ile Leu Ala Ser Gly Ala Ser Asp Ala Ser

35 40 45

cct ctc ttt tat tca atg cat cca tac gtc aaa ccg gag aac tcc aaa 192

Pro Leu Phe Tyr Ser Met His Pro Tyr Val Lys Pro Glu Asn Ser Lys

50 55 60

ttg ctt caa cag ttc gtc cga ggg aag cat gac cgc acc tcg aag gac 240

Leu Leu Gln Gln Phe Val Arg Gly Lys His Asp Arg Thr Ser Lys Asp

65 70 75 80

att gtc tac acg tat gat tct ccc ttc gca caa gac gtt aag cgg aca 288

Ile Val Tyr Thr Tyr Asp Ser Pro Phe Ala Gln Asp Val Lys Arg Thr

85 90 95

atg cgc gag gtg atg aaa ggg agg aac tgg tac gca acc cct ggc ttc 336

Met Arg Glu Val Met Lys Gly Arg Asn Trp Tyr Ala Thr Pro Gly Phe

100 105 110

tgg ctg cgc acc gtt ggg atc atc gcc gtg acg gcc ttt tgc gag tgg 384

Trp Leu Arg Thr Val Gly Ile Ile Ala Val Thr Ala Phe Cys Glu Trp

115	120	125	
cac tgg gct acc acg ggg atg gtg ctg tgg ggc ctg ttg act gga ttc			432
His Trp Ala Thr Thr Gly Met Val Leu Trp Gly Leu Leu Thr Gly Phe			
130	135	140	
atg cac atg cag atc ggc tta tcc atc cag cat gat gcg ggt cac ggg			480
Met His Met Gln Ile Gly Leu Ser Ile Gln His Asp Ala Gly His Gly			
145	150	155	160
gcc atc agc aag aag cct tgg gtc aac gcc ctc ttc gcc tac ggc att			528
Ala Ile Ser Lys Lys Pro Trp Val Asn Ala Leu Phe Ala Tyr Gly Ile			
165	170	175	
gac gtc atc gga tcg tcc cgg tgg att tgg ctg cag tcg cac atc atg			576
Asp Val Ile Gly Ser Ser Arg Trp Ile Trp Leu Gln Ser His Ile Met			
180	185	190	
cgg cac cac acc tac acc aac cag cac ggc ctc gac ctg gat gcg gag			624
Arg His His Thr Tyr Thr Asn Gln His Gly Leu Asp Leu Asp Ala Glu			
195	200	205	
tcg gca gag ccg ttc ctg gtg ttc cac aac tac ccc gcc gca aac acc			672
Ser Ala Glu Pro Phe Leu Val Phe His Asn Tyr Pro Ala Ala Asn Thr			
210	215	220	
gcc cga aag tgg ttc cac cgc ttc cag gct tgg tac atg tac ctt gtg			720
Ala Arg Lys Trp Phe His Arg Phe Gln Ala Trp Tyr Met Tyr Leu Val			
225	230	235	240
ctg ggg gca tac ggg gta tcg ctg gtg tac aac ccg ctc tac att ttc			768
Leu Gly Ala Tyr Gly Val Ser Leu Val Tyr Asn Pro Leu Tyr Ile Phe			
245	250	255	
cgg atg cag cac aat gac acc atc cca gag tct gtc acg gcc atg cgg			816
Arg Met Gln His Asn Asp Thr Ile Pro Glu Ser Val Thr Ala Met Arg			
260	265	270	
gaa aat ggc ttt ctg cgg cgc tac cgc aca ctt gca ttc gtg atg cga			864
Glu Asn Gly Phe Leu Arg Arg Tyr Arg Thr Leu Ala Phe Val Met Arg			

275	280	285	
gct ttc ttc atc ttc cgg acc gca ttc ttg ccc tgg tac ctc act ggg			912
Ala Phe Phe Ile Phe Arg Thr Ala Phe Leu Pro Trp Tyr Leu Thr Gly			
290	295	300	
acc tca ttg ctg atc acc att cct ctg gtg ccc acc gca act ggt gcc			960
Thr Ser Leu Leu Ile Thr Ile Pro Leu Val Pro Thr Ala Thr Gly Ala			
305	310	315	320
ttc ttg acg ttc ttc ttc att ttg tcc cac aat ttt gat ggc tcc gaa			1008
Phe Leu Thr Phe Phe Phe Ile Leu Ser His Asn Phe Asp Gly Ser Glu			
325	330	335	
cgg atc ccc gac aag aac tgc aag gtt aag cga tct gag aag gac gtt			1056
Arg Ile Pro Asp Lys Asn Cys Lys Val Lys Arg Ser Glu Lys Asp Val			
340	345	350	
gag gct gac caa att gac tgg tat cgg gcg cag gtg gag acg tcc tcc			1104
Glu Ala Asp Gln Ile Asp Trp Tyr Arg Ala Gln Val Glu Thr Ser Ser			
355	360	365	
aca tac ggt ggc ccc atc gcc atg ttc ttc act ggc ggt ctc aat ttc			1152
Thr Tyr Gly Gly Pro Ile Ala Met Phe Phe Thr Gly Gly Leu Asn Phe			
370	375	380	
cag atc gag cac cac ctc ttt ccc cgg atg tcg tct tgg cac tac ccc			1200
Gln Ile Glu His His Leu Phe Pro Arg Met Ser Ser Trp His Tyr Pro			
385	390	395	400
ttc gtc cag cag gcg gtc cgg gag tgt tgc gaa cga cat gga gtg cga			1248
Phe Val Gln Gln Ala Val Arg Glu Cys Cys Glu Arg His Gly Val Arg			
405	410	415	
tat gtt ttc tac cct acc atc gtc ggc aac atc atc tcc acc ctg aag			1296
Tyr Val Phe Tyr Pro Thr Ile Val Gly Asn Ile Ile Ser Thr Leu Lys			
420	425	430	
tac atg cat aag gtg ggt gtc gtc cac tgc gtg aag gac gca cag gat			1344

Tyr Met His Lys Val Gly Val Val His Cys Val Lys Asp Ala Gln Asp
 435 440 445

tcc taa 1350

Ser

<210> 303

<211> 449

<212> PRT

<213> Euglena gracilis

<400> 303

Met Ala Leu Ser Leu Thr Thr Glu Gln Leu Leu Glu Arg Pro Asp Leu

1 5 10 15
 Val Ala Ile Asp Gly Ile Leu Tyr Asp Leu Glu Gly Leu Ala Lys Val
 20 25 30

His Gly Gly Gly Asp Leu Ile Leu Ala Ser Gly Ala Ser Asp Ala Ser
 35 40 45

Pro Leu Phe Tyr Ser Met His Pro Tyr Val Lys Pro Glu Asn Ser Lys
 50 55 60

Leu Leu Gln Gln Phe Val Arg Gly Lys His Asp Arg Thr Ser Lys Asp

65 70 75 80
 Ile Val Tyr Thr Tyr Asp Ser Pro Phe Ala Gln Asp Val Lys Arg Thr
 85 90 95

Met Arg Glu Val Met Lys Gly Arg Asn Trp Tyr Ala Thr Pro Gly Phe
 100 105 110

Trp Leu Arg Thr Val Gly Ile Ile Ala Val Thr Ala Phe Cys Glu Trp
 115 120 125

His Trp Ala Thr Thr Gly Met Val Leu Trp Gly Leu Leu Thr Gly Phe

130 135 140
 Met His Met Gln Ile Gly Leu Ser Ile Gln His Asp Ala Gly His Gly
 145 150 155 160

Ala Ile Ser Lys Lys Pro Trp Val Asn Ala Leu Phe Ala Tyr Gly Ile
 165 170 175

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

420	425	430	
Tyr Met His Lys Val Gly Val Val His Cys Val Lys Asp Ala Gln Asp			
435	440	445	
Ser			
<210>	304		
<211>	4070		
<212>	DNA		
<213>	Artificial Sequence		
<220><223>	Plasmid pEgD5M		
<400>	304		
tcgcgcgttt cggtgatgac 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 aagggcgatc ggtgcgggcc tcttcgtat	300		
tacgccagct ggcgaaaagg ggatgtgctg caaggcgatt aagttgggta acgccagggt	360		
tttcccagtc acgacgttgt aaaacgacgg ccagtgccat ggccctgtct ctactaccg	420		
aacagctcct ggagcgacct gatctcgtcg ctatcgatgg tatcctgtac gacctcgagg	480		
gcctggccaa ggtgcatggt ggtggagacc tcattctggc ctctggagcc tccgacgcct	540		
ctccctctt ctactctatg catccctacg tcaagcccga gaactccaag ctctgcagc	600		
aattcgtecg agggaagcat gaccgcacct cgaaggacat tgtctacacg tatgattctc	660		
ccttcgcaca agacgttaa ggcacaatgc gcgaggatg gaaagggagg aactggtacg	720		
caaccctgg ctcttggtg cgcaccgttg ggatcatcgc cgtgacggcc ttttcgagt	780		
ggcactgggc taccacgggg atggtgctgt ggggcctgtt gactggattc atgcacatgc	840		
agatcggtt atccatccag catgatgcgg gtcacggggc catcagcaag aagccttggg	900		
tcaacgcect cttcgctac ggcatcgacg tcateggatc gtcccgttg atttggctgc	960		
agtcgcacat catcgggcac cacacctaca ccaaccagca cggcctcgac ctggatgcgg	1020		
agtcggcaga gccgttcctg gtgttcaca actacccgc cgaaacacc gcccgaagt	1080		
ggttcaccg cttccaggct tggtagatgt acctgtgct gggggcatac ggggtatcgc	1140		
tgggtlataa cccgctctac attttccgga tgcagcaca tgacaccatc ccagagtctg	1200		

tcacggccat gcgggaaaat ggctttctgc ggcgctaccg cacacttgca ttcgtgatgc 1260
 gagctttctt catcttccgg accgcattct tgccttgga cctcactggg acctcattgc 1320
 tgatcaccat tctcttggtg cccaccgcaa ctggtgcctt cttgacgttc ttcttcattt 1380
 tgtccacaa ttttgatggc tccgaacgga tccccgacaa gaactgcaag gtttaagcgat 1440
 ctgagaagga cgttgaggct gaccaaattg actggtatcg ggcgaggtg gagacgtcct 1500
 ccacatacgg tggccccatc gccatgttct tctactggcg tctcaatttc cagatcgagc 1560
 accacctctt tccccggatg tctctttggc actaccctt cgtccagcag gcggtccggg 1620

 agtgttgca acgacatgga gtgcgatatg ttttctaccc taccatcgtc ggcaacatca 1680
 tctccacctt gaagtacatg cataaggtgg gtgtcgtcca ctgctgaag gacgcacagg 1740
 attcctaagc ggccgcaatt cgagctcggg acctcgcaa tgcactaga tatcgatcc 1800
 cgggcccgtc gactgcagag gcctgcacgc aagcttggcg taatcatggt catagctgtt 1860
 tctgtgtga aattgttctc cgctcacaat tccacacaac atacgagccg gaagcataaa 1920
 gtgtaaagc tggggtgcct aatgagttag ctaactcaca ttaattgctg tgcgtcact 1980
 gcccgtttc cagtccggaa acctgtcgtg ccagctgcat taatgaatcg gccaacgcgc 2040

 ggggagaggc ggtttgcgta ttgggcgctc ttccgcttcc tgcgtcactg actcgtcgcg 2100
 ctcggtcgtt cggctgcggc gagcggatc agctcactca aaggcggtaa tacggttctc 2160
 cacagaatca ggggataacg caggaaagaa catgtgagca aaaggccagc aaaaggccag 2220
 gaaccgtaaa aaggccgctg tctggcggtt tttccatagg ctccgcccc ctgacgagca 2280
 tcacaaaaat cgacgtcaa gtcagagggt gcgaaacccg acaggactat aaagatacca 2340
 ggcttttccc cctggaagct cctcgtgctg ctctcctgtt ccgacctgc cgcttaccgg 2400
 atacctgtcc gccttttccc ctccgggaag cgtggcgctt tctcatagct cacgtgtag 2460

 gtatctcagt tccgtgtagg tcttctgctc caagctgggc tgtgtgcacg aacccccgt 2520
 tcagcccgac cgctgcgctt tatccggtta ctatcgtctt gactccaacc cggttaagaca 2580
 cgacttatcg cactggcag cagccactgg taacaggatt agcagagcga ggtatgtagg 2640
 cgggtctaca gacttcttga agtggtggcc taactacggc tacactagaa gaacagtatt 2700
 tggatctgc gctctgctga agccagttac ctccggaaaa agagtggta gctcttgatc 2760
 cggcaaaaa accaccgtg gtagcgggtg tttttttgtt tgcaagcagc agattacgcg 2820
 cagaaaaaaa ggatctcaag aagatccttt gatcttttct acggggtctg acgtcagtg 2880

 gaacgaaaac tcacgttaag ggattttggt catgagatta tcaaaaagga tcttcaccta 2940
 gatcctttta aattaaaaa gaagttttta atcaatctaa agtatatatg agtaaaactg 3000
 gtctgacagt taccaatgct taatcagtga ggcacctatc tcagcgatct gtctatttcg 3060

ttcatccata gttgcctgac tccccgtcgt gtagataact acgatacggg agggcttacc 3120
atctggcccc agtgcctgcaa tgataccgcg agaccacgc tcaccggctc cagatttacc 3180
agcaataaac cagccagccg gaagggccga gcgcagaagt ggtcctgcaa ctttatccgc 3240
ctccatccag tctattaatt gttgccggga agctagagta agtagttcgc cagttaatag 3300

tttgcgcaac gttgttgcca ttgctacagg catcgtgggtg tcacgctcgt cgtttggtat 3360
ggcttcattc agctccggtt cccaacgac aaggcgagtt acatgatccc ccatgttgtg 3420
caaaaaagcg gttagctcct tcggtcctcc gatcgttgtc agaagtaagt tggccgcagt 3480
gttatcactc atggttatgg cagcactgca taattctctt actgtcatgc catccgtaag 3540
atgcttttct gtgactgggtg agtactcaac caagtcatc tgagaatagt gtatgcggcg 3600
accgagttgc tcttgcccgg cgtcaatacg ggataatacc gcgccacata gcagaacttt 3660
aaaagtgtc atcattggaa aacgttcttc ggggcgaaaa ctctcaagga tcttaccgct 3720

gttgagatcc agttcgatgt aaccactcgc tgcaccaac tgatcttcag catcttttac 3780
tttcaccagc gtttctgggt gagcaaaaac aggaaggcaa aatgccgcaa aaaagggaat 3840
aaggcgaca cggaatgtt gaatactcat actcttctt tttcaatatt attgaagcat 3900
ttatcagggt tattgtctca tgagcggata catatttgaa tgtatttaga aaaataaaca 3960
aataggggtt ccgcgcacat ttccccgaaa agtgccacct gacgtctaag aaaccattat 4020
tatcatgaca ttaacctata aaaataggcg taccacgagg ccttttcgtc 4070

<210> 305

<211> 8438

<212> DNA

<213> Artificial Sequence

<220><223> Plasmid pDMW367-5M

<400> 305

ggccgcaagt gtggatgggg aagtgagtgc ccggttctgt gtgcacaatt ggcaatccaa 60
gatggatgga ttcaacacag ggatatagcg agctacgtgg tgggtcgagg atatagcaac 120
ggatatattat gtttgacact tgagaatgta cgatacaagc actgtccaag tacaatacta 180
aacatactgt acatactcat actcgtaccc gggcaacggt ttactttgag tgcagtggct 240
agtgtcttta ctctacagt gtgcaatact gcgtatcata gtctttgatg tatatcgtat 300
tcattcatgt tagttgcgta cgagccggaa gcataaagtg taaagcctgg ggtgcctaata 360
gagttagcta actcacatta attgcgttgc gctcactgcc cgctttccag tcgggaaacc 420

tgctgtgcca gctgcattaa tgaatcggcc aacgcgctgg gagaggcggt ttgcgtattg 480

ggcgctcttc cgcttctctg ctactgact cgtgcgctc ggtcgttcgg ctgcggcgag 540
 cggatcagc tactcaaag gcgtaatac ggttatccac agaacagg gataacgcag 600
 gaaagaacat gtgagcaaaa ggccagcaaa aggccaggaa ccgtaaaaag gccgcgttgc 660
 tggcgttttt ccataggctc cgccccctg acgagcatca caaaaatcga cgctcaagtc 720
 agagggtggcg aaacccgaca ggactataaa gataccaggc gtttcccctt ggaagctccc 780
 tcgtgcgctc tctgttccg accctgccgc ttaccggata cctgtccgcc ttttccctt 840

cgggaagcgt ggcgctttct catagctcac gctgtaggta tctcagttcg gtgtaggtcg 900
 ttcgtctcaa gctgggctgt gtgcacgaac ccccgctca gcccgaccgc tgcgccttat 960
 ccggtaaacta tctcttgag tccaaccgg taagacacga cttatcgcca ctggcagcag 1020
 ccactggtaa caggattagc agagcgaggt atgtaggcgg tgctacagag ttcttgaagt 1080
 ggtggcctaa ctacggctac actagaagga cagtatttgg tatctgcgct ctgctgaagc 1140
 cagttacctt cgaaaaaga gttgtagct cttgatccgg caaacaacc accgctggta 1200
 gcggtggttt tttgtttgc aagcagcaga ttacgcgcag aaaaaagga tctcaagaag 1260

atcctttgat cttttctacg gggctcgacg ctcagtgga cgaaaactca cgttaaggga 1320
 ttttggctcat gagattatca aaaaggatct tcacctagat ctttttaaat taaaaatgaa 1380
 gttttaaatc aatctaaagt atatatgagt aaacttggc tgacagttac caatgcttaa 1440
 tcagttaggc acctatctca gcgactgtc tatttcgttc atccatagtt gcctgactcc 1500
 ccgtcgtgta gataactacg atacgggagg gcttaccatc tggccccagt gctgcaatga 1560
 taccgcgaga cccacgtca ccggctccag atttatcagc aataaaccag ccagccggaa 1620
 ggcccgagcg cagaagtggc cctgcaactt tatccgctc catccagtct attaatgtt 1680

gccggaagc tagagtaagt agttcgccag ttaatagttt gcgcaacgtt gttgccattg 1740
 ctacaggcat cgtgggtgca cgctcgctgt ttggtatggc ttcatcagc tccggttccc 1800
 aacgatcaag gcgagtaca tgatcccca tgttgtgcaa aaaagcgggt agtccttcg 1860
 gtctccgat cgttgtcaga agtaagttgg ccgcagtgtt atcactcatg gttatggcag 1920
 cactgcataa ttctcttact gtcatgcat ccgtaagatg cttttctgtg actggtgagt 1980
 actcaaccaa gtcattctga gaatagtga tgcggcgacc gagttgctct tgccccgct 2040
 caatacggga taataccgcg ccacatagca gaactttaaa agtgcctatc attgaaaaac 2100

gttcttcggg gcgaaaactc tcaaggatct taccgtgtt gagatccagt tcgatgtaac 2160
 ccactcgtgc acccaactga tcttcagcat cttttacttt caccagcgtt tctgggtgag 2220
 caaaaacagg aaggcaaat gccgcaaaaa agggaataag ggcgacacgg aaatgttgaa 2280
 tactcatact ctctctttt caatattatt gaagcattta tcagggttat tgtctcatga 2340

gcggatacat atttgaatgt atttagaaaa ataaacaaat aggggttccg cgcacatttc 2400
 cccgaaaagt gccacctgac gcgcccgtga gcggcgcat aagcgcgcg ggtgtggtgg 2460
 ttacgcgcag cgtgaccgct acacttgcca gcgcccctagc gcccgtcct ttcgctttct 2520

tcccttcttt tctcgccacg ttcgcggct tccccgtca agctctaaat cgggggctcc 2580
 ctttagggtt ccgatttagt gctttacggc acctcgacc caaaaaactt gattagggtg 2640
 atggttcacg tagtgggcca tcgccctgat agacggtttt tcgcccttg acgttggagt 2700
 ccacgttctt taatagtga ctcttggtcc aaactggaac aacactcaac cctatctcgg 2760
 tctattcttt tgatttataa gggattttgc cgatttcggc ctattggtta aaaaatgagc 2820
 tgatttaaca aaaatttaac gcgaatttta acaaaatatt aacgcttaca atttcattc 2880
 gccattcagg ctgcgcaact gttgggaagg gcgatcgggt cgggcctctt cgctattacg 2940

ccagctggcg aaaggggat gtgctgcaag gcgattaagt tgggtaacgc cagggttttc 3000
 ccagtcacga cgttgtaaaa cgacggccag tgaattgtaa tacgactcac tataggcgca 3060
 attgggtacc gggccccccc tcgaggtcga tgggtgctgat aagcttgata tcgaattcat 3120
 gtcacacaaa ccgattctcg cctcaaggaa acctaatct acatccgaga gactgccgag 3180
 atccagtcta cactgattaa ttttcgggcc aataatttaa aaaaatcgtg ttataataa 3240
 ttatatgtat tatatatata catcatgatg atactgacag tcatgtccca ttgctaaata 3300
 gacagactcc atctgccgc tccaactgat gtcttcaata ttttaagggt catctcgcat 3360

tgtttaataa taaacagact ccatctaccg cctccaatg atgttctcaa aatatattgt 3420
 atgaacttat ttttattact tagtattatt agacaactta cttgctttat gaaaaacact 3480
 tcctatttag gaacaattt ataatggcag ttcgttcatt taacaattta tgtagaataa 3540
 atgttataaa tgcgtatggg aaatcttaaa tatggatagc ataaatgata tctgcattgc 3600
 ctaattcgaa atcaacagca acgaaaaaaaa tcccttgtag aacataaata gtcacgaga 3660
 aatatcaact atcaaagaac agctattcac acgttactat tgagattatt attggacgag 3720
 aatcacacac tcaactgtct ttctctcttc tagaaataga ggtacaagta tgtactattc 3780

tcattgttca tacttctagt catttcaccc cacatatcc ttggatttct ctccaatgaa 3840
 tgacattcta tcttgcaaat tcaacaatta taataagata taccaaagta gcggtatagt 3900
 ggcaatcaaa aagcttctct ggtgtgcttc tcgtatttat ttttattcta atgatccatt 3960
 aaaggatat atttatttct tgttatataa tccttttggt tattacatgg gctggataga 4020
 taaaggattt ttgatttaat tttttgctta aattcaatcc cccctcgttc agtgtcaact 4080
 gtaatggtag gaaattacca tacttttgaa gaagcaaaaa aatgaaaga aaaaaaaaa 4140

cgtatttcca ggtagacgt tccgcagaat ctagaatgcg gtatgcggta cattgttctt	4200
cgaacgtaaa agttgcgctc cctgagatat tgtacatitt tgctttttaca agtacaagta	4260
catcgtacaa ctatgtacta ctgttgatgc atccacaaca gtttgttttg tttttttttg	4320
tttttttttt ttctaagtat tcattaccgc tatgtatacc tacttgtact ttagtaagc	4380
cgggttattg gcgttcaatt aatcatagac ttatgaatct gcacgggtgtg cgctgcgagt	4440
tacttttagc ttatgcatgc tacttgggtg taatattggg atctgttcgg aaatcaacgg	4500
atgctcaatc gatttcgaca gtaattaatt aagtcataca caagtcagct ttcttcgagc	4560
ctcatataag tataagtagt tcaacgtatt agcactgtac ccagcatctc cgtatcgaga	4620
aacacaacaa catgccccat tggacagatc atgcggatac acaggttgtg cagtatcata	4680
catactcgat cagacaggtc gtctgacat catacaagct gaacaagcgc tccatacttg	4740
cacgtctct atatacacag ttaaattaca tatccatagt ctaacctcta acagttaatc	4800
ttctggtaag cctcccagcc agccttctgg tatcgcttgg cctcctcaat aggatctcgg	4860
ttctggccgt acagacctcg gccgacaatt atgatatccg ttccggtaga catgacatcc	4920
tcaacagttc ggtactgtg tccgagagcg tctcccttgt cgtcaagacc caccccgggg	4980
gtcagaataa gccagtcctc agagtcgccc ttaggtcggg tctgggcaat gaagccaacc	5040
acaaactcgg ggtcggatcg ggcaagctca atggtctgct tggagtactc gccagtggcc	5100
agagagccct tgcaagacag ctccggccagc atgagcagac ctctggccag ctctctgttg	5160
ggagaggga ctaggaactc cttgtactgg gagttctctg agtcagagac gtctctctc	5220
ttctgttcag agacagtttc ctccgcacca gctcgcaggc cagcaatgat tccggttccg	5280
ggtacaccgt gggcgttggg gatatcggac cactcggcga ttccgtgaca ccggtactgg	5340
tgttgacag tgttgccaat atctgcgaac ttctgtctc cgaacaggaa gaaacctgac	5400
ttaagagcaa gttccttgag ggggagcaca gtgccggcgt aggtgaagtc gtcaatgatg	5460
tcgatatggg tttgatcat gcacacataa ggtccgacct tatcggcaag ctcaatgagc	5520
tccttggtgg tggtaacatc cagagaagca cacaggttgg ttttcttggc tgccacgagc	5580
ttgagcactc gagcggcaaa ggcggacttg tggacgttag ctcgagcttc gtaggagggc	5640
atcttggtgg tgaagaggag actgaaataa atttagtctg cagaactttt tatcggaacc	5700
ttatctgggg cagtgaagta tatgttatgg taatagtac gagttagtgt aacttataga	5760
tagactggac tatacggcta tcggtccaaa ttagaaagaa cgtcaatggc tctctgggcg	5820
tcgcctttgc cgacaaaaat gtgatcatga tgaaagccag caatgacgtt gcagctgata	5880

ttgttgcgg ccaaccgcgc cgaaaacgca gctgtcagac ccacagcctc caacgaagaa	5940
tgtatcgtca aagtgatcca agcacactca tagttggagt cgtactccaa aggcggcaat	6000
gacgagtcag acagatactc gtcgactcag gcgacgacgg aattcctgca gcccattctgc	6060
agaattcagg agagaccggg ttggcggcgt atttgtgtcc caaaaaacag cccaattgc	6120
cccgagaag acggccaggc cgcctagatg acaaatca caactcacag ctgactttct	6180
gccattgcc ctaggggggg gcctttttat atggccaagc caagctctcc acgtcggttg	6240
ggctgcaccc aacaataaat gggtagggtt gcaccaacaa agggatggga tgggggtag	6300
aagatacgag gataacgggg ctcaatggca caaataagaa cgaatactgc cattaagact	6360
cgtgatccag cgactgacac cattgcatca tctaagggcc tcaaaactac ctcggaactg	6420
ctgcgtgat ctggacacca cagaggttcc gagcacttta ggttgcacca aatgtccac	6480
caggtgcagg cagaaaacgc tggaacagcg tgtacagttt gtcttaacaa aaagtgagg	6540
cgctgaggtc gagcagggtg gtgtgacttg ttatagcctt tagagctgcg aaagcgcgta	6600
tggatttggc tcatcaggcc agattgaggg tctgtggaca catgtcatgt tagtgtactt	6660
caatgcccc ctggatatag ccccgacaat aggcctggc ctcatttttt tgccttcgc	6720
acatttccat tgctcggtac ccacaccttg cttctcctgc acttgccaac ctaatactg	6780
gtttacattg accaaccatc tacaagcggg gggttgtct agggatatata taaacagtgg	6840
ctctcccaat cggttgccag tctctttttt cttttcttc cccacagatt cgaaatctaa	6900
actacacatc acacaatgcc tgttactgac gtccttaagc gaaagtccgg tgtcatcgtc	6960
ggcgacgatg tccgagccgt gagtatccac gacaagatca gtgtcgagac gacgcgtttt	7020
gtgtaatgac acaatccgaa agtcgctagc aacacacact ctctacacaa actaaccag	7080
ctctccatgg cctgtctct cactaccgaa cagctcctgg agcgacctga tctcgtcgt	7140
atcgatggtg tctgtacga cctcgagggc ctggccaagg tgcattggtg tggagacctc	7200
attctggcct ctggagcctc cgacgcctct ccctcttct actctatgca tccctacgtc	7260
aagcccgaga actccaagct cctgcagca ttcgtccgag ggaagcatga ccgcacctcg	7320
aaggacattg tctacacgta tgattctccc ttgcacaag acgttaagcg gacaatgcgc	7380
gaggtgatga aaggaggaa ctggtacgca acccctggct tctggctgcg caccgttggg	7440
atcatcgccg tgacggcctt ttgcgagtgg cactgggcta ccacggggat ggtgctgtgg	7500
ggcctgttga ctggattcat gcacatgcag atcggttat ccatccagca tgatcggggt	7560
cacggggcca tcagcaagaa gccttgggtc aacgcctct tcgcctacgg cattgacgtc	7620
atcggtatgt cccggtggat ttggctgcag tcgcacatca tgcggcacca cacctacacc	7680
aaccagcacg gcctcgacct ggatgcggag tcggcagagc cgttcctggt gttccacaac	7740

taccccgccg caaacaccgc ccgaaagtgg ttccaccgct tccaggcttg gtacatgtac	7800
cttgtgctgg gggcataccg ggtatcgctg gtgtacaacc cgctctacat ttcccgatg	7860
cagcacaatg acaccatccc agagtctgtc acggccatgc gggaaaatgg ctttctgcgg	7920
cgctaccgca cacttgcatt cgtgatgcga gctttcttca tcttccggac cgcattcttg	7980
ccctggtagc tcaactgggac ctcatgtctg atcaccattc ctctgggtgc caccgcaact	8040
ggtgccttct tgacgttctt cttcatcttg tcccacaatt ttgatggctc cgaacggatc	8100
cccgacaaga actgcaaggt taagcgatct gagaaggacg ttgaggctga ccaaattgac	8160
tggtatcggg cgcaggtgga gacgtcctcc acatacggtg gccccatcgc catgttcttc	8220
actggcggtc tcaatttcca gatcgagcac cacctcttcc cccggatgic gtcttggcac	8280
tacccttctg tccagcaggc ggtccgggag tgttgcgaac gacatggagt gcgatatgtt	8340
ttctacccta ccatcgctcg caacatcatc tccacctga agtacatgca taaggtgggt	8400
gtcgtccact gcgtaagga cgcacaggat tcctaagc	8438
<210> 306	
<211> 4070	
<212> DNA	
<213> Artificial Sequence	
<220><223> Plasmid pEgD5M1	
<400> 306	
tcgcgcgttt cggatgatgac ggtgaaaacc tctgacacat gcagctcccg gagacggtea	60
cagcttgtct gtaagcggat gccgggagca gacaagcccg tcagggcgcg tcagcgggtg	120
ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc	180
accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcaggcgcc	240
attcgccatt caggctgcgc aactgttggg aagggcgac ggtgcgggcc tcttcgctat	300
tacgccagct ggcgaaaggg ggatgtgctg caaggcgatt aagttgggta acgccagggt	360
tttccagtc acgacgttgt aaaacgacgg ccagtgccat ggccctgtct ctcaactaccg	420
aacagctcct ggagcgacct gatctcgtcg ctatcgatgg tatcctgtac gacctcgagg	480
gcctggccaa ggtgcatggt ggtggagacc tcattctggc ctctggagcc tccgacgcct	540
ctcccccttt ctactctatg catccctacg tcaagcccga gaactccaag ctctgcagc	600
aattcgtccg agggaagcat gaccgcacct cgaaggacat tgtctacacg tatgattctc	660
ccttcgcaca agacgttaag cggacaatgc gcgaggtgat gaaagggagg aactggtacg	720
caaccctgg cttctggctg cgcaccgttg ggatcatcgc cgtgacggcc ttttgcgagt	780

ggcactgggc taccacgggg atgggtgctgt ggggcctgtt gactggattc atgcacatgc	840
agatcggtt atccatccag catgatgcgg gtcacggggc catcagcaag aagccttggg	900
tcaacgccct cttgcctac ggcatgtacg tcatcggtac gtcccgttg atttggtgc	960
agtgcacat catgcggcac cacacctaca ccaaccagca cggcctcgac ctggatgcgg	1020
agtgcgcaga gccgttcttg gtgttccaca actacccgc cgcaaacacc gcccgaagt	1080
ggttccaccg cttccaggt tggtagatgt acctgtgtt gggggcatac ggggtatgc	1140
tgggtgataa cccgtctac attttccgga tgcagcaca tgacaccatc ccagagtctg	1200
tcacggccat gcgggaaaaa ggctttctgc ggcgtaccg cacacttgca ttcgtgatgc	1260
gagctttctt catcttcgg accgcattct tgcctggta cctcactgg acctcattgc	1320
tgatcacat tctctggtg cccaccgaa ctggtgcctt cttgacgtt tcttcatctt	1380
tgtccacaa tttgatggc tccgaacgga tccccgaca gaactgcaag gttagagct	1440
ctgagaagga cgttgaggt gaccaaattg actggtatcg ggcgcagggt gagacgtcct	1500
ccacatacgg tggccccatc gccatgttct tctactggcg tctcaatttc cagatcgagc	1560
accactctt tccccgatg tegtcttggc actacccctt cgtccagcag gcggtccggg	1620
agtgttcga acgacatgga gtgcgatatg tttctaccc taccatcgtc ggcaacatca	1680
tctccacct gaagtacatg cataaggtgg gtgtcgtcca ctgctgaag gacgcacagg	1740
attcctaagc ggccgcaatt cgagctcgg acctcgcga tgcatctaga tatcggtacc	1800
cgggccctgc gactgcagag gcctgcacg aagcttggcg taatcatggt catagctgtt	1860
tctgtgtga aattgttacc cgtcacat tccacacaa atacgagccg gaagcataaa	1920
gtgtaaagc tggggtgcct aatgagtga ctaactcaca ttaattgcgt tgcgtcact	1980
gcccgtttc cagtcgggaa acctgtctg ccagctgcat taatgaatcg gccaacgcgc	2040
ggggagaggc ggtttgcgt tgggcgtc ttcgcttcc tgcctactg actcgtgcg	2100
ctcggtcgtt cggctgcgg gagcggtatc agctactca aaggcggtaa tacggttatc	2160
cacagaatca ggggataacg caggaaagaa catgtgagca aaaggccagc aaaaggccag	2220
gaaccgtaaa aaggccggt tgcgtgcgt tttccatagg ctcgcccc ctgacgagca	2280
tcacaaaaat cgacgtcaa gtcagaggtg gcgaaaccg acaggactat aaagatacca	2340
ggcgtttccc cctggaagt cctcgtgcg ctctctgtt ccgacctgc cgttaccgg	2400
atactgtcc gcctttctc cttcggaag cgtggcgtt tctcatagct cacgtgtag	2460
gtatctcagt tcggtgtag tcgttcgtc caagctgggc tgtgtgcacg aacccccgt	2520
tcagcccgac cgtgcgctt tatccgtaa ctatctctt ggtccaacc cggtaagaca	2580
cgacttatcg cactggcag cagccactgg taacaggatt agcagagcga ggtatgtagg	2640

cgggtgtctaca gaggttcttga agtgggtggcc taactacggc tacactagaa gaacagtatt 2700
 tggatatctgc gctctgtctga agccagttac ctccggaaaa agagtttgta gctcttgatc 2760
 cggcaaacaa accaccgctg gtagcgggtgg tttttttgtt tgcaagcagc agattacgcg 2820

 cagaaaaaaa ggatctcaag aagatccttt gatcttttct acgggggtctg acgctcagtg 2880
 gaacgaaaaac tcacgttaag ggatttttgg catgagatta tcaaaaagga tcttcaccta 2940
 gatcctttta aattaaaaat gaagttttta atcaatctaa agtatatatg agtaaacctg 3000
 gtctgacagt taccaatgct taatcagtga ggacacctatc tcagcgatct gtctatttcg 3060
 ttcattccata gttgcctgac tccccgtcgt gtagataact acgatacggg agggccttacc 3120
 atctggcccc agtgcctcaa tgataccgcg agaccacgc tcaccggctc cagatttacc 3180
 agcaataaac cagccagccg gaagggccga gcgcagaagt ggtcctgcaa ctttatccgc 3240

 ctccatccag tctattaatt gttgccggga agctagagta agtagttcgc cagttaatag 3300
 tttgcgcaac gttgttgcca ttgctacagg catcgtgggtg tcacgctcgt cgtttggat 3360
 ggcttcattc agctccggtt cccaacgac aaggcgagtt acatgatccc ccatgttgtg 3420
 caaaaaagcg gttagctcct tcggtcctcc gatcgttgte agaagtaagt tggccgcagt 3480
 gttatcactc atggttatgg cagcactgca taattctctt actgtcatgc catccgtaag 3540
 atgcttttct gtgactgggtg agtactcaac caagtcattc tgagaatagt gtatgcggcg 3600
 accgagttgc tcttgccccg cgtcaatacg ggataatacc gcgccacata gcagaacttt 3660

 aaaagtgtc atcattggaa aacgttcttc ggggcgaaaa ctctcaagga tcttaccgt 3720
 gttgagatcc agttcgatgt aaccactcg tgcaccaac tgatcttcag catcttttac 3780
 tttcaccagc gtttctgggt gagcaaaaac aggaaggcaa aatgccgcaa aaaagggaat 3840
 aaggcgaca cggaaatgtt gaatactcat actcttctt tttcaatatt attgaagcat 3900
 ttatcagggt tattgtctca tgagcggata catatttgaa tgtatttaga aaaataaaca 3960
 aataggggtt ccgcgcacat tccccgaaa agtgccacct gacgtctaag aaaccattat 4020
 tatcatgaca ttaacctata aaaataggcg tatcacgagg ccttttcgtc 4070

 <210> 307
 <211> 8438
 <212> DNA
 <213> Artificial Sequence
 <220><223> Plasmid pDMW367-5M1
 <400> 307
 ggccgcaagt gtggatgggg aagtgagtgc ccggttctgt gtgcacaatt ggcaatccaa 60

gatggatgga ttcaacacag ggatatagcg agctacgtgg tggcgagg atatagcaac	120
ggataatttat gtttgacact tgagaatgta cgatacaagc actgtccaag tacaatacta	180
aacataactgt acataactcat actcgtaccc gggcaacggt ttactttgag tgcagtggct	240
agtgccttta ctctgacagt gtgcaatact gcgtatcata gtctttgatg tatatcgtat	300
tcattcatgt tagttgcgta cgagccggaa gcataaagtg taaagcctgg ggtgcctaata	360
gagtgaagcta actcacatta attgcgttgc gctcactgcc cgctttccag tcgggaaacc	420
tgtcgtgccca gctgcattaa tgaatcggcc aacgcgcggg gagaggcggt ttgcgtattg	480
ggcgctcttc cgcttctcgc ctcactgact cgtcgcctc ggtcgttcgg ctgcggcgag	540
cggtatcagc tcaactcaag gcggttaatac ggttatccac agaactcagg gataacgcag	600
gaaagaacat gtgagcaaaa ggccagcaaa aggccaggaa ccgtaaaaag gccgcgttgc	660
tggcgttttt ccatagctc cgccccctg acgagcatca caaaaatcga cgtcgaagtc	720
agagggtggcg aaacccgaca ggactataaa gataccaggc gtttccccct ggaagctccc	780
tcgtgcgctc tctgttccg accctgccgc ttaccggata cctgtccgc tttctccctt	840
cggaagcgt ggcgctttct catagctcac gctgtaggta tctcagttcg gtgtaggtcg	900
ttcgctccaa gctgggctgt gtgcacgaac cccccgtca gcccgaccgc tgcgccttat	960
ccggttaacta tcgtcttgag tccaacccgg taagacacga cttatcgcca ctggcagcag	1020
ccactggtaa caggattagc agagcgaggt atgtaggcgg tgctacagag ttcttgaagt	1080
ggtggcctaa ctacggctac actagaagga cagtatttgg tatctgcgct ctgctgaagc	1140
cagttacctt cgaaaaaga gttggtagct cttgatccgg caaacaacc accgctggta	1200
gcggtggttt tttgtttgc aagcagcaga ttacgcgcag aaaaaagga tctcaagaag	1260
atcctttgat cttttctac gggctcgcg ctcagtggaa cgaaaactca cgttaaggga	1320
ttttggtcat gagattatca aaaaggatct tcacctagat ctttttaaat taaaaatgaa	1380
gttttaaatc aatctaaagt atatatgagt aaacttggc tgacagttac caatgcttaa	1440
tcagtgaaggc acctatctca gcgatctgtc tatttcgttc atccatagtt gcctgactcc	1500
ccgtcgtgta gataactacg atacgggagg gcttaccatc tggccccagt gctgcaatga	1560
taccgcgaga cccacgtca ccggctccag atttatcagc aataaaccag ccagccggaa	1620
gggccgagcg cagaagtggc cctgcaactt tatccgcctc catccagtct attaatgtt	1680
gccgggaagc tagagtaagt agttcgccag ttaatatgtt gcgcaacgtt gttgccattg	1740
ctacaggcat cgtggtgtca cgctcgtcgt ttggtatggc ttcatcagc tccggttccc	1800
aacgatcaag gcgagttaca tgatccccca tgttgtgcaa aaaagcggtt agctccttcg	1860
gtcctccgat cgttgtcaga agtaagtgg ccgcagtgtt atcactcatg gttatggcag	1920

cactgcataa ttctcttact gtcatgccat ccgtaagatg cttttctgtg actggtgagt	1980
actcaaccaa gtcattctga gaatagtgtg tgcggcgacc gagttgctct tgcccggcgt	2040
caatacggga taataccgcg ccacatagca gaactttaaa agtgcctatc attggaaaac	2100
gttcttcggg gcgaaaactc tcaaggatct taccgctgtt gagatccagt tcgatgtaac	2160
ccactcgtgc acccaactga tcttcagcat cttttacttt caccagcgtt tctgggtgag	2220
caaaaacagg aaggcaaaat gccgcaaaaa agggaataag ggcgacacgg aatgttgaa	2280
tactcatact ctctcttttt caatattatt gaagcattta tcagggttat tgtctcatga	2340
gcggatacat atttgaatgt atttagaaaa ataaacaaat aggggttcgg cgacatttc	2400
cccgaagaat gccacctgac gcgcctctga gcggcgcat aagcgcgcg ggtgtggtgg	2460
ttacgcgcag cgtgaccgt acacttgcca gcgccttagc gcccgtcct ttcgtttct	2520
tcccttctct tctgcacag ttcgcgggt tccccgtca agctctaaat cgggggctcc	2580
ctttagggtt ccgatttagt gctttacggc acctcgacc caaaaaactt gattagggtg	2640
atggttcacg tagtgggcca tcgccctgat agacggtttt tcgccctttg acgttgagat	2700
ccacgttctt taatagtga ctcttggttc aaactggaac aacactcaac cctatctcgg	2760
tctattcttt tgatttataa gggattttgc cgatttcggc ctattggtta aaaaatgagc	2820
tgatttaaca aaaatttaac gcgaatttta acaaaatatt aacgcttaca atttcattc	2880
gccattcagg ctgcgcaact gttgggaagg gcgatcggtg cgggcctctt cgctattacg	2940
ccagctggcg aaagggggat gtgctgcaag gcgattaagt tgggtaacgc cagggttttc	3000
ccagtcacga cgttgtaaaa cgacggccag tgaattgtaa tacgactcac tataggcgga	3060
attgggtacc gggcccccc tcgaggtcga tgggtgcat aagcttgata tcgaattcat	3120
gtcacacaaa ccgactctcg cctcaaggaa acctaatct acatccgaga gactgccgag	3180
atccagtcta cactgattaa ttttcgggcc aataatttaa aaaaatcgtg ttatataata	3240
ttatatgtat tatatatata catcatgat atactgacag tcatgtcca ttgctaaata	3300
gacagactcc atctgccgc tccaactgat gttctcaata ttttaagggt catctcgcat	3360
tgtttaataa taaacagact ccatctaccg cctccaaatg atgttctcaa aatatattgt	3420
atgaacttat ttttattact tagtattatt agacaactta cttgctttat gaaaaacact	3480
tcctatttag gaaacaattt ataatggcag ttcgttcatt taacaattta tgtagaataa	3540
atgttataaa tgcgtatggg aaatcttaaa tatggatagc ataaatgata tctgcattgc	3600
ctaattcgaa atcaacagca acgaaaaaaaa tccttgtac aacataaata gtcacgaga	3660

aatatcaact atcaaagaac agctattcac acgttactat tgagattatt attggacgag	3720
aatcacacac tcaactgtct ttctctcttc tagaaataca ggtacaagta tgtactattc	3780
tcattgttca tacttctagt catttcatcc cacatatcc ttggatttct ctccaatgaa	3840
tgacattcta tcttgcaaat tcaacaattia taataagata taccaaagta gcggtatagt	3900
ggcaatcaaa aagcttctct ggtgtgcttc tcgtatttat ttttattcta atgatccatt	3960
aaaggatat atttatttct tgttatataa tccttttggt tattacatgg gctggataca	4020
taaaggattt ttgatttaat tttttgctta aattcaatcc cccctcgctc agtgtcaact	4080
gtaatggtag gaaattacca tacttttgaa gaagcaaaaa aatgaaaga aaaaaaaaaat	4140
cgtatttcca ggtagacgt tccgcagaat ctagaatgcg gtagcggtta cattgttctt	4200
cgaacgtaaa agttgcgctc cctgagatat tgtacatttt tgcttttaca agtacaagta	4260
catcgtacaa ctatgtacta ctgttgatgc atccacaaca gtttgttttg ttttttttg	4320
ttttttttt ttctaagat tcattaccgc tatgtatacc tacttgtact ttagtaagc	4380
cgggttattg gcgttcaatt aatcatagac ttatgaatct gcacggtgtg cgctgcgagt	4440
tacttttagc ttatgcatgc tacttgggtg taatattggg atctgttcgg aatcaacgg	4500
atgctcaatc gatttcgaca gtaattaatt aagtcataca caagtcagct ttcttcgagc	4560
ctcatataag tataagtagt tcaacgtatt agcactgtac ccagcatctc cgtatcgaga	4620
aacacaacaa catgccccat tggacagatc atcgcgatc acaggttgtg cagtatcata	4680
catactcgat cacagaggtc gtctgacat catacaagct gaacaagcgc tccatacttg	4740
cacgtctct atatacacag ttaaattaca tatccatagt ctaacctcta acagttaatc	4800
ttctggttaag cctcccagcc agccttctgg tatcgcttgg cctcctcaat aggatctcgg	4860
ttctggccgt acagacctcg gccgacaatt atgatatccg ttccggtaga catgacatcc	4920
tcaacagttc ggtactgctg tccgagagcg tctcccttgt cgtcaagacc cacccegggg	4980
gtcagaataa gccagtcctc agagtcgcc ttaggtcggg tctgggcaat gaagccaacc	5040
acaaactcgg ggtcggatcg ggcaagctca atggtctgct tggagtactc gccagtggcc	5100
agagagccct tgcaagacag ctcgccagc atgagcagac ctctggccag ctctcgttg	5160
ggagagggga ctaggaactc cttgtactgg gatttctcgt agtcagagac gtctccttc	5220
ttctgttcag agacagtttc ctcgccacca gctcgagc cagcaatgat tccggttccg	5280
ggtacaccgt gggcgttggg gatatcggac cactcggcga ttcggtgaca ccggtactgg	5340
tgcttgacag tgttgccaat atctgcgaac tttctgtcct cgaacaggaa gaaaccgtgc	5400

ttaagagcaa gticcttgag ggggagcaca gtgccggcgt aggtgaagtc gtcaatgatg	5460
tcgatatggg ttttgatcat gcacacataa ggtccgacct tatcggaag ctcaatgagc	5520
tccttggtgg tggtaacatc cagagaagca cacaggttgg ttttcttggc tgccacgagc	5580
ttgagcactc gagcggcaaa ggcggacttg tggacgttag ctcgagcttc gtaggagggc	5640
attttggtgg tgaagaggag actgaaataa atttagtctg cagaactttt tatcggaacc	5700
ttatctgggg cagtgaagta tatgttatgg taatagttaac gagttagttag aacttataga	5760
tagactggac tatacggcta tcggtccaaa ttagaaagaa cgtcaatggc tctctgggcg	5820
tcgcctttgc cgacaaaaat gtgatcatga tgaaagccag caatgacgtt gcagctgata	5880
ttgttgtcgg ccaaccgcgc cgaaaacgca gctgtcagac ccacagcctc caacgaagaa	5940
tgtatcgtca aagtgatcca agcacactca tagttggagt cgtactcca aggcggcaat	6000
gacgagtcag acagatactc gtgcactcag gcgacgacgg aattcctgca gcccatctgc	6060
agaattcagg agagaccggg ttggcggcgt atttgtgtcc caaaaaacag cccaattgc	6120
cccggagaag acggccaggc cgcctagatg acaaatcaa caactcacag ctgactttct	6180
gccattgcca ctaggggggg gcctttttat atggccaagc caagctctcc acgtcggttg	6240
ggctgcaccc aacaataaat gggtaggggt gcaccaacaa agggatggga tgggggtag	6300
aagatacgag gataacgggg ctcaatggca caaataagaa cgaatactgc cattaagact	6360
cgtgatccag cgactgacac cattgcatca tctaagggcc tcaaaactac ctcggaactg	6420
ctgcgctgat ctggacacca cagaggttcc gagcacttta ggttgacca aatgtccac	6480
caggtgcagg cagaaaacgc tggaacagcg tgtacagttt gtcttaacaa aaagtgaggg	6540
cgctgaggtc gagcagggtg gtgtgacttg ttatagcctt tagagctgcg aaagcgcgta	6600
tggatttggc tcatcaggcc agattgaggg tctgtggaca catgtcatgt tagtgtactt	6660
caatcgcccc ctggatatag ccccgacaat aggcctgggc ctcatTTTTT tgccttccgc	6720
acatttccat tgctcggtac ccacaccttg ctctctctgc acttgccaac cttaatactg	6780
gtttacattg accaaccatc tacaagcggg gggttgtct agggatatata taaacagtgg	6840
ctctcccaat cggttgccag tctctTTTTT cctttcttcc cccacagatt cgaaatctaa	6900
actacacatc acacaatgcc tgttactgac gtccttaagc gaaagtcggg tgtcatcgtc	6960
ggcgacgatg tccgagccgt gagtatccac gacaagatca gtgtcgagac gacgcgtttt	7020
gtgtaatgac acaatccgaa agtcgctagc aacacacact ctctacacaa actaaccag	7080
ctctccatgg cctgtctct cactaccgaa cagctcctgg agcgacctga tctcgtcgt	7140
atcgatggta tctgtacga cctcgagggc ctggccaagg tgcatggtgg tggagacctc	7200
attctggcct ctggagcctc cgacgcctct ccctcttct actctatgca tcctacgtc	7260

aagcccgaga aticcaagct cctgcagcaa ttcgtccgag ggaagcatga cgcacctcg 7320
aaggacattg tctacacgta tgattctccc ttcgcacaag acgttaagcg gacaatgcgc 7380
gaggtgatga aagggaggaa ctggtagca acccctggct tctggctgcg caccgttggg 7440
atcatcgccg tgacggcctt ttgcgagtgg cactgggcta ccacggggat ggtgctgtgg 7500

ggcctgttga ctggattcat gcacatgcag atcggcttat ccatccagca tgatgcgggt 7560
cacggggcca tcagcaagaa gccttgggtc aacgccctct tcgcctacgg cattgacgtc 7620
atcggatcgt cccggtggat ttggctgcag tcgcacatca tgcggcacca cacctacacc 7680
aaccagcacg gcctcgacct ggatgcggag tcggcagagc cgttctctgt gtccacaac 7740
taccgcgccc caaacaccgc ccgaaagtgg ttccaccgct tccaggcttg gtacatgtac 7800
cttgtgctgg gggcatacgg ggtatcgctg gtgtacaacc cgctctacat tttccggatg 7860
cagcacaatg acaccatccc agagtctgtc acggccatgc gggaaaatgg ctttctgcgg 7920

cgctaccgca cacttgcatt cgtgatgcga gctttcttca tcttccggac cgcattcttg 7980
ccctggtacc tcaactgggac ctcatgtctg atcaccattc ctctggtgcc caccgcaact 8040
ggtgccttct tgacgttctt cttcatcttg tcccacaatt ttgatggctc cgaacggatc 8100
cccgacaaga atgcaaggt taagagctct gagaaggacg ttgaggctga ccaaattgac 8160
tggtatcggg cgcagggtga gacgtcctcc acatacgggtg gccccatcgc catgttcttc 8220
actggcggtc tcaatttcca gatcgagcac cacctcttcc cccggatgtc gtcttggcac 8280
tacccttctg tccagcaggc ggtccgggag tgttcgaac gacatggagt gcgatatgtt 8340

ttctacccta ccatcgtcgg caacatcatc tccacctga agtacatgca taaggtaggt 8400
gtcgtccact gcgtgaagga cgcacaggat tcctaagc 8438

<210> 308

<211> 449

<212> PRT

<213> *Euglena gracilis*

<400> 308

Met Ala Leu Ser Leu Thr Thr Glu Gln Leu Leu Glu Arg Pro Asp Leu

1 5 10 15

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

20 25 30

His Pro Gly Ser Asp Leu Ile Leu Ala Ser Gly Ala Ser Asp Ala Ser

Pro Leu Phe Tyr Ser Met His Pro Tyr Val Lys Pro Glu Asn Ser Lys
50 55 60
Leu Leu Gln Gln Phe Val Arg Gly Lys His Asp Arg Thr Ser Lys Asp
65 70 75 80
Ile Val Tyr Thr Tyr Asp Ser Pro Phe Ala Gln Asp Val Lys Arg Thr
85 90 95
Met Arg Glu Val Met Lys Gly Arg Asn Trp Tyr Ala Thr Pro Gly Phe
100 105 110
Trp Leu Arg Thr Val Gly Ile Ile Ala Val Thr Ala Phe Cys Glu Trp
115 120 125
His Trp Ala Thr Thr Gly Met Val Leu Trp Gly Leu Leu Thr Gly Phe
130 135 140
Met His Met Gln Ile Gly Leu Ser Ile Gln His Asp Ala Ser His Gly
145 150 155 160
Ala Ile Ser Lys Lys Pro Trp Val Asn Ala Leu Phe Ala Tyr Gly Ile
165 170 175
Asp Val Ile Gly Ser Ser Arg Trp Ile Trp Leu Gln Ser His Ile Met
180 185 190
Arg His His Thr Tyr Thr Asn Gln His Gly Leu Asp Leu Asp Ala Glu
195 200 205
Ser Ala Glu Pro Phe Leu Val Phe His Asn Tyr Pro Ala Ala Asn Thr
210 215 220
Ala Arg Lys Trp Phe His Arg Phe Gln Ala Trp Tyr Met Tyr Leu Val
225 230 235 240
Leu Gly Ala Tyr Gly Val Ser Leu Val Tyr Asn Pro Leu Tyr Ile Phe
245 250 255
Arg Met Gln His Asn Asp Thr Ile Pro Glu Ser Val Thr Ala Met Arg
260 265 270
Glu Asn Gly Phe Leu Arg Arg Tyr Arg Thr Leu Ala Phe Val Met Arg
275 280 285
Ala Phe Phe Ile Phe Arg Thr Ala Phe Leu Pro Trp Tyr Leu Thr Gly

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

355 360 365
 Thr Tyr Gly Gly Pro Ile Ala Met Phe Phe Thr Gly Gly Leu Asn Phe
 370 375 380
 Gln Ile Glu His His Leu Phe Pro Arg Met Ser Ser Trp His Tyr Pro
 385 390 395 400
 Phe Val Gln Gln Ala Val Arg Glu Cys Cys Glu Arg His Gly Val Arg
 405 410 415
 Tyr Val Phe Tyr Pro Thr Ile Val Gly Asn Ile Ile Ser Thr Leu Lys

420 425 430
 Tyr Met His Lys Val Gly Val Val His Cys Val Lys Asp Ala Gln Asp
 435 440 445
 Ser

<210> 309

<211> 8438

<212> DNA

<213> Artificial Sequence

<220><223> Plasmid pDMW369S

<400> 309

ggccgcaagt gtggatgggg aagtgagtgc ccggttctgt gtgcacaatt ggcaatccaa 60
 gatggatgga ttcaacacag ggatatagcg agctacgtgg tgggtcgagg atatagcaac 120
 ggatatattat gtttgacact tgagaatgta cgatacaagc actgtccaag tacaatacta 180

aacatactgt acatactcat actcgtaccc gggcaacggt ttactttgag tgcagtggct 240
 agtgctctta ctctacagt gtgcaatact gcgtatcata gtccttgatg tatatcgat 300

tcattcatgt tagttgcgta cgagccggaa gcataaagt taaagcctgg ggtgccta	360
gagtgagcta actcacatta attgcgttgc gctcactgcc cgctttccag tcgggaaacc	420
tgtcgtgcc gctgcattaa tgaatcgcc aacgcgcggg gagaggcggg ttgcgtattg	480
ggcgctcttc cgcttccctg ctactgact cgctgcgctc ggtcgttcgg ctgcggcgag	540
cggtatcagc tactcaaag gcgtaatac ggttatccac agaalcaggg gataacgcag	600
gaaagaacat gtgagcaaaa ggccagcaaa aggccaggaa ccgtaaaaag gccgcgttgc	660
tggcgttttt ccataggctc cgccccctg acgagcatca caaaaatcga cgctcaagtc	720
agaggtagcg aaacccgaca ggactataaa gataccaggc gtttccccct ggaagctccc	780
tcgtgcgctc tctgttccg accctgccgc ttaccggata cctgtccgcc tttctccctt	840
cgggaagcgt ggcgctttct catagctcac gctgtaggta tctcagttcg gtgtaggtcg	900
ttcgctccaa gctgggctgt gtgcacgaac ccccggttca gcccgaccgc tgcgccttat	960
ccgtaacta tcgtcttgag tccaaccggt taagacacga cttatcgcca ctggcagcag	1020
ccactggtaa caggattagc agagcgaggt atgtaggcgg tgctacagag ttcttgaagt	1080
ggtggcctaa ctacggctac actagaagga cagtatttgg tatctgcgt ctgctgaagc	1140
cagttacctt cggaaaaaga gttggtagct cttgatccgg caaacaacc accgctggt	1200
gcggtggttt tttgtttgc aagcagcaga ttacgcgcag aaaaaagga tctcaagaag	1260
atcctttgat cttttctacg gggctgacg ctacgtggaa cgaaaactca cgttaaggga	1320
ttttggtcat gagattatca aaaaggatct tcacctagat ccttttaaat taaaaatgaa	1380
gttttaaatc aatctaaagt atatatgagt aaacttggtc tgacagttac caatgcttaa	1440
tcagtgaggc acctatctca gcgatctgtc tatttcgttc atccatagtt gcctgactcc	1500
ccgtcgtgta gataactacg atacgggagg gcttaccatc tggccccagt gctgcaatga	1560
taccgcgaga cccacgtca ccggctccag atttatcagc aataaaccag ccagccggaa	1620
gggcccagcg cagaagtggg cctgcaactt tatccgctc catccagtct attaattgtt	1680
gccgggaagc tagagtaagt agttcgccag ttaatatgtt gcgcaacgtt gttgccattg	1740
ctacaggcat cgtgggtgca cgctcgtcgt ttggtatggc ttcattcagc tccggttccc	1800
aacgatcaag gcgagttaca tgatcccca tgttgtgcaa aaaagcggtt agtccttcg	1860
gtcctccgat cgttgtcaga agtaagtgg ccgcagtggt atcactcatg gttatggcag	1920
cactgcataa ttctcttact gtcatccat ccgtaagatg cttttctgtg actggtagt	1980
actcaacaa gtcattctga gaatagtgtg tgcggcgacc gagttgctct tgccccgct	2040
caatacggga taataccggt ccacatagca gaactttaaa agtgetcatc attggaaaac	2100
gttcttcggg gcgaaaactc tcaaggatct taccgtgtt gagatccagt tcgatgtaac	2160

ccactcgtgc acccaactga tcttcagcat cttttacttt caccagcggt tctgggtgag 2220
caaaaacagg aaggcaaaat gccgcaaaaa agggaataag ggcgacacgg aaatgttgaa 2280

tactcatact cttccttttt caatattatt gaagcattta tcagggttat tgtctcatga 2340
gcggatacat atttgaatgt atttagaaaa ataaacaaat aggggttccg cgcacatttc 2400
cccgaaaagt gccacctgac gcgccctgta gcggcgcat aagcgcggcg ggtgtggtgg 2460
ttacgcgcag cgtgaccgct acacttgcca gcgccctage gcccgctcct ttcgctttct 2520
tcccttcttt tctcgccacg ttcgccggct tccccgtca agctctaaat cgggggctcc 2580
ctttagggtt ccgatttagt gctttacggc acctcgaccc caaaaaactt gattagggtg 2640
atggttcacg tagtgggcca tcgccctgat agacgggttt tcgccctttg acgttggagt 2700

ccacgttctt taatagtga ctcttggtcc aaactggaac aacactcaac cctatctcgg 2760
tctattcttt tgatttataa gggattttgc cgatttcggc ctattggtta aaaaatgagc 2820
tgatttaaca aaaatttaac gcgaatttta acaaaatatt aacgcttaca atttcattc 2880
gccattcagg ctgcgcaact gttgggaagg gcgacgggtg cgggcctctt cgctattacg 2940
ccagctggcg aaagggggat gtgctgcaag gcgattaagt tgggtaacgc cagggttttc 3000
ccagtcacga cgttgtaaaa cgacggccag tgaattgtaa tacgactcac tataggcgca 3060
attgggtacc gggccccccc tcgaggtcga tgggtgctgat aagcttgata tcgaattcat 3120

gtcacacaaa ccgatcttcg cctcaaggaa acctaatctt acatccgaga gactgccgag 3180
atccagtcta cactgattaa ttttcgggcc aataatttaa aaaaatcgtg ttatataata 3240
ttatatgtat tatatatata catcatgatg atactgacag tcatgtccca ttgctaaata 3300
gacagactcc atctgccgcc tccaactgat gtcttcaata ttttaagggt catctcgcat 3360
tgtttaataa taaacagact ccatctaccg cctccaaatg atgttctcaa aatatattgt 3420
atgaacttat ttttattact tagtattatt agacaactta cttgctttat gaaaaacact 3480
tcctatttag gaacaattt ataatggcag ttcgttcatt taacaattta tgtagaataa 3540

atgttataaa tgcgtatggg aaatcttaaa tatggatagc ataaatgata tctgcattgc 3600
ctaattcgaa atcaacagca acgaaaaaaaa tcccttgtag aacataaata gtcacgaga 3660
aatatcaact atcaaagaac agctattcac acgttactat tgagattatt attggacgag 3720
aatcacacac tcaactgtct ttctctcttc tagaaatata ggtacaagta tgtactattc 3780
tcattgttca tacttctagt catttcatcc cacatatccc ttggatttct ctccaatgaa 3840
tgacattcta tcttgcaaat tcaacaatta taataagata taccaaagta gcggtatagt 3900

ggcaatcaaa aagcttctct ggtgtgcttc tcgtatttat ttttattcta atgatccatt	3960
aaaggtatat atttatttct tgttatataa tccttttgtt tattacatgg gctggataca	4020
taaaggtatt ttgatttaat tttttgctta aattcaatcc cccctcgctc agtgtcaact	4080
glaatggtag gaaattacca tacttttgaa gaagcaaaaa aaatgaaaga aaaaaaaaaat	4140
cgtattttcca ggtagacgt tccgcagaat ctagaatgcg gtatgcggta cattgttctt	4200
cgaacgtaaa agttgcgctc cctgagatat tgtacatttt tgcttttaca agtacaagta	4260
catcgtaaca ctatgtacta ctgttgatgc atccacaaca gtttgttttg ttttttttg	4320
tttttttttt ttctaattgat tcattaccgc tatgtatacc tacttgtact ttagtaagc	4380
cgggttattg gcgttcaatt aatcatagac ttatgaatct gcacgggttg cgctgcgagt	4440
tacttttagc ttatgcatgc tacttgggtg taatatggg atctgttcgg aaatcaacgg	4500
atgtcaatc gatttcgaca gtaattaatt aagtcataca caagtcagct ttcttcgagc	4560
ctcatataag tataagtagt tcaacgtatt agcactgtac ccagcatctc cgtatcgaga	4620
aacacaacaa catgccccat tggacagatc atgcggatac acaggttgtg cagtatcata	4680
catactgat cagacaggct gtctgacat catacaagct gaacaagcg tccatacttg	4740
cacgtctct atatacacag ttaaattaca tatccatagt ctaacctcta acagttaatc	4800
ttctggttaag cctcccagcc agccttctgg tategttgg cctcctcaat aggatctcg	4860
ttctggccgt acagacctcg gccgacaatt atgatatccg ttccggtaga catgacatcc	4920
tcaacagttc ggtactgctg tccgagagcg tctcccttgt cgtcaagacc cccccgggg	4980
gtcagaataa gccagtcctc agagtcgcc ttaggtcggg tctgggcaat gaagccaacc	5040
acaaactcgg ggtcggatcg ggcaagctca atggtctgct tggagtactc gccagtggcc	5100
agagagccct tgcaagacag ctcggccagc atgagcagac ctctggccag ctctctgttg	5160
ggagagggga ctaggaaactc cttgtactgg gatttctcgt agtcagagac gtctccttc	5220
ttctgttcag agacagtttc ctccgcacca gctcgcaggc cagcaatgat tccggttccg	5280
ggtacaccgt gggcgttggg gatatcggac cactcggcga ttccgtgaca ccggtactgg	5340
tgcttgacag tgttgccaat atctgcgaac tttctgtcct cgaacaggaa gaaaccgtgc	5400
ttaagagcaa gttccttgag ggggagcaca gtgccggcgt aggtgaagtc gtcaatgatg	5460
tcgatatggg ttttgatcat gcacacataa ggtccgacct tatcggcaag ctcaatgagc	5520
tccttgggtg tggtaacatc cagagaagca cacaggttgg ttttcttggc tgccacgagc	5580
ttgagcactc gagcggcaaa ggcggacttg tggacgttag ctcgagcttc gtaggagggc	5640

attttgggtgg tgaagaggag actgaaataa atttagtctg cagaactttt tatcggaacc 5700
 ttatctgggg cagtgaagta tatgttatgg taatagttac gagttagttag aacttataga 5760
 tagactggac tatacggcta tcggtccaaa ttagaaagaa cgtcaatggc tctctgggcg 5820
 tcgcctttgc cgacaaaaat gtgatcatga tgaagccag caatgacgtt gcagctgata 5880
 ttgttgcgg ccaaccgcgc gaaaaacgca gctgtcagac ccacagcctc caacgaagaa 5940
 tgtatcgta aagtgatcca agcacactca tagttggagt cgtactcaa aggcggcaat 6000
 gacgagtcag acagatactc gtcgactcag gcgacgacgg aattcctgca gcccatctgc 6060

 agaattcagg agagaccggg ttggcggcgt atttgtgtcc caaaaaacag cccaattgc 6120
 cccggagaag acggccaggc cgcctagatg acaaatcaa caactcacag ctgactttct 6180
 gccattgcca ctaggggggg gcctttttat atggccaagc caagctctcc acgtcggttg 6240
 ggctgacccc aacaataaat gggtagggtt gcaccaacaa agggatggga tgggggtag 6300
 aagatacgag gataacgggg ctcaatggca caaataagaa cgaatactgc cattaagact 6360
 cgtgatccag cgactgacac cattgcatca tctaagggcc tcaaaactac ctcggaactg 6420
 ctgcgtgat ctggacacca cagaggttcc gagcacttta ggttgacca aatgtccac 6480

 caggtgcagg cagaaaacgc tggaacagcg tgtacagttt gtcttaacaa aaagtgaggg 6540
 cgctgaggtc gagcagggtg gtgtgacttg ttatagcctt tagagctgag aaagcgcgta 6600
 tggatttggc tcatcaggcc agattgaggg tctgtggaca catgcatgt tagtgtactt 6660
 caatgccccc ctggatatag ccccgacaat aggccgtggc ctcatTTTT tgccttccgc 6720
 acatttccat tgctcggtac ccacaccttg cttctcctgc acttgccaac cttaatactg 6780
 gtttacattg accaacatct tacaagcggg gggcttgtct agggatatata taaacagtgg 6840
 ctctcccaat cggttgccag tctctTTTT ctttctttc cccacagatt cgaaatctaa 6900

 actacacatc acacaatgcc tgttactgac gtcttaagc gaaagtccgg tgtcatcgtc 6960
 ggcgacgatg tccgagccgt gagtatccac gacaagatca gtgtcgagac gacgcgtttt 7020
 gtgtaatgac acaatccgaa agtcgctagc aacacacact ctctacaaa actaaccag 7080
 ctctccatgg ctctctccct tactaccgag cagctgctcg agcgaccga cctggttgcc 7140
 atcgacggca ttctctacga tctggaaggt cttgccaagg tccatcccgg atccgacttg 7200
 atcctcgctt ctggtgcctc cgatgcttct cctctgttct actccatgca cccttacgtc 7260
 aagcccgaga actcgaagct gcttcaacag ttcgtgcgag gcaagcacga ccgaacctcc 7320

 aaggacattg tctacaccta cgactctccc ttgacacagg acgtcaagcg aactatgcga 7380
 gaggtcatga aaggtcggaa ctggtatgcc acacctggat tctggctgag aaccgttggc 7440
 atcattgctg tcaccgcctt ttgcgagtgg cactgggcta ctaccggaat ggtgctgtgg 7500

ggctctcttga ctggattcat gcacatgcag atcggcctgt ccattcagca cgaatgcctct 7560
 catggtgccca tcagcaaaaa gccctgggtc aacgtctctt ttgcctacgg catcgacgtc 7620
 attggatcgt ccagatggat ctggctgcag tctcacatca tgcgacatca cacctacacc 7680
 aatcagcatg gtctcgacct ggatgccgag tccgcagaac cattccttgt gtccacaac 7740

tacctgtctg ccaacactgc tcgaaagtgg ttaccgat tccaggcctg gtacatgtac 7800
 ctctgtcttg gagcctacgg cgtttcgtg gtgtacaacc ctctctacat ctccgaatg 7860
 cagcacaacg acaccattcc cgagtctgtc acagccatgc gagagaacgg ctttctgcga 7920
 cggtagcgaa ccttgcatt cgttatgcga gctttcttca tctttcgaac cgccttcttg 7980
 ccctggatc tcactggaac ctccctgtc atcaccattc ctctgggtgc cactgtacc 8040
 ggtgccttcc tcaccttctt ttcatcttg tctcacaact tcgatggctc ggagcgaatc 8100
 cccgacaaga actgcaaggt caagagctcc gagaaggacg ttgaagccga tcagatcgac 8160

tggtagacag cttaggtgga gacctcttcc acctacggtg gaccattgc catgttcttt 8220
 actggcggtc tcaacttcca gatcgagcat cacctcttcc ctgaatgtc gtcttggcac 8280
 tatcccttcg tgcagcaagc tgtagcagag tgttgcgaac gacacggagt tcggtacgtc 8340
 ttctacccta ccattgtggg caacatcatt tccacctca agtacatgca caaagtcggt 8400
 gtggttcaat gtgtcaagga cgctcaggat tcctaagc 8438

<210> 310

<211> 39

<212> DNA

<213> Artificial Sequence

<220><223> Primer

<400> 310

cctgtccatt cagcagcatt tctctcatgg tgccatcag 39

<210> 311

<211> 39

<212> DNA

<213> Artificial Sequence

<220><223> Primer

<400> 311

ctgatggcac catgagagaa atcgtgctga atggacagg 39

<210> 312

<211> 39

<212> DNA

<213> Artificial Sequence

<220><223> Primer

<400> 312

cctgtccatt cagcacgata tgtctcatgg tgccatcag 39

<210> 313

<211> 39

<212> DNA

<213> Artificial Sequence

<220><223> Primer

<400> 313

ctgatggcac catgagacat atcgtgctga atggacagg 39

<210> 314

<211> 36

<212> DNA

<213> Artificial Sequence

<220><223> Primer

<400> 314

gtccattcag cacgatgggt ctcattggtgc catcag 36

<210> 315

<211> 36

<212> DNA

<213> Artificial Sequence

<220><223> Primer

<400> 315

ctgatggcac catgagaacc atcgtgctga atggac 36

<210> 316

<211> 31

<212> DNA

<213> Artificial Sequence

<220><223> Primer

<400> 316

gtccattcag cagattcct ctcattggtgc c 31

<210> 317
 <211> 31
 <212> DNA
 <213> Artificial Sequence
 <220><223> Primer
 <400> 317
 ggcacccatga gaggaatcgt gctgaatgga c 31
 <210> 318
 <211> 33
 <212> DNA
 <213> Artificial Sequence
 <220><223> Primer
 <400> 318
 cattcagcac gatgccgctc atggtgccat cag 33
 <210> 319
 <211> 33
 <212> DNA
 <213> Artificial Sequence
 <220><223> Primer
 <400> 319
 ctgatggcac catgagcggc atcgtgctga atg 33
 <210> 320
 <211> 45
 <212> DNA
 <213> Artificial Sequence
 <220><223> Primer
 <400> 320
 ctgtccattc agcacgatgc caaccatggt gccatcagca aaaag 45
 <210> 321
 <211> 45
 <212> DNA
 <213> Artificial Sequence
 <220><223> Primer
 <400> 321

ctttttgctg atggcaccat ggttgcatc gtgctgaatg gacag 45

<210> 322

<211> 35

<212> DNA

<213> Artificial Sequence

<220><223> Primer

<400> 322

cattcagcac gatgccaccc atggtgccat cagca 35

<210> 323

<211> 35

<212> DNA

<213> Artificial Sequence

<220><223> Primer

<400> 323

tgctgatggc accatgggtg gcatcgtgct gaatg 35

<210> 324

<211> 34

<212> DNA

<213> Artificial Sequence

<220><223> Primer

<400> 324

cattcagcac gatgccggtc atggtgccat cagc 34

<210> 325

<211> 34

<212> DNA

<213> Artificial Sequence

<220><223> Primer

<400> 325

gctgatggca ccatgaccgg catcgtgctg aatg 34

<210> 326

<211> 31

<212> DNA

<213> Artificial Sequence

<220><223> Primer
 <400> 326
 ctgtccattc agcacgaggc ctctcatggt g 31
 <210> 327
 <211> 31
 <212> DNA
 <213> Artificial Sequence
 <220><223> Primer
 <400> 327
 caccatgaga ggcctcgtgc tgaatggaca g 31
 <210> 328
 <211> 1350
 <212> DNA
 <213> *Euglena gracilis*
 <220><221> CDS
 <222> (1)..(1350)
 <400> 328
 atg gct ctc tcc ctt act acc gag cag ctg ctc gag cga ccc gac ctg 48

 Met Ala Leu Ser Leu Thr Thr Glu Gln Leu Leu Glu Arg Pro Asp Leu
 1 5 10 15
 gtt gcc atc gac ggc att ctc tac gat ctg gaa ggt ctt gcc aag gtc 96
 Val Ala Ile Asp Gly Ile Leu Tyr Asp Leu Glu Gly Leu Ala Lys Val
 20 25 30
 cat ccc gga tcc gac ttg atc ctc gct tct ggt gcc tcc gat gct tct 144
 His Pro Gly Ser Asp Leu Ile Leu Ala Ser Gly Ala Ser Asp Ala Ser

 35 40 45
 cct ctg ttc tac tcc atg cac cct tac gtc aag ccc gag aac tcg aag 192
 Pro Leu Phe Tyr Ser Met His Pro Tyr Val Lys Pro Glu Asn Ser Lys
 50 55 60
 ctg ctt caa cag ttc gtg cga ggc aag cac gac cga acc tcc aag gac 240
 Leu Leu Gln Gln Phe Val Arg Gly Lys His Asp Arg Thr Ser Lys Asp
 65 70 75 80

att gtc tac acc tac gac tct ccc ttt gca cag gac gtc aag cga act 288
Ile Val Tyr Thr Tyr Asp Ser Pro Phe Ala Gln Asp Val Lys Arg Thr
85 90 95

atg cga gag gtc atg aaa ggt cgg aac tgg tat gcc aca cct gga ttc 336
Met Arg Glu Val Met Lys Gly Arg Asn Trp Tyr Ala Thr Pro Gly Phe
100 105 110

tgg ctg cga acc gtt ggc atc att gct gtc acc gcc ttt tgc gag tgg 384

Trp Leu Arg Thr Val Gly Ile Ile Ala Val Thr Ala Phe Cys Glu Trp
115 120 125

cac tgg gct act acc gga atg gtg ctg tgg ggt ctc ttg act gga ttc 432
His Trp Ala Thr Thr Gly Met Val Leu Trp Gly Leu Leu Thr Gly Phe
130 135 140

atg cac atg cag atc ggc ctg tcc att cag cac gag gcc tct cat ggt 480
Met His Met Gln Ile Gly Leu Ser Ile Gln His Glu Ala Ser His Gly

145 150 155 160

gcc atc agc aaa aag ccc tgg gtc aac gct ctc ttt gcc tac ggc atc 528
Ala Ile Ser Lys Lys Pro Trp Val Asn Ala Leu Phe Ala Tyr Gly Ile
165 170 175

gac gtc att gga tcg tcc aga tgg atc tgg ctg cag tct cac atc atg 576
Asp Val Ile Gly Ser Ser Arg Trp Ile Trp Leu Gln Ser His Ile Met
180 185 190

cga cat cac acc tac acc aat cag cat ggt ctc gac ctg gat gcc gag 624
Arg His His Thr Tyr Thr Asn Gln His Gly Leu Asp Leu Asp Ala Glu
195 200 205

tcc gca gaa cca ttc ctt gtg ttc cac aac tac cct gct gcc aac act 672
Ser Ala Glu Pro Phe Leu Val Phe His Asn Tyr Pro Ala Ala Asn Thr
210 215 220

gct cga aag tgg ttt cac cga ttc cag gcc tgg tac atg tac ctc gtg 720

Ala Arg Lys Trp Phe His Arg Phe Gln Ala Trp Tyr Met Tyr Leu Val
225 230 235 240

ctt gga gcc tac ggc gtt tcg ctg gtg tac aac cct ctc tac atc ttc 768

Leu Gly Ala Tyr Gly Val Ser Leu Val Tyr Asn Pro Leu Tyr Ile Phe
 245 250 255
 cga atg cag cac aac gac acc att ccc gag tct gtc aca gcc atg cga 816
 Arg Met Gln His Asn Asp Thr Ile Pro Glu Ser Val Thr Ala Met Arg
 260 265 270
 gag aac ggc ttt ctg cga cgg tac cga acc ctt gca ttc gtt atg cga 864
 Glu Asn Gly Phe Leu Arg Arg Tyr Arg Thr Leu Ala Phe Val Met Arg
 275 280 285
 gct ttc ttc atc ttt cga acc gcc ttc ttg ccc tgg tat ctc act gga 912
 Ala Phe Phe Ile Phe Arg Thr Ala Phe Leu Pro Trp Tyr Leu Thr Gly
 290 295 300
 acc tcc ctg ctc atc acc att cct ctg gtg ccc act gct acc ggt gcc 960
 Thr Ser Leu Leu Ile Thr Ile Pro Leu Val Pro Thr Ala Thr Gly Ala
 305 310 315 320
 ttc ctc acc ttc ttt ttc atc ttg tct cac aac ttc gat ggc tcg gag 1008
 Phe Leu Thr Phe Phe Phe Ile Leu Ser His Asn Phe Asp Gly Ser Glu
 325 330 335
 cga atc ccc gac aag aac tgc aag gtc aag agc tcc gag aag gac gtt 1056
 Arg Ile Pro Asp Lys Asn Cys Lys Val Lys Ser Ser Glu Lys Asp Val
 340 345 350
 gaa gcc gat cag atc gac tgg tac aga gct cag gtg gag acc tct tcc 1104
 Glu Ala Asp Gln Ile Asp Trp Tyr Arg Ala Gln Val Glu Thr Ser Ser
 355 360 365
 acc tac ggt gga ccc att gcc atg ttc ttt act ggc ggt ctc aac ttc 1152
 Thr Tyr Gly Gly Pro Ile Ala Met Phe Phe Thr Gly Gly Leu Asn Phe
 370 375 380
 cag atc gag cat cac ctc ttt cct cga atg tcg tct tgg cac tat ccc 1200
 Gln Ile Glu His His Leu Phe Pro Arg Met Ser Ser Trp His Tyr Pro
 385 390 395 400
 ttc gtg cag caa gct gtc cga gag tgt tgc gaa cga cac gga gtt cgg 1248
 Phe Val Gln Gln Ala Val Arg Glu Cys Cys Glu Arg His Gly Val Arg

405	410	415	
tac gtc ttc tac cct acc att gtg ggc aac atc att tcc acc ctc aag			1296
Tyr Val Phe Tyr Pro Thr Ile Val Gly Asn Ile Ile Ser Thr Leu Lys			
420	425	430	
tac atg cac aaa gtc ggt gtg gtt cac tgt gtc aag gac gct cag gat			1344
Tyr Met His Lys Val Gly Val Val His Cys Val Lys Asp Ala Gln Asp			
435	440	445	
tcc taa			1350

Ser

<210> 329

<211> 449

<212> PRT

<213> Euglena gracilis

<400> 329

Met Ala Leu Ser Leu Thr Thr Glu Gln Leu Leu Glu Arg Pro Asp Leu

1 5 10 15

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

20 25 30

His Pro Gly Ser Asp Leu Ile Leu Ala Ser Gly Ala Ser Asp Ala Ser

35 40 45

Pro Leu Phe Tyr Ser Met His Pro Tyr Val Lys Pro Glu Asn Ser Lys

50 55 60

Leu Leu Gln Gln Phe Val Arg Gly Lys His Asp Arg Thr Ser Lys Asp

65 70 75 80

Ile Val Tyr Thr Tyr Asp Ser Pro Phe Ala Gln Asp Val Lys Arg Thr

85 90 95

Met Arg Glu Val Met Lys Gly Arg Asn Trp Tyr Ala Thr Pro Gly Phe

100 105 110

Trp Leu Arg Thr Val Gly Ile Ile Ala Val Thr Ala Phe Cys Glu Trp

115 120 125

His Trp Ala Thr Thr Gly Met Val Leu Trp Gly Leu Leu Thr Gly Phe
130 135 140

Met His Met Gln Ile Gly Leu Ser Ile Gln His Glu Ala Ser His Gly
145 150 155 160

Ala Ile Ser Lys Lys Pro Trp Val Asn Ala Leu Phe Ala Tyr Gly Ile

165 170 175

Asp Val Ile Gly Ser Ser Arg Trp Ile Trp Leu Gln Ser His Ile Met
180 185 190

Arg His His Thr Tyr Thr Asn Gln His Gly Leu Asp Leu Asp Ala Glu
195 200 205

Ser Ala Glu Pro Phe Leu Val Phe His Asn Tyr Pro Ala Ala Asn Thr
210 215 220

Ala Arg Lys Trp Phe His Arg Phe Gln Ala Trp Tyr Met Tyr Leu Val

225 230 235 240

Leu Gly Ala Tyr Gly Val Ser Leu Val Tyr Asn Pro Leu Tyr Ile Phe
245 250 255

Arg Met Gln His Asn Asp Thr Ile Pro Glu Ser Val Thr Ala Met Arg
260 265 270

Glu Asn Gly Phe Leu Arg Arg Tyr Arg Thr Leu Ala Phe Val Met Arg
275 280 285

Ala Phe Phe Ile Phe Arg Thr Ala Phe Leu Pro Trp Tyr Leu Thr Gly

290 295 300

Thr Ser Leu Leu Ile Thr Ile Pro Leu Val Pro Thr Ala Thr Gly Ala
305 310 315 320

Phe Leu Thr Phe Phe Phe Ile Leu Ser His Asn Phe Asp Gly Ser Glu
325 330 335

Arg Ile Pro Asp Lys Asn Cys Lys Val Lys Ser Ser Glu Lys Asp Val
340 345 350

Glu Ala Asp Gln Ile Asp Trp Tyr Arg Ala Gln Val Glu Thr Ser Ser

355 360 365

Thr Tyr Gly Gly Pro Ile Ala Met Phe Phe Thr Gly Gly Leu Asn Phe

370 375 380
 Gln Ile Glu His His Leu Phe Pro Arg Met Ser Ser Trp His Tyr Pro
 385 390 395 400
 Phe Val Gln Gln Ala Val Arg Glu Cys Cys Glu Arg His Gly Val Arg
 405 410 415
 Tyr Val Phe Tyr Pro Thr Ile Val Gly Asn Ile Ile Ser Thr Leu Lys

420 425 430
 Tyr Met His Lys Val Gly Val Val His Cys Val Lys Asp Ala Gln Asp
 435 440 445
 Ser

<210> 330

<211> 1350

<212> DNA

<213> Euglena gracilis

<220><221> CDS

<222> (1)..(1350)

<400> 330

atg gct ctc tcc ctt act acc gag cag ctg ctc gag cga ccc gac ctg 48
 Met Ala Leu Ser Leu Thr Thr Glu Gln Leu Leu Glu Arg Pro Asp Leu
 1 5 10 15

gtt gcc atc gac ggc att ctc tac gat ctg gaa ggt ctt gcc aag gtc 96
 Val Ala Ile Asp Gly Ile Leu Tyr Asp Leu Glu Gly Leu Ala Lys Val
 20 25 30

cat ccc gga tcc gac ttg atc ctc gct tct ggt gcc tcc gat gct tct 144
 His Pro Gly Ser Asp Leu Ile Leu Ala Ser Gly Ala Ser Asp Ala Ser
 35 40 45

cct ctg ttc tac tcc atg cac cct tac gtc aag ccc gag aac tcg aag 192

Pro Leu Phe Tyr Ser Met His Pro Tyr Val Lys Pro Glu Asn Ser Lys
 50 55 60
 ctg ctt caa cag ttc gtg cga ggc aag cac gac cga acc tcc aag gac 240
 Leu Leu Gln Gln Phe Val Arg Gly Lys His Asp Arg Thr Ser Lys Asp

65	70	75	80	
att gtc tac acc tac gac tct ccc ttt gca cag gac gtc aag cga act				288
Ile Val Tyr Thr Tyr Asp Ser Pro Phe Ala Gln Asp Val Lys Arg Thr				
	85	90	95	
atg cga gag gtc atg aaa ggt cgg aac tgg tat gcc aca cct gga ttc				336
Met Arg Glu Val Met Lys Gly Arg Asn Trp Tyr Ala Thr Pro Gly Phe				
	100	105	110	
tgg ctg cga acc gtt ggc atc att gct gtc acc gcc ttt tgc gag tgg				384
Trp Leu Arg Thr Val Gly Ile Ile Ala Val Thr Ala Phe Cys Glu Trp				
	115	120	125	
cac tgg gct act acc gga atg gtg ctg tgg ggt ctc ttg act gga ttc				432
His Trp Ala Thr Thr Gly Met Val Leu Trp Gly Leu Leu Thr Gly Phe				
	130	135	140	
atg cac atg cag atc ggc ctg tcc att cag cac gat gcc gct cat ggt				480
Met His Met Gln Ile Gly Leu Ser Ile Gln His Asp Ala Ala His Gly				
	145	150	155	160
gcc atc agc aaa aag ccc tgg gtc aac gct ctc ttt gcc tac ggc atc				528
Ala Ile Ser Lys Lys Pro Trp Val Asn Ala Leu Phe Ala Tyr Gly Ile				
	165	170	175	
gac gtc att gga tcg tcc aga tgg atc tgg ctg cag tct cac atc atg				576
Asp Val Ile Gly Ser Ser Arg Trp Ile Trp Leu Gln Ser His Ile Met				
	180	185	190	
cga cat cac acc tac acc aat cag cat ggt ctc gac ctg gat gcc gag				624
Arg His His Thr Tyr Thr Asn Gln His Gly Leu Asp Leu Asp Ala Glu				
	195	200	205	
tcc gca gaa cca ttc ctt gtg ttc cac aac tac cct gct gcc aac act				672
Ser Ala Glu Pro Phe Leu Val Phe His Asn Tyr Pro Ala Ala Asn Thr				
	210	215	220	
gct cga aag tgg ttt cac cga ttc cag gcc tgg tac atg tac ctc gtg				720
Ala Arg Lys Trp Phe His Arg Phe Gln Ala Trp Tyr Met Tyr Leu Val				

225	230	235	240	
ctt gga gcc tac ggc gtt tcg ctg gtg tac aac cct ctc tac atc ttc				768
Leu Gly Ala Tyr Gly Val Ser Leu Val Tyr Asn Pro Leu Tyr Ile Phe				
245	250	255		
cga atg cag cac aac gac acc att ccc gag tct gtc aca gcc atg cga				816
Arg Met Gln His Asn Asp Thr Ile Pro Glu Ser Val Thr Ala Met Arg				
260	265	270		
gag aac ggc ttt ctg cga cgg tac cga acc ctt gca ttc gtt atg cga				864
Glu Asn Gly Phe Leu Arg Arg Tyr Arg Thr Leu Ala Phe Val Met Arg				
275	280	285		
gct ttc ttc atc ttt cga acc gcc ttc ttg ccc tgg tat ctc act gga				912
Ala Phe Phe Ile Phe Arg Thr Ala Phe Leu Pro Trp Tyr Leu Thr Gly				
290	295	300		
acc tcc ctg ctc atc acc att cct ctg gtg ccc act gct acc ggt gcc				960
Thr Ser Leu Leu Ile Thr Ile Pro Leu Val Pro Thr Ala Thr Gly Ala				
305	310	315	320	
ttc ctc acc ttc ttt ttc atc ttg tct cac aac ttc gat ggc tcg gag				1008
Phe Leu Thr Phe Phe Phe Ile Leu Ser His Asn Phe Asp Gly Ser Glu				
325	330	335		
cga atc ccc gac aag aac tgc aag gtc aag agc tcc gag aag gac gtt				1056
Arg Ile Pro Asp Lys Asn Cys Lys Val Lys Ser Ser Glu Lys Asp Val				
340	345	350		
gaa gcc gat cag atc gac tgg tac aga gct cag gtg gag acc tct tcc				1104
Glu Ala Asp Gln Ile Asp Trp Tyr Arg Ala Gln Val Glu Thr Ser Ser				
355	360	365		
acc tac ggt gga ccc att gcc atg ttc ttt act ggc ggt ctc aac ttc				1152
Thr Tyr Gly Gly Pro Ile Ala Met Phe Phe Thr Gly Gly Leu Asn Phe				
370	375	380		
cag atc gag cat cac ctc ttt cct cga atg tcg tct tgg cac tat ccc				1200

Gln Ile Glu His His Leu Phe Pro Arg Met Ser Ser Trp His Tyr Pro
 385 390 395 400
 ttc gtg cag caa gct gtc cga gag tgt tgc gaa cga cac gga gtt cgg 1248
 Phe Val Gln Gln Ala Val Arg Glu Cys Cys Glu Arg His Gly Val Arg
 405 410 415
 tac gtc ttc tac cct acc att gtg ggc aac atc att tcc acc ctc aag 1296
 Tyr Val Phe Tyr Pro Thr Ile Val Gly Asn Ile Ile Ser Thr Leu Lys
 420 425 430
 tac atg cac aaa gtc ggt gtg gtt cac tgt gtc aag gac gct cag gat 1344
 Tyr Met His Lys Val Gly Val Val His Cys Val Lys Asp Ala Gln Asp
 435 440 445
 tcc taa 1350
 Ser

<210> 331

<211> 449

<212> PRT

<213> Euglena gracilis

<400> 331

Met Ala Leu Ser Leu Thr Thr Glu Gln Leu Leu Glu Arg Pro Asp Leu
 1 5 10 15
 Val Ala Ile Asp Gly Ile Leu Tyr Asp Leu Glu Gly Leu Ala Lys Val
 20 25 30
 His Pro Gly Ser Asp Leu Ile Leu Ala Ser Gly Ala Ser Asp Ala Ser
 35 40 45
 Pro Leu Phe Tyr Ser Met His Pro Tyr Val Lys Pro Glu Asn Ser Lys
 50 55 60
 Leu Leu Gln Gln Phe Val Arg Gly Lys His Asp Arg Thr Ser Lys Asp
 65 70 75 80
 Ile Val Tyr Thr Tyr Asp Ser Pro Phe Ala Gln Asp Val Lys Arg Thr
 85 90 95
 Met Arg Glu Val Met Lys Gly Arg Asn Trp Tyr Ala Thr Pro Gly Phe

100 105 110
 Trp Leu Arg Thr Val Gly Ile Ile Ala Val Thr Ala Phe Cys Glu Trp

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

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

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

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

Glu Ala Asp Gln Ile Asp Trp Tyr Arg Ala Gln Val Glu Thr Ser Ser
355 360 365

Thr Tyr Gly Gly Pro Ile Ala Met Phe Phe Thr Gly Gly Leu Asn Phe

370 375 380

Gln Ile Glu His His Leu Phe Pro Arg Met Ser Ser Trp His Tyr Pro

385 390 395 400

Phe Val Gln Gln Ala Val Arg Glu Cys Cys Glu Arg His Gly Val Arg

405 410 415

Tyr Val Phe Tyr Pro Thr Ile Val Gly Asn Ile Ile Ser Thr Leu Lys

420 425 430

Tyr Met His Lys Val Gly Val Val His Cys Val Lys Asp Ala Gln Asp

435 440 445

Ser

<210> 332

<211> 1350

<212> DNA

<213> *Euglena gracilis*

<220><221> CDS

<222> (1)..(1350)

<400> 332

atg gct ctc tcc ctt act acc gag cag ctg ctc gag cga ccc gac ctg 48

Met Ala Leu Ser Leu Thr Thr Glu Gln Leu Leu Glu Arg Pro Asp Leu

1 5 10 15

gtt gcc atc gac ggc att ctc tac gat ctg gaa ggt ctt gcc aag gtc 96

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

20 25 30

cat ccc gga tcc gac ttg atc ctc gct tct ggt gcc tcc gat gct tct 144

His Pro Gly Ser Asp Leu Ile Leu Ala Ser Gly Ala Ser Asp Ala Ser

35 40 45

cct ctg ttc tac tcc atg cac cct tac gtc aag ccc gag aac tcg aag 192

Pro Leu Phe Tyr Ser Met His Pro Tyr Val Lys Pro Glu Asn Ser Lys

50	55	60	
ctg ctt caa cag ttc gtg cga ggc aag cac gac cga acc tcc aag gac			240
Leu Leu Gln Gln Phe Val Arg Gly Lys His Asp Arg Thr Ser Lys Asp			
65	70	75	80
att gtc tac acc tac gac tct ccc ttt gca cag gac gtc aag cga act			288
Ile Val Tyr Thr Tyr Asp Ser Pro Phe Ala Gln Asp Val Lys Arg Thr			
	85	90	95
atg cga gag gtc atg aaa ggt cgg aac tgg tat gcc aca cct gga ttc			336
Met Arg Glu Val Met Lys Gly Arg Asn Trp Tyr Ala Thr Pro Gly Phe			
	100	105	110
tgg ctg cga acc gtt ggc atc att gct gtc acc gcc ttt tgc gag tgg			384
Trp Leu Arg Thr Val Gly Ile Ile Ala Val Thr Ala Phe Cys Glu Trp			
	115	120	125
cac tgg gct act acc gga atg gtg ctg tgg ggt ctc ttg act gga ttc			432
His Trp Ala Thr Thr Gly Met Val Leu Trp Gly Leu Leu Thr Gly Phe			
	130	135	140
atg cac atg cag atc ggc ctg tcc att cag cac gat gcc ggt cat ggt			480
Met His Met Gln Ile Gly Leu Ser Ile Gln His Asp Ala Gly His Gly			
145	150	155	160
gcc atc agc aaa aag ccc tgg gtc aac gct ctc ttt gcc tac ggc atc			528
Ala Ile Ser Lys Lys Pro Trp Val Asn Ala Leu Phe Ala Tyr Gly Ile			
	165	170	175
gac gtc att gga tcg tcc aga tgg atc tgg ctg cag tct cac atc atg			576
Asp Val Ile Gly Ser Ser Arg Trp Ile Trp Leu Gln Ser His Ile Met			
	180	185	190
cga cat cac acc tac acc aat cag cat ggt ctc gac ctg gat gcc gag			624
Arg His His Thr Tyr Thr Asn Gln His Gly Leu Asp Leu Asp Ala Glu			
	195	200	205
tcc gca gaa cca ttc ctt gtg ttc cac aac tac cct gct gcc aac act			672

Ser Ala Glu Pro Phe Leu Val Phe His Asn Tyr Pro Ala Ala Asn Thr
210 215 220
gct cga aag tgg ttt cac cga ttc cag gcc tgg tac atg tac ctc gtg 720
Ala Arg Lys Trp Phe His Arg Phe Gln Ala Trp Tyr Met Tyr Leu Val
225 230 235 240
ctt gga gcc tac ggc gtt tcg ctg gtg tac aac cct ctc tac atc ttc 768
Leu Gly Ala Tyr Gly Val Ser Leu Val Tyr Asn Pro Leu Tyr Ile Phe
245 250 255
cga atg cag cac aac gac acc att ccc gag tct gtc aca gcc atg cga 816
Arg Met Gln His Asn Asp Thr Ile Pro Glu Ser Val Thr Ala Met Arg
260 265 270
gag aac ggc ttt ctg cga cgg tac cga acc ctt gca ttc gtt atg cga 864
Glu Asn Gly Phe Leu Arg Arg Tyr Arg Thr Leu Ala Phe Val Met Arg
275 280 285
gct ttc ttc atc ttt cga acc gcc ttc ttg ccc tgg tat ctc act gga 912
Ala Phe Phe Ile Phe Arg Thr Ala Phe Leu Pro Trp Tyr Leu Thr Gly
290 295 300
acc tcc ctg ctc atc acc att cct ctg gtg ccc act gct acc ggt gcc 960
Thr Ser Leu Leu Ile Thr Ile Pro Leu Val Pro Thr Ala Thr Gly Ala
305 310 315 320
ttc ctc acc ttc ttt ttc atc ttg tct cac aac ttc gat ggc tcg gag 1008
Phe Leu Thr Phe Phe Phe Ile Leu Ser His Asn Phe Asp Gly Ser Glu
325 330 335
cga atc ccc gac aag aac tgc aag gtc aag agc tcc gag aag gac gtt 1056
Arg Ile Pro Asp Lys Asn Cys Lys Val Lys Ser Ser Glu Lys Asp Val
340 345 350
gaa gcc gat cag atc gac tgg tac aga gct cag gtg gag acc tct tcc 1104
Glu Ala Asp Gln Ile Asp Trp Tyr Arg Ala Gln Val Glu Thr Ser Ser
355 360 365
acc tac ggt gga ccc att gcc atg ttc ttt act ggc ggt ctc aac ttc 1152
Thr Tyr Gly Gly Pro Ile Ala Met Phe Phe Thr Gly Gly Leu Asn Phe

370 375 380
 cag atc gag cat cac ctc ttt cct cga atg tcg tct tgg cac tat ccc 1200
 Gln Ile Glu His His Leu Phe Pro Arg Met Ser Ser Trp His Tyr Pro
 385 390 395 400

ttc gtg cag caa gct gtc cga gag tgt tgc gaa cga cac gga gtt cgg 1248
 Phe Val Gln Gln Ala Val Arg Glu Cys Cys Glu Arg His Gly Val Arg

405 410 415
 tac gtc ttc tac cct acc att gtg ggc aac atc att tcc acc ctc aag 1296
 Tyr Val Phe Tyr Pro Thr Ile Val Gly Asn Ile Ile Ser Thr Leu Lys

420 425 430
 tac atg cac aaa gtc ggt gtg gtt cac tgt gtc aag gac gct cag gat 1344

Tyr Met His Lys Val Gly Val Val His Cys Val Lys Asp Ala Gln Asp

435 440 445
 tcc taa 1350
 Ser

<210> 333

<211> 449

<212> PRT

<213> Euglena gracilis

<400> 333

Met Ala Leu Ser Leu Thr Thr Glu Gln Leu Leu Glu Arg Pro Asp Leu

1 5 10 15
 Val Ala Ile Asp Gly Ile Leu Tyr Asp Leu Glu Gly Leu Ala Lys Val

20 25 30
 His Pro Gly Ser Asp Leu Ile Leu Ala Ser Gly Ala Ser Asp Ala Ser

35 40 45
 Pro Leu Phe Tyr Ser Met His Pro Tyr Val Lys Pro Glu Asn Ser Lys

50 55 60
 Leu Leu Gln Gln Phe Val Arg Gly Lys His Asp Arg Thr Ser Lys Asp

65 70 75 80

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

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

385 390 395 400
 Phe Val Gln Gln Ala Val Arg Glu Cys Cys Glu Arg His Gly Val Arg
 405 410 415
 Tyr Val Phe Tyr Pro Thr Ile Val Gly Asn Ile Ile Ser Thr Leu Lys
 420 425 430
 Tyr Met His Lys Val Gly Val Val His Cys Val Lys Asp Ala Gln Asp
 435 440 445
 Ser

<210> 334
 <211> 454
 <212> PRT

<213> Euglena anabaena

<400> 334
 Met Ala Thr Ile Ser Leu Thr Thr Glu Gln Leu Leu Glu His Pro Glu
 1 5 10 15
 Leu Val Ala Ile Asp Gly Val Leu Tyr Asp Leu Phe Gly Leu Ala Lys
 20 25 30
 Val His Ala Gly Gly Asn Leu Ile Glu Ala Ala Gly Ala Ser Asp Gly
 35 40 45
 Thr Ala Leu Phe Tyr Ser Met His Pro Gly Val Lys Pro Glu Asn Ser
 50 55 60
 Lys Leu Leu Gln Gln Phe Ala Arg Gly Lys His Glu Arg Ser Ser Lys
 65 70 75 80

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

 Trp Tyr Trp Ile Thr Thr Gly Ala Val Met Trp Gly Ile Phe Thr Gly
 130 135 140
 Tyr Phe His Ser Gln Ile Gly Leu Ala Ile Gln His Asp Ala Ser His
 145 150 155 160
 Gly Ala Ile Ser Lys Lys Pro Trp Val Asn Ala Phe Phe Ala Tyr Gly
 165 170 175
 Ile Asp Ala Ile Gly Ser Ser Arg Trp Ile Trp Leu Gln Ser His Ile
 180 185 190

 Met Arg His His Thr Tyr Thr Asn Gln His Gly Leu Asp Leu Asp Ala
 195 200 205
 Ala Ser Ala Glu Pro Phe Ile Leu Phe His Ser Tyr Pro Ala Thr Asn
 210 215 220
 Ala Ser Arg Lys Trp Tyr His Arg Phe Gln Ala Trp Tyr Met Tyr Ile
 225 230 235 240
 Val Leu Gly Met Tyr Gly Val Ser Met Val Tyr Asn Pro Met Tyr Leu
 245 250 255

 Phe Thr Met Gln His Asn Asp Thr Ile Pro Glu Ala Thr Ser Leu Arg
 260 265 270
 Pro Gly Ser Phe Phe Asn Arg Gln Arg Ala Phe Ala Val Ser Leu Arg
 275 280 285
 Leu Leu Phe Ile Phe Arg Asn Ala Phe Leu Pro Trp Tyr Ile Ala Gly
 290 295 300
 Ala Ser Pro Leu Leu Thr Ile Leu Leu Val Pro Thr Val Thr Gly Ile
 305 310 315 320

 Phe Leu Thr Phe Val Phe Val Leu Ser His Asn Phe Glu Gly Ala Glu

325 330 335
 Arg Thr Pro Glu Lys Asn Cys Lys Ala Lys Arg Ala Lys Glu Gly Lys
 340 345 350
 Glu Val Arg Asp Val Glu Glu Asp Arg Val Asp Trp Tyr Arg Ala Gln
 355 360 365
 Ala Glu Thr Ala Ala Thr Tyr Gly Gly Ser Val Gly Met Met Leu Thr
 370 375 380

Gly Gly Leu Asn Leu Gln Ile Glu His His Leu Phe Pro Arg Met Ser
 385 390 395 400
 Ser Trp His Tyr Pro Phe Ile Gln Asp Thr Val Arg Glu Cys Cys Lys
 405 410 415
 Arg His Gly Val Arg Tyr Thr Tyr Tyr Pro Thr Ile Leu Glu Asn Ile
 420 425 430
 Met Ser Thr Leu Arg Tyr Met Gln Lys Val Gly Val Ala His Thr Ile
 435 440 445

Gln Asp Ala Gln Glu Phe
 450

<210> 335

<211> 1362

<212> DNA

<213> Euglena anabaena UTEX 373

<220><221> CDS

<222> (1)..(1362)

<223> delta-5 desaturase

<300><302> DELTA-5 DESATURASES AND THEIR USE IN MAKING POLYUNSATURATED FATTY
 ACIDS

<310> WO 2008/137532

<311> 2008-05-01

<312> 2008-11-13

<313> (1)..(1362)

<300><302> DELTA-5 DESATURASES AND THEIR USE IN MAKING POLYUNSATURATED FATTY
 ACIDS

<310> US-2008-0274521-A1

<311> 2008-04-29

<312> 2008-11-16

<313> (1)..(1362)

<400> 335

atg gcc acc atc tct ttg act act gag caa ctt tta gaa cac cca gaa 48

Met Ala Thr Ile Ser Leu Thr Thr Glu Gln Leu Leu Glu His Pro Glu

1 5 10 15

ctg gtt gca att gat ggg gtg ttg tac gat ctc ttc gga ctg gcg aaa 96

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

20 25 30

gtg cat cca ggt ggc aac ctc att gaa gcc gcc ggt gcc tcc gac gga 144

Val His Pro Gly Gly Asn Leu Ile Glu Ala Ala Gly Ala Ser Asp Gly

35 40 45

acc gcc ctg ttc tac tcc atg cac cct gga gtg aag cca gag aat tcg 192

Thr Ala Leu Phe Tyr Ser Met His Pro Gly Val Lys Pro Glu Asn Ser

50 55 60

aag ctg ctg cag caa ttt gcc cga ggc aaa cac gaa cga agc tcg aag 240

Lys Leu Leu Gln Gln Phe Ala Arg Gly Lys His Glu Arg Ser Ser Lys

65 70 75 80

gac cca gtg tac acc ttt gac agt ccc ttc gcc cag gat gtc aag cag 288

Asp Pro Val Tyr Thr Phe Asp Ser Pro Phe Ala Gln Asp Val Lys Gln

85 90 95

agc gtt cgg gag gtc atg aag ggg cgc aac tgg tac gcc acg ccc ggc 336

Ser Val Arg Glu Val Met Lys Gly Arg Asn Trp Tyr Ala Thr Pro Gly

100 105 110

ttt tgg ctg cgg acc gcg ctg atc atc gcg tgc act gcc ata ggc gaa 384

Phe Trp Leu Arg Thr Ala Leu Ile Ile Ala Cys Thr Ala Ile Gly Glu

115 120 125

tgg tat tgg atc act acc ggg gca gtg atg tgg ggc atc ttc acc ggc 432

Trp Tyr Trp Ile Thr Thr Gly Ala Val Met Trp Gly Ile Phe Thr Gly

130	135	140	
tac ttc cac agc cag att ggg ttg gcg att caa cac gat gcc tct cac			480
Tyr Phe His Ser Gln Ile Gly Leu Ala Ile Gln His Asp Ala Ser His			
145	150	155	160
gga gcc atc agc aaa aag ccc tgg gtg aac gcc ttt ttc gcc tac ggc			528
Gly Ala Ile Ser Lys Lys Pro Trp Val Asn Ala Phe Phe Ala Tyr Gly			
	165	170	175
atc gac gcc att gga tcc tcc cgc tgg atc tgg ctg cag tcc cac att			576
Ile Asp Ala Ile Gly Ser Ser Arg Trp Ile Trp Leu Gln Ser His Ile			
	180	185	190
atg cgc cac cac acc tac acc aac cag cat ggc ctg gac ctg gac gct			624
Met Arg His His Thr Tyr Thr Asn Gln His Gly Leu Asp Leu Asp Ala			
	195	200	205
gcc tcg gcg gag ccg ttc att ttg ttc cac tcc tac ccg gca aca aat			672
Ala Ser Ala Glu Pro Phe Ile Leu Phe His Ser Tyr Pro Ala Thr Asn			
210	215	220	
gcg tca cga aag tgg tac cat cgg ttc cag gcg tgg tac atg tac atc			720
Ala Ser Arg Lys Trp Tyr His Arg Phe Gln Ala Trp Tyr Met Tyr Ile			
225	230	235	240
gtt ttg ggg atg tat ggt gtg tcg atg gtg tac aat ccg atg tac ttg			768
Val Leu Gly Met Tyr Gly Val Ser Met Val Tyr Asn Pro Met Tyr Leu			
	245	250	255
ttc acg atg cag cac aac gac aca atc cca gag gcc acc tct ctt aga			816
Phe Thr Met Gln His Asn Asp Thr Ile Pro Glu Ala Thr Ser Leu Arg			
	260	265	270
cca gcc agc ttt ttc aac cgg cag cgc gcc ttc gcc gtt tcc ctc cgc			864
Pro Gly Ser Phe Phe Asn Arg Gln Arg Ala Phe Ala Val Ser Leu Arg			
	275	280	285
cta ctg ttc atc ttc cgc aac gcc ttc ctc ccc tgg tac atc gcg ggc			912

Leu Leu Phe Ile Phe Arg Asn Ala Phe Leu Pro Trp Tyr Ile Ala Gly
 290 295 300
 gcc tct ccg ctg ctc acc atc ctg ctg gtg cca acg gtc aca ggc atc 960
 Ala Ser Pro Leu Leu Thr Ile Leu Leu Val Pro Thr Val Thr Gly Ile
 305 310 315 320
 ttc ttg aca ttt gtt ttt gtg ctg tcc cat aac ttt gaa ggc gct gag 1008
 Phe Leu Thr Phe Val Phe Val Leu Ser His Asn Phe Glu Gly Ala Glu
 325 330 335
 cgg acc ccc gaa aag aac tgc aag gcc aaa agg gcc aag gag ggg aag 1056
 Arg Thr Pro Glu Lys Asn Cys Lys Ala Lys Arg Ala Lys Glu Gly Lys
 340 345 350
 gag gtc cgc gat gta gag gag gac cgg gtg gac tgg tac cgg gcg cag 1104
 Glu Val Arg Asp Val Glu Glu Asp Arg Val Asp Trp Tyr Arg Ala Gln
 355 360 365
 gcc gag acc gcg gcg acc tac ggg ggc agc gtc ggg atg atg ctg acc 1152
 Ala Glu Thr Ala Ala Thr Tyr Gly Gly Ser Val Gly Met Met Leu Thr
 370 375 380
 ggc ggt ttg aac ctg cag atc gag cac cac ttg ttc ccc cgc atg tcc 1200
 Gly Gly Leu Asn Leu Gln Ile Glu His His Leu Phe Pro Arg Met Ser
 385 390 395 400
 tct tgg cac tac ccc ttc atc caa gat acg gtg cgg gaa tgt tgc aag 1248
 Ser Trp His Tyr Pro Phe Ile Gln Asp Thr Val Arg Glu Cys Cys Lys
 405 410 415
 cgc cat ggc gtg cgc tac aca tac tac ccg acc atc ctg gag aat ata 1296
 Arg His Gly Val Arg Tyr Thr Tyr Tyr Pro Thr Ile Leu Glu Asn Ile
 420 425 430
 atg tcc acg ctc cgc tac atg cag aag gtg ggc gtg gcc cac aca att 1344
 Met Ser Thr Leu Arg Tyr Met Gln Lys Val Gly Val Ala His Thr Ile
 435 440 445
 cag gat gcc cag gaa ttc 1362
 Gln Asp Ala Gln Glu Phe

450

<210> 336

<211> 454

<212> PRT

<213> Euglena anabaena UTEX 373

<400> 336

Met Ala Thr Ile Ser Leu Thr Thr Glu Gln Leu Leu Glu His Pro Glu

1 5 10 15

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

20 25 30

Val His Pro Gly Gly Asn Leu Ile Glu Ala Ala Gly Ala Ser Asp Gly

35 40 45

Thr Ala Leu Phe Tyr Ser Met His Pro Gly Val Lys Pro Glu Asn Ser

50 55 60

Lys Leu Leu Gln Gln Phe Ala Arg Gly Lys His Glu Arg Ser Ser Lys

65 70 75 80

Asp Pro Val Tyr Thr Phe Asp Ser Pro Phe Ala Gln Asp Val Lys Gln

85 90 95

Ser Val Arg Glu Val Met Lys Gly Arg Asn Trp Tyr Ala Thr Pro Gly

100 105 110

Phe Trp Leu Arg Thr Ala Leu Ile Ile Ala Cys Thr Ala Ile Gly Glu

115 120 125

Trp Tyr Trp Ile Thr Thr Gly Ala Val Met Trp Gly Ile Phe Thr Gly

130 135 140

Tyr Phe His Ser Gln Ile Gly Leu Ala Ile Gln His Asp Ala Ser His

145 150 155 160

Gly Ala Ile Ser Lys Lys Pro Trp Val Asn Ala Phe Phe Ala Tyr Gly

165 170 175

Ile Asp Ala Ile Gly Ser Ser Arg Trp Ile Trp Leu Gln Ser His Ile

180 185 190

Met Arg His His Thr Tyr Thr Asn Gln His Gly Leu Asp Leu Asp Ala

195 200 205
 Ala Ser Ala Glu Pro Phe Ile Leu Phe His Ser Tyr Pro Ala Thr Asn
 210 215 220
 Ala Ser Arg Lys Trp Tyr His Arg Phe Gln Ala Trp Tyr Met Tyr Ile
 225 230 235 240
 Val Leu Gly Met Tyr Gly Val Ser Met Val Tyr Asn Pro Met Tyr Leu
 245 250 255
 Phe Thr Met Gln His Asn Asp Thr Ile Pro Glu Ala Thr Ser Leu Arg
 260 265 270

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

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

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

Gln Asp Ala Gln Glu Phe

450

<210> 337

<211> 1362

<212> DNA

<213> Euglena anabaena UTEX 373

<220><221> CDS

<222> (1)..(1362)

<223> synthetic delta-5 desaturase (codon-optimized for Yarrowia
lipolytica)

<300><302> DELTA-5 DESATURASES AND THEIR USE IN MAKING POLYUNSATURATED FATTY
ACIDS

<310> WO 2008/137532

<311> 2008-05-01

<312> 2008-11-13

<313> (1)..(1362)

<300><302> DELTA-5 DESATURASES AND THEIR USE IN MAKING POLYUNSATURATED FATTY
ACIDS

<310> US-2008-0274521-A1

<311> 2008-04-29

<312> 2008-11-06

<313> (1)..(1362)

<400> 337

atg gcc acc atc tcc ctg act acc gag cag ctc ctg gaa cac ccc gag 48

Met Ala Thr Ile Ser Leu Thr Thr Glu Gln Leu Leu Glu His Pro Glu

1 5 10 15

ctc gtt gcc atc gac gga gtc ctg tac gat ctc ttc ggt ctg gcc aag 96

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

20 25 30

gtg cat cca gga ggc aac ctc atc gaa gct gcc ggt gca tcc gac gga 144

Val His Pro Gly Gly Asn Leu Ile Glu Ala Ala Gly Ala Ser Asp Gly

35

40

45

acc gct ctg ttc tac tcc atg cat cct gga gtc aag cca gag aac tcg	192
Thr Ala Leu Phe Tyr Ser Met His Pro Gly Val Lys Pro Glu Asn Ser	
50 55 60	
aag ctt ctg cag caa ttt gcc cga ggc aag cac gaa cga agc tcc aag	240
Lys Leu Leu Gln Gln Phe Ala Arg Gly Lys His Glu Arg Ser Ser Lys	
65 70 75 80	
gat ccc gtg tac acc ttc gac tct ccc ttt gct cag gac gtc aag cag	288
Asp Pro Val Tyr Thr Phe Asp Ser Pro Phe Ala Gln Asp Val Lys Gln	
85 90 95	
tcc gtt cga gag gtc atg aag ggt cga aac tgg tac gcc act cct ggc	336
Ser Val Arg Glu Val Met Lys Gly Arg Asn Trp Tyr Ala Thr Pro Gly	
100 105 110	
ttc tgg ctg aga acc gca ctc atc atc gct tgt act gcc att ggc gag	384
Phe Trp Leu Arg Thr Ala Leu Ile Ile Ala Cys Thr Ala Ile Gly Glu	
115 120 125	
tgg tac tgg atc aca acc gga gca gtg atg tgg ggt atc ttt act gga	432
Trp Tyr Trp Ile Thr Thr Gly Ala Val Met Trp Gly Ile Phe Thr Gly	
130 135 140	
tac ttc cac tcg cag att ggc ttg gcc att caa cac gat gct tct cac	480
Tyr Phe His Ser Gln Ile Gly Leu Ala Ile Gln His Asp Ala Ser His	
145 150 155 160	
gga gcc atc agc aaa aag ccc tgg gtc aac gcc ttt ttc gct tat ggc	528
Gly Ala Ile Ser Lys Lys Pro Trp Val Asn Ala Phe Phe Ala Tyr Gly	
165 170 175	
atc gac gcc att ggt tcc tct cgt tgg atc tgg ctg cag tcc cac att	576
Ile Asp Ala Ile Gly Ser Ser Arg Trp Ile Trp Leu Gln Ser His Ile	
180 185 190	
atg cga cat cac act tac acc aac cag cat ggc ctc gac ctg gat gct	624
Met Arg His His Thr Tyr Thr Asn Gln His Gly Leu Asp Leu Asp Ala	
195 200 205	
gcc tcg gca gag ccg ttc atc ttg ttc cac tcc tat cct gct acc aac	672

Ala Ser Ala Glu Pro Phe Ile Leu Phe His Ser Tyr Pro Ala Thr Asn	
210 215 220	
gcc tct cga aag tgg tac cac cga ttt cag gcg tgg tac atg tac atc	720
Ala Ser Arg Lys Trp Tyr His Arg Phe Gln Ala Trp Tyr Met Tyr Ile	
225 230 235 240	
gtt ctg gga atg tat ggt gtc tcg atg gtg tac aat ccc atg tac ctc	768
Val Leu Gly Met Tyr Gly Val Ser Met Val Tyr Asn Pro Met Tyr Leu	
245 250 255	
ttc aca atg cag cac aac gac acc att ccc gag gcc act tct ctc aga	816
Phe Thr Met Gln His Asn Asp Thr Ile Pro Glu Ala Thr Ser Leu Arg	
260 265 270	
cca ggc agc ttt ttc aat cgg cag cga gct ttc gcc gtt tcc ctt cga	864
Pro Gly Ser Phe Phe Asn Arg Gln Arg Ala Phe Ala Val Ser Leu Arg	
275 280 285	
ctg ctc ttc atc ttc cga aac gcc ttt ctt ccc tgg tac att gct ggt	912
Leu Leu Phe Ile Phe Arg Asn Ala Phe Leu Pro Trp Tyr Ile Ala Gly	
290 295 300	
gcc tct cct ctg ctc acc att ctt ctg gtg ccc acg gtc aca ggc atc	960
Ala Ser Pro Leu Leu Thr Ile Leu Leu Val Pro Thr Val Thr Gly Ile	
305 310 315 320	
ttc ctc acc ttt gtg ttc gtt ctg tcc cat aac ttc gag gga gcc gaa	1008
Phe Leu Thr Phe Val Phe Val Leu Ser His Asn Phe Glu Gly Ala Glu	
325 330 335	
cgg acc cca gag aag aac tgc aag gcc aaa cga gct aag gaa ggc aag	1056
Arg Thr Pro Glu Lys Asn Cys Lys Ala Lys Arg Ala Lys Glu Gly Lys	
340 345 350	
gag gtc aga gac gtg gaa gag gat cga gtc gac tgg tac cga gca cag	1104
Glu Val Arg Asp Val Glu Glu Asp Arg Val Asp Trp Tyr Arg Ala Gln	
355 360 365	
gcc gag act gct gcc acc tac ggt ggc agc gtg gga atg atg ctt aca	1152

Ala Glu Thr Ala Ala Thr Tyr Gly Gly Ser Val Gly Met Met Leu Thr

370	375	380	
ggc ggt ctc aac ctg cag atc gag cat cac ttg ttt ccc cga atg tcc			1200
Gly Gly Leu Asn Leu Gln Ile Glu His His Leu Phe Pro Arg Met Ser			
385	390	395	400
tct tgg cac tat ccc ttc att caa gac acc gtt cgg gag tgt tgc aag			1248
Ser Trp His Tyr Pro Phe Ile Gln Asp Thr Val Arg Glu Cys Cys Lys			
405	410	415	
cga cat ggc gtc cgt tac aca tac tat cct acc att ctc gag aac atc			1296
Arg His Gly Val Arg Tyr Thr Tyr Tyr Pro Thr Ile Leu Glu Asn Ile			
420	425	430	
atg tcc act ctt cga tac atg cag aag gtg ggt gtt gct cac acc att			1344
Met Ser Thr Leu Arg Tyr Met Gln Lys Val Gly Val Ala His Thr Ile			
435	440	445	
cag gat gcc cag gag ttc			1362

Gln Asp Ala Gln Glu Phe

450
 <210> 338
 <211> 454
 <212> PRT
 <213> Euglena anabaena UTEX 373
 <400> 338

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

65 70 75 80
 Asp Pro Val Tyr Thr Phe Asp Ser Pro Phe Ala Gln Asp Val Lys Gln
 85 90 95
 Ser Val Arg Glu Val Met Lys Gly Arg Asn Trp Tyr Ala Thr Pro Gly

 100 105 110
 Phe Trp Leu Arg Thr Ala Leu Ile Ile Ala Cys Thr Ala Ile Gly Glu
 115 120 125
 Trp Tyr Trp Ile Thr Thr Gly Ala Val Met Trp Gly Ile Phe Thr Gly
 130 135 140
 Tyr Phe His Ser Gln Ile Gly Leu Ala Ile Gln His Asp Ala Ser His
 145 150 155 160
 Gly Ala Ile Ser Lys Lys Pro Trp Val Asn Ala Phe Phe Ala Tyr Gly

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

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

 290 295 300
 Ala Ser Pro Leu Leu Thr Ile Leu Leu Val Pro Thr Val Thr Gly Ile
 305 310 315 320

Phe Leu Thr Phe Val Phe Val Leu Ser His Asn Phe Glu Gly Ala Glu
325 330 335

Arg Thr Pro Glu Lys Asn Cys Lys Ala Lys Arg Ala Lys Glu Gly Lys
340 345 350

Glu Val Arg Asp Val Glu Glu Asp Arg Val Asp Trp Tyr Arg Ala Gln

355 360 365
Ala Glu Thr Ala Ala Thr Tyr Gly Gly Ser Val Gly Met Met Leu Thr
370 375 380

Gly Gly Leu Asn Leu Gln Ile Glu His His Leu Phe Pro Arg Met Ser
385 390 395 400

Ser Trp His Tyr Pro Phe Ile Gln Asp Thr Val Arg Glu Cys Cys Lys
405 410 415

Arg His Gly Val Arg Tyr Thr Tyr Tyr Pro Thr Ile Leu Glu Asn Ile

420 425 430
Met Ser Thr Leu Arg Tyr Met Gln Lys Val Gly Val Ala His Thr Ile
435 440 445

Gln Asp Ala Gln Glu Phe
450

<210> 339

<211> 8357

<212> DNA

<213> Artificial Sequence

<220><223> Plasmid pZuFmEaD5S-A(S)

<400> 339

catggccacc atctccctga ctaccgagca gctcctggaa caccgccgagc tcgttgccat 60
cgacggagtc ctgtacgata tcttcgggtct ggccaaggtc catgccggag gcaacctcat 120

cgaagctgcc ggtgcatccg acggaaccgc tctgttctac tccatgcata ctggagtcaa 180
gccagagaac tcgaagcttc tgcagcaatt tgcccagggc aagcacgaac gaagctccaa 240
ggatcccggtg tacaccttcg actctccctt tgctcaggac gtcaagcagt ccgttcgaga 300
ggtcatgaag ggtcgaaact ggtacgccac tcttggtctc tggctgagaa ccgcactcat 360
catcgcttgt actgccattg gcgagtggta ctggatcaca accggagcag tgatgtgggg 420
tatctttact ggatacttcc actcgcagat tggcttggtc attcaacacg atgcttctca 480

cggagccatc agcaaaaagc cctgggtcaa cgcctttttc gcttatggca tcgacgcat 540

 tggttcctct cgttggatct ggctgcagtc ccacattatg cgacatcaca cttacaccaa 600
 ccagcatggc ctgcacctgg atgctgcctc ggacagagccg ttcactttgt tccactccta 660
 tcctgctacc aacgcctctc gaaagtggta ccaccgattt caggcgtggt acatgtacat 720
 cgttctggga atgtatgggt tctcgatggt gtacaatccc atgtacctct tcacaatgca 780
 gcacaacgac accattcccg aggccacttc tctcagacca ggacgctttt tcaatcggca 840
 gcgagctttc gccgtttccc ttcgactgct cttcatcttc cgaaacgcct ttcttcctg 900
 gtacattgct ggtgcctctc ctctgctcac cattctctg gtgcccacgg tcacaggcat 960

 cttectcacc ttigtgttcg ttctgtccca taacttcgag ggagccgaac ggaccccaga 1020
 gaagaactgc aaggccaaac gagctaagga aggcaaggag gtcagagacg tggaagagga 1080
 tcgagtcgac tggtaaccgag cacaggccga gactgctgcc acctacggtg gcagcgtggg 1140
 aatgatgctt acaggcggtc tcaacctgca gatcgagcat cacttgtttc cccgaatgct 1200
 ctcttggcac tatcccttca ttcaagacac cgttcgggag tgttgcaagc gacatggcgt 1260
 ccgttacaca tactatccta ccattctcga gaacatcatg tccactcttc gatacatgca 1320
 gaaggtgggt gttgctcaca ccattcagga tgcccaggag ttctaagcgg ccgaagtgt 1380

 ggatggggaa gtgagtcccc ggttctgtgt gcacaattgg caatccaaga tggatggatt 1440
 caacacaggg atatagcgag ctacgtgggt gtgcgaggat atagcaacgg atatttatgt 1500
 ttgacacttg agaatgtacg atacaagcac tgtccaagta caatactaaa catactgtac 1560
 atactcatac tcgtaccggg gcaacggttt cacttgagtg cagtggctag tgctcttact 1620
 cgtacagtgt gcaatactgc gtatcatagt ctttgatgta tatcgtattc attcatgtta 1680
 gttgcgtacg agccggaagc ataaagtgt aagcctgggg tgcctaatga gtgagctaac 1740
 tcacattaat tgcgttgccg tcaactgccc ctttccagtc gggaaacctg tcgtgccagc 1800

 tgcattaatg aatcgcccaa cgcgcgggga gaggcggttt gcgtattggg cgctcttccg 1860
 cttectcgtc cactgactcg ctgcgctcgg tcgttcggct gcggcgagcg gtatcagctc 1920
 actcaaaggc ggtaatacgg ttatccacag aatcagggga taacgcagga aagaacatgt 1980
 gagcaaaagg ccagcaaaag gccaggaacc gtaaaaaggc cgcgttgctg gcgtttttcc 2040
 ataggctccg cccccctgac gagcatcaca aaaatcgacg ctcaagttag aggtggcgaa 2100
 acccgacagg actataaaga taccaggcgt ttccccctgg aagctccctc gtgcgctctc 2160
 ctgttccgac cctgccgctt accggatacc tgtccgctt tctcccttcg ggaagcgtgg 2220

cgctttctca tagctcacgc tgtaggtatc tcagttcggg gtaggtcggt cgctccaagc 2280
 tgggctgtgt gcacgaaccc cccgttcagc ccgaccgtg cgccttatcc ggtaactatc 2340
 gtcttgagtc caaccggta agacacgact tatcgccact ggcagcagcc actggttaaca 2400
 ggattagcag agcgaggat gttagcggtg ctacagagtt cttgaagtgg tggcctaact 2460
 acggctacac tagaaggaca gtatttggtg tctgcgctct gctgaagcca gttaccttcg 2520
 gaaaaagagt tggtagctct tgatccggca aacaaaccac cgctggtagc ggtggttttt 2580
 ttgtttgcaa gcagcagatt acgcgcagaa aaaaaggatc tcaagaagat cctttgatct 2640

tttctacggg gtctgacgt cagtggaaacg aaaactcacg ttaagggatt ttggtcatga 2700
 gattatcaaa aaggatcttc acctagatcc ttttaatta aaaatgaagt tttaaatcaa 2760
 tctaaagtat atatgagtaa acttgggtctg acagttacca atgcttaatc agtgaggcac 2820
 ctatctcagc gatctgtcta tttcgttcat ccatagtgc ctgactcccc gtcgtgtaga 2880
 taactacgat acgggagggc ttaccatctg gccccagtgc tgcaatgata ccgcgagacc 2940
 cacgctcacc ggctccagat ttatcagcaa taaaccagcc agccggaagg gccgagcgca 3000
 gaagtggtec tgcaacttta tccgcctcca tccagtctat taattgttgc cgggaagcta 3060

gagtaagtag ttcgccagtt aatagtttgc gcaacgttgt tgccattgct acaggcatcg 3120
 tgggtgcacg ctctcgcttt ggtatggctt cattcagctc cgggtcccaa cgatcaaggc 3180
 gagttacatg atcccccatg ttgtgcaaaa aagcggtttag ctcttcgggt cctccgatcg 3240
 ttgtcagaag taagtggccc gcagtgttat cactcatggt tatggcagca ctgcataatt 3300
 ctcttactgt catgccatcc gtaagatgct tttctgtgac tggtagtac tcaaccaagt 3360
 cattctgaga atagtgtatg cggcgaccga gtgtctcttg cccggcgta atacgggata 3420
 ataccgcgcc acatagcaga actttaaaag tgctcatcat tggaaaacgt tcttcggggc 3480

gaaaactctc aaggatctta ccgctgttga gatccagttc gatgtaaccc actcgtgcac 3540
 ccaactgac ttcagcatct ttactttca ccagcgttcc tgggtgagca aaaacaggaa 3600
 ggcaaaatgc cgcaaaaag ggaataaggg cgacacggaa atgttgaata ctcatactct 3660
 tcctttttca atattattga agcatttatc agggttattg tctcatgagc ggatacatat 3720
 ttgaatgtat ttagaaaaat aaacaaatag gggttccgcg cacatttccc cgaaaagtgc 3780
 cacctgacgc gccctgtagc ggcgcattaa gcgcggcggg tgtggtggtt acgcgcagcg 3840
 tgaccgctac acttgccagc gccctagcgc ccgctccttt cgtttcttc ccttccttc 3900

tcgccacgtt cgccggcttt ccccgtaag ctctaaatcg ggggctccct ttagggttcc 3960
 gatttagtgc ttacggcac ctgcaccca aaaaacttga ttaggggtgat ggttcacgta 4020
 gtgggccatc gccctgatag acggtttttc gcccttgac gttggagtcc acgttcttta 4080

atagtggact ctgtttccaa actggaacaa cactcaaccc tatctcggtc tttctttttg 4140
 atttataagg gattttgccg atttcggcct attgggttaa aaatgagctg atttaacaaa 4200
 aatttaacgc gaattttaac aaaatattaa cgcttacaat ttccattcgc cattcaggct 4260
 gcgcaactgt tgggaagggc gatcgggtcg ggctctctcg ctattacgcc agctggcgaa 4320

 agggggatgt gctgcaaggc gattaagttg ggtaacgcca gggttttccc agtcacgacg 4380
 ttgtaaaacg acggccagtg aattgtaata cgactcacta tagggcgaat tgggtaccgg 4440
 gccccccctc gaggtcgatg gtgtcgataa gcttgatata gaattcatgt cacacaaacc 4500
 gatcttcgcc tcaaggaaac ctaattctac atccgagaga ctgccgagat ccagtctaca 4560
 ctgattaatt ttcgggcaaa taatttaaaa aaatcgtgtt atataatatt atatgtatta 4620
 tataatata tcatgatgat actgacagtc atgtccatt gctaaataga cagactccat 4680
 ctgccgcctc caactgatgt tctcaatatt taagggtca tctcgattg ttttaataata 4740

 aacagactcc atctaccgcc tccaaatgat gtctcaaaa tatattgtat gaacttattt 4800
 ttattactta gtattattag acaacttact tgctttatga aaaacacttc ctatttagga 4860
 aacaatttat aatggcagtt cggttcattta acaatttatg tagaataaat gttataaatg 4920
 cgtatgggaa atcttaata tggatagcat aaatgatata tgcattgcct aattcgaaat 4980
 caacagcaac gaaaaaaatc cttgtacaa cataaatagt catcgagaaa tatcaactat 5040
 caaagaacag ctattcacac gttactattg agattattat tggacgagaa tcacacactc 5100
 aactgtcttt ctctcttcta gaaatacagg tacaagtatg tactattctc attgttcata 5160

 ctctagtea tttcatccca catattcctt ggatttctct ccaatgaatg acattctatc 5220
 ttgcaaatc aacaattata ataagatata ccaaagtagc ggtatagtgg caatcaaaaa 5280
 gcttctctgg tgtgcttctc gtatttattt ttattctaata gatccattaa aggtatatat 5340
 ttatttcttg ttatataatc cttttgttta ttacatgggc tggatacata aaggatattt 5400
 gatttaattt ttgtctaaa ttcaatcccc cctcgttcag tgtcaactgt aatggtagga 5460
 aattaccata cttttgaaga agcaaaaaaa atgaaagaaa aaaaaaatcg tatttccagg 5520
 ttagacgttc cgcagaatct agaatgcggt atgcggtaca ttgttcttcg aacgtaaaag 5580

 ttgcgtccc tgagatatg tacatttttg cttttacaag tacaagtaca tcgtacaact 5640
 atgtactact gttgatgat ccacaacagt ttgttttggt tttttttgtt tttttttttt 5700
 ctaatgattc attaccgcta tgtataccia ctgtacttg tagtaagccg ggttattggc 5760
 gttaaatata tcatagactt atgaatctgc acggtgtgcg ctgcgagtta cttttagctt 5820
 atgatgcta cttgggtgta atattgggat ctgttcggaa atcaacggat gctcaatcga 5880
 tttcgacagt aattaattaa gtcatacaca agtcagcttt cttcgagcct catataagta 5940

taagtagttc aacgtattag cactgtaccc agcatctccg tatcgagaaa cacaacaaca	6000
tgccccattg gacagatcat gcggatacac aggttgtgca gtatcataca tactcgatca	6060
gacaggtcgt ctgacatca tacaagctga acaagcgctc catacttgca cgctctctat	6120
atacacagtt aaattacata tccatagtct aacctctaac agttaatctt ctggtaagcc	6180
tcccagccag cttcttggtg tcgcttggtc tctcaatag gatctcggtt ctggccgtac	6240
agacctcggc cgacaattat gatatccgtt ccggtagaca tgacatcctc aacagttcgg	6300
tactgtgtgc cgagagcgtc tcccttgctg tcaagacca ccccgggggt cagaataagc	6360
cagtcctcag agtcgccctt aggtcgggtt tgggcaatga agccaaccac aaactcgggg	6420
tcggatcggg caagctcaat ggtctgcttg gactactcgc cagtggccag agagcccttg	6480
caagacagct cggccagcat gagcagacct ctggccagct tctcgttggg agaggggact	6540
aggaactcct tgtactggga gttctcgtag tcagagacgt cctccttctt ctgttcagag	6600
acagtttctt cggcaccagc tcgcaggcca gcaatgattc cggttccggg tacaccgtgg	6660
gcgttggtga tatcggaaca ctcgcgatt cggtagacacc ggtactggtg ctgacagtg	6720
ttgccaatat ctgcgaactt tctgtcctcg aacaggaaga aaccgtgctt aagagcaagt	6780
tccttgaggg ggagcacagt gccggcgtag gtgaagtcgt caatgatgtc gatatgggtt	6840
ttgatcatgc acacataagg tccgacctta tcggcaagct caatgagctc ctgtgtggtg	6900
gtaacatcca gagaagcaca caggttggtt ttcttggtg ccacgagctt gagcactcga	6960
gcggcaaagg cggacttggt gacgttagct cgagcttcgt aggagggcat ttgtgtggtg	7020
aagaggagac tgaataaat ttagtctgca gaacttttta tcggaacctt atctggggca	7080
gtgaagtata tgttatggtg atagttacga gttagttaa cttatagata gactggacta	7140
tacggctatc ggtccaaatt agaaagaacg tcaatggctc tctgggcgtc gcctttgccg	7200
acaaaaatgt gatcatgatg aaagccagca atgacgttgc agctgatatt gttgtcggcc	7260
aaccgcgcg aaaacgcagc tgtcagacc acagcctcca acgaagaatg tatcgtcaaa	7320
gtgatccaag cacactcata gttggagtcg tactccaaag gcggcaatga cgagtcagac	7380
agatactcgt cgacgtttta acaggttacg cagatctact atagaggaa atttaaatg	7440
ccccggagaa gacggccagg ccgcctagat gacaaattca acaactcaca gctgactttc	7500
tgccattgcc actagggggg ggctttttta tatggccaag ccaagctctc cacgtcggtt	7560
gggctgcacc caacaataaa tgggtagggt tgcaccaaca aagggatggg atggggggtg	7620
gaagatacga ggataacggg gctcaatggc acaataaga acgaatactg ccattaagac	7680

tcgtgatcca gcgactgaca ccattgcata atctaagggc ctcaaaacta cctcggaact 7740
 gctgcgctga tctggacacc acagaggttc cgagcacttt aggttgcacc aaatgtccca 7800
 ccaggtgcag gcagaaaacg ctggaacagc gtgtacagtt tgtcttaaca aaaagtgagg 7860
 gcgctgaggt cgagcagggg ggtgtgactt gttatagcct ttagagctgc gaaagcgcgt 7920
 atggatttgg ctcatcaggc cagattgagg gtctgtggac acatgtcatg ttagtgtact 7980
 tcaatcgccc cctggatata gccccgacaa taggccgtgg cctcatTTTT ttgccttcg 8040
 cacatttcca ttgctcgata cccacacctt gcttctcctg cacttgccaa ccttaatact 8100

ggtttacatt gaccaacatc ttacaagcgg ggggcttgtc taggttatat ataaacagtg 8160
 gtcttcccaa tcggttgcca gtctcttttt tctttcttt cccacagat tcgaaatcta 8220
 aactacacat cacagaattc cgagccgtga gtatccacga caagatcagt gtcgagacga 8280
 cgcgttttgt gtaatgacac aatccgaaag tcgctagcaa cacacactct ctacacaaac 8340
 taaccagct ctggtac 8357

<210> 340

<211> 32

<212> DNA

<213> Artificial Sequence

<220><223> Primer

<400> 340

ttggccattc aacacgaggc ttctcacgga gc 32

<210> 341

<211> 32

<212> DNA

<213> Artificial Sequence

<220><223> Primer

<400> 341

gctccgtgag aagcctcgtg ttgaatggcc aa 32

<210> 342

<211> 44

<212> DNA

<213> Artificial Sequence

<220><223> Primer

<400> 342

gattggcttg gccattcaac acgatttctc tcacggagcc atca 44

<210> 343
 <211> 44
 <212> DNA
 <213> Artificial Sequence
 <220><223> Primer
 <400> 343
 tgatggctcc gtgagagaaa tcgtgttgaa tggccaagcc aatc 44

<210> 344
 <211> 33
 <212> DNA
 <213> Artificial Sequence
 <220><223> Primer
 <400> 344
 gccattcaac acgatgggttc tcacggagcc atc 33

<210> 345
 <211> 33
 <212> DNA
 <213> Artificial Sequence
 <220><223> Primer
 <400> 345
 gatggctccg tgagaacat cgtgttgaat ggc 33

<210> 346
 <211> 44
 <212> DNA
 <213> Artificial Sequence
 <220><223> Primer
 <400> 346
 gattggcttg gccattcaac acgatatgtc tcacggagcc atca 44

<210> 347
 <211> 44
 <212> DNA
 <213> Artificial Sequence
 <220><223> Primer

<400> 347
 tgatggctcc gtgagacata tcgtgttgaa tggccaagcc aatc 44

<210> 348
 <211> 37
 <212> DNA
 <213> Artificial Sequence
 <220><223> Primer
 <400> 348
 ggcttggcca ttcaacacga ttctcttcac ggagcca 37

<210> 349
 <211> 37
 <212> DNA
 <213> Artificial Sequence
 <220><223> Primer
 <400> 349
 tggctccgtg agaggaatcg tgttgaatgg ccaagcc 37

<210> 350
 <211> 44
 <212> DNA
 <213> Artificial Sequence
 <220><223> Primer
 <400> 350
 gattggcttg gccattcaac acgattactc tcacggagcc atca 44

<210> 351
 <211> 44
 <212> DNA
 <213> Artificial Sequence
 <220><223> Primer
 <400> 351
 tgatggctcc gtgagagtaa tcgtgttgaa tggccaagcc aatc 44

<210> 352
 <211> 37
 <212> DNA
 <213> Artificial Sequence

<220><223> Primer
 <400> 352
 ccattcaaca cgatgctgcc cacggagcca tcagcaa 37

<210> 353
 <211> 37
 <212> DNA
 <213> Artificial Sequence
 <220><223> Primer
 <400> 353
 ttgctgatgg ctccgtgggc agcatcgtgt tgaatgg 37

<210> 354
 <211> 32
 <212> DNA
 <213> Artificial Sequence
 <220><223> Primer
 <400> 354
 ggccattcaa cacgatgctt gtcacggagc ca 32

<210> 355
 <211> 32
 <212> DNA
 <213> Artificial Sequence
 <220><223> Primer
 <400> 355
 tggctccgtg acaagcatcg tgttgaatgg cc 32

<210> 356
 <211> 36
 <212> DNA
 <213> Artificial Sequence
 <220><223> Primer
 <400> 356
 cttggccatt caacacgatg ctggtcacgg agccat 36

<210> 357
 <211> 36

<212> DNA
 <213> Artificial Sequence
 <220><223> Primer
 <400> 357
 atggctccgt gaccagcatc gtgttgaatg gccaag 36
 <210> 358
 <211> 45
 <212> DNA
 <213> Artificial Sequence
 <220><223> Primer
 <400> 358
 ttggccattc aacacgatgc taaccacgga gccatcagca aaaag 45
 <210> 359
 <211> 45
 <212> DNA
 <213> Artificial Sequence
 <220><223> Primer
 <400> 359
 ctttttgctg atggctccgt ggtagcatc gtgttgaatg gccaa 45
 <210> 360
 <211> 37
 <212> DNA
 <213> Artificial Sequence
 <220><223> Primer
 <400> 360
 ccattcaaca cgatgctacc cacggagcca tcagcaa 37
 <210> 361
 <211> 37
 <212> DNA
 <213> Artificial Sequence
 <220><223> Primer
 <400> 361
 ttgctgatgg ctccgtgggt agcatcgtgt tgaatgg 37

<210> 362

<211> 1365

<212> DNA

<213> Euglena anabaena

<220><221> CDS

<222> (1)..(1365)

<400> 362

atg gcc acc atc tcc ctg act acc gag cag ctc ctg gaa cac ccc gag 48

Met Ala Thr Ile Ser Leu Thr Thr Glu Gln Leu Leu Glu His Pro Glu

1 5 10 15

ctc gtt gcc atc gac gga gtc ctg tac gat ctc ttc ggt ctg gcc aag 96

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

20 25 30

gtc cat gcc gga ggc aac ctc atc gaa gct gcc ggt gca tcc gac gga 144

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

35 40 45

acc gct ctg ttc tac tcc atg cat cct gga gtc aag cca gag aac tcg 192

Thr Ala Leu Phe Tyr Ser Met His Pro Gly Val Lys Pro Glu Asn Ser

50 55 60

aag ctt ctg cag caa ttt gcc cga ggc aag cac gaa cga agc tcc aag 240

Lys Leu Leu Gln Gln Phe Ala Arg Gly Lys His Glu Arg Ser Ser Lys

65 70 75 80

gat ccc gtg tac acc ttc gac tct ccc ttt gct cag gac gtc aag cag 288

Asp Pro Val Tyr Thr Phe Asp Ser Pro Phe Ala Gln Asp Val Lys Gln

85 90 95

tcc gtt cga gag gtc atg aag ggt cga aac tgg tac gcc act cct ggc 336

Ser Val Arg Glu Val Met Lys Gly Arg Asn Trp Tyr Ala Thr Pro Gly

100 105 110

ttc tgg ctg aga acc gca ctc atc atc gct tgt act gcc att ggc gag 384

Phe Trp Leu Arg Thr Ala Leu Ile Ile Ala Cys Thr Ala Ile Gly Glu

115 120 125

tgg tac tgg atc aca acc gga gca gtg atg tgg ggt atc ttt act gga 432

Trp Tyr Trp Ile Thr Thr Gly Ala Val Met Trp Gly Ile Phe Thr Gly	
130	135 140
tac ttc cac tcg cag att ggc ttg gcc att caa cac gat tcc tct cac	480
Tyr Phe His Ser Gln Ile Gly Leu Ala Ile Gln His Asp Ser Ser His	
145	150 155 160
gga gcc atc agc aaa aag ccc tgg gtc aac gcc ttt ttc gct tat ggc	528
Gly Ala Ile Ser Lys Lys Pro Trp Val Asn Ala Phe Phe Ala Tyr Gly	
	165 170 175
atc gac gcc att ggt tcc tct cgt tgg atc tgg ctg cag tcc cac att	576
Ile Asp Ala Ile Gly Ser Ser Arg Trp Ile Trp Leu Gln Ser His Ile	
	180 185 190
atg cga cat cac act tac acc aac cag cat ggc ctc gac ctg gat gct	624
Met Arg His His Thr Tyr Thr Asn Gln His Gly Leu Asp Leu Asp Ala	
	195 200 205
gcc tcg gca gag ccg ttc atc ttg ttc cac tcc tat cct gct acc aac	672
Ala Ser Ala Glu Pro Phe Ile Leu Phe His Ser Tyr Pro Ala Thr Asn	
	210 215 220
gcc tct cga aag tgg tac cac cga ttt cag gcg tgg tac atg tac atc	720
Ala Ser Arg Lys Trp Tyr His Arg Phe Gln Ala Trp Tyr Met Tyr Ile	
225	230 235 240
gtt ctg gga atg tat ggt gtc tcg atg gtg tac aat ccc atg tac ctc	768
Val Leu Gly Met Tyr Gly Val Ser Met Val Tyr Asn Pro Met Tyr Leu	
	245 250 255
ttc aca atg cag cac aac gac acc att ccc gag gcc act tct ctc aga	816
Phe Thr Met Gln His Asn Asp Thr Ile Pro Glu Ala Thr Ser Leu Arg	
	260 265 270
cca ggc agc ttt ttc aat cgg cag cga gct ttc gcc gtt tcc ctt cga	864
Pro Gly Ser Phe Phe Asn Arg Gln Arg Ala Phe Ala Val Ser Leu Arg	
	275 280 285

ctg ctc ttc atc ttc cga aac gcc ttt ctt ccc tgg tac att gct ggt	912
Leu Leu Phe Ile Phe Arg Asn Ala Phe Leu Pro Trp Tyr Ile Ala Gly	
290 295 300	
gcc tct cct ctg ctc acc att ctt ctg gtg ccc acg gtc aca ggc atc	960
Ala Ser Pro Leu Leu Thr Ile Leu Leu Val Pro Thr Val Thr Gly Ile	
305 310 315 320	
ttc ctc acc ttt gtg ttc gtt ctg tcc cat aac ttc gag gga gcc gaa	1008
Phe Leu Thr Phe Val Phe Val Leu Ser His Asn Phe Glu Gly Ala Glu	
325 330 335	
cgg acc cca gag aag aac tgc aag gcc aaa cga gct aag gaa ggc aag	1056
Arg Thr Pro Glu Lys Asn Cys Lys Ala Lys Arg Ala Lys Glu Gly Lys	
340 345 350	
gag gtc aga gac gtg gaa gag gat cga gtc gac tgg tac cga gca cag	1104
Glu Val Arg Asp Val Glu Glu Asp Arg Val Asp Trp Tyr Arg Ala Gln	
355 360 365	
gcc gag act gct gcc acc tac ggt ggc agc gtg gga atg atg ctt aca	1152
Ala Glu Thr Ala Ala Thr Tyr Gly Gly Ser Val Gly Met Met Leu Thr	
370 375 380	
ggc ggt ctc aac ctg cag atc gag cat cac ttg ttt ccc cga atg tcc	1200
Gly Gly Leu Asn Leu Gln Ile Glu His His Leu Phe Pro Arg Met Ser	
385 390 395 400	
tct tgg cac tat ccc ttc att caa gac acc gtt cgg gag tgt tgc aag	1248
Ser Trp His Tyr Pro Phe Ile Gln Asp Thr Val Arg Glu Cys Cys Lys	
405 410 415	
cga cat ggc gtc cgt tac aca tac tat cct acc att ctc gag aac atc	1296
Arg His Gly Val Arg Tyr Thr Tyr Tyr Pro Thr Ile Leu Glu Asn Ile	
420 425 430	
atg tcc act ctt cga tac atg cag aag gtg ggt gtt gct cac acc att	1344
Met Ser Thr Leu Arg Tyr Met Gln Lys Val Gly Val Ala His Thr Ile	

435 440 445

cag gat gcc cag gag ttc taa 1365

Gln Asp Ala Gln Glu Phe

450

<210> 363

<211> 454

<212> PRT

<213> Euglena anabaena

<400> 363

Met Ala Thr Ile Ser Leu Thr Thr Glu Gln Leu Leu Glu His Pro Glu

1 5 10 15

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

20 25 30

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

35 40 45

Thr Ala Leu Phe Tyr Ser Met His Pro Gly Val Lys Pro Glu Asn Ser

50 55 60

Lys Leu Leu Gln Gln Phe Ala Arg Gly Lys His Glu Arg Ser Ser Lys

65 70 75 80

Asp Pro Val Tyr Thr Phe Asp Ser Pro Phe Ala Gln Asp Val Lys Gln

85 90 95

Ser Val Arg Glu Val Met Lys Gly Arg Asn Trp Tyr Ala Thr Pro Gly

100 105 110

Phe Trp Leu Arg Thr Ala Leu Ile Ile Ala Cys Thr Ala Ile Gly Glu

115 120 125

Trp Tyr Trp Ile Thr Thr Gly Ala Val Met Trp Gly Ile Phe Thr Gly

130 135 140

Tyr Phe His Ser Gln Ile Gly Leu Ala Ile Gln His Asp Ser Ser His

145 150 155 160

Gly Ala Ile Ser Lys Lys Pro Trp Val Asn Ala Phe Phe Ala Tyr Gly

165 170 175

Ile Asp Ala Ile Gly Ser Ser Arg Trp Ile Trp Leu Gln Ser His Ile

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

Met Ser Thr Leu Arg Tyr Met Gln Lys Val Gly Val Ala His Thr Ile

435 440 445

Gln Asp Ala Gln Glu Phe

450

<210> 364

<211> 1365

<212> DNA

<213> Euglena anabaena

<220><221> CDS

<222> (1)..(1365)

<400> 364

atg gcc acc atc tcc ctg act acc gag cag ctc ctg gaa cac ccc gag 48

Met Ala Thr Ile Ser Leu Thr Thr Glu Gln Leu Leu Glu His Pro Glu

1 5 10 15

ctc gtt gcc atc gac gga gtc ctg tac gat ctc ttc ggt ctg gcc aag 96

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

20 25 30

gtc cat gcc gga ggc aac ctc atc gaa gct gcc ggt gca tcc gac gga 144

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

35 40 45

acc gct ctg ttc tac tcc atg cat cct gga gtc aag cca gag aac tcg 192

Thr Ala Leu Phe Tyr Ser Met His Pro Gly Val Lys Pro Glu Asn Ser

50 55 60

aag ctt ctg cag caa ttt gcc cga ggc aag cac gaa cga agc tcc aag 240

Lys Leu Leu Gln Gln Phe Ala Arg Gly Lys His Glu Arg Ser Ser Lys

65 70 75 80

gat ccc gtg tac acc ttc gac tct ccc ttt gct cag gac gtc aag cag 288

Asp Pro Val Tyr Thr Phe Asp Ser Pro Phe Ala Gln Asp Val Lys Gln

85 90 95

tcc gtt cga gag gtc atg aag ggt cga aac tgg tac gcc act cct ggc 336

Ser Val Arg Glu Val Met Lys Gly Arg Asn Trp Tyr Ala Thr Pro Gly

100	105	110	
ttc tgg ctg aga acc gca ctc atc atc gct tgt act gcc att ggc gag			384
Phe Trp Leu Arg Thr Ala Leu Ile Ile Ala Cys Thr Ala Ile Gly Glu			
115	120	125	
tgg tac tgg atc aca acc gga gca gtg atg tgg ggt atc ttt act gga			432
Trp Tyr Trp Ile Thr Thr Gly Ala Val Met Trp Gly Ile Phe Thr Gly			
130	135	140	
tac ttc cac tcg cag att ggc ttg gcc att caa cac gat gct ggt cac			480
Tyr Phe His Ser Gln Ile Gly Leu Ala Ile Gln His Asp Ala Gly His			
145	150	155	160
gga gcc atc agc aaa aag ccc tgg gtc aac gcc ttt ttc gct tat ggc			528
Gly Ala Ile Ser Lys Lys Pro Trp Val Asn Ala Phe Phe Ala Tyr Gly			
165	170	175	
atc gac gcc att ggt tcc tct cgt tgg atc tgg ctg cag tcc cac att			576
Ile Asp Ala Ile Gly Ser Ser Arg Trp Ile Trp Leu Gln Ser His Ile			
180	185	190	
atg cga cat cac act tac acc aac cag cat ggc ctc gac ctg gat gct			624
Met Arg His His Thr Tyr Thr Asn Gln His Gly Leu Asp Leu Asp Ala			
195	200	205	
gcc tcg gca gag ccg ttc atc ttg ttc cac tcc tat cct gct acc aac			672
Ala Ser Ala Glu Pro Phe Ile Leu Phe His Ser Tyr Pro Ala Thr Asn			
210	215	220	
gcc tct cga aag tgg tac cac cga ttt cag gcg tgg tac atg tac atc			720
Ala Ser Arg Lys Trp Tyr His Arg Phe Gln Ala Trp Tyr Met Tyr Ile			
225	230	235	240
gtt ctg gga atg tat ggt gtc tcg atg gtg tac aat ccc atg tac ctc			768
Val Leu Gly Met Tyr Gly Val Ser Met Val Tyr Asn Pro Met Tyr Leu			
245	250	255	
ttc aca atg cag cac aac gac acc att ccc gag gcc act tct ctc aga			816
Phe Thr Met Gln His Asn Asp Thr Ile Pro Glu Ala Thr Ser Leu Arg			
260	265	270	

cca ggc agc ttt ttc aat cgg cag cga gct ttc gcc gtt tcc ctt cga	864
Pro Gly Ser Phe Phe Asn Arg Gln Arg Ala Phe Ala Val Ser Leu Arg	
275 280 285	
ctg ctc ttc atc ttc cga aac gcc ttt ctt ccc tgg tac att gct ggt	912
Leu Leu Phe Ile Phe Arg Asn Ala Phe Leu Pro Trp Tyr Ile Ala Gly	
290 295 300	
gcc tct cct ctg ctc acc att ctt ctg gtg ccc acg gtc aca ggc atc	960
Ala Ser Pro Leu Leu Thr Ile Leu Leu Val Pro Thr Val Thr Gly Ile	
305 310 315 320	
ttc ctc acc ttt gtg ttc gtt ctg tcc cat aac ttc gag gga gcc gaa	1008
Phe Leu Thr Phe Val Phe Val Leu Ser His Asn Phe Glu Gly Ala Glu	
325 330 335	
cgg acc cca gag aag aac tgc aag gcc aaa cga gct aag gaa ggc aag	1056
Arg Thr Pro Glu Lys Asn Cys Lys Ala Lys Arg Ala Lys Glu Gly Lys	
340 345 350	
gag gtc aga gac gtg gaa gag gat cga gtc gac tgg tac cga gca cag	1104
Glu Val Arg Asp Val Glu Glu Asp Arg Val Asp Trp Tyr Arg Ala Gln	
355 360 365	
gcc gag act gct gcc acc tac ggt ggc agc gtg gga atg atg ctt aca	1152
Ala Glu Thr Ala Ala Thr Tyr Gly Gly Ser Val Gly Met Met Leu Thr	
370 375 380	
ggc ggt ctc aac ctg cag atc gag cat cac ttg ttt ccc cga atg tcc	1200
Gly Gly Leu Asn Leu Gln Ile Glu His His Leu Phe Pro Arg Met Ser	
385 390 395 400	
tct tgg cac tat ccc ttc att caa gac acc gtt cgg gag tgt tgc aag	1248
Ser Trp His Tyr Pro Phe Ile Gln Asp Thr Val Arg Glu Cys Cys Lys	
405 410 415	
cga cat ggc gtc cgt tac aca tac tat cct acc att ctc gag aac atc	1296
Arg His Gly Val Arg Tyr Thr Tyr Tyr Pro Thr Ile Leu Glu Asn Ile	
420 425 430	
atg tcc act ctt cga tac atg cag aag gtg ggt gtt gct cac acc att	1344

Met Ser Thr Leu Arg Tyr Met Gln Lys Val Gly Val Ala His Thr Ile

435 440 445

cag gat gcc cag gag ttc taa 1365

Gln Asp Ala Gln Glu Phe

450

<210> 365

<211> 454

<212> PRT

<213> Euglena anabaena

<400> 365

Met Ala Thr Ile Ser Leu Thr Thr Glu Gln Leu Leu Glu His Pro Glu

1 5 10 15

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

20 25 30

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

35 40 45

Thr Ala Leu Phe Tyr Ser Met His Pro Gly Val Lys Pro Glu Asn Ser

50 55 60

Lys Leu Leu Gln Gln Phe Ala Arg Gly Lys His Glu Arg Ser Ser Lys

65 70 75 80

Asp Pro Val Tyr Thr Phe Asp Ser Pro Phe Ala Gln Asp Val Lys Gln

85 90 95

Ser Val Arg Glu Val Met Lys Gly Arg Asn Trp Tyr Ala Thr Pro Gly

100 105 110

Phe Trp Leu Arg Thr Ala Leu Ile Ile Ala Cys Thr Ala Ile Gly Glu

115 120 125

Trp Tyr Trp Ile Thr Thr Gly Ala Val Met Trp Gly Ile Phe Thr Gly

130 135 140

Tyr Phe His Ser Gln Ile Gly Leu Ala Ile Gln His Asp Ala Gly His

145 150 155 160

Gly Ala Ile Ser Lys Lys Pro Trp Val Asn Ala Phe Phe Ala Tyr Gly

165	170	175
Ile Asp Ala Ile Gly Ser Ser Arg Trp Ile Trp Leu Gln Ser His Ile		
180	185	190
Met Arg His His Thr Tyr Thr Asn Gln His Gly Leu Asp Leu Asp Ala		
195	200	205
Ala Ser Ala Glu Pro Phe Ile Leu Phe His Ser Tyr Pro Ala Thr Asn		
210	215	220
Ala Ser Arg Lys Trp Tyr His Arg Phe Gln Ala Trp Tyr Met Tyr Ile		
225	230	235
Val Leu Gly Met Tyr Gly Val Ser Met Val Tyr Asn Pro Met Tyr Leu		
245	250	255
Phe Thr Met Gln His Asn Asp Thr Ile Pro Glu Ala Thr Ser Leu Arg		
260	265	270
Pro Gly Ser Phe Phe Asn Arg Gln Arg Ala Phe Ala Val Ser Leu Arg		
275	280	285
Leu Leu Phe Ile Phe Arg Asn Ala Phe Leu Pro Trp Tyr Ile Ala Gly		
290	295	300
Ala Ser Pro Leu Leu Thr Ile Leu Leu Val Pro Thr Val Thr Gly Ile		
305	310	315
Phe Leu Thr Phe Val Phe Val Leu Ser His Asn Phe Glu Gly Ala Glu		
325	330	335
Arg Thr Pro Glu Lys Asn Cys Lys Ala Lys Arg Ala Lys Glu Gly Lys		
340	345	350
Glu Val Arg Asp Val Glu Glu Asp Arg Val Asp Trp Tyr Arg Ala Gln		
355	360	365
Ala Glu Thr Ala Ala Thr Tyr Gly Gly Ser Val Gly Met Met Leu Thr		
370	375	380
Gly Gly Leu Asn Leu Gln Ile Glu His His Leu Phe Pro Arg Met Ser		
385	390	395
Ser Trp His Tyr Pro Phe Ile Gln Asp Thr Val Arg Glu Cys Cys Lys		
405	410	415

Arg His Gly Val Arg Tyr Thr Tyr Tyr Pro Thr Ile Leu Glu Asn Ile
420 425 430

Met Ser Thr Leu Arg Tyr Met Gln Lys Val Gly Val Ala His Thr Ile
435 440 445

Gln Asp Ala Gln Glu Phe
450

<210> 366

<211> 25

<212> DNA

<213> Artificial Sequence

<220><223> Primer FBAIN-F

<400> 366

caagcggggg gcttgtctag ggtat 25

<210> 367

<211> 40

<212> DNA

<213> Artificial Sequence

<220><223> Primer Y1026

<400> 367

tttgcggccg ctcaactgaat ctttttggtt cccttgtgct 40

<210> 368

<211> 777

<212> DNA

<213> Euglena gracilis

<220><221> CDS

<222> (1)..(777)

<223> mutant delta-9 elongase "EgD9eS-L35G/W132T/I179R"

<400> 368

atg gag gtc gtg aac gaa atc gtc tcc att ggc cag gag gtt ctt ccc 48

Met Glu Val Val Asn Glu Ile Val Ser Ile Gly Gln Glu Val Leu Pro
1 5 10 15

aag gtc gac tat gct cag ctc tgg tct gat gcc tcg cac tgc gag gtg 96

Lys Val Asp Tyr Ala Gln Leu Trp Ser Asp Ala Ser His Cys Glu Val

20	25	30	
ctg tac ggg tcc atc gcc ttc gtc atc ctg aag ttc acc ctt ggt cct			144
Leu Tyr Gly Ser Ile Ala Phe Val Ile Leu Lys Phe Thr Leu Gly Pro			
35	40	45	
ctc gga ccc aag ggt cag tct cga atg aag ttt gtg ttc acc aac tac			192
Leu Gly Pro Lys Gly Gln Ser Arg Met Lys Phe Val Phe Thr Asn Tyr			
50	55	60	
aac ctg ctc atg tcc atc tac tcg ctg ggc tcc ttc ctc tct atg gcc			240
Asn Leu Leu Met Ser Ile Tyr Ser Leu Gly Ser Phe Leu Ser Met Ala			
65	70	75	80
tac gcc atg tac acc att ggt gtc atg tcc gac aac tgc gag aag gct			288
Tyr Ala Met Tyr Thr Ile Gly Val Met Ser Asp Asn Cys Glu Lys Ala			
85	90	95	
ttc gac aac aat gtc ttc cga atc acc act cag ctg ttc tac ctc agc			336
Phe Asp Asn Asn Val Phe Arg Ile Thr Thr Gln Leu Phe Tyr Leu Ser			
100	105	110	
aag ttc ctc gag tac att gac tcc ttc tat ctg ccc ctc atg ggc aag			384
Lys Phe Leu Glu Tyr Ile Asp Ser Phe Tyr Leu Pro Leu Met Gly Lys			
115	120	125	
cct ctg acc acc ttg cag ttc ttt cac cat ctc gga gct cct atg gac			432
Pro Leu Thr Thr Leu Gln Phe Phe His His Leu Gly Ala Pro Met Asp			
130	135	140	
atg tgg ctg ttc tac aac tac cga aac gaa gcc gtt tgg atc ttt gtg			480
Met Trp Leu Phe Tyr Asn Tyr Arg Asn Glu Ala Val Trp Ile Phe Val			
145	150	155	160
ctg ctc aac ggc ttc att cac tgg atc atg tac ggc tac tat tgg acc			528
Leu Leu Asn Gly Phe Ile His Trp Ile Met Tyr Gly Tyr Tyr Trp Thr			
165	170	175	
cga ctg cga aag ctc aag ttc cct atg ccc aag tcc ctg att act tct			576
Arg Leu Arg Lys Leu Lys Phe Pro Met Pro Lys Ser Leu Ile Thr Ser			

180	185	190	
atg cag atc att cag ttc aac gtt ggc ttc tac atc gtc tgg aag tac			624
Met Gln Ile Ile Gln Phe Asn Val Gly Phe Tyr Ile Val Trp Lys Tyr			
195	200	205	
cgg aac att ccc tgc tac cga caa gat gga atg aga atg ttt ggc tgg			672
Arg Asn Ile Pro Cys Tyr Arg Gln Asp Gly Met Arg Met Phe Gly Trp			
210	215	220	
ttt ttc aac tac ttc tac gtt ggt act gtc ctg tgt ctg ttc ctc aac			720
Phe Phe Asn Tyr Phe Tyr Val Gly Thr Val Leu Cys Leu Phe Leu Asn			
225	230	235	240
ttc tac gtg cag acc tac atc gtc cga aag cac aag gga gcc aaa aag			768
Phe Tyr Val Gln Thr Tyr Ile Val Arg Lys His Lys Gly Ala Lys Lys			
245	250	255	
att cag tga			777
Ile Gln			

<210> 369

<211> 258

<212> PRT

<213> Euglena gracilis

<400> 369

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

Tyr Ala Met Tyr Thr Ile Gly Val Met Ser Asp Asn Cys Glu Lys Ala
85 90 95

Phe Asp Asn Asn Val Phe Arg Ile Thr Thr Gln Leu Phe Tyr Leu Ser
100 105 110

Lys Phe Leu Glu Tyr Ile Asp Ser Phe Tyr Leu Pro Leu Met Gly Lys
115 120 125

Pro Leu Thr Thr Leu Gln Phe Phe His His Leu Gly Ala Pro Met Asp
130 135 140

Met Trp Leu Phe Tyr Asn Tyr Arg Asn Glu Ala Val Trp Ile Phe Val
145 150 155 160

Leu Leu Asn Gly Phe Ile His Trp Ile Met Tyr Gly Tyr Tyr Trp Thr
165 170 175

Arg Leu Arg Lys Leu Lys Phe Pro Met Pro Lys Ser Leu Ile Thr Ser
180 185 190

Met Gln Ile Ile Gln Phe Asn Val Gly Phe Tyr Ile Val Trp Lys Tyr
195 200 205

Arg Asn Ile Pro Cys Tyr Arg Gln Asp Gly Met Arg Met Phe Gly Trp
210 215 220

Phe Phe Asn Tyr Phe Tyr Val Gly Thr Val Leu Cys Leu Phe Leu Asn
225 230 235 240

Phe Tyr Val Gln Thr Tyr Ile Val Arg Lys His Lys Gly Ala Lys Lys
245 250 255

Ile Gln

<210> 370

<211> 7769

<212> DNA

<213> Artificial Sequence

<220><223> Plasmid EgD9eS-L35G/W132T/I179R

<400> 370

ggccgcaagt gtggatgggg aagtgagtgc ccggttctgt gtgcacaatt ggcaatccaa 60

gatggatgga ttcaacacag ggatatagcg agctacgtgg tggtgcgagg atatagcaac 120

ggatatttat gtttgacact tgagaatgia cgatacaagc actgtccaag tacaatacta 180

aacatactgt acatactcat actcgtaccc gggcaacggt ttcacttgag tgcagtggct 240

agtgccttta ctctacagt gtgcaatact gcgtatcata gtctttgatg tatatcgat 300

tcattcatgt tagttgcgta cgagccggaa gcataaagtg taaagcctgg ggtgcctaata 360

gagttagcta actcacatta attgcgttgc gctcactgcc cgctttccag tcgggaaacc 420

tgctgtgcca gctgcattaa tgaatcggcc aacgcgcggg gagaggcgggt ttgcgtattg 480

ggcgctcttc cgcttcctcg ctactgact cgctgcgctc ggtcgttcgg ctgcggcgag 540

cggtatcagc tcaactcaaag gcggaatac ggttatccac agaatacagg gataacgcag 600

gaaagaacat gtgagcaaaa ggccagcaaa aggccaggaa ccgtaaaaag gccgcgttgc 660

tggcggtttt ccataggctc cgccccctg acgagcatca caaaaatcga cgctcaagtc 720

agaggtagcg aaacccgaca ggactataaa gataccaggc gtttccccct ggaagctccc 780

tcgtgcgctc tctgtttccg accctgccgc ttaccggata cctgtccgcc tttctccctt 840

cgggaagcgt ggcgctttct catagctcac gctgtaggta tctcagttcg gtgtaggtcg 900

ttcgctccaa gctgggctgt gtgcacgaac ccccggttca gcccgaccgc tgcgccttat 960

ccggttaacta tcgtcttgag tccaacccgg taagacacga cttatcgcca ctggcagcag 1020

ccactggtaa caggattagc agagcgaggt atgtaggcgg tgctacagag ttcttgaagt 1080

ggtggcctaa ctacggctac actagaagga cagtatttgg tatctgcgt ctgctgaagc 1140

cagttacctt cgaaaaaga gttggtagct ctgatccgg caaacaacc accgctggtta 1200

gcggtgggtt ttttgtttgc aagcagcaga ttacgcgcag aaaaaagga tctcaagaag 1260

atcctttgat cttttctac gggctctgacg ctacgtggaa cgaaaactca cgtaaggga 1320

ttttggtcat gagattatca aaaaggatct tcacctagat ctttttaaat taaaaatgaa 1380

gttttaaatc aatctaaagt atatatgagt aaacttggtc tgacagttac caatgcttaa 1440

tcagtgaggc acctatctca gcgatctgtc tatttcgttc atccatagtt gcctgactcc 1500

ccgtcgtgta gataactacg atacgggagg gcttaccatc tggccccagt gctgcaatga 1560

taccgcgaga cccacgtca ccggctccag atttatcagc aataaaccag ccagccggaa 1620

gggccgagcg cagaagtggc cctgcaactt tatccgctc catccagtct attaattgtt 1680

gccgggaagc tagagtaagt agttcgccag ttaatatgtt gcgcaacgtt gttgccattg 1740

ctacaggcat cgtgggtgca cgctcgctgt ttggtatggc ttcatcagc tccggttccc 1800

aacgatcaag gcgagttaca tgatcccca tgttgtgcaa aaaagcgggt agctccttcg 1860

gtcctccgat cgttgtcaga agtaagttgg ccgcagtggt atcactcatg gttatggcag 1920
 cactgcataa ttctcttact gtcatgccat ccgtaagatg cttttctgtg actggtgagt 1980
 actcaaccaa gtcattctga gaatagtgtg tgcggcgacc gagttgctct tgccccggcgt 2040
 caatacggga taataaccgc ccacatagca gaactttaaa agtgctcatc attggaaaac 2100
 gttcttcggg gcgaaaactc tcaaggatct taccgctgtt gagatccagt tcgatgtaac 2160
 ccaactcgtgc acccaactga tcttcagcat cttttacttt caccagcgtt tctgggtgag 2220
 caaaaacagg aaggcaaaat gccgcaaaaa agggaataag ggcgacacgg aaatgttgaa 2280

tactcatact cttccttttt caatattatt gaagcattta tcagggttat tgtctcatga 2340
 gcggatacat atttgaatgt atttagaaaa ataaacaaat aggggttccg cgcacatttc 2400
 cccgaaaagt gccacctgac gcgccctgtg gcggcgcatc aagcgcgcg ggtgtggtgg 2460
 ttacgcgcag cgtgaccgt acacttgcca gcgccctagc gcccgctcct ttcgctttct 2520
 tcccttcctt tctcgccacg ttccgctgtt ttccccgtca agctctaaat cgggggctcc 2580
 ctttaggggt cegatttagt gctttacggc acctcgacc caaaaaactt gattaggggtg 2640
 atgggtcacg tagtgggcca tcgccctgat agacgggttt tcgccctttg acgttggtg 2700

ccacgttctt taatagtga ctcttggtcc aaactggaac aacactcaac cctatctcgg 2760
 tctattcttt tgatttataa gggattttgc cgatttcggc ctattggtta aaaaatgagc 2820
 tgatttaaca aaaatttaac gcgaatttta acaaaatatt aacgcttaca atttcattc 2880
 gccattcagg ctgcgcaact gttgggaagg gcgatcgggt cgggcctctt cgctattacg 2940
 ccagctggcg aaagggggat gtgctgcaag gcgattaagt tgggtaacgc cagggttttc 3000
 ccagtcacga cgttgtaaaa cgacggccag tgaattgtaa tacgactcac tataggcgca 3060
 attgggtacc gggccccccc tcgaggtcga tgggtgtgat aagcttgata tcgaattcat 3120

gtcacacaaa ccgattctcg cctcaaggaa acctaatctt acatccgaga gactgccgag 3180
 atccagtcta cactgattaa ttttcgggcc aataatttaa aaaaatcgtg ttataataa 3240
 ttatatgtat tatatatata catcatgatg atactgacag tcattgtcca ttgctaaata 3300
 gacagactcc atctgccgc tccaactgat gttctcaata ttttaagggt catctcgcat 3360
 tgtttaataa taaacagact ccatctaccg cctccaaatg atgttctcaa aatataattg 3420
 atgaacttat ttttattact tagtattatt agacaactta cttgctttat gaaaaacact 3480
 tcctatttag gaaacaattt ataattggcag ttctgtcatt taacaattta tgtagaataa 3540

atgttataaa tgcgtatggg aaactctaaa tatggatagc ataaatgata tctgcattgc 3600
 ctaattcgaa atcaacagca acgaaaaaaa tcccttgtac aacataaata gtcacgaga 3660
 aatatcaact atcaaagaac agctattcac acgttactat tgagattatt attggacgag 3720

aatcacacac tcaactgtct ttctctcttc tagaaatata ggtacaagta tgtactattc	3780
tcattgttca tacttttagt catttcatcc cacatatcc ttggatttct ctccaatgaa	3840
tgacattcta tcttgcaaat tcaacaatta taataagata taccaaagta gcggtatagt	3900
ggcaatcaaa aagcttctct ggtgtgcttc tcgtatttat ttttattcta atgatccatt	3960
aaaggtatat atttatttct tgttatataa tccttttgtt tattacatgg gctggataca	4020
taaaggtatt ttgatttaat tttttgctta aattcaatcc cccctcgctc agtgtcaact	4080
glaatggtag gaaattacca tacttttgaa gaagcaaaaa aaatgaaaga aaaaaaaaaat	4140
cgtatttcca ggtagacgt tccgcagaat ctagaatgcg gtatgcggta cattgttctt	4200
cgaacgtaaa agttgcgctc cctgagatat tglacatttt tgcttttaca agtacaagta	4260
catcgtacaa ctatgtacta ctgttgatgc atccacaaca gtttgttttg ttttttttg	4320
tttttttttt ttctaattgat tcattaccgc tatgtatacc tacttgtact thtagtaagc	4380
cgggttattg gcgttcaatt aatcatagac ttatgaatct gcacggtgtg cgctgcgagt	4440
tacttttagc ttatgcatgc tacttgggtg taatatggg atctgttcgg aaatcaacgg	4500
atgtcaate gatttcgaca gtaattaatt aagtcataca caagtcagct ttcttcgagc	4560
ctcatataag tataagtagt tcaacgtatt agcactgtac ccagcatctc cgtatcgaga	4620
aacacaacaa catgccccat tggacagatc atgcggatac acaggttgtg cagtatcata	4680
catactgat cagacaggtc gtctgacat catacaagct gaacaagcgc tccatacttg	4740
cacgtctct atatacacag ttaaattaca tatccatagt ctaacctcta acagttaatc	4800
ttctggttaag cctcccagcc agccttctgg tctcgcttgg cctcctcaat aggatctcgg	4860
ttctggccgt acagacctcg gccgacaatt atgatatccg ttccggtaga catgacatcc	4920
tcaacagttc ggtactgtg tccgagagcg tctcccttgt cgtcaagacc cccccgggg	4980
gtcagaataa gccagtctc agagtcgcc ttaggtcggt tctgggcaat gaagccaacc	5040
acaaactcgg ggtcggatcg ggcaagctca atggtctgct tggagtactc gccagtggcc	5100
agagagccct tgcaagacag ctcgccagc atgagcagac ctctggccag ctctctgttg	5160
ggagagggga ctaggaactc cttgtactgg gagttctcgt agtcagagac gtctccttc	5220
ttctgttcag agacagtttc ctcgccacca gctcgaggc cagcaatgat tccggttccg	5280
ggtacaccgt gggcgttggg gatatcggac cactcgcgga ttcggtgaca ccggtactgg	5340
tgcttgacag tgttgccaat atctcgcaac ttctgtcct cgaacaggaa gaaaccgtgc	5400
ttaagagcaa gttccttgag ggggagcaca gtgccgcgt aggtgaagtc gtcaatgatg	5460
tcgatatggg ttttgatcat gcacacataa ggtccgacct tatcggaag ctcaatgagc	5520
tccttgggtg tggtaacatc cagagaagca cacaggttgg ttttcttggc tgccacgagc	5580

ttgagcactc gagcggcaaa ggcggacttg tggacgttag ctcgagcttc gtaggagggc	5640
atTTTggtgg tgaagaggag actgaaataa atttagtctg cagaactttt tatcggaacc	5700
ttatctgggg cagtgaagta tatgttatgg taatagttag gagttagttag aacttataga	5760
tagactggac tatacggcta tcggtccaaa ttgaaagaa cgtcaatggc tctctgggcg	5820
tcgcctttgc cgacaaaaat gtgatcatga tgaaagccag caatgacgtt gcagctgata	5880
ttgttgcgg ccaaccgcgc cgaaaacgca gctgtcagac ccacagcctc caacgaagaa	5940
tgtatcgtca aagtgatcca agcacactca tagttggagt cgtactcca aggcggcaat	6000
gacgagtcag acagatactc gtcgacgttt aaacagtgtg cgcagatcta ctatagagga	6060
acatttaaat tgccccggag aagacggcca ggccgcctag atgacaaatt caacaactca	6120
cagctgactt tctgccattg ccactagggg ggggcctttt tataTggcca agccaagctc	6180
tccacgtcgg ttgggctgca cccaacaata aatgggtagg gttgcaccaa caaagggatg	6240
ggatgggggg tagaagatag gaggataacg gggctcaatg gcacaaataa gaacgaatac	6300
tgccattaag actcgtgac cagcgactga caccattgca tcatctaagg gcctcaaaac	6360
tacctcggaa ctgctgcgt gatctggaca ccacagaggt tccgagcact ttaggttgca	6420
ccaaatgtcc caccaggtgc aggcagaaaa cgctggaaca gcgtgtacag tttgtcttaa	6480
caaaaagtga gggcgtgag gtcgagcagg gtggtgtgac ttgttatagc ctttagagct	6540
gcgaaagcgc gtatggattt ggctcatcag gccagattga gggctctgtg acacatgtca	6600
tgttagtgta cttaacgc cccctggata tagccccgac aataggccgt ggcctcattt	6660
ttttgccttc cgcacatttc cattgctcga taccacacc ttgcttctcc tgcacttgcc	6720
aaccttaata ctggtttaca ttgaccaaca tcttacaagc ggggggcttg tctagggtat	6780
atataaacag tggtctccc aatcggttgc cagtctcttt tttctttct tccccacag	6840
attcgaatc taaactacac atcacagaat tccgagccgt gagtatccac gacaagatca	6900
gtgtcgagac gacgcgtttt gtgtaatgac acaatccgaa agtcgctagc aacacacact	6960
ctctacaaa actaaccag ctctggtagc atggaggtcg tgaacgaaat cgtctccatt	7020
ggccaggagg ttcttcccaa ggtcgactat gctcagctct ggtctgatgc ctgcactgc	7080
gaggtgctgt acgggtccat cgccttcgtc atcctgaagt tcacccttgg tcctctcgga	7140
cccaagggtc agtctcgaat gaagtttgtg ttaccaact acaacctgct catgtccatc	7200
tactcgtgg gctccttct ctctatggcc tacgcatgt acaccattgg tgtcatgtcc	7260
gacaactcgc agaaggcttt cgacaacaat gtcttccgaa tcaccactca gctgttctac	7320

ctcagcaagt tctcagagta cattgactcc ttctatctgc cccatcatggg caagcctctg 7380
accaccttgc agttctttca ccattctcgga gctcctatgg acatgtggct gttctacaac 7440
taccgaaacg aagccgtttg gatctttgtg ctgctcaacg gcttcattca ctggatcatg 7500
tacggctact attggaccgg actgcgaaaag ctcaagttcc ctatgcccaa gtccctgatt 7560
acttctatgc agatcattca gttcaacgtt ggcttctaca tcgtctggaa gtaccggaac 7620
attccctgct accgacaaga tggaatgaga atgtttggct ggtttttcaa ctactttac 7680
gttggtagtg tctgtgtct gttcctcaac ttctacgtgc agacctacat cgtccgaaag 7740

cacaagggag ccaaaaagat tcagtggc 7769

<210> 371

<211> 777

<212> DNA

<213> *Euglena gracilis*

<220><221> CDS

<222> (1)..(777)

<223> mutant delta-9 elongase "EgD9eS-S9D/L35G/Y84C/I179R"

<400> 371

atg gag gtc gtc aac gaa atc gtc gac att ggc cag gag gtt ctt ccc 48
Met Glu Val Val Asn Glu Ile Val Asp Ile Gly Gln Glu Val Leu Pro
1 5 10 15
aag gtc gac tat gct cag ctc tgg tct gat gcc tgc cac tgc gag gtc 96

Lys Val Asp Tyr Ala Gln Leu Trp Ser Asp Ala Ser His Cys Glu Val
20 25 30

ctg tac ggg tcc atc gcc ttc gtc atc ctg aag ttc acc ctt ggt cct 144
Leu Tyr Gly Ser Ile Ala Phe Val Ile Leu Lys Phe Thr Leu Gly Pro
35 40 45

ctc gga ccc aag ggt cag tct cga atg aag ttt gtg ttc acc aac tac 192
Leu Gly Pro Lys Gly Gln Ser Arg Met Lys Phe Val Phe Thr Asn Tyr

50 55 60

aac ctg ctc atg tcc atc tac tgc ctg ggc tcc ttc ctc tct atg gcc 240
Asn Leu Leu Met Ser Ile Tyr Ser Leu Gly Ser Phe Leu Ser Met Ala
65 70 75 80
tac gcc atg tgt acc att ggt gtc atg tcc gac aac tgc gag aag gct 288

Tyr Ala Met Cys Thr Ile Gly Val Met Ser Asp Asn Cys Glu Lys Ala	
85 90 95	
ttc gac aac aat gtc ttc cga atc acc act cag ctg ttc tac ctc agc	336
Phe Asp Asn Asn Val Phe Arg Ile Thr Thr Gln Leu Phe Tyr Leu Ser	
100 105 110	
aag ttc ctc gag tac att gac tcc ttc tat ctg ccc ctc atg ggc aag	384
Lys Phe Leu Glu Tyr Ile Asp Ser Phe Tyr Leu Pro Leu Met Gly Lys	
115 120 125	
cct ctg acc tgg ttg cag ttc ttt cac cat ctc gga gct cct atg gac	432
Pro Leu Thr Trp Leu Gln Phe Phe His His Leu Gly Ala Pro Met Asp	
130 135 140	
atg tgg ctg ttc tac aac tac cga aac gaa gcc gtt tgg atc ttt gtg	480
Met Trp Leu Phe Tyr Asn Tyr Arg Asn Glu Ala Val Trp Ile Phe Val	
145 150 155 160	
ctg ctc aac ggc ttc att cac tgg atc atg tac ggc tac tat tgg acc	528
Leu Leu Asn Gly Phe Ile His Trp Ile Met Tyr Gly Tyr Tyr Trp Thr	
165 170 175	
cga ctg cga aag ctc aag ttc cct atg ccc aag tcc ctg att act tct	576
Arg Leu Arg Lys Leu Lys Phe Pro Met Pro Lys Ser Leu Ile Thr Ser	
180 185 190	
atg cag atc att cag ttc aac gtt ggc ttc tac atc gtc tgg aag tac	624
Met Gln Ile Ile Gln Phe Asn Val Gly Phe Tyr Ile Val Trp Lys Tyr	
195 200 205	
cgg aac att ccc tgc tac cga caa gat gga atg aga atg ttt ggc tgg	672
Arg Asn Ile Pro Cys Tyr Arg Gln Asp Gly Met Arg Met Phe Gly Trp	
210 215 220	
ttt ttc aac tac ttc tac gtt ggt act gtc ctg tgt ctg ttc ctc aac	720
Phe Phe Asn Tyr Phe Tyr Val Gly Thr Val Leu Cys Leu Phe Leu Asn	
225 230 235 240	

ttc tac gtg cag acc tac atc gtc cga aag cac aag gga gcc aaa aag 768

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

245 250 255

att cag tga 777

Ile Gln

<210> 372

<211> 258

<212> PRT

<213> Euglena gracilis

<400> 372

Met Glu Val Val Asn Glu Ile Val Asp Ile Gly Gln Glu Val Leu Pro

1 5 10 15

Lys Val Asp Tyr Ala Gln Leu Trp Ser Asp Ala Ser His Cys Glu Val

20 25 30

Leu Tyr Gly Ser Ile Ala Phe Val Ile Leu Lys Phe Thr Leu Gly Pro

35 40 45

Leu Gly Pro Lys Gly Gln Ser Arg Met Lys Phe Val Phe Thr Asn Tyr

50 55 60

Asn Leu Leu Met Ser Ile Tyr Ser Leu Gly Ser Phe Leu Ser Met Ala

65 70 75 80

Tyr Ala Met Cys Thr Ile Gly Val Met Ser Asp Asn Cys Glu Lys Ala

85 90 95

Phe Asp Asn Asn Val Phe Arg Ile Thr Thr Gln Leu Phe Tyr Leu Ser

100 105 110

Lys Phe Leu Glu Tyr Ile Asp Ser Phe Tyr Leu Pro Leu Met Gly Lys

115 120 125

Pro Leu Thr Trp Leu Gln Phe Phe His His Leu Gly Ala Pro Met Asp

130 135 140

Met Trp Leu Phe Tyr Asn Tyr Arg Asn Glu Ala Val Trp Ile Phe Val

145 150 155 160

Leu Leu Asn Gly Phe Ile His Trp Ile Met Tyr Gly Tyr Tyr Trp Thr
165 170 175
Arg Leu Arg Lys Leu Lys Phe Pro Met Pro Lys Ser Leu Ile Thr Ser
180 185 190
Met Gln Ile Ile Gln Phe Asn Val Gly Phe Tyr Ile Val Trp Lys Tyr
195 200 205
Arg Asn Ile Pro Cys Tyr Arg Gln Asp Gly Met Arg Met Phe Gly Trp
210 215 220
Phe Phe Asn Tyr Phe Tyr Val Gly Thr Val Leu Cys Leu Phe Leu Asn
225 230 235 240
Phe Tyr Val Gln Thr Tyr Ile Val Arg Lys His Lys Gly Ala Lys Lys
245 250 255
Ile Gln

<210> 373

<211> 7769

<212> DNA

<213> Artificial Sequence

<220><223> Plasmid EgD9eS-S9D/L35G/Y84C/I179R

<400> 373

ggccgcaagt gtggatgggg aagtgagtgc cgggttctgt gtgcacaatt ggcaatccaa 60
gatggatgga ttcaacacag ggatatagcg agctacgtgg tgggtcgagg atatagcaac 120
ggatatattat gtttgacact tgagaatgta cgatacaagc actgtccaag tacaatacta 180
aacatactgt acatactcat actcgtaccc gggcaacggg ttactttgag tgcagtggct 240
agtgtcttta ctctacagt gtgcaatact gcgtatcata gtctttgatg tatatcgtat 300
tcattcatgt tagttgcgta cgagccggaa gcataaagtg taaagcctgg ggtgccta 360
gagtgaagcta actcacatta attgcgttgc gctcactgcc cgctttccag tcgggaaacc 420
tgtcgtgcc gctgcattaa tgaatcgcc aacgcgcggg gagaggcggg ttgcgtattg 480
ggcgctcttc cgcttcctcg ctactgact cgctgcgctc ggtcgttcgg ctgcgagcag 540
cggatcagc tcaactcaag gcggaatac ggttatccac agaatcagg gataacgcag 600
gaaagaacat gtgagcaaaa ggccagcaaa aggccaggaa ccgtaaaaag gccgcgttgc 660
tggcgttttt ccataggctc cgccccctg acgagcatca caaaaatcga cgctcaagtc 720

agagggtggcg aaacccgaca ggactataaa gataccaggc gtttcccctt ggaagctccc	780
tcgtgcgctc tctgtttccg accctgccgc ttaccggata cctgtccgcc tttctccctt	840
cgggaagcgt ggcgctttct catagctcac gctgtaggta tctcagttcg gtgtaggtcg	900
ttcgctccaa gctgggctgt gtgcacgaac ccccggttca gcccgaccgc tgcgccttat	960
ccggtaaacta tctgtttgag tccaacccgg taagacacga cttatcgcca ctggcagcag	1020
ccactggtaa caggattagc agagcgaggt atgtaggcgg tgctacagag ttcttgaagt	1080
ggtggcctaa ctacggctac actagaagga cagtatttgg tatctgcgct ctgctgaagc	1140
cagttacctt cggaataaga gttggtagct cttgatccgg caaacaacc accgctggta	1200
gcggtgggtt tttgtttgc aagcagcaga ttacgcgcag aaaaaagga tctcaagaag	1260
atcctttgat cttttctacg gggctctgacg ctcagtggaa cgaaaactca cgttaaggga	1320
ttttggtcat gagattatca aaaaggatct tcacctagat ctttttaaat taaaaatgaa	1380
gttttaaatc aatctaaagt atatatgagt aaacttggtc tgacagttac caatgcttaa	1440
tcagtgaggc acctatctca gcgatctgtc tatttcgttc atccatagtt gcctgactcc	1500
ccgtcgtgta gataactacg atacgggagg gcttaccatc tggccccagt gctgcaatga	1560
taccgcgaga cccacgctca ccgggtccag atttatcagc aataaaccag ccagccggaa	1620
gggccgagcg cagaagtggc cctgcaactt tatccgcctc catccagtct attaatgtt	1680
gccgggaagc tagagtaagt agttcgccag ttaatatgtt gcgcaacgtt gttgccattg	1740
ctacaggcat cgtgggtgca cgctcgtcgt ttggtatggc ttcatcagc tccggttccc	1800
aacgatcaag gcgagttaca tgatccccca tgttgtgcaa aaaagcgggt agctccttcg	1860
gtcctccgat cgttgtcaga agtaagttag ccgcagtggt atcactcatg gttatggcag	1920
cactgcataa ttctcttact gtcatgcat ccgtaaatg cttttctgtg actggtgagt	1980
actcaaccaa gtcattctga gaatagtgtg tgcggcgacc gagttgctct tgccccgcgt	2040
caatacggga taataccgcg ccacatagca gaactttaaa agtgctcatc attgaaaac	2100
gttcttcggg gcgaaaactc tcaaggatct taccgtgtt gagatccagt tcgatgtaac	2160
ccactcgtgc acccaactga tcttcagcat cttttacttt caccagcgtt tctgggtgag	2220
caaaaacagg aaggcaaat gccgcaaaa agggaataag ggcgacacgg aaatgttgaa	2280
tactcatact cttccttttt caatatatt gaagcattta tcagggttat tgtctcatga	2340
gcggatacat atttgaatgt atttagaaaa ataaacaaat aggggttccg cgcacatttc	2400
cccgaataat gccacctgac gcgccctgta gcggcgcat aagcgcgcg ggtgtggtgg	2460

ttacgcgcag cgtgaccgct acacttgcca gcgccctagc gcccgctcct ttcgctttct	2520
tcccttctctt tctcgccacg ttcgccggct tccccgtca agctctaaat cgggggctcc	2580
ctttagggtt ccgatttagt gctttacggc acctcgacce caaaaaactt gattagggtg	2640
atggttcacg tagtgggcca tcgccctgat agacgggttt tcgccctttg acgttggagt	2700
ccacgttctt taatagtga ctcttggtcc aaactggaac aacactcaac cctatctcgg	2760
tctattcttt tgatttataa gggattttgc cgatttcggc ctattggta aaaaatgagc	2820
tgatttaaca aaaatttaac gcgaatttta acaaaatatt aacgcttaca atttcattc	2880
gccattcagg ctgcgcaact gttgggaagg gcgatcgggtg cgggcctctt cgctattacg	2940
ccagctggcg aaagggggat gtgctgcaag gcgattaagt tgggtaacgc cagggttttc	3000
ccagtcacga cgttgtaaaa cgacggccag tgaattgtaa tacgactcac tatagggcga	3060
attgggtacc gggccccccc tcgaggtcga tgggtgcgat aagcttgata tcgaattcat	3120
gtcacacaaa ccgatcttcg cctcaaggaa acctaatctt acatccgaga gactgccgag	3180
atccagtcta cactgattaa ttttcgggcc aataatttaa aaaaatcgtg ttatataata	3240
ttatatgtat tatatatata catcatgatg atactgacag tcatgtccca ttgctaaata	3300
gacagactcc atctgccgcc tccaactgat gttctcaata ttttaaggggt catctcgcat	3360
tgtttaataa taaacagact ccatctaccg cctccaaatg atgttctcaa aatatattgt	3420
atgaacttat ttttattact tagtattatt agacaactta cttgctttat gaaaaacact	3480
tcctatttag gaaacaattt ataatggcag ttcgttcatt taacaattta tgtagaataa	3540
atgttataaa tgcgtatggg aaatcttaaa tatggatagc ataaatgata tctgcattgc	3600
ctaattcgaa atcaacagca acgaaaaaaaa tcccttgtag aacataaata gtcacgaga	3660
aatatcaact atcaaagaac agctattcac acgttactat tgagattatt attggacgag	3720
aatcacacac tcaactgtct ttctctcttc tagaaataca ggtacaagta tgtactattc	3780
tcattgttca tacttctagt catttcatcc cacatatcc ttggatttct ctccaatgaa	3840
tgacattcta tcttgcaaat tcaacaatta taataagata taccaaagta gcggtatagt	3900
ggcaatcaaa aagcttctct ggtgtgcttc tcgtatttat ttttattcta atgatccatt	3960
aaaggtatat atttatttct tgttatataa tccttttggt tattacatgg gctggataca	4020
taaagggtatt ttgatttaaa tttttgctta aattcaatcc cccctcgttc agtgtcaact	4080
glaatggtag gaaattacca tacttttgaa gaagcaaaaa aaatgaaaga aaaaaaaaaat	4140
cgtaattcca ggttagacgt tccgcagaat ctagaatgcg gtatgcggta cattgttctt	4200
cgaacgtaaa agttgcgtc cctgagatat tgtacatttt tgcttttaca agtacaagta	4260
catcgtacaa ctatgtacta ctgttgatgc atccacaaca gtttgttttg ttttttttg	4320

tttttttttt ttctaagat tcattaccgc tatgtatacc tacttgtact ttagtaagc	4380
cgggttattg gcgttcaatt aatcatagac ttatgaatct gcacggtgtg cgctgcgagt	4440
tacttttagc ttatgcatgc tacttgggtg taatatggg atctgttcgg aaatcaacgg	4500
atgctcaatc gatttcgaca gtaattaatt aagtcataca caagtcagct ttcttcgagc	4560
ctcatataag tataagtagt tcaacgtatt agcactgtac ccagcatctc cgtatcgaga	4620
aacacaacaa catgccccat tggacagatc atcgcgatac acaggttgtg cagtatcata	4680
catactcgat cagacaggtc gtctgacat catacaagct gaacaagcgc tccatacttg	4740
cacgtctctc atatacacag ttaaattaca tatccatagt ctaacctcta acagttaatc	4800
ttctggtaag cctcccagcc agccttctgg tatcgcttgg cctcctcaat aggatctcgg	4860
ttctggccgt acagacctcg gccgacaatt atgatatccg ttccggtaga catgacatcc	4920
tcaacagttc ggtactgctg tccgagagcg tctcccttgt cgtcaagacc caccgccggg	4980
gtcagaataa gccagtcctc agagtcgccc ttaggtcggg tctgggcaat gaagccaacc	5040
acaaactcgg ggtcggatcg ggcaagctca atggtctgct tggagtactc gccagtggcc	5100
agagagccct tgcaagacag ctcgccagc atgagcagac ctctggccag ctctcgttg	5160
ggagagggga ctaggaactc cttgtactgg gagttctcgt agtcagagac gtctccttc	5220
ttctgttcag agacagtttc ctcgccacca gctcgcaggc cagcaatgat tccggttccg	5280
ggtacaccgt gggcgttggg gatatcggac cactcggcga ttcggtgaca ccggtactgg	5340
tgcttgacag tgttgccaat atctgcgaac ttctgtcct cgaacaggaa gaaaccgtgc	5400
ttaagagcaa gttccttgag ggggagcaca gtgccggcgt aggtgaagtc gtcaatgatg	5460
tcgatatggg ttttgatcat gcacacataa ggtccgacct tatcggaag ctcaatgagc	5520
tccttggtgg tggtaacatc cagagaagca cacaggttgg ttttcttggc tgccacgagc	5580
ttgagcactc gagcggcaaa ggcggacttg tggacgttag ctcgagcttc gtaggagggc	5640
atcttggtgg tgaagaggag actgaaataa atttagtctg cagaactttt tatcggaacc	5700
ttatctgggg cagtgaagta tatgttatgg taatagttac gagttagtgt aacttataga	5760
tagactggac tatacggcta tcggtccaaa ttagaaagaa cgtcaatggc tctctgggcg	5820
tcgcctttgc cgacaaaaat gtgatcatga tgaaagccag caatgacgtt gcagctgata	5880
ttgttgtcgg ccaaccgcgc cgaaaacgca gctgtcagac ccacagcctc caacgaagaa	5940
tgtatcgtca aagtgatcca agcacactca tagttggagt cgtactcaa aggcggcaat	6000
gacgagtcag acagatactc gtcgacgttt aaacagtgtg cgcagatcta ctatagagga	6060
acatttaaat tgccccggag aagacggcca ggccgcctag atgacaaatt caacaactca	6120
cagctgactt tctgccattg ccactagggg ggggcctttt tataatggcca agccaagctc	6180

tccacgtcgg ttgggctgca cccaacaata aatgggtagg gttgcaccaa caaagggatg 6240

ggatgggggg tagaagatac gaggataacg gggtcfaatg gcacaaataa gaacgaatac 6300

tgccattaag actcgtgac cagcgactga caccattgca tcactaagg gcctcaaaac 6360

tacctcggaa ctgctgcgct gatctggaca ccacagaggt tccgagcact ttaggttgca 6420

ccaaatgtcc caccaggtgc aggcagaaaa cgctggaaca gcgtgtacag ttgtcttaa 6480

caaaaagtga gggcgctgag gtcgagcagg gtggtgtgac ttgttatagc ctttagagct 6540

gcgaaagcgc gtatggattt ggctcatcag gccagattga gggtctgtgg acacatgtca 6600

tgtagtgta cttaaatcgc cccctggata tagccccgac aataggccgt ggcctcattt 6660

ttttgccttc cgcacatttc cattgctcga taccacacc ttgcttctcc tgcacttgcc 6720

aaccttaata ctggtttaca ttgaccaaca tcttacaagc ggggggcttg tctagggtat 6780

atataaacag tggtctctcc aatcggttgc cagtctcttt tttctttct tccccacag 6840

attcgaatc taaactacac atcacagaat tccgagccgt gactatccac gacaagatca 6900

gtgtcgagac gacgcgtttt gtgtaatgac acaatccgaa agtcgctagc aacacacact 6960

ctctacaaa actaaccag ctctggtagc atggaggtcg tgaacgaaat cgtcgacatt 7020

ggccaggagg ttcttcccaa ggtcgactat gctcagctct ggtctgatgc ctgcactgc 7080

gaggtgctgt acgggtccat cgccttcgtc atcctgaagt tcacccttgg tcctctcgga 7140

cccaagggtc agtctcgaat gaagtttggt ttaccaact acaacctgct catgtccatc 7200

tactcgtggt gctccttct ctctatggcc tacgcatgt gtaccattgg tgcattgctc 7260

gacaactgcg agaaggtttt cgacaacaat gtcttcgaa tcaccactca gctgttctac 7320

ctcagcaagt tcctcgagta cattgactcc ttctatctgc cctcatggg caagcctctg 7380

acctggttgc agttctttca ccatctcgga gctctatgg acatgtggct gttctacaac 7440

taccgaaacg aagccgtttg gatctttgtg ctgctcaacg gcttcattca ctggatcatg 7500

tacggctact attggaccgc actgcgaaag ctcaagttcc ctatgcccc gtcctgatt 7560

acttctatgc agatcattca gttaacgtt ggcttctaca tcgtctggaa gtaccggaac 7620

attccctgct accgacaaga tggaatgaga atgtttggct ggtttttcaa ctacttctac 7680

gttggtagtg tcctgtgctt gttcctcaac ttctacgtgc agacctacat cgtccgaaag 7740

cacaaggag ccaaaaagat tcagtgagc 7769

<210> 374

<211> 777

<212> DNA

<213> Euglena gracilis

<220><221> CDS

<222> (1)..(777)

<223> mutant delta-9 elongase "EgD9eS-L35G/Y84C/I179R/Q244N"

<400> 374

atg gag gtc gtg aac gaa atc gtc tcc att ggc cag gag gtt ctt ccc 48

Met Glu Val Val Asn Glu Ile Val Ser Ile Gly Gln Glu Val Leu Pro

1 5 10 15

aag gtc gac tat gct cag ctc tgg tct gat gcc tcg cac tgc gag gtg 96

Lys Val Asp Tyr Ala Gln Leu Trp Ser Asp Ala Ser His Cys Glu Val

20 25 30

ctg tac ggg tcc atc gcc ttc gtc atc ctg aag ttc acc ctt ggt cct 144

Leu Tyr Gly Ser Ile Ala Phe Val Ile Leu Lys Phe Thr Leu Gly Pro

35 40 45

ctc gga ccc aag ggt cag tct cga atg aag ttt gtg ttc acc aac tac 192

Leu Gly Pro Lys Gly Gln Ser Arg Met Lys Phe Val Phe Thr Asn Tyr

50 55 60

aac ctg ctc atg tcc atc tac tcg ctg ggc tcc ttc ctc tct atg gcc 240

Asn Leu Leu Met Ser Ile Tyr Ser Leu Gly Ser Phe Leu Ser Met Ala

65 70 75 80

tac gcc atg tgt acc att ggt gtc atg tcc gac aac tgc gag aag gct 288

Tyr Ala Met Cys Thr Ile Gly Val Met Ser Asp Asn Cys Glu Lys Ala

85 90 95

ttc gac aac aat gtc ttc cga atc acc act cag ctg ttc tac ctc agc 336

Phe Asp Asn Asn Val Phe Arg Ile Thr Thr Gln Leu Phe Tyr Leu Ser

100 105 110

aag ttc ctc gag tac att gac tcc ttc tat ctg ccc ctc atg ggc aag 384

Lys Phe Leu Glu Tyr Ile Asp Ser Phe Tyr Leu Pro Leu Met Gly Lys

115 120 125

cct ctg acc tgg ttg cag ttc ttt cac cat ctc gga gct cct atg gac 432

Pro Leu Thr Trp Leu Gln Phe Phe His His Leu Gly Ala Pro Met Asp

130 135 140

atg tgg ctg ttc tac aac tac cga aac gaa gcc gtt tgg atc ttt gtg 480

Met Trp Leu Phe Tyr Asn Tyr Arg Asn Glu Ala Val Trp Ile Phe Val

145 150 155 160

ctg ctg aac ggc ttc att cac tgg atc atg tac ggc tac tat tgg acc 528

Leu Leu Asn Gly Phe Ile His Trp Ile Met Tyr Gly Tyr Tyr Trp Thr

165 170 175

cga ctg cga aag ctg aag ttc cct atg ccc aag tcc ctg att act tct 576

Arg Leu Arg Lys Leu Lys Phe Pro Met Pro Lys Ser Leu Ile Thr Ser

180 185 190

atg cag atc att cag ttc aac gtt ggc ttc tac atc gtc tgg aag tac 624

Met Gln Ile Ile Gln Phe Asn Val Gly Phe Tyr Ile Val Trp Lys Tyr

195 200 205

cgg aac att ccc tgc tac cga caa gat gga atg aga atg ttt ggc tgg 672

Arg Asn Ile Pro Cys Tyr Arg Gln Asp Gly Met Arg Met Phe Gly Trp

210 215 220

ttt ttc aac tac ttc tac gtt ggt act gtc ctg tgt ctg ttc ctg aac 720

Phe Phe Asn Tyr Phe Tyr Val Gly Thr Val Leu Cys Leu Phe Leu Asn

225 230 235 240

ttc tac gtg aac acc tac atc gtc cga aag cac aag gga gcc aaa aag 768

Phe Tyr Val Asn Thr Tyr Ile Val Arg Lys His Lys Gly Ala Lys Lys

245 250 255

att cag tga 777

Ile Gln

<210> 375

<211> 258

<212> PRT

<213> Euglena gracilis

<400> 375

Met Glu Val Val Asn Glu Ile Val Ser Ile Gly Gln Glu Val Leu Pro

1 5 10 15

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

<210> 376

<211> 7769

<212> DNA

<213> Artificial Sequence

<220><223> Plasmid EgD9eS-L35G/Y84C/I179R/Q244N

<400> 376

ggccgcaagt gtggatgggg aagtgagtgc ccggttctgt gtgcacaatt ggcaatccaa	60
gatggatgga ttcaacacag ggatatagcg agctacgtgg tggatcgagg atatatcaac	120
ggatatattat gtttgacact tgagaatgta cgatacaagc actgtccaag tacaatacta	180
aacatactgt acatactcat actcgtaccc gggcaacggt ttactttgag tgcagtggct	240
agtgtcttta ctctacagt gtgcaatact gcgtatcata gtctttgatg tatatcgtat	300
tcattcatgt tagttgcgta cgagccggaa gcataaagtg taaagcctgg ggtgcctaata	360
gagttagcta actcacatta attgcgttgc gctcactgcc cgctttccag tcgggaaacc	420
tgtcgtgcc gctgcattaa tgaatcgcc aacgcgcggg gagaggcggt ttgcgtattg	480
ggcgtcttc cgcttcctcg ctactgact cgctgcgtc ggtcgttcgg ctgcgagcag	540
cggtatcagc tcaactcaaag gcggttaatac ggttatccac agaatcaggg gataacgcag	600
gaaagaacat gtgagcaaaa ggccagcaaa aggccaggaa ccgtaaaaag gccgcgttgc	660
tggcgttttt ccataggtc cgccccctg acgagcatca caaaaatcga cgctcaagtc	720
agaggtaggc aaacccgaca ggactataaa gataccaggc gtttccccct ggaagctccc	780
tcgtgcgtc tctgttccg accctgccgc ttaccggata cctgtccgc tttctccctt	840
cgggaagcgt ggcgctttct catagctcac gctgtaggta tctcagttcg gtgtaggtcg	900
ttcgctcaa gctgggctgt gtgcacgaac cccccgtca gcccgaccgc tgcgccttat	960
ccggtacta tcgtcttgag tccaaccgg taagacacga cttatcgcca ctggcagcag	1020
ccactggtaa caggattagc agagcgaggt atgtaggcgg tgctacagag ttcttgaagt	1080
ggtggcctaa ctacggctac actagaagga cagtatttgg tatctgcgt ctgctgaagc	1140
cagttacctt cggaaaaaga gttgtagct cttgatccgg caaacaacc accgctggtta	1200
gcggtggttt tttgtttgc aagcagcaga ttacgcgcag aaaaaagga tctcaagaag	1260
atcctttgat cttttctacg gggctgacg ctactggaa cgaaaactca cgtaaggga	1320
ttttggtcat gagattatca aaaaggatct tcacntagat ctttttaaat taaaaatgaa	1380
gttttaaatc aatctaaagt atatatgagt aaacttggc tgacagttac caatgcttaa	1440

tcagtgaggc acctatctca gcgatctgtc tatttcgttc atccatagtt gcctgactcc	1500
ccgtcgtgta gataactacg atacgggagg gcttaccatc tggccccagt gctgcaatga	1560
taccgcgaga cccacgtcca ccggctccag atttatcagc aataaaccag ccagccggaa	1620
gggccgagcg cagaagtggc cctgcaactt tatccgcctc catccagtct attaatgtt	1680
gccgggaagc tagagtaagt agttcgccag ttaatagttt gcgcaacgtt gttgccattg	1740
ctacaggcat cgtgggtgca cgctcgtcgt ttggtatggc ttcatcagc tccggttccc	1800
aacgatcaag gcgagttaca tgatccccc a tgttgtgcaa aaaagcgggt agctccttcg	1860
gtcctccgat cgttgtcaga agtaagttgg ccgcagtgtt atcactcatg gttatggcag	1920
cactgcataa ttctcttact gtcatgccat ccgtaagatg cttttctgtg actggtgagt	1980
actcaaccaa gtcattctga gaatagtga tgccggcgacc gagttgctct tgcgccgct	2040
caatacggga taataccgcg ccacatagca gaactttaaa agtgcctcatc attgaaaaac	2100
gttcttcggg gcgaaaactc tcaaggatct taccgctgtt gagatccagt tcgatgtaac	2160
ccactcgtgc acccaactga tcttcagcat cttttacttt caccagcgtt tctgggtgag	2220
caaaaacagg aaggcaaaat gccgcaaaaa agggaataag ggcgacacgg aaatgttgaa	2280
tactcatact ctctctttt caatattatt gaagcattta tcagggttat tgtctcatga	2340
gcggatacat atttgaatgt atttagaaaa ataaacaaat aggggttccg cgcacatttc	2400
cccgaaaagt gccacctgac gcgcctctga gcggcgcatc aagcgcgcg ggtgtggtgg	2460
ttacgcgcag cgtgaccgt acacttgcca gcgccctagc gcccgctcct ttcgctttct	2520
tcccttctt tctcgccacg ttcccggtt tccccgtca agctctaaat cgggggctcc	2580
ctttagggtt ccgatttagt gctttacggc acctcgaccc caaaaaactt gattagggtg	2640
atggttcacg tagtgggcca tcgcctgat agacggtttt tcgccctttg acgttggagt	2700
ccacgttctt taatagtga ctcttggtcc aaactggaac aacactcaac cctatctcgg	2760
tctattcttt tgatttataa gggattttgc cgatttcggc ctattggtta aaaaatgagc	2820
tgatttaaca aaaatttaac gcgaatttta acaaaatatt aacgcttaca atttcattc	2880
gccattcagg ctgcgcaact gttgggaagg gcgatcgggtg cgggcctctt cgctattacg	2940
ccagctggcg aaagggggat gtgctgcaag gcgattaagt tgggtaacgc cagggttttc	3000
ccagtacga cgttgtaaaa gcacggccag tgaattgtaa tacgactcac tataggcgca	3060
attgggtacc gggcccccc tcgaggctga tgggtgcgat aagcttgata tcgaattcat	3120
gtcacacaaa ccgatcttcg cctcaaggaa acctaatctt acatccgaga gactgccgag	3180
atccagtcta cactgattaa ttttcgggcc aataatttaa aaaaatcgtg ttatataata	3240
ttatatgtat tatatatata catcatgatg atactgacag tcatgtccca ttgctaaata	3300

gacagactcc atctgccgcc tccaactgat gttctcaata ttttaaggggt catctcgcat 3360
 tgtttaataa taaacagact ccatctaccg cctccaatg atgttctcaa aatatattgt 3420
 atgaacttat ttttattact tagtattatt agacaactta cttgctttat gaaaaaact 3480
 tcctatttag gaaacaattt ataatggcag ttcgttcatt taacaattta tgtagaataa 3540

atgttataaa tgcgtatggg aaatcttaaa tatggatagc ataaatgata tctgcattgc 3600
 ctaattcgaa atcaacagca acgaaaaaaaa tcccttgtag aacataaata gtcacgaga 3660
 aatatcaact atcaaagaac agctattcac acgttactat tgagattatt attggacgag 3720
 aatcacacac tcaactgtct ttctctcttc tagaaataca ggtacaagta tgtactattc 3780
 tcattgttca tacttctagt catttcatcc cacatatcc ttggatttct ctccaatgaa 3840
 tgacattcta tcttgcaat tcaacaatta taataagata taccaaagta gcggtatagt 3900
 ggcaatcaaa aagcttctct ggtgtgcttc tcgtatttat ttttattcta atgatccatt 3960

aaaggtatat atttatttct tgttatataa tccttttggt tattacatgg gctggataca 4020
 taaaggtatt ttgatttaat tttttgctta aattcaatcc cccctcgctc agtgtcaact 4080
 gtaatggtag gaaattacca tacttttgaa gaagcaaaaa aaatgaaaga aaaaaaaaaat 4140
 cgtatttcca ggtagacgt tccgcagaat ctagaatgcg gtatgcggtta cattgttctt 4200
 cgaacgtaaa agttgcgctc cctgagatat tgtacatttt tgcttttaca agtacaagta 4260
 catcgtacaa ctatgtacta ctgttgatgc atccacaaca gtttgttttg ttttttttg 4320
 tttttttttt ttctaattgat tcattaccgc tatgtatacc tacttgtact tgtagtaagc 4380

cgggttattg gcgttcaatt aatcatagac ttatgaatct gcacgggtgtg cgctgcgagt 4440
 tacttttagc ttatgcatgc tacttgggtg taatatggg atctgttcgg aaatcaacgg 4500
 atgtcaatc gatttcgaca gtaattaat aagtcataca caagtcagct ttcttcgagc 4560
 ctcatataag tataagtagt tcaacgtatt agcactgtac ccagcatctc cgtatcgaga 4620
 aacacaacaa catgccccat tggacagatc atgcggatc acaggttgtg cagtatcata 4680
 catactcgat cagacaggtc gtctgacat catacaagct gaacaagcgc tccatacttg 4740
 cagctctct atatacacag ttaaattaca tatccatagt ctaacctcta acagttaatc 4800

ttctggtaag cctcccagcc agccttctgg tatcgcttgg cctcctcaat aggatctcgg 4860
 ttctggcgt acagacctg gccgacaatt atgatatccg ttccggtaga catgacatcc 4920
 tcaacagttc ggtactgtct tccgagagcg tctcccttgt cgtcaagacc caccctgggg 4980
 gtcagaataa gccagtcctc agagtcgcc ttaggtcggg tctgggcaat gaagccaacc 5040
 acaaactcgg ggtcggatcg ggcaagctca atgggtctgct tggagtactc gccagtggcc 5100
 agagagccct tgcaagacag ctcgccagc atgagcagac ctctggccag cttctcgttg 5160

ggagagggga ctaggaactc cttgtactgg gaggttctcgt agtcagagac gtcctccttc 5220

ttctgttcag agacagtffc ctcggcacca gctcgcaggc cagcaatgat tccggttccg 5280

ggtacaccgt gggcgttggg gatatacggac cactcggcga ttcggtgaca ccggtactgg 5340

tgcttgacag tgttgccaat atctgcgaac ttctgtcct cgaacaggaa gaaaccgtgc 5400

ttaagagcaa gttccttgag ggggagcaca gtgccggcgt aggtgaagtc gtcaatgatg 5460

tcgatatggg ttttgatcat gcacacataa ggtccgacct tatcggcaag ctcaatgagc 5520

tccttggtgg tggtaacatc cagagaagca cacaggttgg ttttcttggc tgccacgagc 5580

ttgagcactc gagcggcaaa ggccgacttg tggacgttag ctcgagcttc gtaggagggc 5640

atattggtgg tgaagaggag actgaaataa atttagtctg cagaactttt tatcggaaac 5700

ttatctgggg cagtgaagta tatgttatgg taatagttac gagttagtgg aacttataga 5760

tagactggac tatacggcta tcggtcctaaa ttgaaaagaa cgtcaatggc tctctgggcg 5820

tcgcctttgc cgacaaaaat gtgatcatga tgaaagccag caatgacgtt gcagctgata 5880

ttgttgcgg ccaaccgcgc cgaaaacgca gctgtcagac ccacagcctc caacgaagaa 5940

tgtatcgtca aagtgatcca agcacactca tagttggagt cgtactccaa aggcggcaat 6000

gacgagtcag acagatactc gtcgacgttt aaacagtgtg cgcagatcta ctatagagga 6060

acatttaaat tgccccggag aagacggcca ggccgcctag atgacaaatt caacaactca 6120

cagctgactt tcigccattg ccactagggg ggggcctttt tataatggcca agccaagctc 6180

tccacgtcgg ttgggctgca cccaacaata aatgggtagg gttgcacaa caaagggatg 6240

ggatgggggg tagaagatag gaggataacg gggctcaatg gcacaaataa gaacgaatac 6300

tgccattaag actcgtgac cagcgactga caccattgca tcatctaagg gcctcaaaac 6360

tacctcgaa ctgctgcgt gatctggaca ccacagaggt tccgagcact ttaggttgca 6420

ccaaatgtcc caccaggtgc aggcagaaaa cgttggaaca gcgtgtacag ttgtcttaa 6480

caaaaagtga gggcgtgag gtcgagcagg gtggtgtgac ttgttatagc ctttagagct 6540

gcgaaagcgc gtatggattt ggctcatcag gccagattga gggctctgtg acacatgtca 6600

tgtagtgta ctcaatcgc cccctggata tagccccgac aataggccgt ggcctcattt 6660

ttttgcctc cgcacatttc cattgtcga taccacacc ttgcttctcc tgcacttgcc 6720

aaccttaata ctggtttaca ttgaccaaca tcttaacaag ggggggcttg tctagggtat 6780

atataaacag tggctctccc aatcggttgc cagtctcttt tttcctttct tccccacag 6840

attcgaatc taaactacac atcacagaat tccgagccgt gagtatccac gacaagatca 6900

gtgtcgagac gacgcgtttt gtgtaatgac acaatccgaa agtcgctagc aacacacact 6960
 ctctacacaa actaaccgag ctctggtagc atggaggtagc tgaacgaaat cgtctccatt 7020
 ggccaggagg ttcttcccaa ggtcgactat gctcagctct ggtctgatgc ctgcgactgc 7080
 gaggtgctgt acgggtccat cgccttcgtc atcctgaagt tcacccttgg tcctctcgga 7140
 cccaagggtc agtctcgaat gaagtttgtg ttcaccaact acaacctgct catgtccatc 7200
 tactcgctgg gctccttctt ctctatggcc tacgccaatgt gtaccattgg tgtcatgtcc 7260
 gacaactgcg agaaggcttt cgacaacaat gtcttccgaa tcaccactca gctgttctac 7320

ctcagcaagt tcctcgagta cattgactcc ttctatctgc ccctcatggg caagcctctg 7380
 acctggttgc agttctttca ccatctcgga gctcctatgg acatgtggct gtcttacaac 7440
 taccgaaacg aagccgtttg gatctttgtg ctgctcaacg gcttcattca ctggatcatg 7500
 tacggtact attggaccgc actgcgaaag ctcaagttcc ctatgcccaa gtccctgatt 7560
 acttctatgc agatcattca gttcaacgtt ggcttctaca tcgtctggaa gtaccggaac 7620
 attcctgct accgacaaga tggaatgaga atgtttggct gggttttcaa ctacttctac 7680
 gtgggtactg tcctgtgtct gttcttcaac ttctacgtga acacctacat cgtccgaaag 7740

cacaaggag ccaaaaagat tcagttagc 7769

<210> 377

<211> 777

<212> DNA

<213> Euglena gracilis

<220><221> CDS

<222> (1)..(777)

<223> mutant delta-9 elongase "EgD9eS-A21V/L35G/W132T/I179R/Q244N"

<400> 377

atg gag gtc gtg aac gaa atc gtc tcc att ggc cag gag gtt ctt ccc 48

Met Glu Val Val Asn Glu Ile Val Ser Ile Gly Gln Glu Val Leu Pro

1 5 10 15

aag gtc gac tat gtg cag ctc tgg tct gat gcc tcg cac tgc gag gtg 96

Lys Val Asp Tyr Val Gln Leu Trp Ser Asp Ala Ser His Cys Glu Val

20 25 30

ctg tac ggg tcc atc gcc ttc gtc atc ctg aag ttc acc ctt ggt cct 144

Leu Tyr Gly Ser Ile Ala Phe Val Ile Leu Lys Phe Thr Leu Gly Pro

35 40 45

ctc gga ccc aag ggt cag tct cga atg aag ttt gtg ttc acc aac tac 192
Leu Gly Pro Lys Gly Gln Ser Arg Met Lys Phe Val Phe Thr Asn Tyr

50 55 60
aac ctg ctc atg tcc atc tac tcg ctg ggc tcc ttc ctc tct atg gcc 240
Asn Leu Leu Met Ser Ile Tyr Ser Leu Gly Ser Phe Leu Ser Met Ala

65 70 75 80
tac gcc atg tac acc att ggt gtc atg tcc gac aac tgc gag aag gct 288
Tyr Ala Met Tyr Thr Ile Gly Val Met Ser Asp Asn Cys Glu Lys Ala

85 90 95
ttc gac aac aat gtc ttc cga atc acc act cag ctg ttc tac ctc agc 336
Phe Asp Asn Asn Val Phe Arg Ile Thr Thr Gln Leu Phe Tyr Leu Ser

100 105 110
aag ttc ctc gag tac att gac tcc ttc tat ctg ccc ctc atg ggc aag 384
Lys Phe Leu Glu Tyr Ile Asp Ser Phe Tyr Leu Pro Leu Met Gly Lys

115 120 125
cct ctg acc acc ttg cag ttc ttt cac cat ctc gga gct cct atg gac 432

Pro Leu Thr Thr Leu Gln Phe Phe His His Leu Gly Ala Pro Met Asp

130 135 140
atg tgg ctg ttc tac aac tac cga aac gaa gcc gtt tgg atc ttt gtg 480
Met Trp Leu Phe Tyr Asn Tyr Arg Asn Glu Ala Val Trp Ile Phe Val

145 150 155 160
ctg ctc aac ggc ttc att cac tgg atc atg tac ggc tac tat tgg acc 528
Leu Leu Asn Gly Phe Ile His Trp Ile Met Tyr Gly Tyr Tyr Trp Thr

165 170 175
cga ctg cga aag ctc aag ttc cct atg ccc aag tcc ctg att act tct 576
Arg Leu Arg Lys Leu Lys Phe Pro Met Pro Lys Ser Leu Ile Thr Ser

180 185 190
atg cag atc att cag ttc aac gtt ggc ttc tac atc gtc tgg aag tac 624
Met Gln Ile Ile Gln Phe Asn Val Gly Phe Tyr Ile Val Trp Lys Tyr

195	200	205										
cgg aac att ccc tgc tac cga caa gat gga atg aga atg ttt ggc tgg												672
Arg Asn Ile Pro Cys Tyr Arg Gln Asp Gly Met Arg Met Phe Gly Trp												
210	215	220										
ttt ttc aac tac ttc tac gtt ggt act gtc ctg tgt ctg ttc ctc aac												720
Phe Phe Asn Tyr Phe Tyr Val Gly Thr Val Leu Cys Leu Phe Leu Asn												
225	230	235	240									
ttc tac gtg aac acc tac atc gtc cga aag cac aag gga gcc aaa aag												768
Phe Tyr Val Asn Thr Tyr Ile Val Arg Lys His Lys Gly Ala Lys Lys												
245	250	255										
att cag tga												777
Ile Gln												
<210> 378												
<211> 258												
<212> PRT												
<213> Euglena gracilis												
<400> 378												
Met Glu Val Val Asn Glu Ile Val Ser Ile Gly Gln Glu Val Leu Pro												
1 5 10 15												
Lys Val Asp Tyr Val Gln Leu Trp Ser Asp Ala Ser His Cys Glu Val												
20 25 30												
Leu Tyr Gly Ser Ile Ala Phe Val Ile Leu Lys Phe Thr Leu Gly Pro												
35 40 45												
Leu Gly Pro Lys Gly Gln Ser Arg Met Lys Phe Val Phe Thr Asn Tyr												
50 55 60												
Asn Leu Leu Met Ser Ile Tyr Ser Leu Gly Ser Phe Leu Ser Met Ala												
65 70 75 80												
Tyr Ala Met Tyr Thr Ile Gly Val Met Ser Asp Asn Cys Glu Lys Ala												
85 90 95												
Phe Asp Asn Asn Val Phe Arg Ile Thr Thr Gln Leu Phe Tyr Leu Ser												

100 105 110
 Lys Phe Leu Glu Tyr Ile Asp Ser Phe Tyr Leu Pro Leu Met Gly Lys
 115 120 125
 Pro Leu Thr Thr Leu Gln Phe Phe His His Leu Gly Ala Pro Met Asp

 130 135 140
 Met Trp Leu Phe Tyr Asn Tyr Arg Asn Glu Ala Val Trp Ile Phe Val
 145 150 155 160
 Leu Leu Asn Gly Phe Ile His Trp Ile Met Tyr Gly Tyr Tyr Trp Thr
 165 170 175
 Arg Leu Arg Lys Leu Lys Phe Pro Met Pro Lys Ser Leu Ile Thr Ser
 180 185 190
 Met Gln Ile Ile Gln Phe Asn Val Gly Phe Tyr Ile Val Trp Lys Tyr

 195 200 205
 Arg Asn Ile Pro Cys Tyr Arg Gln Asp Gly Met Arg Met Phe Gly Trp
 210 215 220
 Phe Phe Asn Tyr Phe Tyr Val Gly Thr Val Leu Cys Leu Phe Leu Asn
 225 230 235 240
 Phe Tyr Val Asn Thr Tyr Ile Val Arg Lys His Lys Gly Ala Lys Lys
 245 250 255
 Ile Gln

<210> 379

<211> 7769

<212> DNA

<213> Artificial Sequence

<220><223> Plasmid EgD9eS-A21V/L35G/W132T/I179R/Q244N

<400> 379

ggccgcaagt gtggatgggg aagtgagtgc ccggttctgt gtgcacaatt ggcaatccaa 60
 gatggatgga ttcaacacag ggatatagcg agctacgtgg tgggtgcgagg atatagcaac 120
 ggatatattat gtttgacact tgagaatgta cgatacaagc actgtccaag tacaatacta 180
 aacatactgt acatactcat actcgtaccc gggcaacggt ttactttgag tgcagtggt 240
 agtgctctta ctctacagt gtgcaatact gcgtatcata gtccttgatg tatatcgat 300

tcattcatgt tagttgcgta cgagccggaa gcataaagt taaagcctgg ggtgccta	360
gagtgaagcta actcacatta attgcgttgc gctcactgcc cgctttccag tcgggaaacc	420
tgctcgtgcc gctgcattaa tgaatcgcc aacgcgcggg gagaggcggg ttgcgtattg	480
ggcgctcttc cgcttcctcg ctactgact cgctgcgctc ggtcgttcgg ctgcgagag	540
cggtatcagc tcaactcaag gcggaatac ggttatccac agaatacagg gataacgcag	600
gaaagaacat gtgagcaaaa ggccagcaaa aggccaggaa ccgtaaaaag gccgcgttgc	660
tggcggtttt ccataggctc cgccccctg acgagcatca caaaaatcga cgctcaagtc	720
agaggtggcg aaacccgaca ggactataaa gataccaggc gtttccccct ggaagctccc	780
tcgtgcgctc tctgttccg accctgccgc ttaccggata cctgtccgcc tttctccctt	840
cgggaagcgt ggcgctttct catagctcac gctgtaggta tctcagttcg gtgtaggtcg	900
ttcgctccaa gctgggctgt gtgcacgaac ccccggttca gcccgaccgc tgcgccttat	960
ccggtaaacta tcgtctttag tccaaccgg taagacacga cttatcgcca ctggcagcag	1020
ccactggtaa caggattagc agagcgagg atgtaggcgg tgctacagag ttcttgaagt	1080
ggtggcctaa ctacggctac actagaagga cagtatttgg tatctgcgt ctgctgaagc	1140
cagttacctt cgaaaaaga gttggtagct cttgatccgg caaacaacc accgctggta	1200
gcggtggttt tttgtttgc aagcagcaga ttacgcgcag aaaaaagga tctcaagaag	1260
atcctttgat cttttctac gggtctgacg ctacgtggaa cgaaaactca cgtaaggga	1320
ttttggctat gagattatca aaaaggatct tcacctagat ctttttaaat taaaaatgaa	1380
gttttaaatc aatctaaagt atatatgagt aaacttggc tgacagttac caatgcttaa	1440
tcagtgaggc acctatctca gcgatctgtc tatttcgttc atccatagtt gcctgactcc	1500
ccgtcgtgta gataactacg atacgggagg gcttaccatc tggccccagt gctgcaatga	1560
taccgcgaga cccacgtca ccggctccag atttatcagc aataaaccag ccagccggaa	1620
gggccgagcg cagaagtggc cctgcaactt tatccgcctc catccagtct attaattgtt	1680
gccgggaagc tagagtaagt agttcgccag ttaatatgtt gcgcaacgtt gttgccattg	1740
ctacaggcat cgtgggtgca cgctcgtcgt ttggtatggc ttcatcagc tccggttccc	1800
aacgatcaag gcgagttaca tgatcccca tgttgtgcaa aaaagcgggt agctccttcg	1860
gtctccgat cgttgtcaga agtaagtgg ccgcagtggt atcaactatg gttatggcag	1920
cactgcataa ttctcttact gtcatccat ccgtaagatg cttttctgtg actggtagt	1980
actcaacca gtcattctga gaatagtga tgcggcgacc gatttgcct tgcggcgct	2040

caatacggga taataccgcg ccacatagca gaactttaaa agtgctcatc attggaaaac	2100
gttcttcggg gcgaaaactc tcaaggatct taccgtgtt gagatccagt tcgatgtaac	2160
ccactcgtgc acccaactga tcttcagcat cttttacttt caccagcggt tctgggtgag	2220
caaaaacagg aaggcaaaat gccgcaaaaa aggggaataag ggcgacacgg aaatgttgaa	2280
tactcatact cttccttttt caatattatt gaagcattta tcagggttat tgtctcatga	2340
gcggatacat atttgaatgt atttagaaaa ataacaat aggggttcg cgacatttc	2400
cccgaagaat gccacctgac gcgccctgia gcggcgcat aagcgcgcg ggtgtggtgg	2460
ttacgcgcag cgtgaccgt acacttgcca gcgccctagc gcccgctcct ttcgctttct	2520
tcccttcctt tctcgccag ttcgccggt tccccgtca agctctaaat cgggggctcc	2580
ctttagggtt ccgatttagt gctttacggc acctcgacc caaaaaactt gattagggtg	2640
atggttcacg tagtgggcca tcgccctgat agacggtttt tcgccctttg acgttgaggt	2700
ccacgttctt taatagtga cttctgttcc aaactggaac aacactcaac cctatctcgg	2760
tctattcttt tgatttataa gggattttgc cgatttcggc ctattggtta aaaaatgagc	2820
tgatttaaca aaaatttaac gcgaatttta acaaaatatt aacgcttaca atttcattc	2880
gccattcagg ctgcgcaact gttgggaagg gcgatcgggt cgggcctctt cgctattacg	2940
ccagctggcg aaagggggat gtgctgcaag gcgattaagt tgggtaacgc cagggttttc	3000
ccagtcacga cgttgtaaaa cgacggccag tgaattgtaa tacgactcac tataggcgca	3060
attgggtacc gggccccccc tcgaggtcga tgggtgctgat aagcttgata tcgaattcat	3120
gtcacacaaa ccgactctcg cctcaaggaa acctaatct acatccgaga gactgccgag	3180
atccagtcta cactgattaa ttttcgggcc aataattta aaaaatcgtg ttatataata	3240
ttatatgtat tatatatata catcatgatg atactgacag tcattgtcca ttgctaaata	3300
gacagactcc atctgccgc tccaactgat gttctcaata ttttaagggt catctcgcat	3360
tgtttaataa taacagact ccatctaccg cctccaatg atgttctcaa aatatattgt	3420
atgaacttat ttttattact tagtattatt agacaactta cttgctttat gaaaaaact	3480
tcctatttag gaaacaattt ataatggcag ttcgttcatt taacaattta tgtagaataa	3540
atgttataaa tgcgtatggg aaatcttaaa tatggatagc ataaatgata tctgcattgc	3600
ctaattcgaa atcaacagca acgaaaaaaa tcccttgtac aacataaata gtcatcgaga	3660
aatatcaact atcaaagaac agctattcac acgttactat tgagattatt attggacgag	3720
aatcacacac tcaactgtct ttctctcttc tagaaataca ggtacaagta tgtactattc	3780
tcattgttca tacttctagt catttcatcc cacatatcc ttggattttct ctccaatgaa	3840
tgacattcta tcttgcaaat tcaacaattt taataagata taccaaagta gcggtatagt	3900

ggcaatcaaa aagcttctct ggtgtgcttc tcgtatttat ttttattcta atgatccatt 3960
aaaggtatat atttatttct tgttatataa tccttttggt tattacatgg gctggataca 4020
taaaggtatt ttgatttaat tttttgctta aattcaatcc cccctcgttc agtgtcaact 4080
gtaatggtag gaaattacca tacttttgaa gaagcaaaaa aaatgaaaga aaaaaaaaaat 4140

cgtattttcca ggttagacgt tccgcagaat ctagaatgcg gtatgcggtta cattgttctt 4200
cgaacgtaaa agttgcgttc cctgagatat tgtacatttt tgcttttaca agtacaagta 4260
catcgtacaa ctatgtacta ctgttgatgc atccacaaca gtttgttttg ttttttttg 4320
tttttttttt ttctaagtat tcattaccgc tatgtatacc tacttgtact ttagtaagc 4380
cgggttattg gcgttcaatt aatcatagac ttatgaatct gcacgggttg cgctgcgagt 4440
tacttttagc ttatgcatgc tacttgggtg taatattggg atctgttcgg aaatcaacgg 4500
atgctcaatc gatttcgaca gtaattaatt aagtcataca caagtcagct ttcttcgagc 4560

ctcatataag tataagtagt tcaacgtatt agcactgtac ccagcatctc cgtatcgaga 4620
aacacaacaa catgccccat tggacagatc atgcggatac acaggttgtg cagtatcata 4680
catactcgat cagacaggtc gtctgacat cacaagaat gaacaagcgc tccatacttg 4740
cacgtctct atatacacag ttaaattaca tatccatagt ctaacctcta acagttaatc 4800
ttctggtaag cctcccagcc agccttctgg tctcgttgg cctcctcaat aggatctcgg 4860
ttctggccgt acagacctg gccgacaatt atgatatccg ttcgggtaga catgacatcc 4920
tcaacagttc ggtactgctg tccgagagcg tctcccttgt cgtcaagacc cccccgggg 4980

gtcagaataa gccagtcctc agagtcgcc ttaggtcggg tctgggcaat gaagccaacc 5040
acaaactcgg ggtcggatcg ggcaagctca atggtctgct tggagtactc gccagtggcc 5100
agagagccct tgcaagacag ctcggccagc atgagcagac ctctggccag ctctctgttg 5160
ggagagggga ctaggaaactc ctgtacttgg gatttctcgt agtcagagac gtctccttc 5220
ttctgttcag agacagtttc ctcggcacca gctcgcaggc cagcaatgat tccggttccg 5280
ggtacaccgt gggcgttggg gatatcggac cactcggcga ttcggtgaca ccggtactgg 5340
tgcttgacag tgttgccaat atctgcgaac tttctgtcct cgaacaggaa gaaaccgtgc 5400

ttaagagcaa gttccttgag ggggagcaca gtcccgcggt aggtgaagtc gtcaatgatg 5460
tcgatatggg ttttgatcat gcacacataa ggtccgacct tatcggaag ctcaatgagc 5520
tccttgggtg tggtaacatc cagagaagca cacaggttgg ttttcttggc tgccacgagc 5580
ttgagcactc gagcggcaaa ggcggacttg tggacgttag ctcgagcttc gtaggagggc 5640
atcttgggtg tgaaggagg actgaaataa atttagtctg cagaactttt tatcggaacc 5700
ttatctgggg cagtgaagta tatgttatgg taatagttac gagttagtgt aacttataga 5760

tagactggac tatacggcta tcggtccaaa ttagaaagaa cgtcaatggc tctctgggcg	5820
tcgcctttgc cgacaaaaat gtgatcatga tgaaagccag caatgacgtt gcagctgata	5880
ttgtttgtcgg ccaaccgcgc cgaaaacgca gctgtcagac ccacagcctc caacgaagaa	5940
tgtatcgtca aagtgatcca agcacactca tagttggagt cgtactccaa aggcggcaat	6000
gacgagtcag acagatactc gtcgacgttt aaacagtgtg cgcagatcta ctatagagga	6060
acatttaaat tgccccggag aagacggcca ggccgcctag atgacaaatt caacaactca	6120
cagctgactt tctgccattg ccactagggg ggggcctttt tataatggcca agccaagctc	6180
tccacgtcgg ttgggctgca cccaacaata aatgggtagg gttgcaccaa caaagggatg	6240
ggatgggggg tagaagatac gaggataacg gggctcaatg gcacaaataa gaacgaatac	6300
tgccattaag actcgtgatc cagcgactga caccattgca tcatctaagg gcctcaaac	6360
tacctcgaa ctgctgcgt gatctggaca ccacagaggt tccgagcact ttaggttgca	6420
ccaaatgtcc caccaggtgc aggcagaaaa cgctggaaca gcgtgtacag tttgtcttaa	6480
caaaaagtga gggcgtgag gtcgagcagg gtggtgtgac ttgttatagc ctttagagct	6540
gcgaaagcgc gtatggattt ggctcatcag gccagattga gggctctgtg acacatgtca	6600
tgttagtgta ctcaatcgc cccctggata tagccccgac aataggccgt gcctcattt	6660
ttttgccttc cgcacatttc cattgctcga taccacacc ttgcttctcc tgcacttgcc	6720
aaccttaata ctggtttaca ttgaccaaca tcttacaagc ggggggcttg tctagggtat	6780
atataaacag tggtctctcc aatcggttgc cagtctcttt tttctttct tccccacag	6840
attcgaaatc taaactacac atcacagaat tccgagccgt gagtatccac gacaagatca	6900
gtgtcgagac gacgcgtttt gtgtaatgac acaatccgaa agtcgctagc aacacacact	6960
ctctacaaa actaaccag ctctggtagc atggaggtcg tgaacgaaat cgtctccatt	7020
ggccaggagg ttcttccaa ggtcgactat gtgcagctct ggtctgatgc ctgcactgc	7080
gaggtgctgt acgggtccat cgccttcgtc atcctgaagt tcacccttgg tctctcgga	7140
cccaagggtc agtctcgaat gaagtttgtg ttaccaact acaacctgct catgtccatc	7200
tactcgtgg gctccttct ctctatggcc tacgcatgt acaccattgg tgtcatgtcc	7260
gacaactcgc agaaggcttt cgacaacaat gtcttccgaa tcaccactca gctgttctac	7320
ctcagcaagt tctcgagta cattgactcc ttctatctgc cctcatggg caagcctctg	7380
accaccttgc agttctttca ccatctcgga gctcctatgg acatgtggct gttctacaac	7440
taccgaaac aagccgtttg gatctttgtg ctgctcaacg gttcattca ctggatcatg	7500

tacggctact attgacccg actgcgaaag ctcaagttcc ctatgccaa gtccctgatt 7560
 atttctatgc agatcattca gttcaacgtt ggcttctaca tcgtctggaa gtaccggaac 7620
 attccctgct accgacaaga tggaatgaga atgtttggct ggtttttcaa ctacttctac 7680
 gttgggtactg tctgtgtct gttcctcaac ttctacgtga acacctacat cgtccgaaag 7740
 cacaaggag ccaaaaagat tcagtgagc 7769

<210> 380

<211> 5

<212> PRT

<213> Artificial Sequence

<220><223> HDGNH motif

<400> 380

His Asp Gly Asn His

1 5

<210> 381

<211> 5

<212> PRT

<213> Artificial Sequence

<220><223> HDANH motif

<400> 381

His Asp Ala Asn His

1 5

<210> 382

<211> 5

<212> PRT

<213> Artificial Sequence

<220><223> HEFAH motif

<400> 382

His Glu Phe Ala His

1 5

<210> 383

<211> 5

<212> PRT

<213> Artificial Sequence

<220><223> HEFTH motif

<400> 383

His Glu Phe Thr His

1 5

<210> 384

<211> 5

<212> PRT

<213> Artificial Sequence

<220><223> HEMGH motif

<400> 384

His Glu Met Gly His

1 5

<210> 385

<211> 5

<212> PRT

<213> Artificial Sequence

<220><223> HEAGH motif

<400> 385

His Glu Ala Gly His

1 5

<210> 386

<211> 5

<212> PRT

<213> Artificial sequence

<220><223> HDFGH motif

<400> 386

His Asp Phe Gly His

1 5

<210> 387

<211> 5

<212> PRT

<213> Artificial Sequence

<220><223> HDYGH motif

<400> 387

His Asp Tyr Gly His

1 5

<210> 388

<211> 5

<212> PRT

<213> Artificial Sequence

<220><223> HDSCH motif

<400> 388

His Asp Ser Cys His

1 5

<210> 389

<211> 5

<212> PRT

<213> Artificial Sequence

<220><223> HDACH motif

<400> 389

His Asp Ala Cys His

1 5

<210> 390

<211> 5

<212> PRT

<213> Artificial Sequence

<220><223> HDcSH motif

<400> 390

His Asp Cys Ser His

1 5

<210> 391

<211> 5

<212> PRT

<213> Artificial Sequence

<220><223> HDdSH motif

<400> 391

His Asp Asp Ser His

1 5

<210> 392

<211> 5

<212> PRT

<213> Artificial Sequence

<220><223> HDeSH motif

<400> 392

His Asp Glu Ser His

1 5

<210> 393

<211> 5

<212> PRT

<213> Artificial Sequence

<220><223> HDfSH motif

<400> 393

His Asp Phe Ser His

1 5

<210> 394

<211> 5

<212> PRT

<213> Artificial Sequence

<220><223> HDhSH motif

<400> 394

His Asp His Ser His

1 5

<210> 395

<211> 5

<212> PRT

<213> Artificial Sequence

<220><223> HDiSH motif

<400> 395

His Asp Ile Ser His

1 5

<210> 396

<211> 5

<212> PRT

<213> Artificial Sequence

<220><223> HDkSH motif

<400> 396

His Asp Lys Ser His

1 5

<210> 397

<211> 5

<212> PRT

<213> Artificial Sequence

<220><223> HDISH motif

<400> 397

His Asp Leu Ser His

1 5

<210> 398

<211> 5

<212> PRT

<213> Artificial Sequence

<220><223> HDmSH motif

<400> 398

His Asp Met Ser His

1 5

<210> 399

<211> 5

<212> PRT

<213> Artificial Sequence

<220><223> HDnSH motif

<400> 399

His Asp Asn Ser His

1 5

<210> 400

<211> 5

<212> PRT

<213> Artificial Sequence

<220><223> HDpSH motif

<400> 400

His Asp Pro Ser His

1 5

<210> 401

<211> 5

<212> PRT

<213> Artificial Sequence

<220><223> HDqSH motif

<400> 401

His Asp Gln Ser His

1 5

<210> 402

<211> 5

<212> PRT

<213> Artificial Sequence

<220><223> HDrSH motif

<400> 402

His Asp Arg Ser His

1 5

<210> 403

<211> 5

<212> PRT

<213> Artificial Sequence

<220><223> HDtSH motif

<400> 403

His Asp Thr Ser His

1 5

<210> 404

<211> 5

<212> PRT

<213> Artificial Sequence

<220><223> HDvSH motif

<400> 404

His Asp Val Ser His

1 5

<210> 405

<211> 5

<212> PRT

<213> Artificial Sequence

<220><223> HDwSH motif

<400> 405

His Asp Trp Ser His

1 5

<210> 406

<211> 5

<212> PRT

<213> Artificial Sequence

<220><223> HDySH motif

<400> 406

His Asp Tyr Ser His

1 5

<210> 407

<211> 5

<212> PRT

<213> Artificial Sequence

<220><223> HDAcH motif

<400> 407

His Asp Ala Cys His

1 5

<210> 408

<211> 6

<212> PRT

<213> Artificial Sequence

<220><223> HDAdH motif

<400> 408

Val His Asp Ala Asp His

1 5
 <210> 409
 <211> 5
 <212> PRT
 <213> Artificial Sequence
 <220><223> HD AeH motif
 <400> 409

His Asp Ala Glu His

1 5
 <210> 410
 <211> 5
 <212> PRT
 <213> Artificial Sequence
 <220><223> HD AfH motif
 <400> 410

His Asp Ala Phe His

1 5
 <210> 411
 <211> 5
 <212> PRT
 <213> Artificial Sequence
 <220><223> HD AhH motif
 <400> 411

His Asp Ala His His

1 5
 <210> 412
 <211> 5
 <212> PRT

<213> Artificial Sequence
 <220><223> HD AiH motif
 <400> 412

His Asp Ala Ile His

1 5
 <210> 413

<211> 5
 <212> PRT
 <213> Artificial Sequence
 <220><223> HDAkH motif
 <400> 413

His Asp Ala Lys His
 1 5

<210> 414
 <211> 5
 <212> PRT
 <213> Artificial Sequence
 <220><223> HDA1H motif
 <400> 414

His Asp Ala Leu His
 1 5

<210> 415
 <211> 5
 <212> PRT
 <213> Artificial Sequence
 <220><223> HDAmH motif
 <400> 415

His Asp Ala Met His
 1 5

<210> 416
 <211> 5
 <212> PRT
 <213> Artificial Sequence
 <220><223> HDAnH motif
 <400> 416

His Asp Ala Asn His
 1 5

<210> 417
 <211> 5
 <212> PRT

<213> Artificial Sequence

<220><223> HDApH motif

<400> 417

His Asp Ala Pro His

1 5

<210> 418

<211> 5

<212> PRT

<213> Artificial Sequence

<220><223> HDAqH motif

<400> 418

His Asp Ala Gln His

1 5

<210> 419

<211>

5

<212> PRT

<213> Artificial Sequence

<220><223> HDArH motif

<400> 419

His Asp Ala Arg His

1 5

<210> 420

<211> 5

<212> PRT

<213> Artificial Sequence

<220><223> HDAtH motif

<400> 420

His Asp Ala Thr His

1 5

<210> 421

<211> 5

<212> PRT

<213> Artificial Sequence

<220><223> HDAvH motif

<400> 421

His Asp Ala Val His

1 5

<210> 422

<211> 5

<212> PRT

<213> Artificial Sequence

<220><223> HDAwH motif

<400> 422

His Asp Ala Trp His

1 5

<210> 423

<211> 5

<212> PRT

<213> Artificial Sequence

<220><223> HDAyH motif

<400> 423

His Asp Ala Tyr His

1 5

<210> 424

<211> 5

<212> PRT

<213> Artificial Sequence

<220><223> HDxxH motif

<220><221> misc_feature

<222> (3)..(4)

<223> Xaa can be any naturally occurring amino acid

<400> 424

His Asp Xaa Xaa His

1 5

<210> 425

<211> 4

<212> PRT

<213> Artificial Sequence

<220><223> HGGG motif

<400> 425

His Gly Gly Gly

1

<210> 426

<211> 4

<212> PRT

<213> Artificial Sequence

<220><223> HHGG motif

<400> 426

His His Gly Gly

1

<210> 427

<211> 4

<212> PRT

<213> Artificial Sequence

<220><223> HPGS motif

<400> 427

His Pro Gly Ser

1

<210> 428

<211> 4

<212> PRT

<213> Artificial Sequence

<220><223> HAGG motif

<400> 428

His Ala Gly Gly

1

<210> 429

<211

> 5

<212> PRT

<213> Artificial Sequence

<220><223> HDGSH motif

<400> 429

His Asp Gly Ser His

1 5

<210> 430

<211> 5

<212> PRT

<213> Artificial Sequence

<220><223> HDSSH motif

<400> 430

His Asp Ser Ser His

1 5

<210> 431

<211> 5

<212> PRT

<213> Artificial Sequence

<220><223> HDAAH motif

<400> 431

His Asp Ala Ala His

1 5

<210> 432

<211> 5

<212> PRT

<213> Artificial Sequence

<220><223> HDAGH motif

<400> 432

His Asp Ala Gly His

1 5

<210> 433

<211> 5

<212> PRT

<213> Artificial Sequence

<220><223> HEASH motif

<400> 433

His Glu Ala Ser His

1 5

<210> 434

<211> 5

<212> PRT

<213> Artificial Sequence

<220><223> HDXSH motif

<220><221> misc_feature

<222> (3)..(3)

<223> Xaa can be any naturally occurring amino acid

<400> 434

His Asp Xaa Ser His

1 5

<210> 435

<211> 5

<212> PRT

<213> Artificial Sequence

<220><223> HDAXH motif

<220><221> misc_feature

<222> (4)..(4)

<223> Xaa can be any naturally occurring amino acid

<400> 435

His Asp Ala Xaa His

1 5