(19)대한민국특허청(KR)
(12) 공개특허공보(A)

(51) 。Int. Cl.
C12P 7/64 (2006.01)
C12N 15/74 (2006.01)

(11) 공개번호 10-2007-0085665
(43) 공개일자 2007년08월27일

(21) 출원번호 10-2007-7012473
(22) 출원일자 2007년06월01일
심사청구일자 없음
번역문 제출일자 2007년06월01일
(86) 국제출원번호 PCT/US2005/040256
국제출원일자 2005년11월03일
(87) 국제공개번호 WO 2006/052871
국제공개일자 2006년05월18일

(30) 우선권주장 60/624,812 2004년11월04일 미국(US)

(71) 출원인 이 아이 듀폰 디 네모아 앤드 캄파니
미합중국 데라웨아주 (우편번호 19898) 윌밍톤시 마아켓트 스트리이트 1007

(72) 발명자 다뮤드, 하워드, 지.
미국 19707 델라웨어주 호케신 켄윅 로드 4
길리즈, 피터, 존
미국 19350 펜실바니아주 란데너그 포트 로얄 코트 501
마쿨, 다니엘, 조셉
미국 19130 펜실바니아주 필라델피아 메러디스 스트리트 2525
피카타지오, 스티븐, 케이.
미국 19350 펜실바니아주 란데너그 메도우 우드 레인 17
라지안티, 제임스, 존
미국 19701 델라웨어주 베어 디사이디들리 레인 15
쉐, 지시옹
미국 19317-9728 펜실바니아주 채드스 포드 하비 레인 111
야다브, 나렌드라, 에스.
미국 19317 펜실바니아주 채드스 포드 크놀 레인 3704
장, 홍시앙
미국 19317 펜실바니아주 채드스 포드 로더필드 레인 3808
주, 퀸, 쿤
미국 19382 펜실바니아주 웨스트 체스터 리비에르 로드 544
세이프, 존, 이.
미국 08001 뉴 저지주 알로웨이 포스트 오피스 박스 129

(74) 대리인 김영
양영준

전체 청구항 수 : 총 46 항

(54) 도코사헥사엔산을 생성하는 야로위아 리폴리티카 균주

(57) 요약

본 발명에는 전체 오일 분획 중 5.6% 초과의 도코사헥사엔산 (DHA, ω-3 다불포화 지방산)을 생성할 수 있는 유질 효모 야로위아 리폴리티카의 유전자 조작된 균주가 기재되어 있다. 이 균주는 이종 데새투라제, 엘론가제 및 아실트랜스퍼라제를 발현하는 다양한 키메라 유전자를 포함하고, 임의로는 DHA의 합성 및 고축적이 가능한 다양한 천연 데새투라제 및 아실트랜스퍼라제 녹아웃을 포함한다. 상기 숙주 세포 내에서 DHA를 생성하는 방법으로서, 생성 숙주 세포가 청구된다.

대표도

도 1

특허청구의 범위

청구항 1.

ω-3/ω-6 지방산 생합성 경로의 유전자인

a) Δ6 데새투라제(desaturase)를 코딩하는 하나 이상의 유전자;

b) $C_{18/20}$ 엘론가제(elongase)를 코딩하는 하나 이상의 유전자;

c) Δ5 데새투라제를 코딩하는 하나 이상의 유전자;

d) Δ17 데새투라제를 코딩하는 하나 이상의 유전자;

e) $C_{20/22}$ 엘론가제를 코딩하는 하나 이상의 유전자; 및

f) Δ4 데새투라제를 코딩하는 하나 이상의 유전자

를 포함하는 유전자 풀(gene pool)을 포함하는 백그라운드 야로위아 종(*Yarrowia* sp.)을 포함하는, 도코사헥사엔산의 생성을 위한 재조합 생성 숙주 세포.

청구항 2.

ω-3/ω-6 지방산 생합성 경로의 유전자인

a) Δ15 데새투라제를 코딩하는 하나 이상의 유전자;

b) Δ6 데새투라제를 코딩하는 하나 이상의 유전자;

c) $C_{18/20}$ 엘론가제를 코딩하는 하나 이상의 유전자;

d) Δ5 데새투라제를 코딩하는 하나 이상의 유전자;

e) $C_{20/22}$ 엘론가제를 코딩하는 하나 이상의 유전자; 및

f) Δ4 데새투라제를 코딩하는 하나 이상의 유전자

를 포함하는 유전자 풀을 포함하는 백그라운드 야로위아 종을 포함하는, 도코사헥사엔산의 생성을 위한 재조합 생성 숙주 세포.

청구항 3.

ω-3/ω-6 지방산 생합성 경로의 유전자인

a) Δ9 엘론가제를 코딩하는 하나 이상의 유전자;

b) Δ8 데새투라제를 코딩하는 하나 이상의 유전자;

c) Δ5 데새투라제를 코딩하는 하나 이상의 유전자;

d) Δ17 데새투라제를 코딩하는 하나 이상의 유전자;

e) $C_{20/22}$ 엘론가제를 코딩하는 하나 이상의 유전자; 및

f) Δ4 데새투라제를 코딩하는 하나 이상의 유전자

를 포함하는 유전자 풀을 포함하는 백그라운드 야로위아 종을 포함하는, 도코사헥사엔산의 생성을 위한 재조합 생성 숙주 세포.

청구항 4.

ω-3/ω-6 지방산 생합성 경로의 유전자인

a) Δ15 데새투라제를 코딩하는 하나 이상의 유전자;

b) Δ9 엘론가제를 코딩하는 하나 이상의 유전자;

c) Δ8 데새투라제를 코딩하는 하나 이상의 유전자;

d) Δ5 데새투라제를 코딩하는 하나 이상의 유전자;

e) $C_{20/22}$ 엘론가제를 코딩하는 하나 이상의 유전자; 및

f) Δ4 데새투라제를 코딩하는 하나 이상의 유전자

를 포함하는 유전자 풀을 포함하는 백그라운드 야로위아 종을 포함하는, 도코사헥사엔산의 생성을 위한 재조합 생성 숙주 세포.

청구항 5.

제1항 내지 제4항 중 어느 한 항에 있어서, 상기 유전자 풀이 Δ12 데새투라제를 코딩하는 하나 이상의 유전자를 임의로 포함하는 것인 재조합 생성 숙주.

청구항 6.

제5항에 있어서, 상기 백그라운드 야로위아 종이 Δ12 데새투라제 활성을 갖는 폴리펩티드를 코딩하는 임의의 천연 유전자가 결여된 것인 재조합 생성 숙주.

청구항 7.

제1항 내지 제4항 중 어느 한 항에 있어서, 상기 ω-3/ω-6 지방산 생합성 경로 유전자 중 하나 이상이 서열 210 내지 221로 이루어진 군으로부터 선택된 핵산 서열을 갖는 프로모터 서열의 조절하에 있는 것인 재조합 생성 숙주.

청구항 8.

제5항에 있어서, 상기 Δ12 데새투라제가 서열 29, 31, 33, 35, 36, 37, 39, 41, 43, 45, 46, 48 내지 50으로 이루어진 군으로부터 선택된 아미노산 서열을 갖는 것인 재조합 생성 숙주.

청구항 9.

제1항 또는 제2항에 있어서, 상기 Δ6 데새투라제가 서열 2 및 5로 이루어진 군으로부터 선택된 아미노산 서열을 갖고; 상기 $C_{18/20}$ 엘론가제가 서열 23 및 26으로 이루어진 군으로부터 선택된 아미노산 서열을 갖는 것인 재조합 생성 숙주 세포.

청구항 10.

제3항 또는 제4항에 있어서, 상기 Δ9 엘론가제가 서열 70 및 23으로 이루어진 군으로부터 선택된 아미노산 서열을 갖고; 상기 Δ8 데새투라제가 서열 78, 80 및 82로 이루어진 군으로부터 선택된 아미노산을 갖는 것인 재조합 생성 숙주 세포.

청구항 11.

제1항 내지 제4항 중 어느 한 항에 있어서, 상기 Δ5 데새투라제를 코딩하는 하나 이상의 유전자가 효소성 기질로서

a) 리놀렌산 및 디호모-γ-리놀렌산;

b) α-리놀렌산 및 에이코사테트라엔산; 및

c) 리놀렌산, 디호모-γ-리놀렌산, α-리놀렌산 및 에이코사테트라엔산

으로 이루어진 군으로부터 선택된 2종 이상의 지방산에 결합하는 이작용성 Δ5/Δ6 데새투라제 폴리펩티드를 코딩하는 것인 재조합 생성 숙주.

청구항 12.

제11항에 있어서, 상기 이작용성 Δ5/Δ6 데새투라제 폴리펩티드가 서열 15 및 18로 이루어진 군으로부터 선택된 아미노산 서열을 갖는 것인 재조합 생성 숙주.

청구항 13.

제1항 내지 제4항 중 어느 한 항에 있어서, 상기 Δ5 데새투라제가 서열 7, 9, 12, 15 및 18로 이루어진 군으로부터 선택된 아미노산 서열을 갖고; 상기 $C_{20/22}$ 엘론가제가 서열 101 및 103으로 이루어진 군으로부터 선택된 아미노산 서열을 갖고; 상기 Δ4 데새투라제가 서열 105 및 107로 이루어진 군으로부터 선택된 아미노산 서열을 갖는 것인 재조합 생성 숙주 세포.

청구항 14.

제1항 또는 제3항에 있어서, 상기 Δ17 데새투라제가 서열 20에 기재된 아미노산 서열을 갖는 것인 재조합 생성 숙주 세포.

청구항 15.

제1항 또는 제3항에 있어서, 상기 유전자 풀이 Δ15 데새투라제를 코딩하는 하나 이상의 유전자를 임의로 포함하는 것인 재조합 생성 숙주 세포.

청구항 16.

제2항 또는 제4항에 있어서, 상기 Δ15 데새투라제를 코딩하는 상기 하나 이상의 유전자가 효소성 기질로서의 올레산 및 리놀레산 둘다에 결합하는 이작용성 Δ15/Δ12 데새투라제 폴리펩티드를 코딩하는 것인 재조합 생성 숙주.

청구항 17.

제16항에 있어서, 상기 이작용성 Δ15/Δ12 데새투라제 폴리펩티드가 서열 52에 기재된 아미노산 서열을 갖는 것인 재조합 생성 숙주.

청구항 18.

제2항 또는 제4항에 있어서, 상기 Δ15 데새투라제가 서열 52, 54, 56, 58, 60, 62 및 64 내지 68로 이루어진 군으로부터 선택된 아미노산 서열을 갖는 것인 재조합 생성 숙주.

청구항 19.

제1항 내지 제4항 중 어느 한 항에 있어서, 상기 유전자 풀이 임의로

a) Δ9 데새투라제를 코딩하는 하나 이상의 유전자;

b) $C_{16/18}$ 엘론가제를 코딩하는 하나 이상의 유전자; 및

c) $C_{14/16}$ 엘론가제를 코딩하는 하나 이상의 유전자

로 이루어진 군으로부터 선택된 ω-3/ω-6 지방산 생합성 경로 유전자를 포함하는 것인 재조합 생성 숙주.

청구항 20.

제19항에 있어서, 상기 $C_{16/18}$ 엘론가제가 서열 84, 87 및 95로 이루어진 군으로부터 선택된 아미노산 서열을 갖고; 상기 $C_{14/16}$ 엘론가제가 서열 98에 기재된 아미노산 서열을 갖는 것인 재조합 생성 숙주.

청구항 21.

제1항 내지 제4항 중 어느 한 항에 있어서, 상기 유전자 풀이 임의로

a) 디아실글리세롤 아실트랜스퍼라제 (DGAT1);

b) 디아실글리세롤 아실트랜스퍼라제 (DGAT2);

c) 인지질:디아실글리세롤 아실트랜스퍼라제 (PDAT);

d) 아실-CoA:1-아실 리소포스파티딜콜린 아실트랜스퍼라제 (LPCAT);

e) 글리세롤-3-포스페이트 아실트랜스퍼라제 (GPAT); 및

f) 리소포스파티드산 아실트랜스퍼라제 (LPAAT)

로 이루어진 군으로부터 선택된 아실트랜스퍼라제를 코딩하는 하나 이상의 유전자를 포함하는 것인 재조합 생성 숙주.

청구항 22.

제21항에 있어서, 상기 디아실글리세롤 아실트랜스퍼라제 (DGAT1)가 서열 123 및 125 내지 129로 이루어진 군으로부터 선택된 아미노산 서열을 갖고; 상기 디아실글리세롤 아실트랜스퍼라제 (DGAT2)가 서열 131, 133, 135 및 137로 이루어진 군으로부터 선택된 아미노산 서열을 갖고; 상기 인지질:디아실글리세롤 아실트랜스퍼라제 (PDAT)가 서열 118에 기재된 아미노산 서열을 갖고; 상기 글리세롤-3-포스페이트 아실트랜스퍼라제 (GPAT)가 서열 139에 기재된 아미노산 서열을 갖고; 상기 리소포스파티드산 아실트랜스퍼라제 (LPAAT)가 서열 109, 111, 113 및 116으로 이루어진 군으로부터 선택된 아미노산 서열을 갖고; 상기 아실-CoA:1-아실 리소포스파티딜콜린 아실트랜스퍼라제 (LPCAT)가 서열 121에 기재된 아미노산 서열을 갖는 것인 재조합 생성 숙주.

청구항 23.

ω-3/ω-6 지방산 생합성 경로의 유전자인

a) Δ6 데새투라제를 코딩하는 하나 이상의 유전자;

b) $C_{18/20}$ 엘론가제를 코딩하는 하나 이상의 유전자;

c) Δ5 데새투라제를 코딩하는 하나 이상의 유전자;

d) Δ17 데새투라제를 코딩하는 하나 이상의 유전자;

e) $C_{20/22}$ 엘론가제를 코딩하는 하나 이상의 유전자;

f) Δ4 데새투라제를 코딩하는 하나 이상의 유전자;

g) $C_{16/18}$ 엘론가제를 코딩하는 하나 이상의 유전자; 및

h) Δ12 데새투라제를 코딩하는 하나 이상의 유전자

를 포함하는 유전자 풀을 포함하는 백그라운드 야로위아 종을 포함하고;

상기 백그라운드 야로위아 종이 오로티딘-5'-포스페이트 데카르복실라제 (Ura3), 이소프로필 말레이트 데히드로게나제 (Leu2), 아실-CoA 옥시다제 (Pox3), 아실-CoA 옥시다제 (Pox2), Δ12 데새투라제 및 리파제 1 (Lip1)로 이루어진 군으로부터 선택된 효소를 코딩하는 임의의 천연 유전자가 결여된 것인,

도코사헥사엔산 생성을 위한 재조합 생성 숙주 세포.

청구항 24.

ω-3/ω-6 지방산 생합성 경로의 유전자인

a) i. Δ6 데새투라제를 코딩하는 하나 이상의 유전자; $C_{18/20}$ 엘론가제를 코딩하는 하나 이상의 유전자; Δ5 데새투라제를 코딩하는 하나 이상의 유전자; Δ17 데새투라제를 코딩하는 하나 이상의 유전자; $C_{20/22}$ 엘론가제를 코딩하는 하나 이상의 유전자; 및 Δ4 데새투라제를 코딩하는 하나 이상의 유전자;

ii. Δ9 엘론가제를 코딩하는 하나 이상의 유전자; Δ8 데새투라제를 코딩하는 하나 이상의 유전자; Δ5 데새투라제를 코딩하는 하나 이상의 유전자; Δ17 데새투라제를 코딩하는 하나 이상의 유전자; $C_{20/22}$ 엘론가제를 코딩하는 하나 이상의 유전자; 및 Δ4 데새투라제를 코딩하는 하나 이상의 유전자;

iii. Δ15 데새투라제를 코딩하는 하나 이상의 유전자; Δ6 데새투라제를 코딩하는 하나 이상의 유전자; $C_{18/20}$ 엘론가제를 코딩하는 하나 이상의 유전자; Δ5 데새투라제를 코딩하는 하나 이상의 유전자; $C_{20/22}$ 엘론가제를 코딩하는 하나 이상의 유전자; 및 Δ4 데새투라제를 코딩하는 하나 이상의 유전자; 및

iv. Δ15 데새투라제를 코딩하는 하나 이상의 유전자; Δ9 엘론가제를 코딩하는 하나 이상의 유전자; Δ8 데새투라제를 코딩하는 하나 이상의 유전자; Δ5 데새투라제를 코딩하는 하나 이상의 유전자; $C_{20/22}$ 엘론가제를 코딩하는 하나 이상의 유전자; 및 Δ4 데새투라제를 코딩하는 하나 이상의 유전자

로 이루어진 군으로부터 선택된 하나 이상의 유전자 세트;

b) i. Δ12 데새투라제;

ii. Δ9 데새투라제;

iii. $C_{14/16}$ 엘론가제; 및

iv. $C_{16/18}$ 엘론가제

로 이루어진 군으로부터 선택된 효소를 코딩하는 하나 이상의 유전자; 및

c) (i) DGAT1, DGAT2 및 PDAT로 이루어진 군으로부터 선택된 디아실글리세롤 아실트랜스퍼라제;

(ii) 아실-CoA:1-아실 리소포스파티딜콜린 아실트랜스퍼라제 (LPCAT);

(iii) 글리세롤-3-포스페이트 아실트랜스퍼라제 (GPAT);

(iv) 리소포스파티드산 아실트랜스퍼라제 (LPAAT);

(v) 포스포리파제 C; 및

(vi) 포스포리파제 A_2

로 이루어진 군으로부터 선택된 효소를 코딩하는 하나 이상의 유전자

를 포함하는 유전자 풀을 포함하는 백그라운드 야로위아 종을 포함하며,

여기서,

(1) 상기 백그라운드 야로위아 종이 Δ12 데새투라제 활성을 갖는 폴리펩티드를 코딩하는 임의의 천연 유전자가 결여되고;

(2) 상기 백그라운드 야로위아 종이 리파제 1 (Lip1-), 퍼옥시좀 아실 CoA 옥시다제 ACO3 (Pox3-), 아실-CoA 옥시다제 2 (Pox2-), 오로티딘-5'-포스페이트 데카르복실라제 (Ura3-), 사카로핀 데히드로게나제 (Lys5-), 리파제 2 (Lip2-) 및 이소프로필 말레이트 데히드로게나제 (Leu2-)로 이루어진 군으로부터 선택된 효소를 코딩하는 임의의 천연 유전자가 결여된 것인,

도코사헥사엔산 생성을 위한 재조합 생성 숙주 세포.

청구항 25.

제1항 내지 제4항 중 어느 한 항에 있어서, 전체 지방산의 백분율로서 약 5% 이상의 도코사헥사엔산을 포함하는 미생물성 오일을 생성하는 재조합 생성 숙주.

청구항 26.

제25항에 있어서, 도코사헥사엔산을 포함하는 미생물성 오일을 생성하고, 상기 미생물성 오일이 임의의 γ-리놀레산이 결여된 것인 재조합 생성 숙주.

청구항 27.

a) 도코사헥사엔산을 포함한 미생물성 오일을 생성하는, 제1항 내지 제4항 중 어느 한 항에 따른 생성 숙주를 배양하는 단계; 및

b) 임의로 단계 a)의 미생물성 오일을 회수하는 단계

를 포함하는, 도코헥사엔산을 포함한 미생물성 오일을 생성하는 방법.

청구항 28.

제27항의 방법에 의해 생성된 미생물성 오일.

청구항 29.

제28항에 있어서, 약 5% 이상의 도코사헥사엔산을 함유하는 미생물성 오일.

청구항 30.

제28항에 있어서, 임의의 γ-리놀레산이 결여된 미생물성 오일.

청구항 31.

제28항에 있어서, 리놀레산, γ-리놀렌산, 에이코사디엔산, 디호모-γ-리놀레산, 아라키돈산, α-리놀렌산, 스테아리돈산, 에이코사트리엔산, 에이코사테트라엔산, 도코사펜타엔산 및 도코사헥사엔산으로 이루어진 군으로부터 선택된 지방산을 포함하는 블렌딩 오일.

청구항 32.

제27항의 방법에 의해 생성된 유효량의 미생물성 오일을 포함하는 식품.

청구항 33.

제32항에 있어서, 식품 유사물, 육류 제품, 곡물 제품, 베이킹 식품, 스낵 식품 및 유제품으로 이루어진 군으로부터 선택된 식품.

청구항 34.

제27항의 방법에 의해 생성된 유효량의된 미생물성 오일을 포함하는, 의료용 식품, 식이 보조제, 유아용 조제식 및 약제로 이루어진 군으로부터 선택된 제품.

청구항 35.

제27항의 방법에 의해 생성된 유효량의 미생물성 오일을 포함하는 동물 사료.

청구항 36.

제35항에 있어서, 애완 동물 사료, 반추 동물 사료, 가금류 사료 및 양식용 사료로 이루어진 군으로부터 선택된 동물 사료.

청구항 37.

유효량의 미생물성 오일을 포함하고, 임의로 제1항 내지 제4항 중 어느 한 항의 재조합 숙주를 포함하는 효모 생물체를 포함하는 동물 사료.

청구항 38.

제37항에 있어서, 상기 효모 생물체가 단백질, 지질, 탄수화물, 비타민, 미네랄 및 핵산으로 이루어진 군으로부터 선택된 사료 영양분을 포함하는 것인 동물 사료.

청구항 39.

제27항의 방법에 의해 생성된 미생물성 오일과 식품을 배합하는 것을 포함하는, 도코사헥사엔산이 보충된 식품의 제조 방법.

청구항 40.

제27항의 방법에 의해 생성된 미생물성 오일과, 의료용 식품, 식이 보조제, 유아용 조제식 및 약제로 이루어진 군으로부터 선택된 제품을 배합하는 것을 포함하는, 도코사헥사엔산이 보충된 상기 제품의 제조 방법.

청구항 41.

제27항의 방법에 의해 생성된 미생물성 오일과 동물 사료를 배합하는 것을 포함하는, 도코사헥사엔산이 보충된 동물 사료의 제조 방법.

청구항 42.

제37항의 동물 사료와, 사료 영양분을 포함하는 효모 생물체를 배합하는 것을 포함하는, 도코헥사엔산을 포함하는 동물 사료에 사료 영양분을 보충하는 방법.

청구항 43.

제42항에 있어서, 상기 사료 영양분이 단백질, 지질, 탄수화물, 비타민, 미네랄 및 핵산으로 이루어진 군으로부터 선택된 것인 방법.

청구항 44.

제27항의 방법에 의해 생성된 도코사헥사엔산 함유 미생물성 오일을 인간 또는 동물이 소모가능하거나 사용가능한 형태로 제공하는 것을 포함하는, 에이코사펜타엔산이 풍부한 식이 보충제를 인간, 동물 또는 양식 유기체에게 제공하는 방법.

청구항 45.

제27항의 방법에 의해 생성된 도코사헥사엔산 함유 미생물성 오일을 인간 또는 동물이 소모가능하거나 사용가능한 형태로 제공하여 도코사헥사엔산의 결핍을 치료하는 것을 포함하는, 동물 또는 인간에서 도코사헥사엔산의 결핍을 치료하는 방법.

청구항 46.

ATCC 기탁번호 ATCC PTA-7187의 야로위아 리폴리티카(*Yarrowia lipolytica*) Y3000인, 도코사헥사엔산의 생성에 유용한 재조합 생성 숙주.

명세서

기술분야

본 발명은 생명공학 분야에 속한다. 보다 구체적으로, 본 발명은 고농도의 도코사헥사엔산 (ω-3 다불포화 지방산)을 생성할 수 있는 유질 효모 야로위아 리폴리티카(*Yarrowia lipolytica*)의 유전자 조작된 균주에 관한 것이다.

배경기술

도코사헥사엔산 (DHA; 시스-4,7,10,13,16,19-도코사헥사엔산; C22:6 ω-3)은 뇌 기능의 성장, 기능 발달 및 건강 유지에 필수적이며, 요람기로부터 노화까지 인생 전반에 요구된다 [Horrocks, L.A. and Y.K. Yeo. Pharmacol. Res. 40(3) :211-225 (1999)]. DHA 결핍은 태아기 알콜 증후군, 주의력 결핍 과잉활동 장애, 낭포성 섬유증, 페닐케톤뇨증, 단극적 우울증, 공격성 적대성 및 부신백질이영양증과 연관된다. 반대로, 증가된 DHA 섭취는 유익함을 나타내거나, 염증성 장애 (예를 들어, 류마티스 관절염), II형 당뇨병, 고혈압, 아테롬성 동맥경화증, 우울증, 심근경색증, 혈전증, 일부 암에서 유리한 효과를 갖고, 알쯔하이머병(Alzheimer's disease)과 같은 퇴행성 장애의 발병을 예방하는 데 유리한 효과를 갖는다.

어류 (예를 들어, 연어, 송어, 고등어)는 이들이 DHA와 같은 장쇄 지방산을 천연적으로 고농도로 함유하기 때문에 DHA의 중요한 공급원이다. 많은 연구 [미국 정부가 발행한 2005년 식사 가이드라인 자문 위원회 보고서(2005 Dietary Guidelines Advisory Committee Report for Americans, part D, section 4 (coordinated by the U.S. Dept. of Health & Human Services and the U.S. Dept. of Agriculture))에 보고됨]에 근거하여, 미국 심장 협회(the American Heart Association), 국제 콜레스테롤 교육 프로그램(the National Cholesterol Education Program), 세계 건강 기구(the World Health Association), 유럽 심장 학회(the European Society for Cardiology), 미국 당뇨 협회(the American Diabetes Association) 및 영양분에 대한 영국 과학 자문 위원회(the United Kingdom Scientific Advisory Committee on Nutrition)는 모두 이렇게 시사된 심장보호 효과를 위해 1주에 2회 식사의 생선 (각 식사 시 하루에 DHA 및 에이코사펜타엔산 (EPA, C20:5 ω-3) 약 450 mg 정도의 동량을 제공함)을 권장한다. 이와 같이, DHA는 기능성 식품, 유아용 영양분, 벌크 영양소 및 동물 건강과 관련된 다양한 제품에 혼입된다.

도코사펜타엔산 (DPA, C22:5 ω-3)의 생리학적 기능은 아직 공지되어 있지 않지만, 이 지방산는 DHA의 대사 전구체이고 EPA를 연장시킨 바로 하류 생성물이다. DPA는 또한 그 함량이 매우 낮기는 하지만 어유에도 함유된 것으로 공지되어 있다. DPA에 대해 단지 공지된 기능은 제약 제제를 뇌로 운반하기 위해 담체로서 유용성을 갖는다는 것이다 (일본 특허 공개 (공개) 제61-204136호 (1986)). 그러나, DHA 결핍을 보충할 때 DPA가 증가하는 것으로 공지되어 있으므로, DPA는 동물체 내 생리학적 역할을 수행할 수 있는 것으로 예상된다 [Homayoun et al., J. Neurochem., 51 :45 (1988); Hamm et al., Biochem. J., 245:907 (1987); 및 Rebhung et al., Biosci. Biotech. Biochem., 58:314 (1994)]. 따라서, DPA 및 DHA 모두는 중요한 ω-3 지방산으로서 여겨져야만 한다. 당업자는 DHA에 대한 본원에서의 교시가 또한 광범위하게 적용가능하고 DPA 제조에 관한 본원의 교시가 미래의 목적하는 생성물이 될 것임을 인지할 것이다.

DHA는 여러 유형의 어유 및 해양 플랑크톤에서 천연적으로 발견되지만, 이러한 ω-3 지방산의 공급은 늘어나는 수요를 충족시키기에는 불충분할 것이라고 예측된다. 어유는 매우 불균질한 조성 (이 때문에, EPA를 풍부하게 하기 위해서는 과도한 정제가 요구됨), 불쾌한 맛 및 냄새 (이는 경제적으로 제거하기가 어렵고, 어유가 식품 성분으로서 허용되지 않게 함)를 가지며, 중금속 오염물의 환경적인 생체 축적 및 (기후, 질환 또는 남획으로 인한) 입수성의 변동이 쉽게 일어난다.

어유에 대한 대체물로서, DHA는 또한 미생물에서 생성될 수 있다. 일반적으로, 미생물성 오일 생성은 오일 합성 (이는 통상적인 세포 대사의 과정에서 일어남)을 가능하게 하는 적합한 배양 배지에서 EPA를 천연 합성할 수 있는 적절한 미생물을 배양한 후, 발효 배지로부터 미생물을 분리하고, 세포내 오일의 회수를 위한 처리하는 것을 포함한다. 특이적인 미생물 유기체 [예를 들어, 시조키트륨(*Schizochytrium*) 종 (미국 특허 제5,340,742호; 미국 특허 제6,582,941호); 울케니아 (*Ulkenia*) (미국 특허 제6,509,178호); 수도모나스(*Pseudomonas*) 종 YS-180 (U.S. 6,207,441); 트라우스토키트륨 (*Thraustochytrium*) 속 균주 LFF1 (미국 특허 제2004/0161831 A1호); 크립테코디늄 코니(*Crypthecodinium cohnii*) (미국 특허 제2004/0072330 A1호; 문헌 [de Swaaf, M. E. et al. Biotechnol Bioeng. 81 (6):666-72 (2003)] 및 [Appl Microbiol Biotechnol. 61 (1):40-3 (2003)]); 에밀리아니아(*Emiliania*) 종 (일본 특허 출원 (공개) 제5-308978호 (1993)); 및 자포노키트륨(*Japonochytrium*) 종 (ATCC #28207; 일본 특허 출원 (공개) 제199588/1989호)]를 기초로 하는 여러 수많은 방법이 존재한다. 추가로, 하기 미생물: 비브리오 마리누스(*Vibrio marinus*) (심해로부터 단리된 박테리아; ATCC #15381); 미세조류 시클로텔라 크립티카(*Cyclotella cryptica*) 및 이소크리시스 갈바나(*Isochrysis galbana*); 및 편모충 곰팡이, 예를 들어 트라우스토키트륨 아우레움(*Thraustochytrium aureum*) (ATCC #34304; 문헌 [Kendrick, Lipids, 27:15 (1992)]), 및 ATCC #28211, ATCC #20890 및 ATCC #20891로 지칭되는 트라우스토키트륨 종은 DHA를 생산하는 능력이 있는 것으로 공지되어 있다. 그리고, 이들 몇몇 방법이 다양한 제한의 결과로서 산업적 사용성에 대하

여 허용되지 않더라도, DHA의 상업적 제조를 위한 적어도 3가지의 상이한 발효 과정: DHASCO™ (매릴랜드주 콜롬비아 소재의 마르텍 바이오사이언시스 코포레이션(Martek Biosciences Corporation)) 제조용 크립테코디늄 코니의 발효; DHAGold (마르텍 바이오사이언시스 코포레이션)로서 종래 공지된 오일 제조용 시조키트륨 종의 발효; 및 DHActive™ (독일 프랑크푸르트 소재의 뉴트리노바(Nutrinova)) 제조용 울케니아 종의 발효가 존재한다. 이러한 성공에도 불구하고, 이들 각각의 방법은 발효가 미생물 자체의 천연 능력에 의존적이기 때문에 생성된 오일 조성물의 특징을 조절하거나 또는 오일의 수율을 실질적으로 개선시키는 능력이 없어서 적용하기 어렵다.

따라서, 재조합 수단을 이용하는 DHA의 미생물 생성은 천연 미생물 공급원으로부터의 생성에 비해 여러 이점이 있을 것으로 예상된다. 예를 들어, 숙주의 천연 발생 미생물 지방산 프로파일은 숙주에서 새로운 생합성 경로의 도입 및/또는 원치않는 경로의 저해에 의해 변경되어, 원하는 PUFA (또는 그의 접합된 형태)의 생성 수준을 증가시키고, 원치않는 PUFA의 생성을 감소시킬 수 있기 때문에, 오일 생성을 위한 바람직한 특징을 갖는 재조합 미생물을 사용할 수 있다. 둘째로, 재조합 미생물은 특정 용도를 가질 수 있는 특정한 형태의 PUFA를 제공할 수 있다. 그리고, 마지막으로, 미생물성 오일 생성은 배양 조건의 조절에 의해, 현저하게는 미생물에서 발현되는 효소를 위한 특정한 기질 공급원을 제공함으로써, 또는 화합물 첨가/유전자 조작에 의해 원치않는 생화학적 경로를 저해함으로써 조작될 수 있다. 따라서, 예를 들어 그렇게 생성된 ω-3 지방산과 ω-6 지방산의 비율을 변형시킬 수 있거나, 다른 PUFA 하류 또는 상류 생성물의 유의한 축적이 없이 특이적 PUFA (예컨대, DHA)의 생성을 유전자 조작할 수 있다.

DHA의 미생물성 생성은 먼저 중간체 지방산인 EPA의 합성이 요구된다. 그리고, 미생물에서 생성되는 대부분의 DHA는 Δ6 데새투라제(desaturase)/Δ6 엘론가제(elongase) 경로 (조류, 이끼류, 진균, 선충류 및 인간에서 우세하게 발견됨)를 통해 합성되며, 여기서, 1) Δ12 데새투라제의 작용에 의해 올레산이 LA로 전환되고, 2) 임의로, Δ15 데새투라제의 작용에 의해 LA는 ALA로 전환되고, 3) Δ6 데새투라제의 작용에 의해 LA는 GLA로 전환되고/되거나 ALA는 STA로 전환되고, 3) $C_{18/20}$ 엘론가제의 작용에 의해 GLA가 DGLA로 전환되고/되거나 STA가 ETA로 전환되고, 3) Δ5 데새투라제의 작용에 의해 DGLA가 ARA로 전환되고/되거나 ETA가 EPA로 전환되고, 4) 임의로, Δ17 데새투라제의 작용에 의해 ARA가 EPA로 전환된다 (도 1). 그러나, EPA의 생합성을 위한 별법의 Δ9 엘론가제/Δ8 데새투라제 경로는 일부 유기체, 예컨대 유글레나 종에서 작동하며, 이는 C_{20} PUFA의 형성을 위한 주요 경로이다 (문헌 [Wallis, J.G., and Browse, J. Arch. Biochem. Biophys. 365:307-316 (1999)]; 제WO 00/34439호 및 [Qi, B. et al. FEBS Letters. 510:159-165 (2002)]). 이 경로에서는, 1) Δ9 엘론가제에 의해 LA 및 ALA가 각각 EDA 및 ETrA로 전환되고, 2) Δ8 데새투라제에 의해 EDA 및 ETrA가 각각 DGLA 및 ETA로 전환되고, 3) 상기 기재된 바와 같이 DGLA 및 ETA가 궁극적으로 EPA로 전환된다. EPA의 합성에 따라, $C_{20/22}$ 엘론가제는 기질을 DPA로 전환시킨 다음, Δ4 데새투라제에 의해 탈포화시켜 DHA를 수득하는 데 작용한다.

상기 문헌은 ω-3/ω-6 PUFA 생합성 경로의 다양한 부분을 사카로마이세스 세레비지애(*Saccharomyces cerevisiae*) (비-유질 효모)에 도입한 수많은 최근의 예시적인 방법을 보고한다. 구체적으로, 문헌 [Dyer, J. M. et al., Appl. Eniv. Microbiol., 59:224-230 (2002)]은 리놀렌산의 합성을 보고하였고, 넛존(Knutzon) 등의 미국 특허 제6,136,574호는 리놀렌산 (LA), γ-리놀렌산 (GLA), ALA 및 스테아리돈산 (STA)의 생성을 유도하였고, 문헌 [Domergue, F. et al., Eur. J. Biochem. 269:4104-4113 (2002)]는 EPA의 생성을 기재하였고, 문헌 [Pereira, S. L. et al., Biochem. J. 384:357-366 (2004)]는 최초로 DHA를 생성하였다 (EPA가 기질로서 공급되는 경우 전체 지방산의 3.8%). 그러나, 이러한 성공에도 불구하고, DHA의 시판량 (즉, 전체 지방산에 대하여 5 내지 30% 초과)을 경제적으로 생성할 수 있는 복잡한 대사성 유전자 조작은 보고된 바가 없었다. 추가로, 이러한 조작을 위해 가장 적절한 숙주 유기체의 선택과 관련하여 상당한 편차가 존재한다.

최근에, 피카타지오(Picataggio) 등의 제WO 2004/101757호 및 동시-계류중인 미국 특허 출원 제60/624812호에서는 PUFA, 예컨대 ARA, EPA 및 DHA의 생성을 위한 바람직한 미생물 부류로서 유질 효모, 구체적으로 야로위아 리폴리티카 (이전에는 칸디다 리폴리티카(*Candida lipolytica*)로서 분류됨)의 유용성에 대해 연구하였다. 유질 효모는 천연적인 오일 합성 및 축적이 가능한 효모로서 정의되며, 여기서, 오일 축적은 세포 건조 중량의 약 80%까지일 수 있다. 이들 유기체는 ω-6 및 ω-3 지방산 생성이 천연적으로 결핍되어 있음에도 불구하고 (천연적으로 생성되는 PUFA는 18:2 지방산 (및 덜 통상적으로는, 18:3 지방산)으로 제한되기 때문), 피카타지오 등의 상기 문헌에서는 비교적 간단한 유전자 조작 접근법을 이용하여 야로위아 리폴리티카에서 (전체 지방산의) 1.3% ARA 및 1.9% EPA의 생성을 입증하였고, 더 복잡한 대사성 유전자 조작을 이용하여 28% 이하의 EPA 생성을 입증하였다. 그러나, 상기 특정한 숙주 유기체에서 더욱 복잡한 대사성 유전자 조작에 의해 DHA를 경제적이고 상업적으로 생성한 적은 없다.

본 출원인은 Δ6 데새투라제/Δ6 엘론가제 경로를 이용하여 전체 오일 분획 중 5% 초과의 DHA를 생성할 수 있는 다양한 야로위아 리폴리티카 균주의 유전자 조작에 의해 상기 문제점을 해결하였다. 이들 유질 효모에서 DHA 생성률을 더욱 증가시키기 위해 추가의 대사성 유전자 조작 및 발효 방법뿐만 아니라, Δ9 엘론가제/Δ8 데새투라제 경로를 통해 DHA를 생성할 수 있는 방법 (이에 따라 GLA가 결여된 DHA-함유 오일을 생성함)을 제공한다.

<발명의 요약>

본 발명은 도코사헥사엔산 (DHA)을 생성하기 위해 유전자 조작된 재조합 생성 숙주, 이를 제조하는 방법 및 본 발명의 제조합 숙주에 의해 생성된 미생물성 오일을 함유한 식품 공급 제품에 관한 것이다.

따라서, 본 발명의 한 실시양태는, ω-3/ω-6 지방산 생합성 경로의 유전자인

a) Δ6 데새투라제를 코딩하는 하나 이상의 유전자;

b) $C_{18/20}$ 엘론가제를 코딩하는 하나 이상의 유전자;

c) Δ5 데새투라제를 코딩하는 하나 이상의 유전자;

d) Δ17 데새투라제를 코딩하는 하나 이상의 유전자;

e) $C_{20/22}$ 엘론가제를 코딩하는 하나 이상의 유전자; 및

f) Δ4 데새투라제를 코딩하는 하나 이상의 유전자

를 포함하는 유전자 풀을 포함하는 백그라운드 야로위아 종을 포함하는, 도코사헥사엔산의 생성을 위한 재조합 생성 숙주 세포를 제공한다.

또다른 실시양태에서, 본 발명은 ω-3/ω-6 지방산 생합성 경로의 유전자인

a) Δ15 데새투라제를 코딩하는 하나 이상의 유전자;

b) Δ6 데새투라제를 코딩하는 하나 이상의 유전자;

c) $C_{18/20}$ 엘론가제를 코딩하는 하나 이상의 유전자;

d) Δ5 데새투라제를 코딩하는 하나 이상의 유전자;

e) $C_{20/22}$ 엘론가제를 코딩하는 하나 이상의 유전자; 및

f) Δ4 데새투라제를 코딩하는 하나 이상의 유전자

를 포함하는 유전자 풀을 포함하는 백그라운드 야로위아 종을 포함하는, 도코사헥사엔산 생성을 위한 재조합 생성 숙주 세포를 제공한다.

또다른 실시양태에서, 본 발명은 ω-3/ω-6 지방산 생합성 경로의 유전자인

a) Δ9 엘론가제를 코딩하는 하나 이상의 유전자;

b) Δ8 데새투라제를 코딩하는 하나 이상의 유전자;

c) Δ5 데새투라제를 코딩하는 하나 이상의 유전자;

d) Δ17 데새투라제를 코딩하는 하나 이상의 유전자;

e) $C_{20/22}$ 엘론가제를 코딩하는 하나 이상의 유전자; 및

f) Δ4 데새투라제를 코딩하는 하나 이상의 유전자

를 포함하는 유전자 풀을 포함하는 백그라운드 야로위아 종을 포함하는, 도코사헥사엔산 생성을 위한 재조합 생성 숙주 세포를 제공한다.

별법의 실시양태에서, 본 발명은 ω-3/ω-6 지방산 생합성 경로의 유전자인

a) Δ15 데새투라제를 코딩하는 하나 이상의 유전자;

b) Δ9 엘론가제를 코딩하는 하나 이상의 유전자;

c) Δ8 데새투라제를 코딩하는 하나 이상의 유전자;

d) Δ5 데새투라제를 코딩하는 하나 이상의 유전자;

e) $C_{20/22}$ 엘론가제를 코딩하는 하나 이상의 유전자; 및

f) Δ4 데새투라제를 코딩하는 하나 이상의 유전자

를 포함하는 유전자 풀을 포함하는 백그라운드 야로위아 종을 포함하는, 도코사헥사엔산 생성을 위한 재조합 생성 숙주 세포를 제공한다.

바람직한 실시양태에서, 본 발명의 생성 숙주는 임의로는 Δ12 데새투라제를 코딩하는 하나 이상의 유전자를 포함한다. 또다른 바람직한 실시양태에서, 본 발명은 전체 지방산에 대한 백분율로서 약 5% 이상의 도코사헥사엔산을 갖는 미생물성 오일을 생성하는 제조합 생성 숙주를 제공한다.

또다른 실시양태에서, 본 발명은

a) 도코사헥사엔산을 포함하는 미생물성 오일을 생성하는 본 발명의 생성 숙주를 배양하는 단계; 및

b) 임의로 단계 a)의 미생물성 오일을 회수하는 단계

를 포함하는, 도코사헥사엔산을 포함하는 미생물성 오일을 생성하는 방법을 제공한다.

또다른 실시양태에서, 본 발명은 상기 방법으로 제조된 미생물성 오일 및 본 발명의 생성 숙주를 제공한다. 바람직한 실시양태에서, 본 발명의 미생물성 오일은 DHA를 함유하지만 임의의 γ-리놀레산은 함유하지 않는다.

또다른 실시양태에서, 본 발명은 본 발명의 방법으로 생성된 유효량의 미생물성 오일을 포함하는 식품을 제공한다. 다르게는, 본 발명은 본 발명의 방법으로 생성된 유효량의 미생물성 오일을 포함하는, 의료용 식품, 식이 보조제; 유아용 조제식 및 약제로 이루어진 군으로부터 선택된 제품을 제공한다.

다르게는, 본 발명은 본 발명의 방법으로 생성된 유효량의 미생물성 오일을 포함하는 동물 사료를 제공한다.

또다른 실시양태에서, 본 발명은 도코사헥사엔산을 함유하는 본 발명의 방법에 의해 생성된 미생물성 오일을 인간 또는 동물이 소모가능하거나 사용가능한 형태로 제공하는 것을 포함하는, 에이코사펜타엔산이 풍부한 인간, 동물 또는 양식 유기체 식이 보충제를 제공하는 방법을 제공한다.

다르게는, 본 발명은 도코사헥사엔산을 함유하는 본 발명의 방법에 의해 생성된 미생물성 오일을 인간 또는 동물이 소모가능하거나 사용가능한 형태로 제공하는 것을 포함하는, 동물 또는 인간에서 도코사헥사엔산의 결핍을 치료하는 방법을 제공한다.

<미생물 기탁>

하기 미생물은 미국 20110-2209 버지니아주 마나사스 유니버시티 불러바드 10801에 소재하는 아메리칸 타입 컬쳐 콜렉션 (ATCC)에 기탁되었으며, 기탁 번호 및 기탁일은 하기와 같다.

미생물	기탁 번호	기탁일
플라스미드 pY89-5	ATCC PTA-6048	2004년 6월 4일
야로위아 리폴리티카 Y2047	ATCC PTA-7186	2005년 10월 26일
야로위아 리폴리티카 Y2201	ATCC PTA-7185	2005년 10월 26일
야로위아 리폴리티카 Y2096	ATCC PTA-7184	2005년 10월 26일
야로위아 리폴리티카 Y3000	ATCC PTA-7187	2005년 10월 26일

발명의 상세한 설명

본원에 인용된 모든 특허, 특허 출원 및 공보는 그의 전문이 본원에 참고로 포함된다. 구체적으로, 하기하는 출원인의 양수인의 동시 계류중인 출원을 포함한다:

미국 특허 출원 제10/840478호 (2004년 5월 6일 출원),

미국 특허 출원 제10/840579호 (2004년 5월 6일 출원),

미국 특허 출원 제10/840325호 (2004년 5월 6일 출원),

미국 특허 출원 제10/869630호 (2004년 6월 16일 출원),

미국 특허 출원 제10/882760호 (2004년 7월 1일 출원),

미국 특허 출원 제10/985109호 (2004년 11월 10일 출원),

미국 특허 출원 제10/987548호 (2004년 11월 12일 출원),

미국 특허 출원 제60/624812호 (2004년 11월 4일 출원),

미국 특허 출원 제11/024545호 및 제11/024544호 (2004년 12월 29일 출원),

미국 특허 출원 제60/689031호 (2005년 6월 9일 출원),

미국 특허 출원 제11/183664호 (2005년 7월 18일 출원),

미국 특허 출원 제11/185301호 (2005년 7월 20일 출원),

미국 특허 출원 제11/190750호 (2005년 7월 27일 출원),

미국 특허 출원 제11/225354호 (2005년 9월 13일 출원),

미국 특허 출원 제10/253,882호 (2005년 10월 19일 출원),

미국 특허 출원 제11/254,173호 (2005년 10월 19일 출원),

미국 특허 출원 제10/253,882호 (2005년 10월 19일 출원),

미국 특허 출원 제11/254,173호 (2005년 10월 19일 출원).

본 발명에 따라, 본 출원인은 5% 초과의 도코사헥사엔산 (DHA, 22:6, ω-3)을 생성할 수 있는 야로위아 리폴리티카의 생성 숙주 균주를 제공한다. 이러한 특정 다불포화 지방산 (PUFA)의 축적은 고수준의 재조합 발현을 위한 유질 효모 숙주내에서 Δ6 데새투라제, $C_{18/20}$ 엘론가제, Δ5 데새투라제, Δ17 데새투라제, $C_{20/22}$ 엘론가제 또는 Δ4 데새투라제 활성을 갖는 단백질을 포함하는, 기능적 ω-3/ω-6 지방산 생합성 경로의 도입에 의해 달성된다. 따라서, 본 발명은 야로위아 리폴리티카를 유전자 조작하여 DHA 및 그의 유도체를 상업적으로 생성할 수 있다는 것을 증명한다. 생성 방법 또한 청구한다.

본 발명은 여러 용도를 갖는다. 본원에 개시된 방법에 의해 제조된 PUFA 또는 그의 유도체는 정맥내 영양 공급을 행하고 있는 환자를 위한 또는 영양 불량의 예방 또는 치료를 위한 식이 대체물 또는 보충제, 특히 유아용 조제식으로서 사용될 수 있다. 별법으로, 정제된 PUFA (또는 그의 유도체)를 일반적인 사용시에 수용자가 식이 보조를 위해 원하는 양으로 섭취하도록 제제화된 요리용 오일, 지방 또는 마가린에 혼입할 수 있다. PUFA는 또한 유아용 조제식, 영양 보충제 또는 다른 식품에 혼입될 수 있으며, 소염제 또는 콜레스테롤 강하제로서 사용될 수 있다. 임의로, 상기 조성물은 (인간 또는 가축의) 약제 용도로 사용될 수 있다. 이 경우, PUFA는 일반적으로 경구 투여되지만, 성공적으로 흡수될 수 있는 임의의 경로, 예를 들어 비경구 (예컨대, 피하, 근육내 또는 정맥내), 직장, 질 또는 국소 (예컨대, 피부 연고 또는 로션으로서)로 투여될 수 있다.

재조합 수단에 의해 생성된 PUFA로 인간 또는 동물을 보충함으로써, 첨가된 PUFA의 수준뿐 아니라 이들의 대사 결과물이 증가될 수 있다. 예를 들어, DHA로 처리하면, DHA뿐 아니라 DHA의 하류 생성물, 예컨대 에이코사노이드 (즉, 프로스타글란딘, 류코트리엔, 트롬복산)의 수준도 증가될 수 있다. 복잡한 조절 메카니즘을 이용하여 다양한 PUFA를 조합하거나 PUFA의 상이한 접합체를 부가할 수 있으며, 이로써 개체에서 특이적인 PUFA의 원하는 수준을 달성하기 위해 상기 메카니즘을 방해, 조절 또는 극복할 수 있다.

별법의 실시양태에서, 양식용 사료 (즉, 건식 사료, 반-습식 및 습식 사료)는 일반적으로 영양제 조성물의 1% 내지 2% 이상이 ω-3 및/또는 ω-6 PUFA인 것을 요구하기 때문에, 본원에 개시된 방법에 의해 제조된 PUFA 또는 그의 유도체를 상기 제제의 합성에 이용할 수 있다.

정의

본원에서는 수많은 용어 및 약어들이 사용된다. 하기 정의를 제공한다.

"오픈 리딩 프레임(Open reading frame)"은 ORF로 약칭된다.

"폴리머라제 연쇄 반응"은 PCR로 약칭된다.

"아메리칸 타입 컬쳐 콜렉션"은 ATCC로 약칭된다.

"다불포화 지방산"은 PUFA로 약칭된다.

"디아실글리세롤 아실트랜스퍼라제"는 DAG AT 또는 DGAT로 약칭된다.

"인지질:디아실글리세롤 아실트랜스퍼라제"는 PDAT로 약칭된다.

"글리세롤-3-포스페이트 아실트랜스퍼라제"는 GPAT로 약칭된다.

"리소포스파티드산 아실트랜스퍼라제"는 LPAAT로 약칭된다.

"아실-CoA:1-아실 리소포스파티딜콜린 아실트랜스퍼라제"는 "LPCAT"로 약칭된다.

"아실-CoA:스테롤-아실트랜스퍼라제"는 ARE2로 약칭된다.

"디아실글리세롤"은 DAG로 약칭된다.

"트리아실글리세롤"은 TAG로 약칭된다.

"조효소 A"는 CoA로 약칭된다.

"포스파티딜-콜린"은 PC로 약칭된다.

용어 "푸사륨 모닐리포르메"는 "푸사륨 베르티실리오이데스(*Fusarium verticillioides*)"와 동의어이다.

용어 "식품"은 인간 소비에 일반적으로 적합한 임의의 음식을 나타낸다. 전형적인 식품으로는 육류 제품, 곡류 제품, 베이킹 식품, 스낵 식품, 유제품 등이 있으나, 이로 한정되는 것은 아니다.

용어 "기능성 식품"은 함유된 전형적인 영양제를 넘어서는 건강상의 이점을 제공할 수 있는 임의의 변형된 식품 또는 성분을 비롯하여 건강에 좋은 생성물을 잠재적으로 포함하는 식품을 나타낸다. 기능성 식품은 비타민, 허브 및 건강 기능 물질이 강화된 시리얼, 빵 및 음료와 같은 식품을 포함할 수 있다. 기능성 식품은 그의 영양적 가치를 넘어서는 건강상의 이점을 제공하는 물질을 함유하며, 상기 물질은 식품에 천연적으로 존재하거나 의도적으로 첨가된 것이다.

본원에 사용된 용어 "의료용 식품"은 의사의 지시하에 장에서 소비되거나 장으로 투여되도록 제제화되고, 인정된 과학 원리에 기초하여 구별되는 영양 요건이 의료적 평가에 의해 확립된 질환 또는 증상에서 특정한 식이 관리를 위해 의도된 식품을 나타낸다 (문헌 [section 5(b) of the Orphan Drug Act (21 U.S.C. 360ee(b)(3)] 참조). 하기의 경우에만 "의료용 식품"에 해당된다: (i) 경구 섭취 또는 관에 의한 장내 공급에 의해 부분적으로 또는 전적으로 환자에게 공급하기 위한 특이적으로 제제화되고 가공된 제품이다 (천연 상태로 사용되는 천연 발생 음식물과는 반대됨); (ii) 치료적 또는 만성 의료적 요구 때문에 통상적인 음식물 또는 특정한 영양제를 섭취, 소화, 흡수 또는 대사하는 능력이 제한되거나 손상된 환자, 또는 정상적인 식사의 변형만으로는 달성될 수 없는 의학적으로 결정된 다른 특정한 영양 요건을 가진 환자의 식이 관리를 위해 의도된다; (iii) 의학적 평가에 의해 결정되는 바와 같이, 특정한 질환 또는 증상으로 인한 특유한 영양분 요구의 관리를 위해 특이적으로 변형된 영양분 지원을 제공한다; (iv) 의학적 관리하에 사용되도록 의도된다; (v) 지속적인 현행의 의학적 관리를 받는 환자에게만 의도되며, 상기 환자가 다른 것 중에서도 의료용 식품의 사용 지침에 대한 주기적 기준에 대해 의학적 주의를 요한다. 따라서, 식이 보조제 또는 통상적인 식품과는 달리, 구별되는 영양 요건이 확립되어 있는 질환 또는 증상의 특별한 식이 관리를 위해 의도된 의료용 식품은 특정한 질환 또는 증상을 위해 구별되는 영양분 지원을 제공하는 것과 관련하여 과학적으로 타당하게 요구될 수 있다. 의료용 식품은 특정한 식이적 용도를 위한 보다 광범위한 부류의 식품 (예컨대, 저자극성 식품), 및 의료용 식품이 의학적 관리하에 사용된다는 요건에 따라 건강을 요구하는 식품 (예컨대, 식이 보조제)과는 구별된다.

용어 "의료용 영양제"는 본원에서 정의된 바와 같은 의료용 식품으로서, 전형적으로 특정한 식이적 요구를 위해 구체적으로 고안된 강화 음료를 나타낸다. 의료용 영양제는 일반적으로 특정한 의료적 또는 식이적 증상에 대해 집중된 식이 조성물을 포함한다. 시판되는 의료용 영양제의 예로는 엔슈어(Ensure®) 및 부스트(Boost®)가 있으나, 이로 제한되지 않는다.

본원에 사용된 용어 "약제"는 미국에서 판매되는 경우 연방 식품, 약품 및 화장품법의 섹션 505 또는 505에 의해 규제되는 화합물 또는 물질을 의미한다.

용어 "유아용 조제식"은 사람의 젖과 유사하여 오로지 유아에 의해 소비되는 것으로 고안된 식품을 의미한다. 유아용 조제식의 전형적인 예로는 시밀락(Similac®) 및 이소밀(Isomil®)이 있으나, 이로 제한되지 않는다.

용어 "식이 보조제"는 (i) 음식을 보충하기 위해 의도되어, 통상적인 식품이나 식사 또는 음식용 단품으로서는 사용되지 않고, (ii) 한가지 이상의 식이 성분 (예를 들어, 비타민, 미네랄, 허브 또는 다른 식물성 약품, 아미노산, 효소 및 분비선 추출물) 또는 이들의 구성성분을 함유하고, (iii) 환약, 캡슐, 정제 또는 액체로서 입으로 복용하도록 의도되고, (iv) 식이 보조제인 것으로 표시된 제품을 나타낸다.

"식품 유사물"은 육류, 치즈, 우유 등과 같은 대응 식품과 비슷하게 제조되고, 상기 대응 식품의 외관, 맛 및 조직감을 갖도록 의도된 식품-유사 제품이다. 따라서, 본원에 사용된 용어 "식품"은 또한 식품 유사물을 포함한다.

용어 "양식용 사료" 및 "양식사료"는 양식 산업에서 천연 사료를 보충하거나 대체하기 위한 제작 음식 또는 인공 음식 (제제화된 사료)을 나타낸다. 따라서, 양식사료는 양식 물고기 및 갑각류 (즉, 하급 주식용 어류 종 [예를 들어, 민물고기, 예컨대 잉어, 틸라피아 및 메기] 및 명품 또는 한정 시장을 위한 고급 환금 작물 종 [예를 들어, 주로 해양 종 및 양측 회유성 종, 예컨대 새우, 연어, 송어, 방어, 농어, 강성돔 및 참바리])에 유용한 인공적으로 배합된 사료를 나타낸다. 이들 제제화 사료는 양식 종을 위한 영양적으로 완전한 음식을 형성하기 위해 서로를 보충하는 다양한 비율의 여러 성분들로 이루어진다.

용어 "동물 사료"는 동물, 예를 들어 가축 (애완 동물, 농장 동물 등) 또는 어장에서 식품 생산을 위해 길러지는 동물에 의해서만 소비되도록 의도된 사료를 나타낸다.

용어 "사료 영양분"은 본 발명의 제조합 생성 숙주를 포함하는 효모 생물체로부터 유도될 수 있는 단백질, 지질, 탄수화물, 비타민, 미네랄 및 핵산과 같은 영양분을 의미한다.

본원에 사용된 용어 "생물체(biomass)"는 구체적으로 상업적으로 유의한 양으로 재조합 생성 숙주에서 생성된 EPA의 발효에 소모되거나 사용되는 효모 세포 물질을 나타낸다.

용어 "지방산"은 약 C_{12} 내지 C_{22}의 다양한 쇄 길이의 장쇄 지방족산 (알칸산)을 나타낸다 (장쇄 및 단쇄 길이의 산은 모두 공지되어 있다). 우세한 쇄 길이는 C_{16} 내지 C_{22}이다. 지방산의 구조는 "X:Y"의 간단한 표기법으로 표현되며, 여기서, X는 특정 지방산에서 탄소 (C) 원자의 총 개수를 나타내고, Y는 이중 결합의 개수를 나타낸다. "포화 지방산"과 "불포화 지방산", "일불포화 지방산"과 "다불포화 지방산" (또는 "PUFA"), "오메가-6 지방산" (ω-6 또는 n-6)과 "오메가-3 지방산" (ω-3 또는 n-3)의 차이에 대한 더욱 상세한 설명은 제WO 2004/101757호에 제공되어 있다.

본원에서 PUFA를 기재하기 위해 사용된 명명법을 하기 표 3에 나타낸다. "속기 표시"란에서, 탄소의 개수, 이중 결합의 개수, 및 오메가 탄소 (이를 1번으로 함)에서부터 세었을 때 오메가 탄소에 가장 근접한 이중 결합의 위치를 나타내기 위해 오메가-참조 시스템을 사용한다. 하기 표에는 또한 ω-3 및 ω-6 지방산 및 그의 전구체의 일반명, 명세서에 사용될 약어, 및 각 화합물의 화학명이 기재된다.

[표 3]

불포화 지방산 및 전구체의 명명법			
일반명	**약어**	**화학명**	**속기 표시**
미리스트산	--	테트라데칸산	14:0
팔미트산	팔미테이트	헥사데칸산	16:0
팔미트올레산	--	9-헥사데센산	16:1
스테아르산	--	옥타데칸산	18:0
올레산	--	시스-9-옥타데센산	18:1
리놀레산	LA	시스-9,12-옥타데카디엔산	18:2 ω-6
γ-리놀레산	GLA	시스-6,9,12-옥타데카트리엔산	18:3 ω-6
에이코사디엔산	EDA	시스-11,14-에이코사디엔산	20:2 ω-6
디호모-γ-리놀레산	DGLA	시스-8,11,14-에이코사트리엔산	20:3 ω-6
아라키돈산	ARA	시스-5,8,11,14-에이코사테트라엔산	20:4 ω-6
α-리놀렌산	ALA	시스-9,12,15-옥타데카트리엔산	18:3 ω-3
스테아리돈산	STA	시스-6,9,12,15-옥타데카테트라엔산	18:4 ω-3
에이코사트리엔산	ETrA	시스-11,14,17-에이코사트리엔산	20:3 ω-3
에이코사테트라엔산	ETA	시스-8,11,14,17-에이코사테트라엔산	20:4 ω-3
에이코사펜타엔산	EPA	시스-5,8,11,14,17-에이코사펜타엔산	20:5 ω-3
도코사펜타엔산	DPA	시스-7,10,13,16,19-에이코사펜타엔산	22:5 ω-3
도코사헥사엔산	DHA	시스-4,7,10,13,16,19-에이코사헥사엔산	22:6 ω-3

용어 "고수준 DHA 생성"은 미생물 숙주의 전체 지질 중 약 5% 이상의 DHA, 바람직하게는 전체 지질 중 약 10% 이상의 DHA, 더욱 바람직하게는 전체 지질 중 약 15% 이상의 EDHA, 더욱 바람직하게는 전체 지질 중 약 20% 이상의 DHA, 더욱 바람직하게는 전체 지질 중 약 25-30% 이상의 DHA를 생성하는 것을 나타낸다. 따라서, DHA의 구조적 형태에는 제한이 없지만, 예를 들어 DHA는 유리 지방산 또는 에스테르화된 형태, 예컨대 아실글리세롤, 인지질, 술포지질 또는 당지질로서 전체 지질 중에 존재할 수 있다.

용어 "임의의 GLA가 결여된"은 약 0.1% 아래까지 검출가능한 수준을 갖는 장비를 이용하여 GC 분석에 의해 측정하였을 때, 미생물 숙주의 전체 지질 중에 임의의 검출가능한 GLA가 결여된 것을 나타낸다.

용어 "필수 지방산"은 특정한 필수 지방산을 새로 합성할 수 없는 유기체가 생존을 위해 섭취해야만 하는 특정한 PUFA를 나타낸다. 예를 들어, 포유동물은 필수 지방산 LA (18:2, ω-6) 및 ALA (18:3, ω-3)를 합성할 수 없다. 다른 필수 지방산으로는 GLA (ω-6), DGLA (ω-6), ARA (ω-6), EPA (ω-3) 및 DHA (ω-3)가 있다.

"미생물성 오일" 또는 "단세포 오일"은 미생물 (예컨대, 조류, 유질 효모 및 섬사상 진균)에 의해 그의 일생동안 천연적으로 생산되는 오일이다. 용어 "오일"은 25℃에서 액체이고 보통 다불포화된 지질을 나타낸다. 반대로, 용어 "지방"은 25℃에서 고체이고 보통 포화된 지질을 나타낸다.

"지질체"는 보통 특정한 단백질 및 단층의 인지질에 의해 둘러싸인 지질 액적을 나타낸다. 이들 기관은 대부분의 유기체가 중성 지질을 전달/저장하는 부위이다. 지질체는 TAG-생합성 효소를 함유하는 소포체의 마이크로도메인으로부터 발생하는 것으로 생각되며, 이들의 합성 및 크기는 특정한 단백질 성분에 의해 조절되는 것으로 여겨진다.

"중성 지질"은 세포에서 저장 지방 및 오일로서 지질체 중에서 흔히 발견되는 지질을 나타내며, 상기 지질은 대개 세포의 pH에서 하전된 기를 전혀 갖지 않기 때문에 그렇게 불린다. 일반적으로, 이들은 수 친화성이 전혀 없이 비-극성이다. 중성 지질은 일반적으로 글리세롤과 지방산의 모노-, 디- 및/또는 트리에스테르이며, 이들은 각각 모노아실글리세롤, 디아실글리세롤 또는 TAG (또는 집합적으로 아실글리세롤)로도 지칭된다. 아실글리세롤로부터 유리 지방산을 방출시키기 위해 가수분해 반응이 일어나야 한다.

용어 "트리아실글리세롤", "오일" 및 "TAG"는 3개의 지방 아실 잔기가 글리세롤 분자에 에스테르화된 중성 지질을 나타낸다 (상기 용어들은 본 명세서 내에서 서로 교환가능하게 사용될 것이다). 이러한 오일은 장쇄 PUFA 뿐 아니라, 단쇄 포화 및 불포화 지방산 및 장쇄 포화 지방산을 포함할 수 있다. 따라서, "오일 생합성"은 일반적으로 세포에서 TAG를 합성하는 것을 나타낸다.

용어 "아실트랜스퍼라제"는 아미노-아실기 이외의 기의 전달을 담당하는 효소 (EC 2.3.1.-)를 나타낸다.

용어 "DAG AT"는 디아실글리세롤 아실트랜스퍼라제 (아실-CoA-디아실글리세롤 아실트랜스퍼라제 또는 디아실글리세롤 O-아실트랜스퍼라제로도 공지됨) (EC 2.3.1.20)를 나타낸다. 이 효소는 아실-CoA 및 1,2-디아실글리세롤을 TAG 및 CoA로 전환시키는 것을 담당한다 (이는 TAG 생합성의 최종 단계와 관련이 있음). 두가지 부류의 DAG AT 효소인 DGAT1 및 DGAT2가 존재한다. DGAT1 부류는 아실-CoA:콜레스테롤 아실트랜스퍼라제 (ACAT) 유전자 부류와 상동성을 공유하는 반면, DGAT2는 관련이 없다 (문헌 [Lardizabal et al., J. Biol Chem. 276(42):38862-38869 (2001)]).

용어 "PDAT"는 인지질:디아실글리세롤 아실트랜스퍼라제 효소 (EC 2.3.1.158)를 나타낸다. 이 효소는 인지질의 sn-2 위치로부터 1,2-디아실글리세롤의 sn-3 위치로 아실기를 전달하여 리소인지질 및 TAG을 생성하는 것을 담당한다 (이는 TAG 생합성의 최종 단계와 관련이 있음). 이 효소는 아실-CoA-독립성 메카니즘을 통해 TAG를 합성한다는 점에서 DGAT (EC 2.3.1.20)와는 상이하다.

용어 "ARE2"는 아실-CoA + 스테롤 = CoA + 스테롤 에스테르의 반응을 촉매하는 아실-CoA:스테롤-아실트랜스퍼라제 효소 (EC 2.3.1.26; 스테롤-에스테르 신타제 2 효소로도 공지됨)를 나타낸다.

용어 "GPAT"는 *gpat* 유전자에 의해 코딩되며, 아실-CoA 및 sn-글리세롤 3-포스페이트를 CoA 및 1-아실-sn-글리세롤 3-포스페이트로 전환시키는 (인지질 생합성의 제1 단계) 글리세롤-3-포스페이트 O-아실트랜스퍼라제 효소 (E.C. 2.3.1.15)를 나타낸다.

용어 "LPAAT"는 리소포스파티드산-아실트랜스퍼라제 효소 (EC 2.3.1.51)를 나타낸다. 이 효소는 아실-CoA 기를 1-아실-sn-글리세롤 3-포스페이트 (즉, 리소포스파티드산) 상에 전달하여 CoA 및 1,2-디아실-sn-글리세롤 3-포스페이트 (포스파티드산)를 생성하는 것을 담당한다. 상기 용어는 또한 아실-CoA:1-아실-sn-글리세롤-3-포스페이트 2-O-아실트랜스퍼라제, 1-아실-sn-글리세롤-3-포스페이트 아실트랜스퍼라제 및/또는 1-아실글리세롤포스페이트 아실트랜스퍼라제 (AGAT로 약칭됨)로서의 LPAAT를 나타낸다.

용어 "LPCAT"는 아실-CoA:1-아실 리소포스파티딜-콜린 아실트랜스퍼라제를 나타낸다. 이 효소는 CoA와 포스파티딜콜린 (PC) 사이의 아실기 교환을 담당한다. 여기서, 상기 용어는 또한 CoA와 다른 인지질, 예컨대 리소포스파티드산 (LPA) 사이의 아실 교환과 관련된 효소를 나타낸다.

"전체 지질 및 오일 분획 중 PUFA 백분율(%)"은 이들 분획 중 전체 지방산에 대한 PUFA의 백분율을 나타낸다. 용어 "전체 지질 분획" 또는 "지질 분획"은 모두 유질 유기체 내의 모든 지질 (즉, 중성 및 극성), 예컨대 포스파티딜콜린 (PC) 분획, 포스파티딜에탄올아민 (PE) 분획 및 트리아실글리세롤 (TAG 또는 오일) 분획에 위치하는 지질의 합계를 나타낸다. 그러나, 용어 "지질" 및 "오일"은 본 명세서 내에서 서로 교환가능하게 사용될 것이다.

용어 "포스파티딜콜린" 또는 "PC"는 세포막의 주요 구성성분인 인지질을 나타낸다. PC의 화학 구조는 일반적으로 콜린 분자, 포스페이트기 및 글리세롤을 포함하는 것으로 기재될 수 있으며, 여기서, 지방 아실 쇄는 R 기로서 글리세롤 분자의 sn-1 및 sn-2 위치에 부착된다.

용어 "PUFA 생합성 경로 효소"는 PUFA의 생합성과 관련된 임의의 하기 효소 (및 하기 효소를 코딩하는 유전자)를 나타낸다: Δ4 데새투라제, Δ5 데새투라제, Δ6 데새투라제, Δ12 데새투라제, Δ15 데새투라제, Δ17 데새투라제, Δ9 데새투라제, Δ8 데새투라제, Δ9 엘론가제, $C_{14/16}$ 엘론가제, $C_{16/18}$ 엘론가제, $C_{18/20}$ 엘론가제 및/또는 $C_{20/22}$ 엘론가제.

용어 "ω-3/ω-6 지방산 생합성 경로"는 적절한 조건하에 발현될 때, ω-3 및 ω-6 지방산 중 하나 또는 둘다의 생성을 촉매하는 효소를 코딩하는 유전자의 집합을 나타낸다. 전형적으로 ω-3/ω-6 지방산 생합성 경로와 관련된 유전자는 하기 효소 중 일부 또는 모두를 코딩한다: Δ12 데새투라제, Δ6 데새투라제, $C_{18/20}$ 엘론가제, $C_{20/22}$ 엘론가제, Δ9 엘론가제, Δ5 데새투라제, Δ17 데새투라제, Δ15 데새투라제, Δ9 데새투라제, Δ8 데새투라제 및 Δ4 데새투라제. 대표적인 경로를 도 1에 도시하는데, 도 1은 다양한 중간체를 통해 올레산을 DHA로 전환시키는 것을 제공하며, 이는 ω-3 및 ω-6 지방산 둘다가 공통된 공급원으로부터 생성될 수 있는 방법을 입증한다. 상기 경로는 천연적으로 두가지 부분으로 나뉘어지며, 한 부분은 ω-3 지방산만을 생성할 것이고, 다른 부분은 ω-6 지방산만을 생성할 것이다. ω-3 지방산만을 생성하는 부분은 본원에서 ω-3 지방산 생합성 경로로 지칭되는 반면, ω-6 지방산만을 생성하는 부분은 본원에서 ω-6 지방산 생합성 경로로 지칭될 것이다.

본원에서 ω-3/ω-6 지방산 생합성 경로의 문맥에서 사용된 용어 "기능적"은 상기 경로의 일부 (또는 모든) 유전자가 활성 효소를 발현하여, 생체내 가수분해 또는 기질 전환을 일으키는 것을 의미한다. 수많은 지방산 생성물은 상기 경로의 유전자의 하위 집합의 발현만을 필요로 할 것이기 때문에, "ω-3/ω-6 지방산 생합성 경로" 또는 "기능적 ω-3/ω-6 지방산 생합성 경로"가 상기 문단에 열거된 모든 유전자를 필요로 한다는 것을 암시하는 것은 아님을 이해해야 한다.

용어 "ω-6 Δ6 데새투라제/Δ6 엘론가제 경로"는 최소한 Δ6 데새투라제, $C_{18/20}$ 엘론가제, Δ5 데새투라제, Δ17 데새투라제, $C_{20/22}$ 엘론가제 및 Δ4 데새투라제의 유전자를 포함하는 DHA 지방산 생합성 경로를 나타낼 것이다. 용어 "ω-3 Δ6 데새투라제/Δ6 엘론가제 경로"는 최소한 Δ15 데새투라제, Δ6 데새투라제, $C_{18/20}$ 엘론가제, Δ5 데새투라제, $C_{20/22}$ 엘론가제 및 Δ4 데새투라제의 유전자를 포함하는 DHA 지방산 생합성 경로를 나타낼 것이다. 용어 "Δ6 데새투라제/Δ6 엘론가제 경로의 조합"은 최소한 Δ15 데새투라제, Δ6 데새투라제, $C_{18/20}$ 엘론가제, Δ5 데새투라제, Δ17 데새투라제, $C_{20/22}$ 엘론가제 및 Δ4 데새투라제의 유전자를 포함하는 DHA 지방산 생합성 경로를 나타낼 것이다. 마지막으로, 용어 "Δ6 데새투라제/Δ6 엘론가제 경로"는 포괄적으로 상기 기재된 Δ6 데새투라제/Δ6 엘론가제 경로 중 임의의 하나 (또는 그 이상)을 나타낼 것이다.

관련된 방식으로, 용어 "ω-6 Δ9 엘론가제/Δ8 데새투라제 경로"는 최소한 Δ9 엘론가제, Δ8 데새투라제, Δ5 데새투라제, Δ17 데새투라제, $C_{20/22}$ 엘론가제 및 Δ4 데새투라제의 유전자를 포함하는 DHA 지방산 생합성 경로를 나타낼 것이다. 용어 "ω-3 Δ9 엘론가제/Δ8 데새투라제 경로"는 최소한 Δ15 데새투라제, Δ9 엘론가제, Δ8 데새투라제, Δ5 데새투라제, $C_{20/}$

$_{22}$ 엘론가제 및 Δ4 데새투라제의 유전자를 포함하는 DHA 지방산 생합성 경로를 나타낼 것이다. 용어 "Δ9 엘론가제/Δ8 데새투라제 경로의 조합"은 최소한 Δ15 데새투라제, Δ9 엘론가제, Δ8 데새투라제, Δ5 데새투라제, Δ17 데새투라제, $C_{20/22}$ 엘론가제 및 Δ4 데새투라제의 유전자를 포함하는 DHA 지방산 생합성 경로를 나타낼 것이다. 그리고, 용어 "Δ9 엘론가제/Δ8 데새투라제 경로"는 포괄적으로 상기 기재된 Δ9 엘론가제/Δ8 데새투라제 경로 중 임의의 하나 (또는 그 이상)을 나타낼 것이다.

용어 "데새투라제"는 1개 이상의 지방산을 탈포화시켜, 즉 이중 결합을 도입시켜, 관심을 가진 지방산 또는 전구체를 생성시킬 수 있는 폴리펩티드를 나타낸다. 구체적인 지방산을 나타내기 위해 본 명세서에서 오메가-참조 시스템을 사용하였지만, 델타-시스템을 이용하여 기질의 카르복실 말단에서부터 세어 데새투라제의 활성을 표시하는 것이 더욱 편리하다. 본원에서 특히 관심을 갖는 것은 1) 분자의 카르복실-말단에서부터 세었을 때 8번째와 9번째 탄소 원자 사이에서 지방산을 탈포화시키고, 예를 들어 EDA에서 DGLA 및/또는 ETrA에서 ETA로의 전환을 촉매하는 Δ8 데새투라제; 2) LA에서 GLA로 및/또는 ALA에서 STA로의 전환을 촉매하는 Δ6 데새투라제; 3) DGLA에서 ARA로 및/또는 ETA에서 EPA로의 전환을 촉매하는 Δ5 데새투라제; 4) DPA에서 DHA로의 전환을 촉매하는 Δ4 데새투라제; 5) 올레산에서 LA로의 전환을 촉매하는 Δ12 데새투라제; 6) LA에서 ALA로 및/또는 GLA에서 STA로의 전환을 촉매하는 Δ15 데새투라제; 7) ARA에서 EPA 및/또는 DGLA에서 ETA로의 전환을 촉매하는 Δ17 데새투라제; 및 8) 팔미테이트에서 팔미트올레산 (16:1)으로 및/또는 스테아레이트에서 올레산 (18:1)으로의 전환을 촉매하는 Δ9 데새투라제이다.

본 발명의 Δ15 데새투라제와 관련하여, 용어 "이작용성"은 효소성 기질로서 올레산 및 LA 둘다를 사용하는 능력을 갖는 폴리펩티드를 의미한다. 유사하게, 본 발명의 Δ5 데새투라제와 관련하여, 용어 "이작용성"은 (1) DGLA 및 ETA로 이루어진 군으로부터 선택된 하나 이상의 효소성 기질; 및 (2) LA 및 ALA로 이루어진 군으로부터 선택된 하나 이상의 효소성 기질을 사용하는 능력을 갖는 폴리펩티드를 의미한다. "효소성 기질"은 폴리펩티드가 활성 부위에서 기질과 결합하여, 그에 대해 각각의 방식으로 작용하는 것을 의미한다.

용어 "엘론가제 시스템"은 상기 엘론가제 시스템이 작용하는 지방산 기질보다 2개의 탄소가 더 긴 지방산을 생성하도록 지방산 탄소 쇄의 신장을 담당하는 4가지 효소의 집합을 나타낸다. 보다 구체적으로, 신장 과정은 CoA가 아실 담체인 지방산 신타제와 연관되어 일어난다 (문헌 [Lassner et al., The Plant Cell 8:281-292 (1996)]). 기질-특이적이며 속도-제한적인 제1 단계에서는, 말로닐-CoA가 장쇄 아실-CoA와 축합되어 CO_2 및 β-케토아실-CoA (이의 아실 부분은 2개의 탄소 원자만큼 신장되었음)를 생성한다. 후속 반응에서는, β-히드록시아실-CoA로 환원되고, 에노일-CoA로 탈수되고, 제2 환원에 의해 신장된 아실-CoA가 생성된다. 엘론가제 시스템에 의해 촉매되는 반응의 예로는 GLA에서 DGLA로의 전환, STA에서 ETA로의 전환, 및 EPA에서 DPA로의 전환이 있다.

목적을 위해, 제1 축합 반응 (즉, 말로닐-CoA에서 β-케토아실-CoA로의 전환)을 촉매하는 효소는 포괄적으로 "엘론가제"로 지칭될 것이다. 일반적으로, 엘론가제의 기질 선택성은 다소 광범위하지만, 쇄 길이와 불포화 정도 및 유형 모두에 의해 구별된다. 따라서, 엘론가제는 상이한 특이성을 가질 수 있다. 예를 들어, $C_{14/16}$ 엘론가제는 C_{14} 기질 (예컨대, 미리스트산)을 이용할 것이고, $C_{16/18}$ 엘론가제는 C_{16} 기질 (예컨대, 팔미테이트)을 이용할 것이고, $C_{18/20}$ 엘론가제는 C_{18} 기질 (예컨대, GLA, STA)을 이용할 것이고, $C_{20/22}$ 엘론가제는 C_{20} 기질 (예컨대, EPA)을 이용할 것이다. 마찬가지 방식으로, Δ9 엘론가제는 LA에서 EDA로의 전환 및 ALA에서 ETrA로의 전환을 촉매할 수 있다. 일부 엘론가제가 광범위한 특이성을 가지며, 따라서, 단일 효소가 여러 엘론가제 반응을 촉매할 수 있다 (예컨대, $C_{16/18}$ 엘론가제 및 $C_{18/20}$ 엘론가제 둘다로서 작용할 수 있다)는 점을 주의하는 것이 중요하다. 바람직한 실시양태에서, 적합한 숙주를 지방산 엘론가제를 위한 유전자로 형질전환시킴으로써 지방산 엘론가제의 특이성을 실험적으로 측정하고, 숙주의 지방산 프로파일에 대한 그의 효과를 측정하는 것이 가장 바람직하다.

용어 "고친화도 엘론가제" 또는 "EL1S" 또는 "ELO1"는 바람직하게는 GLA (엘론가제 반응 [즉, Δ6 엘론가제]의 생성물로서 DGLA와 함께)에 대한 기질 특이성을 갖는 $C_{18/20}$ 엘론가제를 나타낸다. 이러한 한 엘론가제는 제WO 00/12720호에 기재되어 있고, 본원에서 서열 22 및 23로서 제공된다. 그러나, 본 발명자들은 상기 효소가 18:2 (LA) 및 18:3 (ALA)에 대해서도 일부 활성을 갖는다는 것을 밝혀냈고, 따라서, 서열 23은 Δ9 엘론가제 활성 (그의 Δ6 엘론가제 활성 이외에도)을 나타낸다. 따라서, 본원에서 서열 23로서 제공된 $C_{18/20}$ 엘론가제는 예를 들어 이소크리시스 갈바나 Δ9 엘론가제 (서열 70) 대신에, 본원에 기재된 Δ6 데새투라제/Δ6 엘론가제 경로 내에서 및 Δ9 엘론가제/Δ8 데새투라제 경로 내에서 모두 기능할 수 있는 것으로 결론지어 진다.

용어 "EL2S" 또는 "ELO2"는 바람직하게는 GLA (엘론가제 반응의 생성물로서 DGLA와 함께) 및/또는 STA (엘론가제 반응의 생성물로서 STA와 함께)에 대한 기질 특이성을 갖는 $C_{18/20}$ 엘론가제를 나타낸다. 이러한 한 엘론가제는 미국 특허 제6,677,145호에 기재되어 있으며, 본원에서 서열 25 및 26으로서 제공된다.

용어 "ELO3"는 *elo3* 유전자 (서열 86)에 의해 코딩되는 모르티에렐라 알피나 $C_{16/18}$ 지방산 엘론가제 효소 (본원에서 서열 87로서 제공됨)를 나타낸다. 용어 "YE2"는 본원에서 서열 94로서 제공된 유전자에 의해 코딩되는 야로위아 리폴리티카 $C_{16/18}$ 지방산 엘론가제 효소 (본원에서 서열 95로서 제공됨)를 나타낸다. 본원에서 보고된 데이타에 기초하면, ELO3 및 YE2 둘다 팔미테이트 (16:0)에서 스테아르산 (18:0)으로의 전환을 우세하게 촉매한다.

용어 "YE1"은 본원에서 서열 97로 제공된 유전자에 의해 코딩되는 야로위아 리폴리티카 $C_{14/16}$ 지방산 엘론가제 효소 (본원에서 서열 98로 제공됨)를 나타낸다. 본원에서 보고된 데이타에 기초하면, YE2는 미리스트산 (14:0)에서 팔미테이트 (16:0)로의 전환을 우세하게 촉매한다.

용어 "전환 효율" 및 "기질 전환율(%)"은 특정 효소 (예컨대, 데새투라제 또는 엘론가제)가 기질을 생성물로 전환시킬 수 있는 효율을 나타낸다. 전환 효율은 하기 식에 따라 측정된다.

([생성물]/[기질+ 생성물])×100

여기서, '생성물'은 즉시 생성물 및 그로부터 유래된 경로에서의 모든 생성물을 나타낸다.

용어 "유질"은 에너지 공급원을 지질 형태로 저장하는 경향이 있는 유기체를 나타낸다 (문헌 [Weete, In: Fungal Lipid Biochemistry, 2nd Ed., Plenum, 1980]). 일반적으로, 이들 미생물의 세포성 오일 함량은 S형 곡선을 따르는데, 지질의 농도는 성장의 대수기 말기 또는 정지기 초기에서 최고에 도달할 때까지 증가한 다음, 정지기 말기 및 사멸기 동안에 점진적으로 감소한다 (문헌 [Yongmanitchai and Ward, Appl. Environ. Microbiol. 57:419-25 (1991)]).

용어 "유질 효모"는 오일을 생성할 수 있는 효모로서 분류되는 미생물을 나타낸다. 일반적으로, 유질 미생물의 세포성 오일 또는 트리아실글리세롤 함량은 S형 곡선을 따르는데, 지질의 농도는 성장의 대수기 말기 또는 정지기 초기에서 최고에 도달할 때까지 증가한 다음, 정지기 후기 및 사멸기 동안에 점진적으로 감소한다 (문헌 [Yongmanitchai and Ward, Appl. Environ. Microbiol. 57:419-25 (1991)]). 오일을 건조 세포 중량의 약 25% 넘게 축적하는 유질 미생물은 드물지 않다. 유질 효모의 예로는 야로위아, 칸디다, 로도토룰라(*Rhodotorula*), 로도스포리듐(*Rhodosporidium*), 크립토코커스(*Cryptococcus*), 트리코스포론(*Trichosporon*) 및 리포마이세스(*Lipomyces*) 속이 있으나, 이로 제한되지 않는다.

용어 "발효가능한 탄소 공급원"은 에너지를 유도하기 위해 미생물이 대사하는 탄소 공급원을 의미한다. 전형적인 본 발명의 탄소 공급원으로는 단당류, 올리고당류, 다당류, 알칸, 지방산, 지방산 에스테르, 모노글리세리드, 디글리세리드, 트리글리세리드, 이산화탄소, 메탄올, 포름알데히드, 포르메이트 및 탄소-함유 아민이 있으나, 이로 제한되지 않는다.

본원에 사용된 "단리된 핵산 단편"은 임의로 합성, 비-천연 또는 변경된 뉴클레오티드 염기를 함유하는 단일- 또는 이중-가닥 RNA 또는 DNA의 중합체이다. DNA 중합체 형태의 단리된 핵산 단편은 cDNA, 게놈 DNA 또는 합성 DNA의 하나 이상의 절편으로 이루어질 수 있다.

아미노산 또는 뉴클레오티드 서열의 "실질적인 부분"은 당업자에 의한 서열의 수동적 평가에 의해, 또는 컴퓨터-자동화 서열 비교 및 BLAST (Basic Local Alignment Search Tool; 문헌 [Altschul, S. F., et al., J. Mol. Biol. 215:403-410 (1993)])와 같은 알고리즘을 이용한 확인에 의해, 폴리펩티드 또는 유전자를 추정적으로 확인하기 위해, 상기 폴리펩티드의 아미노산 서열 또는 상기 유전자의 뉴클레오티드 서열을 충분히 포함하는 부분이다. 일반적으로, 공지된 단백질 또는 유전자에 대한 상동성으로서 폴리펩티드 또는 핵산 서열을 추정적으로 확인하기 위해서는, 10개 이상의 연속 아미노산 또는 30개 이상의 뉴클레오티드로 이루어진 서열이 필요하다. 더욱이, 뉴클레오티드 서열과 관련하여, 유전자 확인 (예컨대, 써던 혼성화) 및 단리 (예컨대, 박테리아 콜로니 또는 박테리오파지 플라크의 동일계내 혼성화)의 서열-의존성 방법에서는 20 내지 30개의 연속 뉴클레오티드를 포함하는 유전자-특이적 올리고뉴클레오티드 프로브가 사용될 수 있다. 또한, 12 내지 15개의 염기로 이루어진 짧은 올리고뉴클레오티드는 PCR에서 증폭 프라이머로서 사용되어, 상기 프라이머를 포함하는 특정 핵산 단편을 수득할 수 있다. 따라서, 뉴클레오티드 서열의 "실질적인 부분"은 서열을 포함하는 핵산 단편을 구체적으로 확인하고/하거나 단리하기 위해 상기 서열을 충분히 포함한다.

용어 "상보적인"은 서로 혼성화될 수 있는 뉴클레오티드 염기들 사이의 관계를 기재하기 위해 사용된다. 예를 들어, DNA와 관련하여, 아데노신은 티민에 대해 상보적이고, 시토신은 구아닌에 대해 상보적이다.

"코돈 다의성(degeneracy)"은 코딩된 폴리펩티드의 아미노산 서열에 영향을 주지 않으면서, 뉴클레오티드 서열의 다양성을 허용하는 유전자 코드의 성질을 나타낸다. 당업자는 주어진 아미노산을 특정하기 위해 뉴클레오티드 코돈을 사용하는 데 있어서 특정한 숙주 세포에 의해 나타나는 "코돈-편향성(codon-bias)"을 잘 알 고 있다. 따라서, 숙주 세포에서의 개선된 발현을 위해 유전자를 합성할 때, 코돈 사용 빈도가 숙주 세포에서 바람직한 코돈 사용 빈도에 접근하도록 유전자를 고안하는 것이 바람직하다.

DNA 서열과 관련하여 "화학적으로 합성된"은 구성 뉴클레오티드들을 시험관내에서 조립한 것을 의미한다. DNA의 수동적인 화학적 합성은 널리 확립된 절차를 이용하여 달성할 수 있거나, 또는 자동화된 화학적 합성은 시판되는 수많은 기계를 이용하여 수행할 수 있다. "합성 유전자"는 당업자에게 공지된 절차를 이용하여 화학적으로 합성된 올리고뉴클레오티드 빌딩 블록으로부터 조립될 수 있다. 이들 빌딩 블록을 라이게이션하고 어닐링하여 유전자 절편을 형성한 다음, 이들을 효소적으로 조립하여 전체 유전자를 구축한다. 따라서, 뉴클레오티드 서열의 최적화에 기초하여 최적의 유전자 발현을 위해 유전자들을 맞추어서, 숙주 세포의 코돈 편향성을 반영할 수 있다. 당업자는 숙주가 우호적인 이들 코돈에 대해 코돈 사용이 편향되는 경우 유전자 발현이 성공적일 수 있음을 이해한다. 바람직한 코돈의 결정은 서열 정보를 입수할 수 있는 숙주 세포로부터 유래된 유전자의 개관에 기초할 수 있다.

"유전자"는 특정한 단백질을 발현하는 핵산 단편을 나타내며, 이는 코딩 영역 단독을 나타내거나, 코딩 서열의 이전 (5' 비-코딩 서열) 및 이후 (3' 비-코딩 서열)의 조절 서열을 포함할 수 있다. "천연 유전자"는 그의 조절 서열과 함께 천연적으로 발견되는 유전자를 나타낸다. "키메라 유전자"는 천연적으로 함께 발견되지 않는 조절 및 코딩 서열을 포함하는, 천연 유전자가 아닌 임의의 유전자를 나타낸다. 따라서, 키메라 유전자는 상이한 공급원으로부터 유래된 조절 서열 및 코딩 서열, 또는 동일한 공급원으로부터 유래된 조절 서열 및 코딩 서열을 포함할 수 있으나, 천연적으로 발견되는 것과는 상이한 방식으로 정렬된다. "내인성 유전자"는 유기체 게놈 내의 천연적인 위치에 있는 천연 유전자를 나타낸다. "외래" 유전자는 유전자 전달에 의해 숙주 유기체로 도입된 유전자를 나타낸다. 외래 유전자는 비-천연 유기체에 삽입된 천연 유전자, 천연 숙주 내의 새로운 위치로 도입된 천연 유전자, 또는 키메라 유전자를 포함할 수 있다. "트랜스진"은 형질전환 절차에 의해 게놈에 도입된 유전자이다. "코돈-최적화된 유전자"는 숙주 세포의 바람직한 코돈 사용 빈도를 모방하도록 고안된 코돈 사용 빈도를 갖는 유전자이다.

"코딩 서열"은 특정한 아미노산 서열을 코딩하는 DNA 서열을 나타낸다. "적합한 조절 서열"은 코딩 서열의 상류 (5' 비-코딩 서열), 내부, 또는 하류 (3' 비-코딩 서열)에 위치하며, 관련된 코딩 서열의 전사, RNA 가공 또는 안정성, 또는 번역에 영향을 미치는 뉴클레오티드 서열을 나타낸다. 조절 서열은 프로모터, 번역 리더 서열, 인트론, 폴리아데닐화 인식 서열, RNA 가공 부위, 이펙터(effector) 결합 부위, 및 스템-루프(stem-loop) 구조체를 포함할 수 있다.

"프로모터"는 코딩 서열 또는 기능적 RNA의 발현을 조절할 수 있는 DNA 서열을 나타낸다. 일반적으로, 코딩 서열은 프로모터 서열에 대해 3'에 위치한다. 프로모터는 전체가 천연 유전자로부터 유래될 수 있거나, 천연적으로 발견되는 상이한 프로모터로부터 유래된 상이한 요소들로 이루어질 수 있거나, 심지어 합성 DNA 절편을 포함할 수 있다. 당업자라면, 상이한 프로모터가 상이한 조직 또는 세포 유형에서, 상이한 발달 단계에서, 또는 상이한 환경적인 또는 생리학적 조건에 반응하여, 유전자의 발현을 지시할 수 있다는 것을 이해할 것이다. 유전자가 대부분의 세포 유형에서 대부분 발현될 수 있게 하는 프로모터는 보통 "구성적(constitutive) 프로모터"로서 지칭된다. 추가로, 많은 경우 조절 서열의 정확한 경계가 완전히 정의되지 않았기 때문에, 상이한 길이의 DNA 단편이 동일한 프로모터 활성을 가질 수 있다는 것이 인식될 것이다.

용어 "GPAT 프로모터" 또는 "GPAT 프로모터 영역"은 발현에 필요하며 *gpat* 유전자에 의해 코딩되는 글리세롤-3-포스페이트 O-아실트랜스퍼라제 효소 (예컨대 2.3.1.15)의 'ATG' 번역 개시 코돈 이전의 5' 상류 비번역 영역을 나타낸다. 적합한 야로위아 리폴리티카 GPAT 프로모터 영역의 예는 미국 특허 출원 제11/225354호에 기재되어 있다.

용어 "GPD 프로모터" 또는 "GPD 프로모터 영역"은 발현에 필요하며 *gpd* 유전자에 의해 코딩되는 글리세르알데히드-3-포스페이트 데히드로게나제 효소 (E.C. 1.2.1.12)의 'ATG' 번역 개시 코돈 이전의 5' 상류 비번역 영역을 나타낸다. 적합한 야로위아 리폴리티카 GPD 프로모터 영역의 예는 제WO 2005/003310호에 기재되어 있다.

용어 "GPM 프로모터" 또는 "GPM 프로모터 영역"은 발현에 필요하며 *gpm* 유전자에 의해 코딩되는 포스포글리세레이트 뮤타제 효소 (EC 5.4.2.1)의 'ATG' 번역 개시 코돈 이전의 5' 상류 비번역 영역을 나타낸다. 적합한 야로위아 리폴리티카 GPM 프로모터 영역의 예는 제WO 2005/003310호에 기재되어 있다.

용어 "FBA 프로모터" 또는 "FBA 프로모터 영역"은 발현에 필요하며 *fba1* 유전자에 의해 코딩되는 프럭토스-비스포스페이트 알돌라제 효소 (E.C. 4.1.2.13)의 'ATG' 번역 개시 코돈 이전의 5' 상류 비번역 영역을 나타낸다. 적합한 야로위아 리폴리티카 FBA 프로모터 영역의 예는 제WO 2005/049805호에 기재되어 있다.

용어 "FBAIN 프로모터" 또는 "FBAIN 프로모터 영역"은 발현에 필요한 *fba1* 유전자의 'ATG' 번역 개시 코돈 이전의 5' 상류 비번역 영역과 함께 *fba1* 유전자의 인트론을 갖는 5' 코딩 영역의 부분을 나타낸다. 적합한 야로위아 리폴리티카 FBAIN 프로모터 영역의 예는 제WO 2005/049805호에 기재되어 있다.

용어 "GPDIN 프로모터" 또는 "GPDIN 프로모터 영역"은 발현에 필요한 *gpd* 유전자의 'ATG' 번역 개시 코돈 이전의 5' 상류 비번역 영역과 함께 *gpd* 유전자의 인트론을 갖는 5' 코딩 영역의 부분을 나타낸다. 적합한 야로위아 리폴리티카 GPDIN 프로모터 영역의 예는 미국 특허 출원 제11/183664호에 기재되어 있다.

용어 "YAT1 프로모터" 또는 "YAT1 프로모터 영역"은 *yat1* 유전자에 의해 코딩되며 발현에 필요한 암모늄 수송자 효소 (TC 2.A.49; 진뱅크(GenBank) 관리 번호 XM_504457)의 'ATG' 번역 개시 코돈 이전의 5' 상류 비번역 영역을 나타낸다. 적합한 야로위아 리폴리티카 YAT1 프로모터 영역의 예는 미국 특허 출원 제11/185301호에 기재되어 있다.

용어 "EXP1 프로모터" 또는 "EXP1 프로모터 영역"은 발현에 필요하며 야로위아 리폴리티카 "YALI0C12034g" 유전자 (진뱅크 관리 번호 XM_501745)에 의해 코딩되는 단백질의 'ATG' 번역 개시 코돈 이전의 5' 상류 비번역 영역을 나타낸다. sp | Q12207 사카로마이세스 세레비지애 비고전적 방출 단백질 2 (그의 기능은 절단가능한 신호 서열이 결여된 단백질의 신규한 방출 경로와 관련이 있음)에 대한 "YALI0C12034g"의 유의한 상동성에 기초하여, 본원에서는 상기 유전자를 *exp1* 유전자라고 지칭하고, 그가 코딩하는 단백질을 EXP1라고 지칭한다. 적합한 야로위아 리폴리티카 EXP1 프로모터 영역의 예는 서열 221로서 기재되어 있지만, 이는 제한하려는 의도는 아니다. 당업자라면, EXP1 프로모터 서열의 정확한 경계가 완전히 정의되지 않았기 때문에, 증가되거나 감소된 길이의 DNA 단편이 동일한 프로모터 활성을 가질 수 있다는 것을 인식할 것이다.

용어 "프로모터 활성"은 프로모터의 전사 효율의 평가를 나타낼 것이다. 이는 예를 들어 직접적으로는 프로모터로부터 mRNA 전사체의 양을 측정함으로써 (예컨대, 노던 블럿 또는 프라이머 신장 방법), 또는 간접적으로는 프로모터로부터 발현된 유전자 생성물의 양을 측정함으로써 결정될 수 있다.

"인트론"은 대부분의 진핵세포의 유전자 서열 (코딩 영역, 5' 비-코딩 영역, 또는 3' 비-코딩 영역)에서 발견되는 비-코딩 DNA의 서열을 나타낸다. 그들의 완전한 기능은 알려지지 않았지만, 일부 인핸서가 인트론에 위치한다 (문헌 [Giacopelli F. et al., Gene Expr. 11: 95-104 (2003)]). 이들 인트론 서열은 전사는 되지만, mRNA가 단백질로 번역되기 전에 예비-mRNA 전사체 내로부터 제거된다. 이러한 인트론 제거 과정은 인트론의 어느 한 측에 있는 서열 (엑손)의 자체-스플라이싱에 의해 일어난다.

용어 "인핸서"는 인접한 진핵세포 프로모터로부터의 전사 수준을 상승시켜 유전자 전사를 증가시킬 수 있는 시스-조절 서열을 나타낸다. 인핸서는 DNA의 여러 10 킬로베이스에 걸쳐 프로모터에 대해 작용할 수 있고, 그들이 조절하는 프로모터에 대해 5' 또는 3'일 수 있다. 인핸서는 또한 인트론 내부에 존재할 수 있다.

용어 "3' 비-코딩 서열" 및 "전사 종결자"는 코딩 서열의 하류에 위치하는 DNA 서열을 나타낸다. 이는 폴리아데닐화 인식 서열, 및 mRNA 가공 또는 유전자 발현에 영향을 미칠 수 있는 조절 신호를 코딩하는 다른 서열을 포함한다. 폴리아데닐화 신호는 보통 폴리아데닐산 구역을 mRNA 전구체의 3' 말단에 부가하는 것에 영향을 미치는 것을 특징으로 한다. 3' 영역은 관련된 코딩 서열의 전사, RNA 가공 또는 안정성, 또는 번역에 영향을 미칠 수 있다.

"RNA 전사체"는 DNA 서열의 RNA 폴리머라제-촉매된 전사의 결과 생성된 생성물을 나타낸다. RNA 전사체가 DNA 서열의 완벽한 상보적인 카피인 경우에는 1차 전사체로 지칭되거나, 또는 상기 1차 전사체의 전사후 가공으로부터 유래된 RNA 서열일 수 있는 경우에는 성숙 RNA로서 지칭된다. "전령 RNA" 또는 "mRNA"는 인트론이 없으며 세포에 의해 단백질로 번역될 수 있는 RNA를 나타낸다. "cDNA"는 mRNA에 상보적이며 그로부터 유래된 이중-가닥 DNA를 나타낸다. "센스" RNA는 mRNA를 포함하며 세포에 의해 단백질로 번역될 수 있는 RNA 전사체를 나타낸다. "안티센스 RNA"는 표적 1차 전사체 또는 mRNA의 전체 또는 일부에 대해 상보적이고, 표적 유전자의 발현을 차단하는 RNA 전사체를 나타낸다 (미국 특허 제5,107,065호, 제WO 99/28508호). 안티센스 RNA는 특정한 유전자 전사체의 임의의 부분, 즉 5' 비-코딩 서열, 3' 비-코딩 서열, 또는 코딩 서열에 대해 성보적일 수 있다. "기능적 RNA"는 번역되지는 않지만 세포 과정에 대해 효과를 갖는 안티센스 RNA, 리보자임 RNA, 또는 다른 RNA를 나타낸다.

용어 "작동가능하게 연결된"은 하나의 기능이 다른 것에 의해 영향을 받도록 단일 핵산 단편과 핵산 서열이 회합되는 것을 나타낸다. 예를 들어, 프로모터는 코딩 서열의 발현에 영향을 미칠 수 있을 때 상기 코딩 서열에 작동가능하게 연결된다 (즉, 상기 코딩 서열은 프로모터의 전사 조절하에 있다). 코딩 서열은 센스 또는 안티센스 배향으로 조절 서열에 작동가능하게 연결된다.

본원에 사용된 용어 "발현"은 본 발명의 핵산 단편으로부터 유래된 센스 (mRNA) 또는 안티센스 RNA의 전사 및 안정한 축적을 나타낸다. 발현은 또한 mRNA에서 폴리펩티드로의 번역을 나타낸다.

"성숙" 단백질은 번역후 가공된 폴리펩티드, 즉 1차 번역 생성물에 존재하는 임의의 프리- 또는 프로펩티드가 제거된 것을 나타낸다. "전구체" 단백질은 mRNA 번역의 1차 생성물, 즉 프리- 및 프로펩티드가 여전히 존재하는 것을 나타낸다. 프리- 및 프로펩티드는 세포내 국부화 신호일 수 있지만, 이로 제한되지 않는다.

용어 "레콤비나제"는 DNA 구조를 변경하기 위해 부위-특이적 재조합을 수행하는 효소를 나타내며, 이는 트랜스포사제, 람다 통합/제거 효소뿐 아니라, 부위-특이적 레콤비나제를 포함한다.

"레콤비나제 부위" 또는 "부위-특이적 레콤비나제 서열"은 레콤비나제가 인식하여 결합하는 DNA 서열을 나타낸다. 이는 기능이 유지되는 한 야생형 또는 돌연변이체 레콤비나제 부위일 수 있으며, 레콤비나제 효소는 여전히 상기 부위를 인식하여 DNA 서열에 결합하고, 인접한 두개의 레콤비나제 부위 사이의 재조합을 촉매할 수 있다.

"형질전환"은 핵산 분자를 숙주 유기체에 전달하여 유전적으로 안정한 내력을 나타내는 것이다. 예를 들어, 핵산 분자는 자율적으로 복제하는 플라스미드일 수 있거나, 또는 숙주 유기체의 게놈에 통합될 수 있다. 형질전환된 핵산 단편을 함유하는 숙주 유기체는 "트랜스제닉" 또는 "재조합" 또는 "형질전환된" 유기체로서 지칭된다.

용어 "플라스미드", "벡터" 및 "카세트"는 세포의 중심적인 대사의 일부분이 아닌 유전자를 보유하며, 보통 환형의 이중-가닥 DNA 단편의 형태를 갖는 별개의 염색체 요소를 나타낸다. 이러한 요소는 임의의 공급원으로부터 유래된 선형 또는 환형의 단일- 또는 이중-가닥 DNA 또는 RNA의 자율 복제 서열, 게놈 통합 서열, 파지 또는 뉴클레오티드 서열일 수 있으며, 여기서, 수많은 뉴클레오티드 서열이 특유한 구축물 내로 접합 또는 재조합되고, 상기 구축물은 적절한 3' 비번역 서열과 함께 선택된 유전자 생성물을 위한 프로모터 단편 및 DNA 서열을 세포 내로 도입할 수 있다. "발현 카세트"는 외래 유전자를 함유하고, 상기 외래 유전자 이외에도 외래 숙주에서 상기 유전자의 발현을 증가시킬 수 있는 요소를 갖는 특정한 벡터를 나타낸다.

용어 "상동성 재조합"은 (교차되는 동안) 2개의 DNA 분자 사이의 DNA 단편의 교환을 나타낸다. 교환된 단편은 상기 2개의 DNA 분자 사이에서 동일한 뉴클레오티드 서열 부위 (즉, "상동성 영역")에 의해 플랭킹된다. 용어 "상동성 영역"은 상동성 재조합에 참여하는 서로 상동성인 핵산 단편 상의 뉴클레오티드 서열의 신장부를 나타낸다. 효과적인 상동성 재조합은 일반적으로 이들 상동성 영역의 길이가 약 10 bp 이상이고, 바람직하게는 약 50 bp 이상인 경우에 일어난다. 전형적으로, 재조합을 위해 의도된 단편은 표적화된 유전자 파괴 또는 대체가 필요한 2개 이상의 상동성 영역을 함유한다.

용어 "서열 분석 소프트웨어"는 뉴클레오티드 또는 아미노산 서열의 분석에 유용한 임의의 컴퓨터 알고리즘 또는 소프트웨어 프로그램을 나타낸다. "서열 분석 소프트웨어"는 시판되는 것이거나, 독자적으로 개발된 것일 수 있다. 전형적인 서열 분석 소프트웨어로는 1) GCG 집합의 프로그램 (위스콘신 팩키지 버전 9.0(Wisconsin Package Version 9.0), 제네틱 컴퓨터 그룹(Genetics Computer Group; GCG), 미국 위스콘신주 매디슨 소재); 2) BLASTP, BLASTN, BLASTX (문헌 [Altschul et al., J. Mol. Biol. 215:403-410 (1990)]); 3) DNASTAR (디엔에이스타, 인크(DNASTAR, Inc.), 미국 위스콘신주 매디슨 소재); 4) 시퀀쳐(Sequencher) (진 코드스 코포레이션(Gene Codes Corporation), 미국 미시건주 앤 아버 소재); 및 5) 스미스-워터맨(Smith-Waterman) 알고리즘이 도입된 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])이 있으나, 이로 제한되는 것은 아니다. 본원의 문맥상, 서열 분석 소프트웨어를 분석을 위해 이용하는 경우, 달리 명시하지 않는 한 분석 결과는 참조한 프로그램의 "디폴트 값"을 기초로 할 것이라는 것을 이해할 것이다. 본원에 사용된 "디폴트 값"은 처음 초기화시에 소프트웨어에 원래 내장되어 있는 임의의 집합의 값 또는 파라미터를 의미할 것이다.

용어 "보존된 도메인" 또는 "모티프"는 진화론적으로 관련이 있는 단백질의 정렬된 서열을 따라 특정한 위치에서 보존된 아미노산의 집합을 의미한다. 다른 위치의 아미노산은 상동성 단백질들 사이에서 다를 수 있는 반면, 특정한 위치에서 고도로 보존된 아미노산은 단백질의 구조, 안정성 또는 활성에 필수적인 아미노산을 나타낸다. 이들은 단백질 상동체 족의

정렬된 서열에서 고도의 보존성에 의해 확인되기 때문에, 새로 결정된 서열을 갖는 단백질이 이전에 확인된 단백질 족에 속하는지를 결정하는 식별자 또는 "사인"으로서 사용될 수 있다. Δ15 데새투라제 활성을 갖는 진균 단백질임을 표시하는 모티프는 서열 223로 제공되고, Δ12 데새투라제 활성을 갖는 진균 단백질임을 표시하는 모티프는 서열 224으로 제공된다.

본원에서 이용된 표준 재조합 DNA 및 분자 클로닝 기술은 당업계에 공지되어 있고, 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)] (이후 "마니아티스"); [Silhavy, T. J., Bennan, M. L. and Enquist, L. W., Experiments with Gene Fusions, Cold Spring Harbor Laboratory: Cold Spring Harbor, NY (1984)] 및 [Ausubel, F, M. et al., Current Protocols in Molecular Biology, published by Greene Publishing Assoc, and Wiley-Interscience (1987)]에 기재되어 있다.

DHA 생성을 위한 바람직한 미생물 숙주: 야로위아 리폴리티카

본 발명자들에 의한 작업 이전에는 (피카타지오(Picataggio) 등의 제WO 2004/101757호 참조), PUFA를 위한 생성 플랫폼으로서 사용하기에 적합한 미생물 부류로서 유질 효모를 조사한 적이 없었다. 전형적으로 유질 효모로서 확인된 족으로는 야로위아, 칸디다, 로도토룰라, 로도스포리듐, 크립토코커스, 트리코스포론 및 리포마이세스가 있으나, 이로 제한되지 않는다. 보다 구체적으로, 오일-합성 효모의 예로는 로도스포리듐 토룰리데스(*Rhodosporidium toruloides*), 리포마이세스 스타르케이이(*Lipomyces starkeyii*), 리포마이세스 리포페루스(*L. lipoferus*), 칸디다 레브카우피(*Candida revkaufi*), 칸디다 풀케리마(*C. pulcherrima*), 칸디다 트로피칼리스(*C. tropicalis*), 칸디다 유틸리스(*C. utilis*), 트리코스포론 풀란스(*Trichosporon pullans*), 트리코스포론 쿠타네움(*T. cutaneum*), 로도토룰라 글루티누스(*Rhodotorula glutinus*), 로도토룰라 그라미니스(*R. graminis*) 및 야로위아 리폴리티카 (이전에는 칸디다 리폴리티카로 분류됨)가 있으나, 이로 제한되지 않는다.

유질 효모는 DHA의 경제적이고 상업적인 생성을 위한 숙주 유기체로서 용이하게 사용될 수 있는 여러 성질을 갖는 것으로 고려되었다. 첫째, 상기 유기체는 천연적으로 오일을 합성 및 축적할 수 있는 것들로 한정되며, 여기서, 상기 오일은 세포 건조 중량의 약 25% 초과, 더욱 바람직하게는 세포 건조 중량의 약 30% 초과, 가장 바람직하게는 세포 건조 중량의 약 40%를 차지할 수 있다. 둘째, 오일 함량이 높은 유질 효모를 성장시키는 기술은 널리 개발되어 있다 (예를 들어, 제EP 0 005 277B1호; 문헌 [Ratledge, C, Prog. Ind. Microbiol. 16:119-206 (1982)] 참조). 그리고, 이들 유기체는 과거에 다양한 목적을 위해 상업적으로 사용되었다. 예를 들어, 야로위아 리폴리티카의 다양한 균주는 역사적으로 이소시트레이트 리아제 (제DD 259637호); 리파제 (제SU 1454852호, 제WO 2001083773호, 제DD 279267호); 폴리히드록시알카노에이트 (제WO 2001088144호); 시트르산 (제RU 2096461호, 제RU2090611호, 제DD 285372호, 제DD 285370호, 제DD 275480호, 제DD 227448호, 제PL 160027호); 에리트리톨 (제EP 770683호); 2-옥소글루타르산 (제DD 267999호); γ-데카락톤 (미국 특허 제6,451,565호, 제FR 2734843호); γ-도데카락톤 (제EP 578388호); 및 피루브산 (제JP 09252790호)의 제조 및 생산에 사용되었다.

유질 효모로서 분류된 유기체들 중에서, 본원의 목적상 야로위아 리폴리티카가 바람직한 미생물 숙주로서 선택되었다. 이러한 선별은, ω-3 지방산을 TAG 분획에 혼입시킬 수 있는 유질 균주를 입수할 수 있었고, 상기 유기체를 유전자 조작할 수 있었으며, 식품 등급의 시트르산의 일반적으로 안전한 것으로 인식된 (Generally Recognized As Safe ("GRAS"), 미국 식약청에 따라) 공급원으로서 이전에 상기 종을 사용하였다는 지식에 기초하였다. 추가의 실시양태에서, 지질 함량이 높고 (건조 중량%로서 측정) 부피 생산율이 높은 (g/Lh^{-1}로서 측정) 야생형 균주의 확인을 위해 표적화된 예비 연구 때문에, ATCC #20362, ATCC #8862, ATCC #18944, ATCC #76982 및/또는 LGAM S(7)1로 지칭되는 야로위아 리폴리티카 균주가 가장 바람직하다 (문헌 [Papanikolaou S., and Aggelis G., Bioresour. Technol. 82(1):43-9 (2002)]).

제WO 2004/101757호에 기재된 바와 같이, 이전에 야로위아 리폴리티카는 ω-3/ω-6 생합성 경로를 코딩하는 유전자의 도입 및 발현에 의해 각각 1.3% ARA 및 1.9% EPA를 생성하도록 유전자 조작되었다. 보다 구체적으로, 두가지 상이한 DNA 발현 구축물 (ARA 합성을 위해 Δ6 데새투라제, Δ5 데새투라제 및 고-친화도 PUFA $C_{18/20}$ 엘론가제, 또는 EPA 합성을 위해 Δ6 데새투라제, Δ5 데새투라제, 고-친화도 PUFA $C_{18/20}$ 엘론가제 및 코돈-최적화된 Δ17 데새투라제를 포함함)를 따로 형질전환시키고, 효소 오로티딘-5'-포스페이트 데카르복실라제 (EC 4.1.1.23)를 코딩하는 야로위아 리폴리티카 염색체 URA3 유전자에 통합시켰다. 적절한 기질을 공급한 숙주 세포의 GC 분석에 의해 ARA 및 EPA의 생성을 검출하였다. 이 작업은 유질 숙주가 ω-6 및 ω-3 지방산을 생성하도록 유전자 조작되는 능력에 대한 개념을 증명하는데는 적합하

였지만, DHA 생성을 입증하거나, 또는 전체 오일 분획 중 5% 초과의 DHA, 더욱 바람직하게는 전체 오일 분획 중 10% 초과의 DHA, 훨씬 더욱 바람직하게는 전체 오일 분획 중 15 내지 20%의 DHA, 가장 바람직하게는 전체 오일 분획 중 25 내지 30%의 DHA를 합성할 수 있도록 복잡한 대사성 유전자를 조작하거나 제안하는데는 실패하였다.

동시 계류 중인 미국 특허 출원 제60/624812호에서는, 야로위아 리폴리티카를 사용한 복잡한 대사성 유전자 조작이 (1) EPA의 합성 및 고축적을 허용하는, 바람직한 데새투라제 및 엘론가제를 확인하기 위해; (2) 오메가 지방산을 저장 지질 풀로 전달하는 아실트랜스퍼라제의 활성을 조작하기 위해; (3) 강력한 프로모터, 다중 카피에서의 발현 및/또는 코돈-최적화를 사용하여 데새투라제, 엘론가제 및 아실트랜스퍼라제를 과발현시키기 위해; (4) EPA의 전반적인 축적을 감소시키는 PUFA 생합성 경로 내에서 특정 유전자의 발현을 하향-조절하기 위해; 그리고 (5) EPA 제조에 영향을 미치는 경로 및 광범위한 조절자를 조작하기 위해 수행한다. 이것은 야로위아 리폴리티카의 하나의 특정 재조합 균주에서 28% 이하의 EPA를 생성하였다.

본 출원에서, 유사한 복잡한 대사성 유전자 조작은 야로위아 리폴리티카 재조합 균주에서 전체 오일 분획 중 5% 초과의 DHA을 생성하기 위해 수행되었다. 보다 구체적으로, 균주는 Δ6 데새투라제/Δ6 엘론가제 경로를 사용하여 유전학적으로 유전자 조작되었고; 별법 실시양태에서, 형질전환 균주를 유전학적으로 유전자 조작하여 Δ9 엘론가제/Δ8 데새투라제 경로에 사용함으로써, GLA가 결여된 고-DHA 오일을 생성할 수 있다. 사용된 대사성 유전자 조작의 측면은 상기 유질 효모에서 DHA 생성을 유의하게 증가시키기 위해 수행될 수 있는 추가 유전자 조작 및 발효 방법과 함께 하기에 논의될 것이다.

개관: 지방산 및 트리아실글리세롤의 미생물성 생합성

일반적으로, 유질 미생물에서의 지질 축적은 성장 배지에 존재하는 전반적인 탄소:질소 비율에 반응하여 유발된다. 유질 미생물에서 유리 팔미테이트 (16:0)의 새로운 합성을 유도하는 상기 과정은 제WO 2004/101757호에 상세하게 기재되어 있다. 팔미테이트는 더욱 장쇄의 포화 및 불포화 지방산 유도체의 전구체이고, 이는 엘론가제 및 데새투라제의 작용에 의해 형성된다. 예를 들어, Δ9 데새투라제의 작용에 의해 팔미테이트가 그의 불포화 유도체 [팔미트올레산 (16:1)]로 전환되고, 유사하게, 팔미테이트가 $C_{16/18}$ 지방산 엘론가제에 의해 신장되어 스테아르산 (18:0)을 형성하고, 이는 Δ9 데새투라제에 의해 그의 불포화 유도체로 전환되어 올레산 (18:1)을 수득할 수 있다.

TAG (지방산에 대한 1차 저장 단위)는 1) 아실트랜스퍼라제를 통해 아실-CoA의 한 분자가 글리세롤-3-포스페이트로 에스테르화되어 리소포스파티드산을 생성하고; 2) 아실트랜스퍼라제를 통해 아실-CoA의 제2 분자가 에스테르화되어 1,2-디아실글리세롤 포스페이트 (흔히 포스파티드산으로 확인됨)를 수득하고; 3) 포스파티드산 포스파타제에 의해 포스페이트를 제거하여 1,2-디아실글리세롤 (DAG)을 수득하고; 4) 아실트랜스퍼라제의 작용에 의해 제3 지방산을 부가하여 TAG를 형성하는 것을 포함하는 일련의 반응 (도 2)에 의해 형성된다.

포화 및 불포화 지방산 및 단쇄 및 장쇄 지방산을 비롯한 광범위한 지방산을 TAG에 도입시킬 수 있다. 아실트랜스퍼라제에 의해 TAG에 도입될 수 있는 지방산의 일부 제한적인 예로는 카프르산 (10:0), 라우르산 (12:0), 미리스트산 (14:0), 팔미트산 (16:0), 팔미트올레산 (16:1), 스테아르산 (18:0), 올레산 (18:1), 바센산 (18:1), LA, 엘레오스테아르산 (18:3), ALA, GLA, 아라키드산 (20:0), EDA, ETrA, DGLA, ETA, ARA, EPA, 베헨산 (22:0), DPA, DHA, 리그노세르산 (24:0), 네르본산 (24:1), 세로트산 (26:0) 및 몬탄산 (28:0) 지방산이 있다. 본 발명의 바람직한 실시양태에서, DHA를 TAG에 도입하는 것이 가장 바람직하다.

DHA, ω-3 지방산의 생합성

올레산이 DHA로 전환되는 대사 과정은 탄소 원자의 첨가에 의한 탄소 쇄의 신장, 및 이중 결합의 첨가에 의한 분자의 탈포화를 포함한다. 이는 소포체 막에 존재하는 일련의 특정한 탈포화 및 신장 효소를 필요로 한다. 그러나, 도 1 및 하기에 기재된 바와 같이, DHA 생성에 대한 여러 별법의 경로가 존재한다 (모든 경우에서, DHA 생성이 EPA 합성을 요구함).

구체적으로, 모든 경로는 Δ12 데새투라제의 작용에 의해 올레산에서 LA (18:2) (첫번째 ω-6 지방산)로의 초기 전환을 포함한다. 이어서, EPA 생합성을 위한 "ω-6 Δ6 데새투라제/Δ6 엘론가제 경로"를 이용하여 (이로써 EPA 생합성은 주로 ω-6 지방산의 형성을 통해 일어남), PUFA를 다음에 따라 형성한다: (1) Δ6 데새투라제의 작용에 의해 LA가 GLA로 전환되고; (2) $C_{18/20}$ 엘론가제의 작용에 의해 GLA가 DGLA로 전환되고; (3) Δ5 데새투라제의 작용에 의해 DGLA가 ARA로 전환되고; (4) Δ17 데새투라제의 작용에 의해 ARA가 EPA로 전환된다. 별법으로, EPA 생합성이 주로 "ω-3 Δ6 데새투라제/Δ6 엘론가제 경로"를 통해 일어나는 경우에는, (1) Δ15 데새투라제의 작용에 의해 LA가 ALA (첫번째 ω-3 지방산)로 전환되고; (2) Δ6 데새투라제의 작용에 의해 ALA가 STA로 전환되고; (3) $C_{18/20}$ 엘론가제의 작용에 의해 STA가 ETA로 전환되

고; (4) Δ5 데새투라제의 작용에 의해 ETA가 EPA로 전환된다. 임의로, ω-6 및 ω-3 지방산의 조합물은 EPA가 생성되기 전에, Δ17 데새투라제의 작용에 의해 DGLA로부터 ETA를 제조할 때, 또는 Δ15 데새투라제 및 Δ17 데새투라제 둘다를 Δ6 데새투라제, $C_{18/20}$ 엘론가제 및 Δ5 데새투라제와 함께 동시 발현시킬 때 합성될 수 있다.

EPA의 생합성에 대한 별법의 경로는 Δ9 엘론가제 및 Δ8 데새투라제를 이용한다. 보다 구체적으로, "ω-6 Δ9 엘론가제/Δ8 데새투라제 경로"를 통해, Δ9 엘론가제의 작용에 의해 LA가 EDA로 전환된 다음, Δ8 데새투라제가 EDA를 DGLA로 전환시킨다. 상기 기재된 바와 같이, Δ5 데새투라제의 작용에 의한 DGLA의 후속적인 탈포화에 의해 ARA를 수득하고, ARA는 Δ17 데새투라제의 작용에 의해 EPA로 직접 전환된다. 반대로, "ω-3 Δ9 엘론가제/Δ8 데새투라제 경로"를 이용하여, 먼저 Δ15 데새투라제의 작용에 의해 LA가 ALA로 전환된다. 그 다음, Δ9 엘론가제의 작용에 의해 ALA가 ETrA로 전환됨 후, Δ8 데새투라제가 ETrA를 ETA로 전환시킨다. Δ5 데새투라제의 작용에 의해 ETA가 후속적으로 탈포화되어 EPA를 수득한다.

EPA 합성 시에, $C_{20/22}$ 엘론가제는 기질을 DPA로 전환시킨다. 이어서, DPA는 Δ4 데새투라제의 작용에 의해 DHA로 전환된다.

명확히 하기 위해, 이들 각 경로 및 구별되는 특징들을 하기 표에 요약하였다.

[표 4]

EPA 생합성을 위한 대안적인 생합성 경로		
명칭	**EPA를 위한 최소 필요 유전자***	**경로**
ω-6 Δ6 데새투라제/Δ6 엘론가제 경로	Δ6D, $C_{18/20}$ ELO, Δ5D, Δ17D, $C_{20/22}$ ELO, Δ4D	PUFA **생성물**의 ω-3/ω-6 비를 개선시킴
ω-3 Δ6 데새투라제/Δ6 엘론가제 경로	Δ15D, Δ6D, $C_{18/20}$ ELO, Δ5D, $C_{20/22}$ ELO, Δ4D	후속적인 PUFA 생합성을 위한 **기질**의 ω-3/ω-6 비를 개선시키고, GLA가 결여된 오일을 생성함
Δ6 데새투라제/Δ6 엘론가제 경로의 조합	Δ15D, Δ6D, $C_{18/20}$ ELO, Δ5D, Δ17D, $C_{20/22}$ ELO, Δ4D	---
ω-6 Δ9 엘론가제/Δ8 데새투라제 경로	Δ9 ELO, Δ8D, Δ5D, Δ17D, $C_{20/22}$ ELO, Δ4D	PUFA **생성물**의 ω-3/ω-6 비를 개선시킴
ω-3 Δ9 엘론가제/Δ8 데새투라제 경로	Δ15D, Δ9 ELO, Δ8D, Δ5D, $C_{20/22}$ ELO, Δ4D	후속적인 PUFA 생합성을 위한 **기질**의 ω-3/ω-6 비를 개선시키고, GLA가 결여된 오일을 생성함
Δ9 엘론가제/Δ8 데새투라제 경로의 조합	Δ15D, Δ9 ELO, Δ8D, Δ5D, Δ17D, $C_{20/22}$ ELO, Δ4D	---
* 약어: "D" = 데새투라제; "ELO" = 엘론가제		

DHA 합성을 위한 미생물 유전자의 선별

DHA의 생성을 위해 야로위아 리폴리티카에 도입될 필요가 있는 특별한 기능들은 숙주 세포 (및 그의 천연 PUFA 프로파일 및/또는 데새투라제/엘론가제 프로파일), 기질의 유용성, 및 원하는 최종 생성물에 의해 좌우된다. 천연 숙주 세포와 관련하여, 야로위아 리폴리티카가 천연적으로 18:2 지방산을 생성할 수 있어서, 천연 Δ12 데새투라제 (서열 28 및 29; 제WO 2004/104167호 참조)를 보유한다는 것이 공지되어 있다. 원하는 최종 생성물과 관련하여, 생성되는 오일의 최종 지방산 프로파일 (즉, 고 DHA 오일의 최종 조성물 중 % GLA)의 면에서 Δ9 엘론가제/Δ8 데새투라제 경로 발현과는 반대되는 Δ6 데새투라제/Δ6 엘론가제 경로 발현의 결과를 상기 기재하였다.

따라서, 일부 실시양태에서는, Δ6 데새투라제/Δ6 엘론가제 경로를 통해 DHA를 생성하는 것이 바람직할 것이다. 따라서, 최소한 하기 유전자가 숙주 유기체에 도입되어 DHA 생합성을 위해 발현되어야 한다: Δ6 데새투라제, $C_{18/20}$ 엘론가제, Δ5

데새투라제, 및 Δ17 데새투라제 또는 Δ15 데새투라제 중 하나 (또는 둘다), $C_{20/22}$ 엘론가제 및 Δ4 데새투라제. 더욱 바람직한 실시양태에서, 숙주 균주는 추가로 Δ9 데새투라제, Δ12 데새투라제, $C_{14/16}$ 엘론가제 및 $C_{16/18}$ 엘론가제 중 하나 이상을 포함한다.

별법의 실시양태에서, DHA를 GLA의 동시 합성 없이 생성하는 것이 바람직하다 (따라서, Δ9 엘론가제/Δ8 데새투라제 경로의 발현이 필요하다). 이 전략은 최소한 하기 유전자가 숙주 유기체에 도입되어 DHA 생합성을 위해 발현되는 것을 필요로 한다: Δ9 엘론가제, Δ8 데새투라제, Δ5 데새투라제, 및 Δ17 데새투라제 또는 Δ15 데새투라제 중 하나 (또는 둘다), $C_{20/22}$ 엘론가제 및 Δ4 데새투라제. 더욱 바람직한 실시양태에서, 숙주 균주는 추가로 Δ9 데새투라제, Δ12 데새투라제, $C_{14/16}$ 엘론가제 및 $C_{16/18}$ 엘론가제 중 하나 이상을 포함한다.

당업자라면 DHA 생합성에 필요한 각각의 효소를 코딩하는 다양한 후보 유전자를 확인할 수 있을 것이다. 유용한 데새투라제 및 엘론가제 서열은 임의의 공급원으로부터 유래될 수 있는데, 예를 들어, 천연 공급원 (박테리아, 조류, 진균, 식물, 동물 등)으로부터 단리되거나, 반-합성 경로를 통해 제조되거나, 새로이 합성될 수 있다. 숙주에 도입되는 데새투라제 및 엘론가제 유전자의 특정한 공급원은 본 발명에서 중요하지는 않지만, 데새투라제 또는 엘론가제 활성을 갖는 특정한 폴리펩티드를 선택하는데 있어서 고려할 점으로는 1) 폴리펩티드의 기질 특이성; 2) 폴리펩티드 또는 그의 성분이 속도-제한적 효소인지의 여부; 3) 데새투라제 또는 엘론가제가 원하는 PUFA의 합성에 필수적인지의 여부; 및/또는 4) 폴리펩티드에 필요한 보조인자가 있다. 발현된 폴리펩티드는 바람직하게는 숙주 세포에서 그가 위치하는 생화학적 환경과 상용성인 파라미터를 갖는다. 예를 들어, 폴리펩티드는 숙주 세포에서 기질에 대해 다른 효소와 경쟁할 수 있어야 한다. 따라서, 폴리펩티드의 K_M 및 특정한 활성의 분석은 주어진 숙주 세포에서 PUFA 생성을 변형시키기 위해 주어진 폴리펩티드의 적합성을 측정하는 것으로 고려된다. 특정 숙주 세포에서 사용된 폴리펩티드는 의도된 숙주 세포에 존재하는 생화학적 조건하에 기능할 수 있지만, 원하는 PUFA를 변형시킬 수 있는 데새투라제 또는 엘론가제 활성을 갖는 임의의 폴리펩티드일 수 있는 것이다.

추가의 실시양태에서, 각각의 특정한 데새투라제 및/또는 엘론가제의 전환 효율을 고려하는 것이 유용할 것이다. 보다 구체적으로, 대부분의 효소는 기질에서 생성물로의 100% 전환 효율로 기능하지 않기 때문에, 숙주 세포에서 생성된 비정제된 오일의 최종 지질 프로파일은 전형적으로 원하는 DHA 뿐만 아니라 다양한 상류 중간체 PUFA로 구성되는 다양한 PUFA의 혼합물 (예컨대, 100% DHA 오일과 대조적임)일 것이다. 따라서, 각 효소의 전환 효율의 고려는 DHA의 생합성을 최적화할 때 중요한 변수이며, 이는 생성물의 최종 원하는 지질 프로파일의 측면에서 고려되어야 한다.

상기 각각을 고려하여, 적절한 데새투라제 및 엘론가제 활성을 갖는 후보 유전자는 공개적으로 입수가능한 문헌 (예컨대, 진뱅크), 특허 문헌, 및 PUFA를 생성하는 능력을 갖는 미생물의 실험적 분석에 따라 확인될 수 있다. 예를 들어, 하기 진뱅크 관리 번호는 DHA 생합성에 유용한 공개적으로 입수가능한 유전자의 예를 나타낸다: AY131238, Y055118, AY055117, AF296076, AF007561, L11421, NM_031344, AF465283, AF465281, AF110510, AF465282, AF419296, AB052086, AJ250735, AF126799, AF126798 (Δ6 데새투라제); AF390174 (Δ9 엘론가제); AF139720 (Δ8 데새투라제); AF199596, AF226273, AF320509, AB072976, AF489588, AJ510244, AF419297, AF07879, AF067654, AB022097 (Δ5 데새투라제); AAG36933, AF110509, AB020033, AAL13300, AF417244, AF161219, AY332747, AAG36933, AF110509, AB020033, AAL13300, AF417244, AF161219, X86736, AF240777, AB007640, AB075526, AP002063 (Δ12 데새투라제); NP_441622, BAA18302, BAA02924, AAL36934 (Δ15 데새투라제); AF338466, AF438199, E11368, E11367, D83185, U90417, AF085500, AY504633, NM_069854, AF230693 (Δ9 데새투라제); AY630574, AY332747, AY278558, AF489589 (Δ4 데새투라제); 및 NP_012339, NP_009963, NP_013476, NP_599209, BAB69888, AF244356, AAF70417, AAF71789, AF390174, AF428243, NP_955826, AF206662, AF268031, AY591335, AY591336, AY591337, AY591338, AY605098, AY605100, AY630573 ($C_{14/16}$, $C_{16/18}$, $C_{18/20}$ 및 $C_{20/22}$ 엘론가제). 유사하게, 특허 문헌은 PUFA 생성과 관련된 유전자의 추가의 여러 DNA 서열 (및/또는 상기 여러 유전자와 관련된 상세한 내용 및 그의 단리 방법)을 제공한다 [예를 들어, 제WO 02/077213호 (Δ9 엘론가제); 제WO 00/34439호 및 제WO 04/057001호 (Δ8 데새투라제); 미국 특허 제5,968,809호 (Δ6 데새투라제); 미국 특허 제5,972,664호 및 미국 특허 제6,075,183호 (Δ5 데새투라제); 제WO 94/11516호, 미국 특허 제5,443,974호, 제WO 03/099216호 및 제WO 05/047485호 (Δ12 데새투라제); 제WO 93/11245호 (Δ15 데새투라제); 제WO 91/13972호 및 미국 특허 제5,057,419호 (Δ9 데새투라제); 미국 출원 제2003/0196217호 (Δ17 데새투라제); 및 제WO 02/090493호 (Δ4 데새투라제); 및, 제WO 00/12720호, 미국 특허 제6,403,349호, 미국 특허 제6,677,145호, 미국 특허 제2002/0139974 A1호, 미국 특허 2004/0111763호 ($C_{14/16}$, $C_{16/18}$, $C_{18/20}$ 및 $C_{20/22}$ 엘론가제)]. 이들 각 특허 및 출원은 그의 전문이 본원에 참고로 포함된다.

상기 예는 제한하려고 의도된 것이 아니며, 상이한 공급원으로부터 유래된 (1) Δ6 데새투라제, $C_{18/20}$ 엘론가제, Δ5 데새투라제, 및 Δ17 데새투라제 또는 Δ15 데새투라제 중 하나 (또는 둘다), $C_{20/22}$ 엘론가제 및 Δ4 데새투라제를 코딩하는 수많은 다른 유전자 (및 임의로 Δ9 데새투라제, Δ12 데새투라제, $C_{14/16}$ 엘론가제 및/또는 $C_{16/18}$ 엘론가제를 코딩하는 다른 유전자); 또는 (2) Δ9 엘론가제, Δ8 데새투라제, Δ5 데새투라제, 및 Δ17 데새투라제 또는 Δ15 데새투라제 중 하나 (또는 둘다), $C_{20/22}$ 엘론가제 및 Δ4 데새투라제를 코딩하는 수많은 다른 유전자 (및 임의로 Δ9 데새투라제, Δ12 데새투라제, $C_{14/16}$ 엘론가제 및/또는 $C_{16/18}$ 엘론가제를 코딩하는 다른 유전자)를 코딩하는 수많은 다른 유전자들이 야로위아 리폴리티카에 도입시키기에 적합할 것이다.

DHA 합성에 바람직한 유전자

야로위아 리폴리티카에서의 발현에 적합할 수 있는 데새투라제 및 엘론가제를 광범위하게 선택할 수 있지만, 본 발명의 바람직한 실시양태에서는 데새투라제 및 엘론가제가 하기 (또는 그의 유도체)로부터 선택된다:

[표 5]

야로위아 리폴리티카에서 EPA 생성합을 위한 바람직한 데새투라제 및 엘론가제			
ORF	유기체	참조문헌	서열번호
Δ6 데새투라제	모르티에렐라 알피나	진뱅크 관리 번호 AF465281; 미국 특허 제5,968,809호	서열 1, 서열 2
Δ6 데새투라제	모르티에렐라 알피나	진뱅크 관리 번호 AB070555	서열 4, 서열 5
$C_{18/20}$ 엘론가제 ("EL01")	모르티에렐라 알피나	진뱅크 관리 번호 AX464731; 제WO 00/12720호	서열 22, 서열 23
$C_{18/20}$ 엘론가제 ("EL02")	트라우스토키트륨 아우레움	미국 특허 제6,677,145호	서열 25, 서열 26
Δ9 엘론가제	이소크리시스 갈바나	진뱅크 관리 번호 AF390174	서열 69, 서열 70
Δ8 데새투라제	유글레나 그라실리스	동시 계류 중인 미국 특허 출원 제11/166993호	서열 77, 서열 78
Δ5 데새투라제	모르티에렐라 알피나	진뱅크 관리 번호 AF067654; 미국 특허 제6,075,183호	서열 6, 서열 7
Δ5 데새투라제	이소크리시스 갈바나	제WO 02/081668 A2호	서열 8, 서열 9
Δ5 데새투라제	호모 사피엔스	진뱅크 관리 번호 NP_037534	서열 11, 서열 12
Δ5/Δ6 데새투라제	다니오 레리오	진뱅크 관리 번호 AF309556	서열 14, 서열 15
Δ5/Δ6 데새투라제	다니오 레리오	진뱅크 관리 번호 BC068224	서열 16
Δ5/Δ6 데새투라제	다니오 레리오	--	서열 17, 서열 18
Δ17 데새투라제	사프롤레그니아 디클리나	미국 특허 출원 제US 2003/0196217 A1호	서열 19, 서열 20

[표 5a]

Δ4 데새투라제	트라우스토키트륨 아우레움	진뱅크 관리 번호 AAN75707	서열 104, 서열 105

$C_{20/22}$ 엘론가제	오스트레오코쿠스 타우리	진뱅크 관리 번호 AY591336	서열 100, 서열 101
$C_{16/18}$ 엘론가제 ("YE2")	야로위아 리폴리티카	--	서열 94, 서열 95
$C_{16/18}$ 엘론가제 ("EL03")	모르티에렐라 알피나	--	서열 86, 서열 87
$C_{16/18}$ 엘론가제 (rEL02)	라투스 노르베기쿠스	진뱅크 관리 번호 AB071986	서열 83, 서열 84
$C_{14/16}$ 엘론가제 ("YE1")	야로위아 리폴리티카	--	서열 97, 서열 98
Δ12 데새투라제	야로위아 리폴리티카	제WO 2004/104167호	서열 28, 서열 29
Δ12 데새투라제	모르티에렐라 이자벨리나	진뱅크 관리 번호 AF417245	서열 30, 서열 31
Δ12 데새투라제 (Fm d12)	푸사륨 모닐리포르메	제WO 2005/047485호	서열 32, 서열 33
Δ12 데새투라제 (An d12)	아스페르길루스 니둘란스	아스페르길루스 니둘란스 게놈 프로젝트에서 콘티그 1.15 (스캐폴드 1); AAG36933; 제WO 2005/047485호	서열 34, 서열 35
Δ12 데새투라제	아스페르길루스 플라부스	진뱅크 관리 번호 AY280867 (버전 AY280867.1; gi:30721844); 제WO 2005/047485호	서열 36
Δ12 데새투라제 (Afd12p)	아스페르길루스 푸미가투스	아스페르길루스 푸미가투스 게놈 프로젝트에서 AFA.133c 344248:345586 리버스 (AfA5C5.001c); 제WO 2005/047485호	서열 37
Δ12 데새투라제 (Mg d12)	마그나포르테 그리세아	마그나포르테 그리세아 게놈 프로젝트에서 콘티그 2.375 중 로커스 MG01985.1; 제WO 2005/047485호	서열 38, 서열 39

[표 5b]

Δ12 데새투라제 (Nc d12)	뉴로스포라 크라싸	진뱅크 관리 번호 AABX01000374; 제WO 2005/047485호	서열 40, 서열 41
Δ12 데새투라제 (Fg d12)	푸사륨 그라미네아륨	푸사륨 그라미네아륨 게놈 프로젝트에서 콘티그 1.233; 제WO 2005/047485호	서열 42, 서열 43
Δ12 데새투라제 (Mad12)	모르티에렐라 알피나	진뱅크 관리 번호 AB020033	서열 44, 서열 45
Δ12 데새투라제 (Skd12)	사카로마이세스 클루이베리	진뱅크 관리 번호 BAD08375	서열 46
Δ12 데새투라제 (Kld12p)	클루이베로마이세스 락티스	"효모 프로젝트 게놀레부레스(*Genolevures*)"의 클루이베로마이세스 락티스 데이타베이스의 KllaOB:35614..36861 안티센스 (m)으로부터의 gnl|GLV|KLLV0B00473g ORF (프랑스 탈렌스 세덱스 LaBRI 바이오포르마틱스 센터)	서열 47, 서열 48
Δ12 데새투라제 (Cad12p)	칸디다 알비칸스	진뱅크 관리 번호 EAK94955	서열 49
Δ12 데새투라제 (Dhd12p)	데바리오마이세스 한세니이 CBS767D	진뱅크 관리 번호 CAG90237	서열 50
Δ15 데새투라제 (Fm d15)	푸사륨 모닐리포르메	제WO 2005/047479호	서열 51, 서열 52

Δ15 데새투라제 (An d15)	아스페르길루스 니둘란스	아스페르길루스 니둘란스 게놈 프로젝트에서 콘티그 1.122 (스캐폴드 9); 제WO 2005/047479호	서열 53, 서열 54
Δ15 데새투라제 (Mg d15)	마그나포르테 그리세아	마그나포르테 그리세아 게놈 프로젝트에서 콘티그 2.1597 중 로커스 MG08474.1; 제WO 2005/047479호	서열 55, 서열 56

[표 5c]

Δ15 데새투라제 (Nc d15)	뉴로스포라 크라싸	진뱅크 관리 번호 AABX01000577; 제WO 2005/047479호	서열 57, 서열 58
Δ15 데새투라제 (Fg d15)	푸사륨 그라미네아륨	푸사륨 그라미네아륨 게놈 프로젝트에서 콘티그 1.320; 제WO 2005/047479호	서열 59, 서열 60
Δ15 데새투라제 (Mad15)	모르티에렐라 알피나	진뱅크 관리 번호 AB182163	서열 61, 서열 62
Δ15 데새투라제 (Kld15p)	클루이베로마이세스 락티스	진뱅크 관리 번호 XM_451551	서열 63, 서열 64
Δ15 데새투라제 (Cad15p)	칸디다 알비칸스	진뱅크 관리 번호 EAL03493	서열 65
Δ15 데새투라제 (Skd15)	사카로마이세스 클루이베리	진뱅크 관리 번호 BAD11952	서열 66
Δ15 데새투라제 (Dhd15p)	데바리오마이세스 한세니이 CBS767	진뱅크 관리 번호 CAG88182	서열 67
Δ15 데새투라제 (Afd15p)	아스페르길루스 푸미가투스	진뱅크 관리 번호 EAL85733	서열 68
* 주의: 아스페르길루스 푸미가투스 게놈 프로젝트는 유니버시티 오브 맨체스터의 공동 연구기관인 생어 인스티튜트(Sanger Institute) 및 더 인스티튜트 오브 게놈 리써치(The Institute of Genome Research (TIGR))에서 후원되고, 아스페르길루스 니둘란스 게놈 프로젝트는 매리랜드주 캠브리지 소재의 센터 포 게놈 리써치(Center for Genome Research (CGR))에서 후원되고, 마그나포르테 그리세아 게놈 프로젝트는 CGR 및 인터내셔널 라이스 블라스트 게놈 협회(International Rice Blast Genome Consortium)에서 후원하고, 푸사륨 그라미네아륨 게놈 프로젝트는 CGR 및 인터내셔널 지베렐라 제애 게노믹스 협회(International Gibberella zeae Genomics Consortium (IGGR)에 의해 후원되었다.			

본 발명자들은 야로위아 리폴리티카에서 발현될 때 각 효소의 기질 특이성 및/또는 기질 선택성을 측정하거나 확인하기 위해 다양한 엘론가제에 대해 상당한 분석을 수행하였다. 예를 들어, 두 가지 야로위아 리폴리티카 엘론가제의 코딩 서열을 공개적으로 입수하고, 각각의 단백질이 추정적 장쇄 지방-아실 엘론가제로서 해석되거나, 다른 지방산 엘론가제에 대해 유의한 상동성을 공유하였지만, 이들 효소의 기질 특이성은 측정된 적이 없었다. 본원에서 수행한 분석에 기초하여, YE1은 C_{16} 지방산을 생성하는 기질로서 C_{14} 지방산을 우세하게 사용하는 지방산 엘론가제 (즉, $C_{14/16}$ 엘론가제)인 것으로 긍정적으로 결정되었고, YE2는 C_{18} 지방산을 생성하는 기질로서 C_{16} 지방산을 우세하게 사용하는 지방산 엘론가제 (즉, $C_{16/18}$ 엘론가제)임이 결정되었다. 이와 관련하여, 신규한 모르티에렐라 알피나 ELO3 유전자를 확인할 때, 다른 지방산 엘론가제에 대한 상동성으로서 서열을 특성화하였다. 그러나, $C_{16/18}$ 엘론가제로서 ELO3의 특이성을 확인하기 위해서는 지질 프로파일 분석이 필요하였다.

Δ12 데새투라제와 관련하여, 본 발명자들은 놀랍게도 야로위아 리폴리티카에서 18:2를 생성하는데 있어서 푸사륨 모닐리포르메 Δ12 데새투라제 (서열 32에 의해 코딩됨)가 천연 야로위아 리폴리티카 Δ12 데새투라제에 비해 더 높은 효율로 기능한다는 것을 발견하였다 (제WO 2005/047485호 참조). 구체적으로, 야로위아 리폴리티카에서 TEF 프로모터의 조절하에 푸사륨 모닐리포르메 Δ12 데새투라제의 발현은, 이전에 TEF 프로모터의 조절하에 야로위아 리폴리티카 Δ12 데새투라제를 코딩하는 키메라 유전자에 의해 수득되는 것 (LA의 생성물 축적률 59%)보다 18:2를 더 높은 수준 (LA의 생성물 축

적률 68%)으로 생성하는 것으로 측정되었다. 이는 ([18:2+ 18:3]/[18:1+ 18:2+ 18:3])×100으로 계산되는 기질 전환율(%)의 차이가 각각 85% 및 74%인 것에 상응한다. 이러한 결과에 기초하여, 본 발명의 진균 푸사륨 모닐리포르메 Δ12 데새투라제의 발현은 야로위아 리폴리티카의 고농도 DHA-생산 균주의 유전자 조작 수단으로서 다른 공지된 Δ12 데새투라제에 비해 바람직하다 (그러나, 당업자는 푸사륨 모닐리포르메 Δ12 데새투라제의 활성이 예를 들어 코돈-최적화에 따라 야로위아 리폴리티카에서 개선될 수 있다고 예측할 것이다).

별법으로, 야로위아 리폴리티카에서 개선된 효율로 기능할 수 있는 새로운 5가지 Δ12 데새투라제가 최근에 확인되었다. 구체적으로, 사카로마이세스 클루이베리 Δ12 데새투라제 (진뱅크 관리 번호 BAD08375)가 문헌 [Watanabe et al., Biosci. Biotech. Biocheml. 68(3):721-727 (2004)]에 기재되었고, 모르티에렐라 알피나로부터의 Δ12 데새투라제 (진뱅크 관리 번호 AB182163)는 문헌 [Sakuradani et al., Eur. J. Biochem. 261 (3):812-820 (1999)]에 기재되었다. 이들 서열 및 상기 기재된 방법을 이용하여, 추가의 3가지 Δ12 데새투라제가 본 발명자들에 의해 확인되었다: 클루이베로마이세스 락티스 gnl | GLV | KLLA0B00473g ORF (서열 48), 칸디다 알비칸스 진뱅크 관리 번호 EAK94955 (서열 49) 및 데바리오마이세스 한세니이 CBS767 진뱅크 관리 번호 CAG90237 (서열 50). 야로위아 리폴리티카에서 이들 추가의 Δ12 데새투라제의 과발현은 LA의 생성을 증가시키는 수단으로서 유용할 수 있으며, 이로써 다른 하류 PUFA (예컨대, DHA)의 생성을 증가시킬 수 있다.

또다른 바람직한 실시양태에서, 푸사륨 모닐리포르메 (서열 51 및 52)는 ALA의 생성을 증가시키는데 바람직한 Δ15 데새투라제인데, 그 이유는 이 특정한 Δ15 데새투라제가 이전에 공지된 Δ15 데새투라제에 비하여 여러가지 독특한 특징을 갖고 있기 때문이다. 첫째, 푸사륨 모닐리포르메 Δ15 데새투라제는 그의 상당한 Δ12 데새투라제 활성 (따라서 이작용성 효소로서 특징지워짐)에 의해 구별된다. 이전의 연구에서, 서열 52을 코딩하는 키메라 유전자로 형질전환된 야로위아 리폴리티카의 Δ12 데새투라제-파괴된 균주가 LA를 ALA로 96% ([18:3]/[18:2+ 18:3]×100로서 계산된 기질 전환율(%)) 전환시킬 뿐 아니라, 올레산을 LA로 24% (([18:2+ 18:3]/[18:1+ 18:2+ 18:3])×100으로 계산된 기질 전환율(%)) 전환시킬 수 있다는 것이 측정되었다. 둘째로, 푸사륨 모닐리포르메 Δ15 데새투라제는, 다른 이종 발현된 Δ15 데새투라제에 대해 기재된 것에 비해 [예컨대, 비-유질 효모 사카로마이세스 세레비지애에서 캐노르합디티스 엘레강스 Δ15 데새투라제를 발현할 때 ALA의 생성물 축적률(%)이 단지 4.1%인 반면에 (문헌 [Meesapyodsuk et al., Biochem. 39:11948-11954 (2000)]), 사카로마이세스 세레비지애에서 브라시카 나푸스 Δ15 데새투라제를 발현할 때 ALA의 생성물 축적률(%)은 단지 1.3%이었음 (문헌 [Reed., D.W. et al., Plant Physiol. 122:715-720 (2000)])], 야로위아 리폴리티카에서 발현될 때 ALA를 매우 고농도로 합성가능하게 한다 [즉, 서열 52을 코딩하는 키메라 유전자로 형질전환된 야로위아 리폴리티카가 형질전환된 숙주 세포에서의 전체 지방산에 비해 ALA의 생성물 축적률(%)이 31%인 것으로 입증되었으며, 이는 ALA로의 전환 효율 ([18:3]/[18:2+ 18:3]×100로서 계산)이 83%와 동등함]. 마지막으로, 푸사륨 모닐리포르메 Δ15 효소는 18:2의 하류 ω-6 유도체에 대해 비교적 광범위한 기질 특이성을 갖는다. 구체적으로, Δ15 데새투라제는 GLA에서 STA로의 전환, DGLA에서 ETA로의 전환, 및 ARA에서 EPA로의 전환을 촉매할 수 있다.

바람직한 Δ15 데새투라제로서 푸사륨 모닐리포르메 Δ15 효소가 확인되었으나, 6가지 새로운 Δ15 데새투라제가 야로위아 리폴리티카에서 개선된 효율로 기능할 수 있다는 것이 최근 확인되었다. 구체적으로, 사카로마이세스 클루이베리 Δ15 데새투라제 (진뱅크 관리 번호 BAD11952; Skd15)는 문헌 [Oura et al., Microbiol. 150:1983-1990 (2004)]에 기재되어 있고, 모르티에렐라 알피나 Δ15 데새투라제 (진뱅크 관리 번호 AB182163; Mad15)는 문헌 [Sakuradani et al., Appl. Microbiol. Biotechnol. 66:648-654 (2005)]에 기재되어 있다. 각각 이전에 확인된 사카로마이세스 클루이베리 및 모르티에렐라 알피나 Δ12 데새투라제에 대한 그들의 근접한 상동성에 이은 기능적 활성의 측정에 기초하여 상기 두 서열의 일부가 확인되었기 때문에, 이들 두 쌍의 단백질은 푸사륨 모닐리포르메, 아스페르길루스 니둘란스, 마그나포르테 그리세아, 뉴로스포라 크라싸 및 푸사륨 그라미네아륨과 유사한 밀접한 관련이 있는 진균 Δ12 및 Δ15 데새투라제의 추가의 예로 제공된다 (상기 표 참조). 상기 발견은 진균 Δ12 데새투라제-유사 서열의 "쌍"이 Δ15 데새투라제 활성을 갖는 하나의 단백질 및 Δ12 데새투라제 활성을 갖는 하나의 단백질을 포함한다는 본 발명자들의 이전의 가설을 더욱더 뒷받침한다 (제WO 2005/047480호 및 제WO 2005/047485호 참조). 따라서, Δ12 데새투라제-유사 단백질의 유사한 "쌍"이 클루이베로마이세스 락티스, 칸디다 알비칸스, 데바리오마이세스 한세니이 CBS767 및 아스페르길루스 푸미가투스에서 확인되었고, 예상된 바와 같이, 각 쌍의 한 구성원은 이전에 확인된 사카로마이세스 클루이베리 Δ12 데새투라제 (Skd12)에 대해 더욱 근접하게 정렬하고, 나머지는 Skd15에 더욱 근접하게 정렬하였다 (도 3A). 따라서, 이 분석에 기초하여, 본 발명자들은 야로위아 리폴리티카에서의 과발현이 ω-3 지방산의 생성을 증가시키는데 유용할 수 있는 추정적 진균 Δ15 데새투라제로서 클루이베로마이세스 락티스 진뱅크 관리 번호 XM_451551, 데바리오마이세스 한세니이 CBS767 진뱅크 관리 번호 CAG88182, 칸디다 알비칸스 진뱅크 관리 번호 EAL03493 및 아스페르길루스 푸미가투스 진뱅크 관리 번호 EAL85733을 확인하였다.

추가의 실시양태에서에서, 본 발명자들은 Δ12 데새투라제 활성과 대조적으로 Δ15 데새투라제 활성을 갖는 진균 서열을 쉽게 구별하는 수단을 확인하였다. 구체적으로, Mad12, Skd12, Nc d12, Fm d12, Mg d12, An d12, Fg d12, Dhd12p,

Kld12p, Cad12p, Afd12p, Mad15, Skd15, Nc d15, Fm d15, Mg d15, An d15, Fg d15, Dhd15p, Kld15p, Cad15p 및 Afd15p로 이루어진 아미노산 서열 정렬을 분석할 때 (상기 표 참조), 모든 진균 Δ15 또는 Δ12 데새투라제가 Fm d15의 위치 102 (서열 52)에 상응하며 고도로 보존된 His 박스 1 ("HECGH"; 서열 222)로부터 떨어져 있는 단지 3개의 아미노산 잔기인 위치에서 각각 Ile 또는 Val 아미노산 잔기를 함유한다는 것이 명백해졌다 (표 6).

[표 6]

진균 Δ 12 및 Δ 15 데새투라제의
보존된 His 박스 1주변의 아미노산 정렬

상응하는 아미노산 잔기	모티프	데새투라제
서열 45 의 107-118	WVLAHECGHQSF	Mad12
서열 46 의 116-127	WVLAHECGHQAF	Skd12
서열 41의 153-164	WVLAHECGHQAF	Nc d12
서열33 의 149-160	WVIAHECGHGAF	Fm d12
서열 39의 160-171	WVLAHECGHQAF	Mg d12
서열 35의 143-154	WVLAHECGHQAF	An d12
서열 43의 130-141	WVIAHECGHGAF	Fg d12
서열 48 의 106-117	WVLAHECGHQAF	Kld12p
서열 49 의 135-146	WVLAHECGHQAF	Cad12p
서열 50 의 120-131	WVLAHECGHQAF	Dhd12p
서열 37의 142-153	WVLAHECGHQAF	Afd12p
서열 62 의 105-116	WILAHECGHGAF	Mad15
서열 66 의 117-128	WILAHECGHSAF	Skd15
서열 58의 119-130	WILAHECGHGAF	Nc d15
서열 52의 101-112	WILGHECGHGAF	Fm d15
서열 56의 95-106	WILAHECGHGAF	Mg d15
서열 54의 88-99	WILAHECGHGAF	An d15
서열 60의 101-112	WILGHECGHGAF	Fg d15
서열 64 의 117-128	WILAHECGHGAF	Kld15p
서열 65 의 130-141	WILAHECGHGAF	Cad15p
서열 67 의 132-143	WILAHECGHGAF	Dhd15p
서열 68 의 94-105	WILAHECGHGAF	Afd15p

본 발명자들은 상기 위치에 있는 Ile 및 Val이 각각 진균 데새투라제에서 Δ15 및 Δ12 데새투라제 특이성의 결정자인 것으로 결론지었다. 보다 구체적으로, 본 발명자들은 상응하는 잔기(들) (즉, 또는 모티프 IXXHECGH [서열 223])에서 Ile를 갖는 임의의 진균 Δ12 데새투라제-유사 단백질이 Δ15 데새투라제일 것이고, 상응하는 잔기(들) (즉, 또는 모티프 VXXHECGH [서열 224])에서 Val을 갖는 임의의 진균 Δ12 데새투라제-유사 단백질이 Δ12 데새투라제일 것이라고 제안한다. 따라서, 상기 단일 루이신/발린 아미노산은 미래의 진균 데새투라제가 확인되고 해석되는 것으로 고려될 중요한 잔기일 것이다. 또한, 상기 위치에서 Ile에서 Val으로의 변화로 인한 돌연변이가 진균 Δ12 데새투라제-유사 단백질을 코딩하는 유전자 (예컨대, 본원에서 서열 52로 기재된 푸사륨 모닐리포르메 데새투라제)에서, 예를 들어 Δ12 탈포화에 대한 효소 특이성을 변경시키고, 반대로, 상기 위치에서 Val에서 Ile로의 변화로 인한 돌연변이가 예를 들어 Δ15 탈포화에 대한 효소 특이성을 변경시킬 것으로 고려된다.

바람직한 실시양태에서, 다양한 Δ5 데새투라제가 이용될 특정한 경로에 따라 DHA 생성을 위한 숙주 세포에서 발현하는데 가장 이로운 것으로 선택될 수 있다. 구체적으로, ω-6 Δ6 데새투라제/Δ6 엘론가제 경로 또는 ω-6 Δ9 엘론가제/Δ8 데새투라제 경로를 발현할 때, 모르티에렐라 알피나, 이소크리시스 갈바나 및 호모 사피엔스 Δ5 데새투라제가 바람직하다. 반대로, ω-3 Δ6 데새투라제/Δ6 엘론가제 경로 또는 ω-3 Δ9 엘론가제/Δ8 데새투라제 경로를 이용하는 것이 바람직한 경우에는 (이로써 ω-3 PUFA의 합성이 우세함), 예를 들어 피토프테라 메가스페르마(*Phytopthera megasperma*) 또는 다니오 레리오로부터의 ω-3-우호성 Δ5 데새투라제를 사용하는 것이 유리할 수 있다. 하스팅스(Hastings) 등은 원래 사카로마이세스 세레비지애에서 다니오 레리오 cDNA (진뱅크 관리 번호 AF309556)의 발현이 이작용성 Δ6 및 Δ5 데새투라제 활성 (ω-6 기질에 비해 ω-3 기질에 대해 명확한 우호성을 가지며, Δ5 데새투라제에 비해 Δ6 데새투라제 활성이 약간 더 높음)을 나타낸다고 보고하였다. 그 후, 본 발명자들은 ORF의 위치 984에서 1 bp (T) 결실 (결과적으로 돌연변이 없음) 및 위치 1171에서 1 bp 치환 (G에서 A로) (결과적으로 V에서 M으로 아미노산 변화)에서 차이가 나는, 진뱅크 관리 번호 AF309556의 동족체로서 진뱅크 관리 번호 BC068224를 확인하였다. 그 후, 진뱅크 관리 번호 BC068224의 V1171 M 돌연변이를 제외하고는 진뱅크 관리 번호 AF309556과 동일한 돌연변이체 단백질 (본원에서 "Drd6/d5(M)"로서 확인됨)을 생성하였다. 본 발명자들의 초기 연구에서 사카로마이세스 세레비지애에서 Drd6/d5(M)의 발현이 진뱅크 관리 번호 AF309556에 비해 약 50% 더 낮은 활성을 나타내는 것으로 결정되었지만, ETA를 제조하는 야로위아 균주의 발현은 이

작용성 Δ5/Δ6 데새투라제가 훨씬 더 ω-3-특이적이었음을 확인하였다. 따라서, 상기 효소 (본원에서 서열 18으로 확인됨), 또는 유사한 기질 특이성을 갖는 효소가 ω-3 PUFA의 증가된 합성을 위한 ω-3 Δ6 데새투라제/Δ6 엘론가제 경로 또는 ω-3 Δ9 엘론가제/Δ8 데새투라제 경로의 발현시에 바람직하다.

물론, 본 발명의 별법의 실시양태에서, 서열 2, 5, 7, 9, 12, 15, 18, 20, 23, 26, 29, 31, 33, 35-37, 39, 41, 43, 45, 46, 48-50, 52, 54, 56, 58, 60, 62, 64-68, 70, 78, 84, 87, 95, 98, 101 및 105에 의해 코딩되는 데새투라제 및 엘론가제와 실질적으로 동일한 다른 DNA 또한 야로위아 리폴리티카에서 DHA의 생성을 위해 사용될 수 있다. "실질적으로 동일한"은 선택된 폴리펩티드, 또는 상기 아미노산 서열을 코딩하는 핵산 서열에 대해 80%, 90% 또는 95% 이상의 상동성으로 증가한 우호성을 나타내는 아미노산 서열 또는 핵산 서열을 위해 의도된다. 폴리펩티드의 경우, 비교 서열의 길이는 일반적으로 16개 이상의 아미노산, 바람직하게는 20개 이상의 아미노산, 가장 바람직하게는 35개 이상의 아미노산이다. 핵산의 경우, 비교 서열의 길이는 일반적으로 50개 이상의 뉴클레오티드, 바람직하게는 60개 이상의 뉴클레오티드, 더욱 바람직하게는 75개 이상의 뉴클레오티드, 가장 바람직하게는 110개 이상의 뉴클레오티드이다.

전형적으로, 상동성은 서열 분석 소프트웨어를 이용하여 측정되고, 여기서 용어 "서열 분석 소프트웨어"는 뉴클레오티드 또는 아미노산 서열의 분석에 사용되는 임의의 컴퓨터 알고리즘 또는 소프트웨어 프로그램을 나타낸다. "서열 분석 소프트웨어"는 시판되는 것이거나 독자적으로 개발될 수 있다. 전형적인 서열 분석 소프트웨어로는 1) GCG 집합의 프로그램 (위스콘신 팩키지 버전 9.0, 제네틱 컴퓨터 그룹, 미국 위스콘신주 매디슨 소재); 2) BLASTP, BLASTN, BLASTX (문헌 [Altschul et al., J. Mol. Biol. 215:403-410 (1990)]); 3) DNASTAR (디엔에이스타, 인크, 미국 위스콘신주 매디슨 소재); 및 4) 스미스-워터맨 알고리즘이 도입된 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])이 있으나, 이로 제한되는 것은 아니다. 본원의 문맥상, 서열 분석 소프트웨어를 분석을 위해 이용하는 경우, 달리 명시하지 않는 한 분석 결과는 참조한 프로그램의 "디폴트 값"을 기초로 할 것이라는 것을 이해할 것이다. 본원에 사용된 "디폴트 값"은 처음 초기화시에 소프트웨어에 원래 내장되어 있는 임의의 집합의 값 또는 파라미터를 의미할 것이다. 일반적으로, 이러한 컴퓨터 소프트웨어는 다양한 치환, 결실 및 다른 변형에 대한 상동성 정도를 분배함으로써 유사한 서열을 매칭시킨다.

더욱 바람직한 실시양태에서, 서열 2, 9, 12, 20, 23, 26, 70, 78, 84, 101 및 105에 기재된 것과 실질적으로 동일한 데새투라제 및 엘론가제를 코딩하는 코돈-최적화된 유전자를 사용한다. 구체적으로, 당업자에게 공지된 바와 같이, 천연 유전자의 코돈을 선택된 숙주 미생물에서 최적의 유전자 발현을 위한 코돈으로 대체시켜서 코딩된 mRNA의 번역 효율을 증가시킴으로써 이종 유전자의 발현을 증가시킬 수 있다. 따라서, 외래 숙주에서 발현되는 특정 폴리펩티드를 코딩하는 유전자의 일부분을 변형시켜서, 변형된 폴리펩티드가 다른 숙주에 의해 바람직한 코돈을 사용하도록 하는 것이 종종 유용하며, 숙주 우호성 코돈의 사용은 상기 폴리펩티드를 코딩하는 외래 유전자의 발현을 실질적으로 개선시킬 수 있다.

일반적으로, 숙주 우호성 코돈은 관심을 가진 특정 숙주 종 내에서 단백질에서의 코돈 사용을 측정하고 (바람직하게는 최다량으로 발현된 것), 가장 높은 빈도로 사용된 것을 측정함으로써 결정될 수 있다. 이어서, 관심을 가진 폴리펩티드 (예컨대, 데새투라제, 엘론가제, 아실트랜스퍼라제)의 코딩 서열은 상기 숙주 종에서 바람직한 코돈을 전부 또는 일부 사용하여 합성될 수 있다. 모든 (또는 일부) DNA 또한 전사된 mRNA에서 존재하는 2차 구조체의 임의의 불안정화 서열 또는 영역을 제거함으로써 합성될 수 있다. 그리고, 모든 (또는 일부) DNA 또한 원하는 숙주 세포에서 염기 조성을 더욱 바람직한 것으로 변경시킴으로써 합성될 수 있다.

추가로, 번역 개시 코돈 'ATG' 주변의 뉴클레오티드 서열은 효모 세포의 발현에 영향을 미치는 것으로 확인되었다. 원하는 폴리펩티드가 효모에서 불량하게 발현되는 경우에는, 외생 유전자의 뉴클레오티드 서열이 효율적인 효모 번역 개시 서열을 포함하여 최적의 유전자 발현을 달성하도록 변형될 수 있다. 효모에서의 발현의 경우, 이는 비효율적으로 발현되는 유전자를 프래임 내에서 내인성 효모 유전자, 바람직하게는 고도로 발현되는 유전자에 융합시킴으로써 부위-지정 돌연변이 유발에 의해 수행될 수 있다. 별법으로, 야로위아 리폴리티카에 대해 본원에서 입증된 바와 같이, 숙주에서 컨센서스 (consensus) 번역 개시 서열을 측정하여, 상기 서열을 관심을 가진 숙주에서의 최적의 발현을 위해 이종 유전자로 유전자 조작할 수 있다.

본 발명에서, 표 5의 여러 데새투라제 및 엘론가제 유전자를 상기 기재된 숙주 우호성을 기준으로 하여 야로위아 리폴리티카에서의 발현을 위해 코돈-최적화하였다. 이는 먼저 야로위아 리폴리티카 코돈 사용 프로파일을 측정하고 (제WO 04/101757호 참조), 바람직한 코돈을 확인함으로써 가능해진다. 이어서, 야로위아 리폴리티카에서의 유전자 발현을 더욱 최적화시키기 위해, 'ATG' 개시 코돈 주변의 컨센서스 서열 (즉, 'MAMMATGNHS' (서열 155), 사용된 핵산 다의성 코드는 M=A/C; S=C/G; H=A/C/T; 및 N=A/C/G/T임)을 측정하였다. 하기 표 7은 야로위아 리폴리티카에서 발현될 때 천연 및 코돈-최적화된 유전자의 활성을 비교하여, 각각의 코돈-최적화된 유전자에 대해 상세하게 제공한다. %Sub. Conv.는 "기질 전환율(%)"의 약어이고 및 코돈-Opt.는 "코돈-최적화된"의 약어이다.

[표 7]

야로위아 리폴리티카에서 DHA 생합성을 위한 가장 바람직한 코돈-최적화된 데새투라제 및 엘론가제					
천연 유전자	천연 유전자 %Sub.Conv.	코돈-Opt. 유전자에서 변형된 전체 염기	코돈-Opt. 유전자 %Sub.Conv.	참조문헌	코돈-Opt. 서열 번호
모르티에렐라 알피나 Δ6 데새투라제 (진뱅크 관리 번호 AF465281)	30%	1374 bp의 152 (144 코돈에 상응)	42%	제WO 04/101753호	서열 3
모르티에렐라 알피나 고친화성 $C_{18/20}$ 엘론가제 (진뱅크 관리 번호 AX464731)	30%	957 bp의 94 (85 코돈에 상응)	47%	제WO 04/101753호	서열 24
트라우스토키트륨 아우레움 $C_{18/20}$ 엘론가제 ("EL02")	33%	817 bp의 114 (108 코돈에 상응)	46%	--	서열 27
사프롤레그니아 디클리나 Δ17 데새투라제 (미국 특허 출원 제2003/0196217 A1호)	23%	1077 bp의 127 (117 코돈에 상응)	45%	동시 계류 중인 미국 특허 출원 제10/840478호	서열 19
이소크리시스 갈바나 Δ9 엘론가제	--	789 bp의 126 (123 코돈에 상응)	30%	--	서열 71
유글레나 그라실리스 Δ8 데새투라제	--	1263 bp의 207 (192 코돈에 상응)	75%	동시 계류 중인 미국 특허 출원 제11/166993호	서열 81
이소크리시스 갈바나 Δ5 데새투라제	7%	1323 bp의 203 (193 코돈에 상응)	32%	--	서열 10
호모 사피엔스 Δ5 데새투라제 (진뱅크 관리 번호 NP_037534)	--	1335 bp의 227 (207 코돈에 상응)	30%	--	서열 13
트라우스토키트륨 아우레움 Δ4 데새투라제	---	1546 bp의 170 (166 코돈에 상응)	20%	--	서열 106
오스트레오코쿠스 타우리 $C2_{0/22}$ 엘론가제	---	903 bp의 160 (147 코돈에 상응)	67%	--	서열 102
라투스 노르베기쿠스 $C_{16/18}$ 엘론가제 (진뱅크 관리 번호 AB071986)	--	792 bp의 127 (125 코돈에 상응)	43%	--	서열 85

본 발명의 추가의 별법의 실시양태에서, 서열 3, 10, 13, 19, 24, 27, 71, 81, 85, 102 및 106에 제시된 바람직한 데새투라제 및 엘론가제와 실질적으로 동일하지 않은 다른 DNA 또한 본 발명의 목적을 위해 사용될 수 있다. 예를 들어, 본 발명에 따라 야로위아 리폴리티카에 도입되기에 유용한 Δ6 데새투라제 폴리펩티드를 코딩하는 DNA 서열은 GLA 또는 STA를 생성하는 능력을 가진 미생물로부터 수득될 수 있다. 이러한 미생물로는 예를 들어 모르티에렐라, 코니디오볼루스(*Conidiobolus*), 피티움, 피토파토라(*Phytophathora*), 페니실륨(*Penicillium*), 포르피리듐, 코이도스포륨(*Coidosporium*), 무코르(*Mucor*), 푸사륨, 아스페르길루스, 로도토룰라 및 엔토모프토라(*Entomophthora*) 속에 속하는 것들이 있다. 포르피리듐 속 중에서, 특히 포르피리듐 크루엔툼(*P. cruentum*)이 관심의 대상이다. 모르티에렐라 속 중에서, 특히 모르티에렐라 엑스텐시아(*M. extensia*), 모르티에렐라 엑시구아, 모르티에렐라 히그로필라, 모르티에렐라 라만니아나 바르. 안굴리스포라(*M. ramanniana var. angulispora*) 및 모르티에렐라 알피나가 관심의 대상이다. 무코르 속 중에서, 무코르 시르시넬로이데스(*M. circinelloides*) 및 무코르 자바니쿠스(*M. javanicus*)가 관심의 대상이다.

별법으로, 모르티에렐라 알피나 Δ6 데새투라제와 실질적으로 동일하지만, 예를 들어 분자의 카르복실 말단으로부터 6번째 탄소에서 지방산 분자가 탈포화될 수 있는 관련 데새투라제 또한 본 발명에서 Δ6 데새투라제로서 유용할 수 있으며, 상

기 데새투라제는 여전히 효과적으로 LA를 GLA로 및/또는 ALA를 STA로 전환시킬 것으로 추정된다. 따라서, 관련 데새투라제 및 엘론가제는 본원에 개시된 데새투라제 및 엘론가제와 실질적으로 동일하게 기능할 수 있는 그들의 능력에 의해 확인될 (또는 생성될) 수 있다.

상기 제안된 바와 같이, 또다른 실시양태에서 당업자는 본원의 목적에 적합한, 예를 들어 Δ12 데새투라제 및 Δ6 데새투라제 활성 둘다를 갖는 융합 단백질을 생성할 수 있다. 이는 연결 링커를 이용하여 Δ12 데새투라제와 Δ6 데새투라제를 융합시킴으로써 가능해질 것이다. Δ12 데새투라제 또는 Δ6 데새투라제는 융합 단백질의 N-말단 부분에 있을 수 있다. 적절한 링커 분자를 고안하고 합성하는 수단은 당업자에게 용이하게 공지되어 있으며, 예를 들어 상기 링커는 알라닌 또는 리신 아미노산의 신장체일 수 있고, 융합 효소의 활성에는 영향을 미치지 않을 것이다.

마지막으로, 서열을 합성하고 서열을 함께 가져오는 방법이 문헌에 확립되어 있음은 당업계에 널리 공지되어 있다. 따라서, 천연 발생 데새투라제 및/또는 엘론가제 유전자의 돌연변이를 달성하기 위해, 시험관내 돌연변이유발 및 선별, 부위-지정 돌연변이유발, 화학적 돌연변이유발, "유전자 뒤섞기(shuffling)" 방법 또는 다른 수단을 이용할 수 있다. 이는 숙주 세포에서 기능하기에 더욱 바람직한 물리적 및 동력학적 파라미터 (예컨대, 긴 반감기 또는 원하는 PUFA의 높은 생성률)를 가지며 생체내에서 데새투라제 또는 엘론가제 활성을 갖는 폴리펩티드의 생성을 허용할 것이다.

요약하면, 야로위아 리폴리티카에서 DHA 생성에 적합한 PUFA 생합성 경로 효소를 코딩하는 바람직한 데새투라제 및 엘론가제 유전자의 서열이 제시되지만, 이들 유전자는 본 발명을 제한하려는 의도는 아니다. 본 발명의 목적에 적합한 PUFA 생합성 경로 효소를 코딩하는 수많은 다른 유전자를 다양한 공급원 (예컨대, 적절한 데새투라제 또는 엘론가제 활성을 갖는, 야생형의 코돈-최적화된 합성 및/또는 돌연변이체 효소)으로부터 단리할 수 있다. 이들 별법의 데새투라제는 하기의 능력을 특징으로 한다: 1) 분자의 카르복실-말단에서부터 세었을 때 17번째 및 18번째 탄소 원자 사이의 지방산의 탈포화, 및 ARA에서 EPA로 및/또는 DGLA에서 ETA로의 전환을 촉매 (Δ17 데새투라제); 2) LA에서 GLA로 및/또는 ALA에서 STA로의 전환을 촉매 (Δ6 데새투라제); 3) DGLA에서 ARA로 및/또는 ETA에서 EPA로의 전환을 촉매 (Δ5 데새투라제); 4) 올레산에서 LA로의 전환을 촉매 (Δ12 데새투라제); 5) LA에서 ALA로의 전환을 촉매 (Δ15 데새투라제); 6) EDA에서 DGLA로 및/또는 ETrA에서 ETA로의 전환을 촉매 (Δ8 데새투라제); 7) DPA에서 DHA로의 전환을 촉매 (Δ4 데새투라제); 및/또는 8) 팔미테이트에서 팔미트올레산로 및/또는 스테아레이트에서 올레산으로의 전환을 촉매 (Δ9 데새투라제). 유사한 방식으로, 본원의 목적에 적합한 엘론가제는 특정한 공급원으로부터의 것으로 제한되지 않고, 대신에, 본원의 목적에 사용되는 효소는 엘론가제가 작용하는 기질에 비해 2개의 탄소만큼 지방산 탄소 쇄를 신장시켜서, 모노- 또는 다불포화 지방산을 생성하는 능력을 특징으로 한다. 보다 구체적으로, 이들 엘론가제는 하기 능력을 특징으로 한다: 1) LA에서 EDA로 및/또는 ALA에서 ETrA로의 신장 (Δ9 엘론가제); 2) C_{18} 기질을 신장시켜 C_{20} 생성물을 생성 ($C_{18/20}$ 엘론가제); 3) C_{14} 기질을 신장시켜 C_{16} 생성물을 생성 ($C_{14/16}$ 엘론가제); 4) C_{16} 기질을 신장시켜 C_{18} 생성물을 생성 ($C_{16/18}$ 엘론가제); 및/또는 5) C_{20} 기질을 신장시켜 C_{22} 생성물을 생성 ($C_{20/22}$ 엘론가제). 또한, 일부 엘론가제는 광범위한 기질 특이성의 결과 여러 엘론가제 반응을 촉매할 수 있다는 것을 주목하는 것이 중요하다.

<u>아실트랜스퍼라제 및 TAG 생합성의 최종 단계에서 이들의 역할</u>

아실트랜스퍼라제는 TAG의 생합성과 밀접한 관계가 있다. 문헌 [D. Sorger and G. Daum, Appl. Microbiol. Biotechnol. 61:289-299 (2003)] 및 [H. Muellner and G. Daum, Acta Biochimica Polonica, 51 (2):323-347 (2004)]에서는 TAG 합성과 관련된 유전자 및 그를 유도하는 대사 중간체에 대한 상세한 설명을 비롯하여, 효모에서의 TAG 생합성에 대해 포괄적으로 간단히 고찰되었다. 상기 문헌의 저자들은 상이한 부류의 진핵세포 아실트랜스퍼라제 유전자 족에 대해 요약하였지만, 이들은 또한 지질 입자에서 TAG 생합성 및 천연 지질 형성의 조절적 측면이 명확하지 않음을 확인하였다.

중성 지질 합성을 유도하는 아실-CoA-의존성 또는 독립성 에스테르화 반응과 관련된 네 가지 진핵세포 아실트랜스퍼라제 유전자 족이 확인되었다:

(1) 아실-CoA:콜레스테롤 아실트랜스퍼라제 (ACAT) 족, EC 2.3.1.26 (일반적으로 스테롤 아실트랜스퍼라제로 공지됨). 이 족의 유전자는 아실-CoA 및 스테롤을 CoA 및 스테롤 에스테르로 전환시키는 것을 담당하는 효소를 포함한다. 이 족은 또한 TAG 생합성의 최종 단계와 관련된 DGAT1을 포함한다.

(2) 레시틴:콜레스테롤 아실트랜스퍼라제 (LCAT) 족, EC 2.3.1.43. 이 족의 유전자는 포스파티딜콜린 및 스테롤을 스테롤 에스테르 및 1-아실글리세로포스포콜린으로 전환시키는 것을 담당한다. 이 족은 또한 인지질의 sn-2 위치의 아실기를 1,2-디아실글리세롤의 sn-3 위치로 전달하여 TAG를 생합성하는 것과 관련된 인지질:디아실글리세롤 아실트랜스퍼라제 (PDAT) 효소를 포함한다.

(3) 디아실글리세롤 아실트랜스퍼라제 (DAG AT) 족, EC 2.3.1.20. 이 족의 유전자 (DGAT2 포함)는 TAG 생합성의 최종 단계와 관련이 있다.

(4) 글리세롤-3-포스페이트 아실트랜스퍼라제 및 아실-CoA 리소포스파티드산 아실트랜스퍼라제 (GPAT/LPAAT) 족. GPAT (E.C. 2.3.1.15) 단백질은 TAG 생합성의 첫번째 단계를 담당하는 반면, LPAAT (E.C. 2.3.1.51) 효소는 TAG 생합성의 두번째 단계와 관련이 있다. 이 족은 또한 인지질과 CoA 사이에서의 아실 교환을 촉매하는 리소포스파티딜콜린 아실트랜스퍼라제 (LPCAT)를 포함한다.

또한, 이들 네 가지 아실트랜스퍼라제 유전자 족은 중성 지질 형성을 위한 중복 생합성 시스템을 나타내고, 상이한 조절, 대안적인 국부화, 및 상이한 기질 특이성의 결과인 것으로 여겨진다 (상기 문헌 [H. Muellner and G. Daum]). 이들 네 가지 각 유전자 족은 야로위아 리폴리티카에서 5% 초과의 DHA를 합성하도록 하는 대사성 유전자 조작과 관련하여 중요하게 본원에서 논의될 것이다.

각종 아실트랜스퍼라제의 기능성

야로위아 리폴리티카에 존재하는 이러한 많은 아실트랜스퍼라제들 사이의 상호작용은 도 4에 개략적으로 도시하였다. TAG 생합성의 직접적인 메카니즘에 우선 초점을 맞추면, 이 과정의 제1 단계는 아실-CoA 1개 분자를 GPAT를 통해 sn-글리세롤-3-포스페이트로 에스테르화하여 리소포스파티드산 (LPA) (및 부산물로서의 CoA)를 생성하는 것이다. 이어서, 리소포스파티드산은 아실-CoA의 또다른 분자의 에스테르화에 의해 포스파티드산 (PA) (및 부산물로서의 CoA)으로 전환되며, 이 반응은 LPAAT에 의해 촉매된다. 이어서, 포스파티드산 포스파타제가 포스파티드산에서 포스페이트기를 제거하여 1,2-디아실글리세롤 (DAG)을 생성한다. 그 후, DAG AT (예컨대, DGAT1, DGAT2 또는 PDAT)에 의해서 DAG의 sn-3 위치에 세번째 지방산이 부가되어 TAG가 형성된다.

과거, DGAT1은 TAG 합성에 특이적으로 관여하여 아실-CoA기를 DAG로 전달하여 TAG를 형성함으로써 아실-CoA와 DAG를 TAG와 CoA로 전환시키는 반응을 촉매하는 효소에 불과하다고 여겨졌다. DGAT1은 ACAT에 상동성인 것으로 알려져 있었으나, 최근의 연구를 통해, ACAT 유전자 족과는 관련이 없는 새로운 족의 DAG AT 효소가 밝혀졌다. 따라서, 이제는 명칭을 달리하여 ACAT 유전자 족과 관련이 있는 DAG AT 효소 (DGAT1 족) 및 관련이 없는 DAG AT 효소 (DGAT2 족)을 구별한다 [Lardizabal et al., J. Biol. Chem. 276(42):38862-38869 (2001)]. DGAT2 족의 구성원은 진핵생물의 모든 주요 종 (진균, 식물, 동물 및 원시(basal) 진핵생물)에서 확인된 바 있다.

훨씬 더 최근에, 문헌 [Dahlqvist et al. (Proc. Nat. Acad. Sci. (USA) 97:6487-6492 (2000)] 및 문헌 [Oelkers et al. (J. Biol. Chem. 275:15609-15612 (2000)]은 TAG 합성이 아실-CoA가 없어도 아실-CoA와 무관한 메카니즘을 통해 일어날 수 있음을 발견하였다. 구체적으로, PDAT가 포스파티딜콜린 기질의 sn-2 위치에서 아실기를 제거하여 DAG로 전달함으로써 TAG가 생성된다. 이 효소는 LCAT 족과 구조적으로 관련이 있고, PDAT의 기능이 DGAT2만큼 잘 특징규명되지는 않았지만 PDAT는 일부 오일종자 식물에서 인지질로부터 "독특한(unusual)" 지방산을 제거하는데 중요한 역할을 할 것이라고 추측되어 왔다 [Banas, A. et al., Biochem. Soc. Trans. 28(6):703-705 (2000)].

사카로마이세스 세레비지애에서의 TAG 합성과 관련하여, 3가지 경로가 기재된 바 있다 [Sandager, L. et al., J. Biol. Chem. 277(8):6478-6482 (2002)). 첫째, TAG는 DAG 및 아실-CoA로부터 DGAT2 (DGA1 유전자에 의해 코딩됨)의 활성에 의해 주로 합성된다. 그러나 보다 최근에는, PDAT (LRO1 유전자에 의해 코딩됨) 또한 확인된 바 있다. 최종적으로, 2종의 아실-CoA:스테롤-아실트랜스퍼라제 (ARE1 유전자 및 ARE2 유전자에 의해 코딩됨)는 아실-CoA 및 스테롤을 이용하여 스테롤 에스테르 (및 소량의 TAG, 문헌 [Sandager et al., Biochem. Soc. Trans. 28(6):700-702 (2000)] 참조)를 생성한다고 알려져 있다. PDAT와 DGAT2는 함께 사카로마이세스 세레비지애 중 오일 생합성의 대략 95%를 담당한다.

DGAT1, DGAT2, PDAT 및 ARE2를 코딩하는 여러가지 공개적으로 입수가능한 서열 (하기 참조)을 기초로, 본 발명자들은 야로위아 리폴리티카에서 DGAT1 (서열 122), DGAT2 (서열 130, 132 및 134 [여기서, 서열 130는 서열 132 및 서열 134에서 제공되는 것과 같은 2종 이상의 추가의 네스티드 ORF를 함유함. 서열 134에 의해 코딩되는 ORF는 다른 공지된 DGAT 효소와 높은 정도의 유사성을 보유하며 서열 134에서의 파괴는 천연 유전자의 DGAT 기능을 소실시키기 때문에, 서열 135의 폴리펩티드가 DGAT 기능성을 보유함이 확인됨]), PDAT (서열 117) 및 ARE2 (서열 119)를 코딩하는 유전자들을 단리하고 특징규명하였다. 그러나, 사카로마이세스 세레비지애에서 개발된 모델 (여기서는 PDAT와 DGAT2가 오일 생합성의 대략 95%를 담당함)과는 달리, 야로위아 리폴리티카의 PDAT, DGAT2 및 DGAT1은 오일 생합성을 최대 약 95% 이하로 담당 (ARE2가 오일 생합성에 추가로 약간 기여할 수도 있음)하는 것으로 발견되었다.

야로위아 리폴리티카의 TAG 분획에서 DHA를 축적하는데 중요한 기능을 할 수 있는 최종 아실트랜스퍼라제 효소는 LPCAT이다. 도 4에 나타난 바와 같이, 이 효소 (EC 2.3.1.23)는 sn-포스파티딜콜린의 sn-2 위치에서의 2원 아실 교환을 담당하여 ω-6 및 ω-3 PUFA 생합성을 증진시킨다는 가설이 있다. 이 가설은 하기 사람들의 연구를 기초로 한다: (1) 스팀느 에스.(Stymne S.) 및 에이.케이.스토바르트(A.K. Stobart) [Biochem J. 223(2):305-14(1984)]: 이들은 LPCAT가 아실-CoA 풀과 포스파티딜콜린 (PC) 풀 사이의 교환에 영향을 미칠 것이라는 가설을 세웠음; (2) 도메르구, 에프.(Domergue, F.) 등 [J. Bio. Chem 278:35115 (2003)]: 이들은 효모에서 PC의 sn-2 위치에 GLA가 축적되고 ARA가 효율적으로 합성될 수 없는 것은 아실-CoA 풀에서 일어나는 PUFA 생합성에 신장 단계가 포함되어 있으나 Δ5 및 Δ6 탈포화 단계는 PC의 sn-2 위치에서 우세하게 일어나기 때문이라고 제안하였음; (3) 아바디, 에이.(Abbadi, A.) 등 [The Plant Cell, 16:2734-2748 (2004)]: 이들은 트랜스제닉 오일종자 식물에서의 PUFA 축적 제약에 관한 분석을 기초로 하여, LPCAT가 Δ6 데새투라제/Δ6 엘론가제 경로의 성공적인 재구성에 중대한 역할을 한다고 제안하였음; (4) 제WO 2004/076617 A2호 (렌쯔, 에이.(Renz, A.) 등): 이들은 사카로마이세스 세레비지애에서 유전자 도입된 Δ6 데새투라제/Δ6 엘론가제 경로의 신장 효율을 실질적으로 개선시키는, 캐노르합디티스 엘레강스로부터의 LPCAT를 코딩하는 유전자 (T06E8.1)를 제공하였음. 데새투라제는 지질-커플링된 지방산 (sn-2 아실 PC)에서 이중 결합의 도입을 촉매하지만 엘론가제는 오로지 CoA 에스테르화된 지방산 (아실-CoA)의 신장만을 촉매하기 때문에, 본 발명자들은 LPCAT가 인지질과 아실-CoA 풀 사이에서 새로 합성된 지방산이 효율적이고 지속적으로 교환될 수 있게 한다고 결론을 내렸다.

DHA 합성을 위한 이종 아실트랜스퍼라제 유전자의 선별

야로위아 리폴리티카에서 천연적으로 생산된 PUFA는 18:2 지방산 (및 덜 통상적으로는, 18:3 지방산)으로 한정되기 때문에, GPAT, LPAAT (즉, LPAAT1 또는 LPAAT2), DGAT1, DGAT2, PDAT 및 LPCAT를 코딩하는 숙주 유기체의 천연 유전자는 18:3 이상의 길이를 갖는 지방산을 포함하는 TAG (예컨대, DHA)를 효율적으로 합성하는데 어려움이 있을 수 있다. 따라서, 일부 경우에서는 이종 (또는 "외래") 아실트랜스퍼라제가 천연 효소보다 바람직할 수 있다.

수많은 아실트랜스퍼라제 유전자가 여러 유기체에서 확인되었고, 간행물 및 특허 문헌에 기술되어 있다. 예를 들어, 하기 하는 진뱅크 관리 번호는 지질 생합성에 유용한 공개적으로 입수가능한 아실트랜스퍼라제 유전자의 예이다: CQ891256, AY441057, AY360170, AY318749, AY093169, AJ422054, AJ311354, AF251795, Y00771, M77003 (GPAT); Q93841, Q22267, Q99943, 015120, Q9NRZ7, Q9NRZ5, Q9NUQ2, O35083, Q9D1E8, Q924S1, Q59188, Q42670, P26647, P44848, Q9ZJN8, O25903 Q42868, Q42870, P26974, P33333, Q9XFW4, CQ891252, CQ891250, CQ891260, CQ891258, CQ891248, CQ891245, CQ891241, CQ891238, CQ891254, CQ891235 (LPAAT); AY445635, BC003717, NM_010046, NM_053437, NM_174693, AY116586, AY327327, AY327326, AF298815 및 AF164434 (DGAT1); 및 NC_001147 [유전자좌 NP_014888], NM_012079, NMJ 27503, AF051849, AJ238008, NM_026384, NM_010046, AB057816, AY093657, AB062762, AF221132, AF391089, AF391090, AF129003, AF251794 및 AF164434 (DGAT2); P40345, O94680, NP_596330, NP_190069 및 AB006704 [gi:2351069] (PDAT). 유사하게, 특허 문헌은 TAG 생성에 관여하는 유전자의 많은 추가 DNA 서열 (및/또는 상기 여러 유전자 및 이들의 단리 방법에 관한 상세 사항)을 제공한다 [예를 들어, 미국 특허 제5,210,189호, 제WO 2003/025165호 (GPAT), 유럽 특허 제1144649 A2호, 유럽 특허 제1131438호, 미국 특허 제5,968,791호, 미국 특허 제6,093,568호, 제WO 2000/049156호 및 제WO 2004/087902호 (LPAAT), 미국 특허 제6,100,077호, 미국 특허 제6,552,250호, 미국 특허 제6,344,548호, 미국 특허 출원 제2004/0088759 A1호 및 미국 특허 출원 제20040078836 A1호 (DGAT1), 미국 특허 출원 제2003/124126호, 제WO 2001/034814호, 미국 특허 출원 제2003/115632호, 미국 특허 출원 제2003/0028923호 및 미국 특허 출원 제2004/0107459호 (DGAT2), 제WO 2000/060095호 (PDAT) 및 제WO 2004/076617 A2호 (LPCAT)].

상기 예는 제한하고자 하는 것이 아니며, 다른 공급원으로부터 유래된 DGAT1, DGAT2, PDAT, GPAT, LPCAT 및 LPAAT를 코딩하는 수많은 다른 유전자가 야로위아 리폴리티카로의 도입에 적합하다. 예를 들어, 본 발명자들은 모르티에렐라 알피나 (서열 124 및 125), 뉴로스포라 크라싸 (서열 126), 지베렐라 제애 PH-1 (서열 127), 마그나포르테 그리세아 (서열 128) 및 아스페르길루스 니둘란스 (서열 129)로부터 신규 DGAT1을 확인하였고, 모르티에렐라 알피나로부터 신규 DGAT2 (서열 136 및 137), GPAT (서열 138 및 139), LPAAT1 (서열 108 및 109) 및 LPAAT2 (서열 110 및 111)를 확인하였다.

DHA 합성에 바람직한 아실트랜스퍼라제 유전자

야로위아 리폴리티카에서의 발현에 적합할 수 있는 아실트랜스퍼라제는 광범위하게 선택될 수 있지만, 본 발명의 바람직한 실시양태에서는 유의한 양의 더욱 장쇄의 ω-6 (예컨대, ARA) 및/또는 ω-3 (예컨대, EPA, DHA) PUFA를 생산하는 유기체로부터의 DGAT1, DGAT2, PDAT, GPAT, LPAAT 및 LPCAT가 선택된다. 따라서, 하기 효소 (또는 그의 유도체)가 특히 바람직하다:

[표 8]

DHA를 고도로 생성하는 야로위아 리폴리티카 균주에서의 발현에 바람직한 이종 아실트랜스퍼라제			
ORF	유기체	참조문헌	서열번호
DGAT1	모르티에렐라 알피나	동시 계류 중인 미국 특허 출원 제11/024544호	서열 124, 서열 125
DGAT2	모르티에렐라 알피나	동시 계류 중인 미국 특허 출원 제11/024545호	서열 136, 서열 137
GPAT	모르티에렐라 알피나	-	서열 138, 서열 139
LPAAT1	모르티에렐라 알피나	-	서열 108, 서열 109
LPAAT2	모르티에렐라 알피나	동시 계류 중인 미국 특허 출원 제60/689031호	서열 110, 서열 111
LPCAT	캐노르합디티스 엘레강스	클론 T06E8.1, 제WO 2004/076617 A2호	서열 121

본원에 개시한 본 발명을 한정하려는 것은 아니었지만, 모르티에렐라 알피나는 이종 아실트랜스퍼라제의 바람직한 공급원으로서 선택되었는데, 이는 상기 천연 유기체가 ARA를 전체 지방산 (TFA)의 50% 초과의 농도로 생산할 수 있기 때문이다. 유사한 방식으로, 캐노르합디티스 엘레강스는 TFA의 최대 20 내지 30%를 EPA로 생산할 수 있다.

물론, 본 발명의 별법의 실시양태에서, 서열 108 내지 111, 121, 124, 125 및 136 내지 139에 의해 코딩되는 아실트랜스퍼라제와 실질적으로 동일한 다른 DNA 역시 야로위아 리폴리티카에서의 이종 발현에 사용되어 TAG 분획 중 EPA의 생성 및 축적을 용이하게 할 수 있다. 더욱 바람직한 실시양태에서, 서열 108 내지 111, 121, 124, 125 및 136 내지 139에 기재된 서열과 실질적으로 동일한 아실트랜스퍼라제를 코딩하는 코돈-최적화된 유전자가 사용된다.

외래 유전자의 발현을 위한 일반적인 발현 시스템, 카세트, 벡터 및 형질전환

외래 단백질의 고수준 발현을 지시하는 조절 서열, 예컨대 DHA의 고수준 생성을 유도하는 조절 서열을 함유하는 발현 벡터 및 미생물 발현 시스템은 당업자에게 공지되어 있다. 이들 중 임의의 것을 사용하여 바람직한 데새투라제, 엘론가제 및 아실트랜스퍼라제를 코딩하는 키메라 유전자를 구축할 수 있다. 이어서, 코딩된 효소의 고수준 발현을 제공하는 표준 형질전환 방법을 이용하여 이들 키메라 유전자를 야로위아 리폴리티카에 도입할 수 있다.

숙주 세포의 형질전환에 유용한 벡터 또는 DNA 카세트는 당업계에 공지되어 있다. 상기 구축물에 존재하는 서열들을 구체적으로 선택하는 것은, 원하는 발현 생성물, 숙주 세포의 특성, 및 형질전환된 세포와 형질전환되지 않은 세포를 분리하기 위해 제안된 수단에 따라 달라진다. 그러나 전형적으로, 상기 벡터 또는 카세트는 관련 유전자(들)의 전사 및 번역을 지시하는 서열, 선별가능한 마커 및 자율 복제 또는 염색체 통합을 허용하는 서열을 함유한다. 적합한 벡터는 전사 개시를 조절하는 유전자 (예컨대, 프로모터)인 영역 5' 및 전사 종결을 조절하는 DNA 단편 (즉, 종결자)인 영역 3'를 포함한다. 두가지 조절 영역 모두가 형질전환된 숙주 세포로부터의 유전자로부터 유래된 경우가 가장 바람직하지만, 이러한 조절 영역이 생성 숙주로서 선택된 특정 종에 천연인 유전자로부터 유래되어야 하는 것은 아님을 이해해야 한다.

2종 이상의 유전자가 별도의 복제 벡터로부터 발현되는 경우, 각 벡터는 상이한 선별 수단을 갖는 것이 바람직하며, 안정적인 발현을 유지하고 구축물 중 요소들의 재배열을 방지하기 위해서 다른 구축물과의 상동성이 없어야 한다. 조절 영역의 현명한 선택, 선별 수단 및 도입된 구축물의 증식 방법을 실험적으로 결정하여, 도입된 모든 유전자가 원하는 생성물의 합성을 위해 제공되기에 필요한 수준으로 발현되도록 할 수 있다.

관심 유전자(들)을 포함하는 구축물은 임의의 표준 기술을 이용하여 숙주 세포에 도입할 수 있다. 이들 기술로는 형질전환법 (예컨대, 리튬 아세테이트 형질전환 [Methods in Enzymology, 194:186-187 (1991)]), 원형질체 융합법, 볼리스트 충격(bolistic impact), 전기천공법, 미세주입법, 또는 관심 유전자(들)을 숙주 세포에 도입하는 임의의 다른 방법 등이 있다. 야로위아 리폴리티카에 적용할 수 있는 더욱 구체적인 교시 문헌으로는, 미국 특허 제4,880,741호 및 동 제5,071,764호 및 문헌 [Chen, D. C. et al. (Appl Microbiol Biotechnol. 48(2):232-235 (1997))] 등이 있다.

편의상, 본원에서는 임의의 방법으로 조작되어 DNA 서열 (예컨대, 발현 카세트)이 도입된 숙주 세포를 "형질전환" 또는 "재조합"된 것이라고 지칭할 것이다. 형질전환된 숙주는 발현 구축물의 카피를 1개 이상 보유할 것이며, 유전자가 게놈에 통합되었는지, 증폭되었는지 또는 다중 카피수를 갖는 염색체외 요소에 존재하는지 여부에 따라 2개 또는 그 이상일 수 있다. 형질전환된 숙주 세포는 제WO 2004/101757호 및 제WO 2005/003310호에 기재된 것과 같은 다양한 선별 기술로 확인할 수 있다.

본원에서 사용하기에 바람직한 선별 방법은 카나마이신, 하이그로마이신 및 아미노 글리코시드 G418에 대한 내성, 및 또한 우라실, 루이신 또는 히스티딘이 없는 배지에서의 성장 능력이다. 별법의 실시양태에서, 5-플루오로오로트산 (5-플루오로우라실-6-카르복실산 1수화물, "5-FOA")이 효모 Ura^- 돌연변이체의 선별에 사용된다. 상기 화합물은 오로티딘 5'-모노포스페이트 데카르복실라제 (OMP 데카르복실라제)를 코딩하는 기능적 URA3 유전자를 보유하는 효모 세포에 독성이고, 따라서, 이러한 독성을 기초로, 5-FOA는 Ura^- 돌연변이체 효모 균주의 선별 및 확인에 특히 유용하다 [Bartel, P. L. and Fields, S., Yeast 2-Hybrid System, Oxford University: New York, v. 7, pp 109-147, 1997].

본원에서 이용되는 별법의 바람직한 선별 방법은 술포닐우레아 내성을 기초로 하는, 야로위아 리폴리티카에 대한 우성의 비-항생제 마커에 의존한다. 상기 기술은 또한 일반적으로 반수체, 이배체, 이수체(異數體) 또는 이형접합체일 수 있는 다른 산업상의 효모 균주에도 적용될 수 있다. 산업상의 효모 균주를 위한 유전적 형질전환 시스템의 개발에 있어서 (1) 천연적으로 영양요구성인 균주가 거의 없고, 자발적이거나 유도된 영양요구성 돌연변이체의 단리가 균주의 배수성으로 인해 곤란하며, (2) 항생제 내성 유전자를 보유하는 유전자 변형된 유기체의 방출에 대한 제약으로 인해서 항생제 내성 마커의 사용이 균주의 상업적 이용을 제한할 수 있다는 2가지 주요 한계를 극복할 것이라 예측된다. 푸이그(Puig) 등 [J. Agric. Food Chem. 46:1689-1693 (1998)]은 표적 균주를 유전자 조작하여 유리딘에 대한 영양요구성으로 만들고 이후에는 URA3 마커를 사용하여 관심 형질을 도입하는 것을 기초로 하여 이러한 한계를 극복하는 방법을 개발하였지만, 상기 전략은 불필요한 작업이 요구되어 지나치게 소모적이다.

본원에 개시된, 야로위아 리폴리티카를 형질전환시키기 위한 신규한 술포닐우레아 내성 선별 마커는 외래 유전자가 아닌 돌연변이체 천연 유전자에 의존한다. 따라서, 따라서, 영양요구성일 필요도 없으며 영양요구성을 초래하지도 않고 야생형 균주의 형질전환을 초래하지도 않는다. 보다 구체적으로, 마커 유전자 (서열 243)는 술포닐 우레아 제초제 내성을 부여하는 단일 아미노산 변화 (W497L)를 갖는 천연 아세토히드록시산 신타제 (AHAS 또는 아세토락테이트 신타제, E.C. 4.1.3.18)이다. AHAS는 분지쇄 아미노산의 생합성 경로에서 공통적인 첫번째 효소이며, 술포닐우레아 및 이미다졸리논 제초제의 표적이다. W497L 돌연변이는 사카로마이세스 세레비지애에서의 작업에 근거하여 공지되어 있다 ([Falco, S. C, et al., Dev. Ind. Microbiol. 30:187-194 (1989)], [Duggleby, R.G., et. al. Eur. J. Biochem. 270:2895 (2003)]). 다음과 같은 결과를 나타낸 초기 시험을 통해, 야로위아 세포가 제초제에 천연적으로 내성이 있지는 않음이 결정되었다: 1) 제초제 흡수가 불량하거나 흡수하지 않음, 2) AHAS의 천연적인 제초제-내성 형태가 존재함, 및/또는 3) 제초제-불활성화 메카니즘을 이용함. 이로써, 형질전환체의 선별 수단으로서 돌연변이체 AHAS 유전자 (서열 243)의 합성 및 사용이 가능해졌다.

선별 마커를 재사용하는 추가의 방법은 부위-특이적 레콤비나제 시스템에 의존한다. 간략하게 설명하면, 상기 부위-특이적 재조합 시스템은 (1) 특징적인 DNA 서열 [예컨대 LoxP]을 갖는 재조합 부위, 및 (2) 2개 이상의 재조합 부위가 동일 DNA 분자상에서 소정의 간격을 두고 동일 방향으로 배향되어 있는 경우에 DNA 서열에 특이적으로 결합하여 DNA 서열들 사이의 재조합 (즉, 제거)을 촉매하는 레콤비나제 효소 [예컨대 Cre]의 2가지 요소로 구성된다. 바람직한 선별 마커 쌍이 여러 회의 순차적인 형질전환에 사용되도록 이것을 "재순환"시키는 것이 가능하기 때문에, 이 방법은 선별 수단으로서 유용성이 있다.

구체적으로, 숙주 게놈에 삽입되는 것이 바람직한 표적 유전자 (예컨대, 데새투라제, 엘론가제, 아실트랜스퍼라제), 및 또한 재조합 부위들로 측접(flnaking)되어 있는 제1 선별 마커 (예컨대, Ura3, 하이그로마이신 포스포트랜스퍼라제 [HPT])를 포함하는 통합 구축물을 제조한다. 형질전환 및 형질전환체의 선별 후, 제2 선별 마커 (예컨대, 술포닐우레아 내성 [AHAS]) 및 게놈에 도입된 부위-특이적 재조합 부위를 인식하는데 적합한 레콤비나제를 보유하는 복제 플라스미드를 도

입함으로써 제1 선별 마커를 염색체로부터 제거해 낸다. 제2 마커를 보유하는 형질전환체의 선별 및 숙주 게놈으로부터의 제1 선별 마커 제거를 확인한 후, 상기 복제 플라스미드는 숙주에 의해 구별없이 관리(cure)된다. 이로써 제1 선별 마커와 제2 선별 마커 어느 것도 보유하지 않는 형질전환체가 생성되며, 따라서 상기 관리된 균주는 또다른 회의 형질전환에 이용될 수 있다. 당업자는 상기 방법이 본 발명에서 사용되는 특정 선별 마커 또는 특정 부위-특이적 재조합 시스템에 한정되는 것이 아님을 인지할 것이다.

야로위아 리폴리티카 중 외래 유전자의 과발현

당업자에게 공지된 바와 같이, 유전자 (예컨대, 데새투라제)를 클로닝 벡터에 삽입하는 것만으로는 이것이 필요한 수준으로 성공적으로 발현될 것을 보장하지 못한다. 전사, 번역, 단백질 안정성, 산소 제한 및 숙주 세포로부터의 분비 측면을 조절하는 수많은 상이한 유전자 요소들을 조작하는 것이 바람직할 수 있다. 보다 구체적으로, 유전자 발현은 다음을 변경함으로써 조절될 수 있다: 관련 전사 프로모터 및 종결자 서열의 성질, 클로닝된 유전자의 카피 수, 상기 유전자가 플라스미드에 보유되는지 또는 숙주 세포의 게놈에 통합되는지의 여부, 합성된 외래 단백질의 최종적인 세포내 위치, 숙주 유기체 중에서의 번역 효율, 숙주 세포 중 클로닝된 유전자 단백질의 내재적 안정성, 및 클로닝된 유전자 내에서 숙주 세포의 바람직한 코돈 사용 빈도에 도달할만큼의 코돈 사용. 이러한 과발현 방법 중 여러가지가 이하에서 논의될 것이고, 본 발명에서 예를 들어 야로위아 리폴리티카 중 데새투라제, 엘론가제 및 아실트랜스퍼라제의 과발현을 위한 수단으로서 유용하다.

원하는 유전자(들)의 발현은 mRNA 또는 코딩된 단백질로부터 불안정화 서열을 제거/결실시키거나, 또는 상기 mRNA에 안정화 서열을 부가함으로써 더 강력한 프로모터 (조절형 또는 구성적 프로모터)의 사용을 통해 전사 수준에서 증가되어 발현이 증가될 수 있다 (미국 특허 제4,910,141호).

원하는 숙주 세포에서 데새투라제, 엘론가제 및 아실트랜스퍼라제 유전자의 발현을 구동시키는데 유용한 개시 조절 영역 또는 프로모터는 매우 다양하며 당업자에게 알려져 있다. 사실상, 야로위아 리폴리티카에서 이들 유전자의 발현을 지시할 수 있는 임의의 프로모터가 본 발명에 적합하다. 숙주 세포에서의 발현은 일시적이거나 안정적인 방식으로 수행될 수 있다. 일시적 발현은 관심 유전자에 작동가능하게 연결된 조절가능한 프로모터의 활성을 유도하여 달성될 수 있고, 별법으로, 안정적인 발현은 관심 유전자에 작동가능하게 연결된 구성적 프로모터를 이용하여 달성될 수 있다. 예를 들어, 숙주 세포가 효모인 경우에는 특히 상기 숙주 종으로부터 유래된, 효모 세포에서 기능적인 전사 영역 및 번역 영역이 제공된다. 전사 개시 조절 영역은 예를 들어 1) 해당 경로 중의 유전자, 예컨대 알콜 데히드로게나제, 글리세르알데히드-3-포스페이트-데히드로게나제, 포스포글리세레이트 뮤타제, 프럭토스-비스포스페이트 알돌라제, 포스포글루코스-이소머라제, 포스포글리세레이트 키나제, 글리세롤-3-포스페이트 O-아실트랜스퍼라제 등, 또는 2) 조절가능한 유전자, 예컨대 산 포스파타제, 락타제, 메탈로티오네인, 글루코아밀라제, 번역 신장 인자 EF1-α (TEF) 단백질 (미국 특허 제6,265,185), 리보솜 단백질 S7 (미국 특허 제6,265,185), 암모늄 수송자 단백질, 방출 단백질 등으로부터 얻을 수 있다. 구성적 전사를 원하는지 또는 유도된 전사를 원하는지의 여부, 관심 ORF의 발현에 있어서의 프로모터 효율, 구축의 용이성 등에 따라, 수많은 조절 서열 중 임의의 것을 사용할 수 있다. 상기한 예는 본원에 개시한 본 발명을 제한하려는 것이 아니다.

당업자가 알고 있는 바와 같이, 다양한 프로모터의 활성 비교에는 여러가지 방법이 이용가능하다. 이러한 유형의 비교는, 각 프로모터를 이들 프로모터에 ω-6 및 ω-3 지방산의 생성에 유용한 키메라 유전자의 구축이 필요한 추가의 용도에 사용하기 위한 프로모터 강도 결정을 용이하게 하는데 유용하다. 따라서, 리포터 유전자 (즉, β-글루쿠로니다제 (GUS)를 코딩하는 이. 콜라이(*E. coli*) 유전자)의 발현을 기초로 하여 프로모터 활성을 간접적으로 정량하는 것이 유용할 수 있다. 별법의 실시양태에서, 때로는 더욱 정량적인 수단을 이용하여 프로모터 활성을 정량화하는 것이 유용할 수 있다. 적합한 방법 중 하나는 실시간 PCR (실시간 PCR 이용에 대하여 일반적으로 고찰하기 위해서는, 문헌 [Ginzinger, D. J., Experimental Hematology, 30:503-512 (2002)] 참조)을 이용하는 것이다. 실시간 PCR은 형광성 리포터의 검출 및 정량화를 기초로 한다. 이 신호는 반응물 중 PCR 생성물의 양에 직접 비례하여 증가한다. 각 주기에서 형광 방출의 양을 기록함으로써, PCR 생성물의 양에 있어서의 최초의 유의한 증가가 표적 주형의 초기 양과 상관관계가 있는 지수 성장기 동안의 PCR 반응을 모니터링할 수 있다. 앰플리콘을 정량적으로 검출하기 위한 2가지 일반적 방법이 있다: (1) 형광성 프로브의 사용, 또는 (2) DNA-결합제 (예컨대, SYBR-그린 I, 에티듐 브로마이드)의 사용. 상대적인 유전자 발현 비교를 위해, 내인성 조절을 내부 기준물 (예컨대, 염색체에 의해 코딩되는 16S rRNA 유전자)로서 사용하여, 각각의 실시간 PCR 반응에 첨가되는 DNA의 총량에 있어서의 차이에 대해 표준화할 수 있도록 할 필요가 있다. 실시간 PCR에 대한 구체적인 방법은 당업계에 문서화되어 있다. 예를 들어 문헌 [Real Time PCR Special Issue (Methods, 25(4):383-481 (2001)]을 참조한다).

실시간 PCR 반응 후에, 기록된 형광 강도를 이용하여 하기 방법을 통해 주형의 양을 정량한다: 1) 절대적 표준 방법 (알려진 양의 기준물, 예컨대 시험관내 번역된 RNA (cRNA)를 사용함), 2) 상대적 표준 방법 (알려진 양의 표적 핵산을, 디자인된 검정법의 각 운행에 포함시킴), 또는 3) 유전자 발현의 상대적 정량화를 위한 비교 C_T 방법 ($\Delta\Delta C_T$) (표적 서열의 상대

적 양을, 선택된 임의의 기준물 값과 비교하고, 그 결과를 기준 값에 대한 상대치로서 나타냄). 비교 C_T 방법은, 우선 표적의 C_T 값과 표준물의 C_T 값 사이의 차이 (ΔC_T)를 결정하는 것이 필요하다: $\Delta C_T = C_T$ (표적) - C_T (표준물). 정량할 각 샘플에 대해 이 값을 계산하고, 한 샘플을 각각의 비교가 수행될 기준물로서 선택해야 한다. 비교 $\Delta\Delta C_T$ 계산은 각 샘플의 ΔC_T와 기준선의 ΔC_T 사이의 차이를 알아낸 후에, 이들 값을 식 $2^{-\Delta\Delta CT}$에 따른 절대값으로 변환시키는 것을 포함한다.

야로위아 리폴리티카에서의 발현에 적합할 수 있는 프로모터의 선택폭은 매우 다양하지만, 본 발명의 바람직한 실시양태에서는 프로모터가 하기 표 9에 기재한 프로모터 (또는 그의 유도체)로부터 선택된다:

[표 9]

야로위아 리폴리티카에서의 과발현에 바람직한 천연 프로모터					
프로모터 명칭	위치*	천연 유전자	활성 "등급"	참조문헌	서열번호
TEF	---	번역 신장 인자 EF1-α	1	미국 특허 제 6,265,185호 (뮬러 (Muller) 등), 진뱅크 관리 번호 AF054508	서열 218
GPD	-968 bp 내지 +3 bp	글리세르알데히드-3-포스페이트-데히드로게나제	2	제WO 2005/003310호	서열 210
GPM	-875 bp 내지 +3 bp	포스포-글리세레이트 뮤타제	1	제WO 2005/003310호	서열 212
FBA	-1001 bp 내지 -1 bp	프럭토스-비스포스페이트 알돌라제	4	제WO 2005/049805호	서열 213
FBAIN	-804 bp 내지 +169 bp (102 bp 인트론 [+64 내지 +165] 포함)	프럭토스-비스포스페이트 알돌라제	7	제WO 2005/049805호	서열 214
FBAINm	-804 bp 내지 +169 bp, 변형됨***	프럭토스-비스포스페이트 알돌라제	5	제WO 2005/049805호	서열 215
GPDIN	-973 bp 내지 +201 bp (146 bp 인트론 [+49 내지 +194] 포함)	글리세르알데히드-3-포스페이트-데히드로게나제	3	동시 계류 중인 미국 특허 출원 제11/183664호	서열 211
GPAT	-1130 bp 내지 +3 bp	글리세롤-3-포스페이트-O-아실트랜스퍼라제	5	동시 계류 중인 미국 특허 출원 제11/225354호	서열 216
YAT1	-778 bp 내지 -1 bp	암모늄 수송자 효소	6	동시 계류 중인 미국 특허 출원 제11/185301호	서열 217
EXP1	-1000 bp 내지 -1 bp	방출 단백질	6	----	서열 221

* 천연 유전자에 대한 위치임.
여기서, 'ATG' 번역 개시 코돈의 'A' 위치를 +1로 함.

*** FBAINm 프로모터는 FBAIN 프로모터의 변형된 형태로, FBAINm이 FBAIN 프로모터의 ATG 번역 개시 코돈과 인트론 사이에 52 bp 결실부를 가지며 (따라서, N-말단의 22개 아미노산만을 포함함) 인트론 뒤에 새로운 번역 컨센서스 모티프를 가짐. 추가로, 발현될 유전자의 코딩 영역과 융합될 경우에 FBAIN 프로모터는 융합 단백질을 생성하지만, FBAINm 프로모터는 그러한 융합 단백질을 생성하지 않음.

GPM의 활성은 TEF과 거의 동일하지만, GPD, FBA, FBAIN, FBAINm, GPDIN, GPAT, YAT1 및 EXP1의 활성은 모두가 TEF보다 높다 (활성은, "활성 등급"이라는 제목의 컬럼에서 상대적 방식으로 정량화함. 여기서, '1'은 가장 낮은 활성을 갖

는 프로모터에 상응하고 '7'은 가장 높은 활성을 갖는 프로모터에 상응함). 이러한 정량화는 각 프로모터가 리포터로서 β-글루쿠로니다제 (GUS)를 코딩하는 이. 콜라이 유전자 [Jefferson, R.A. Nature. 14;342:837-838 (1989)] 및 야로위아 Xpr 유전자의 3' 영역의 약 100 bp를 보유하는 키메라 유전자의 생성에 사용된 비교 연구를 기초로 한다. 각각의 발현된 구축물에서의 GUS 활성은, 조직화학적 및/또는 형광측정 검정 [Jefferson, R. A. Plant Mol. Biol. Reporter 5:387-405 (1987)] 및/또는 실시간 PCR의 이용으로 측정한 것이다.

YAT1 프로모터는, 본 발명자들에 의해 야로위아 내에서 유질 조건 (즉, 질소 제한)하에 유도가능한 것으로 확인된 최초의 프로모터로서 특징규명되었다는 점에서 독특하다. 구체적으로, YAT1 프로모터는 질소 (예컨대, 최대 약 0.5%의 암모늄 술페이트)를 함유하는 배지 중에서 활성이지만, 상기 프로모터의 활성은 숙주 세포를 질소-제한 조건 (예컨대, 암모늄을 매우 낮은 수준으로 함유하거나 암모늄이 없는 배지 중)하에서 성장시킬 경우에 증가한다. 따라서, 바람직한 배지는 암모늄 술페이트 또는 다른 적합한 암모늄 염을 약 0.1% 미만으로 함유하는 것이다. 더욱 바람직한 실시양태에서, YAT1 프로모터는, 숙주 세포를 탄소:질소의 비율 (즉, C:N)이 높은 배지, 예컨대 약 8 내지 12% 글루코스 및 약 0.1% 이하의 암모늄 술페이트를 함유하는 고농도 글루코스 배지 (HGM)에서 성장시킬 경우에 유도된다. 이들 조건은 또한 유질인 효모 (예컨대, 야로위아 리폴리티카) 중에서 유질성을 유도하기에도 충분하다. 세포 추출물의 GUS 활성을 기초로, 세포를 최소 배지에서 HGM으로 변경시켜 24시간 동안 성장시킬 때 YAT1 프로모터의 활성은 약 37배 증가하였고, HGM 중에서 120시간이 지난 후에는 이 활성이 다소 저하되었으나 질소를 포함하는 최소 배지 중에서의 활성에 비해 여전히 25배 더 높았다 (실시예 1).

물론, 본 발명의 별법의 실시양태에서, 상기 표 9에 기재한 임의의 프로모터 영역으로부터 유래된 다른 프로모터도 야로위아 리폴리티카에서 TAG 분획물 중 DHA의 생성과 축적을 용이하게 하기 위한 이종 발현에 사용될 수 있다. 특히, 상기한 임의의 프로모터의 길이 변형은 동일한 활성을 갖는 돌연변이체 프로모터를 초래할 수 있는데, 이는 이들 조절 서열의 정확한 경계가 아직 완전하게 규명되지 못했기 때문이다. 별법의 실시양태에서, FBAIN 및 GPDIN 프로모터의 인트론 내에 위치한 인핸서를 사용하여 천연 야로위아 프로모터에 비해 활성이 증가된 키메라 프로모터를 생성할 수 있다 (예컨대, 키메라 GPM::FBAIN 및 GPM::GPDIN 프로모터 (서열 219 및 서열 220)는 야로위아 Xpr 유전자 3' 영역의 약 100 bp와 함께 GUS 리포터 유전자의 발현을 구동시킬 경우, GPM 프로모터 단독인 경우에 비해 활성이 증가되었음).

종결 영역은 개시 영역이 얻어진 유전자 또는 상이한 유전자의 3' 영역으로부터 유래될 수 있다. 다수의 종결 영역이 공지되어 있으며 다양한 숙주에서 (이들이 유래된 속과 종이 동일한 경우와 상이한 경우 둘다에 이용될 때) 만족스런 기능을 발휘한다. 종결 영역은 일반적으로 임의의 특정 성질이 아니라 편의성을 위주로 선택된다. 종결 영역은 효모 유전자, 특히 사카로마이세스, 쉬조사카로마이세스(*Schizosaccharomyces*), 칸디다, 야로위아 또는 클루이베로마이세스로부터 유래되는 것이 바람직하다. γ-인터페론 및 α-2 인터페론을 코딩하는 포유동물 유전자의 3'-영역도 효모에서 기능하는 것으로 공지되어 있다. 또한, 종결 조절 영역도 바람직한 숙주에 천연인 각종 유전자로부터 유래될 수도 있다. 임의로, 종결 부위는 필요치 않을 수 있지만, 이것이 포함되는 것이 가장 바람직하다. 제한하려는 것은 아니지만, 본원의 개시 내용에 유용한 종결 영역으로는 야로위아 리폴리티카 세포외 프로테아제 (XPR (진뱅크 관리 번호 M17741)) 3' 영역의 약 100 bp, 아실-coA 옥시다제 (Aco3 (진뱅크 관리 번호 AJ001301 및 CAA04661), Pox3 (진뱅크 관리 번호 XP_503244)) 종결자, Pex20 (진뱅크 관리 번호 AF054613) 종결자, Pex16 (진뱅크 관리 번호 U75433) 종결자, Lip1 (진뱅크 관리 번호 Z50020) 종결자, Lip2 (진뱅크 관리 번호 AJ012632) 종결자, 및 3-옥소아실-coA 티올라제 (OCT, 진뱅크 관리 번호 X69988) 종결자 등이 있다.

상기한 데새투라제, 엘론가제 및/또는 아실트랜스퍼라제 유전자의 추가의 카피 (즉, 1개 초과의 카피)가 야로위아 리폴리티카에 도입되어 DHA 생산 및 축적을 증가시킬 수 있다. 구체적으로, 추가의 카피의 유전자가 단일 발현 구축물 내에 클로닝될 수 있고/있거나; 플라스미드 카피수를 증가시키거나 클로닝된 유전자를 게놈에 다중 통합시킴으로써 추가의 카피의 클로닝된 유전자(들)이 숙주 세포에 도입될 수 있다 (하기 참조). 예를 들어 한 실시양태에서, 야로위아 리폴리티카의 균주 (즉, 균주 Y3000)는 3개 카피의 Δ12 데새투라제, 2개 카피의 Δ6 데새투라제, 4개 카피의 $C_{18/20}$ 엘론가제, 5개 카피의 Δ5 데새투라제, 3개 카피의 Δ17 데새투라제, 3개 카피의 $C_{16/18}$ 엘론가제, 1개 카피의 $C_{20/22}$ 엘론가제 및 1개 카피의 Δ4 데새투라제를 포함하는 키메라 유전자를 야로위아 게놈에 도입 및 통합시켜서 DHA를 5% 초과로 생산하도록 유전자 조작되었다.

일반적으로, 일단 유질 효모 중에서의 발현에 적합한 DNA (예컨대, 프로모터, ORF 및 종결자를 포함하는 키메라 유전자)가 수득되면, 이것을 숙주 세포 중에서 자가 복제가 가능한 플라스미드 벡터에 위치시키거나, 또는 숙주 세포의 게놈에 직접 통합시킨다. 발현 카세트의 통합은 숙주 게놈 내에서 무작위로 일어날 수도 있고, 또는 숙주 유전자좌와의 재조합을 표

적화하기에 충분할 만큼 숙주 게놈에 상동성인 영역을 함유하는 구축물을 사용하여 표적화될 수도 있다. 본 발명에서 제한되지는 않지만, 전사 조절 영역 및 번역 조절 영역의 전부 또는 일부는 구축물이 표적화되는 내인성 유전자좌에 의해 제공될 수 있다.

본 발명에서, 야로위아 리폴리티카에서 유전자를 발현시키는 바람직한 방법은 선형 DNA를 숙주의 게놈으로 통합시키는 것이고, 게놈 내 여러 위치에 통합시키는 것이 유전자의 고수준 발현을 원하는 경우에 특히 유용할 수 있다. 이러한 목적을 위해, 게놈 내에 다중 카피로 존재하는 서열을 동정하는 것이 바람직하다.

슈미드-베르거(Schmid-Berger) 등 [J. Bact. 176(9):2477-2482 (1994)]은 야로위아 리폴리티카에서 최초의 레트로트란스포손-유사 요소 Ylt1을 발견하였다. 이 레트로트란스포손은 제타 영역이라고 불리는 긴 말단 반복부 (LTR; 각각 대략 700 bp 길이)의 존재를 특징으로 한다. Ylt1 및 솔로 제타 요소들은 게놈 내에서 각각 35개 이상의 카피/게놈 및 50 내지 60개 카피/게놈으로 분산되어 존재하였다. 이들 두 요소 모두가 상동성 재조합 부위로서 기능하는 것으로 결정되었다. 추가로, 쥬렛젝(Juretzek) 등의 연구 [Yeast 18:97-113 (2001)]는 플라스미드를 효모 게놈의 반복적 영역에 표적화함으로써 (양 말단에 LTR 제타 영역을 갖는 선형 DNA를 사용함), 적은 카피 수의 플라스미드 형질전환체를 사용하여 달성되는 발현에 비해 유전자 발현이 크게 증가할 수 있음을 입증하였다. 따라서, 제타-지시된 통합은 플라스미드 DNA를 야로위아 리폴리티카에 다중 통합시켜서 유전자 발현이 고수준이 되게 하는 수단으로서 이상적일 수 있다. 그러나 불행히도, 야로위아 리폴리티카의 모든 균주가 제타 영역을 보유하는 것은 아니다 (예컨대, ATCC #20362로 동정된 균주). 균주에 이러한 영역이 없는 경우, 발현 카세트를 포함하는 플라스미드 DNA를 별법의 유전자좌에 통합시켜서 발현 카세트를 원하는 카피 수로 만드는 것도 가능하다. 예를 들어, 바람직한 별법의 유전자좌로는 Ura3 유전자좌 (진뱅크 관리 번호 AJ306421), Leu2 유전자의 유전자좌 (진뱅크 관리 번호 AF260230), Lys5 유전자 (진뱅크 관리 번호 M34929), Aco2 유전자의 유전자좌 (진뱅크 관리 번호 AJ001300), Pox3 유전자의 유전자좌 (Pox3 (진뱅크 관리 번호 XP_503244) 또는 Aco3 (진뱅크 관리 번호 AJ001301)), Δ12 데새투라제 유전자의 유전자좌 (서열 28), Lip1 유전자의 유전자좌 (진뱅크 관리 번호 Z50020) 및/또는 Lip2 유전자의 유전자좌 (진뱅크 관리 번호 AJ012632) 등이 있다.

유리하게도, Ura3 유전자는 5-FOA 선별 (상기 문헌)을 병행하여 반복적으로 사용될 수 있다. 보다 구체적으로, 천연 Ura3 유전자를 우선 넉아웃(knockout)시켜서 Ura- 표현형을 갖는 균주를 생성할 수 있고, 이때의 선별은 5-FOA 내성을 기초로 한다. 이어서, 다중 키메라 유전자 및 신규한 Ura3 유전자의 클러스터(cluster)를 야로위아 게놈의 상이한 유전자좌에 통합시킴으로써 Ura+ 표현형을 갖는 신규 균주를 생성할 수 있다. 도입된 Ura3 유전자가 넉아웃된 경우에는 이후의 통합에 의해 새로운 Ura3- 균주 (다시 5-FOA 선별을 이용하여 동정함)를 생성할 수 있다. 따라서, Ura3 유전자 (5-FOA 선별을 병행함)를 여러 회의 형질전환에서 선별 마커로서 이용할 수 있고, 이에 따라 유전자 변형물이 야로위아 게놈에 간편한 방식으로 쉽게 통합되도록 할 수 있다.

몇몇 예에서는, 본 발명의 단백질을 상이한 세포내 구획 (예컨대, 아실-CoA 풀 *vs.* 포스파티딜콜린 풀)으로 지시하는 것이 유용할 것이다. 본원에 기재한 목적상, DHA는 유리 지방산으로서 존재하거나 또는 에스테르화된 형태, 예컨대 아실글리세롤, 인지질, 술포지질 또는 당지질로서 존재할 수 있다. DHA 생합성을 허용하는 폴리펩티드를 코딩하는 상기 키메라 유전자가 적절한 세포내 표적화 서열을 포함하도록 추가로 유전자 조작될 수 있음이 이해된다.

쥬렛젝 등 [Yeast, 18:97-113 (2001)]은 야로위아 리폴리티카 내에서 통합된 플라스미드 카피수의 안정성이 개개의 형질전환체, 수용성 균주 및 사용된 표적화 플랫폼에 따라 달라진다는 것에 주목하였다. 따라서, 당업자는 원하는 발현 수준 및 패턴을 보이는 균주를 수득하기 위해서 다중 형질전환체가 스크리닝되어야 함을 인지할 것이다. 이러한 스크리닝은 DNA 블럿의 써던 분석 [Southern, J. Mol. Biol. 98:503 (1975)], mRNA 발현의 노던 분석 [Kroczek, J. Chromatogr. Biomed. Appl., 618(1-2):133-145 (1993)], 단백질 발현의 웨스턴 분석, PUFA 생성물의 표현형 분석 또는 GC 분석으로 수행될 수 있다.

요약하면, 상기 수단 각각은 야로위아 리폴리티카에서 특정 유전자 생성물 (예컨대, 데새투라제, 엘론가제, 아실트랜스퍼라제)의 발현을 증가시키는데 유용하며, 생명공학 분야의 당업자는 DHA를 많이 생산할 수 있는 가장 적절한 조합의 방법들을 쉽게 선택할 수 있을 것이다.

DHA 생산 증가를 위한 경로의 조작

상기 방법은 개개의 이종 유전자 발현을 상향 조절하는데 유용하지만, 야로위아 리폴리티카에서 DHA 생산을 증가시키는 것은 훨씬 더 복잡한 일이고 다양한 대사 경로를 조화롭게 조작하는 것이 요구될 수 있다. PUFA 생합성 경로의 조작이 우선 이루어진 후에 TAG 생합성 경로 및 TAG 분해 경로에서의 바람직한 조작이 뒷따를 것이다.

앞서 기재한 바와 같이, DHA를 전체 오일 분획의 5% 초과로, 또는 더욱 바람직하게는 DHA를 전체 오일 분획의 10% 초과로, 또는 훨씬 더욱 바람직하게는 DHA를 전체 오일 분획의 15 내지 20% 초과로, 또는 가장 바람직하게는 DHA를 전체 오일 분획의 25 내지 30% 초과로 생산하는 야로위아 리폴리티카 균주의 구축에는 (1) 적어도 Δ6 데새투라제/Δ6 엘론가제 경로를 발현하기 위한 Δ6 데새투라제, $C_{18/20}$ 엘론가제, Δ5 데새투라제 및 Δ17 데새투라제 또는 Δ15 데새투라제 (또는 둘다), $C_{20/22}$ 엘론가제 및 Δ4 데새투라제; 또는 (2) 적어도 Δ9 엘론가제/Δ8 데새투라제 경로를 발현하기 위한 Δ9 엘론가제, Δ8 데새투라제, Δ5 데새투라제 및 Δ17 데새투라제 또는 Δ15 데새투라제 (또는 둘다), $C_{20/22}$ 엘론가제 및 Δ4 데새투라제가 필요하다. 그러나, 이들 중 어느 실시양태에서도 숙주 균주에 Δ9 데새투라제, Δ12 데새투라제, $C_{14/16}$ 엘론가제 및/또는 $C_{16/18}$ 엘론가제를 추가로 포함시키는 것이 바람직할 수 있다.

일부 경우에서는, 천연 야로위아 리폴리티카 Δ12 데새투라제를 푸사륨 모닐리포르메 Δ12 데새투라제로 대체하는 것이 유리한 것으로 입증될 수 있으며, 이는 푸사륨 모닐리포르메 Δ12 데새투라제가 증가된 기질 전환율(%)을 나타내기 때문이다 (제WO 2005/047485호). 보다 구체적으로, 이들 2종의 Δ12 데새투라제 모두가 올레산이 LA로 전환되는 것을 촉매하지만, 상기 2종의 효소는 전반적인 특이성 (이에 따라 각 효소의 기질 전환율(%)이 영향을 받음)이 상이하다. 본 발명자들은 푸사륨 모닐리포르메 Δ12 데새투라제가 LA를 포스포티딜콜린 기질의 sn-2 위치 (이로써 Δ6 데새투라제에 의한 이후의 반응을 용이하게 함)에 이동시키는 능력(loading capacity)이 야로위아 리폴리티카 Δ12 데새투라제보다 더 높다고 결정하였다. 이것을 기초로, 푸사륨 모닐리포르메 Δ12 데새투라제의 과발현과 야로위아 리폴리티카 Δ12 데새투라제의 넉아웃은 DHA로의 추후 전환을 위한 생성물을 증가시킬 수 있다.

일부 실시양태에서, 숙주 유기체의 천연 DAG AT의 활성을 조절함으로써 야로위아 리폴리티카 숙주의 지질 및 오일 내 PUFA의 백분율을 조작하는 것이 유용할 수 있다. 구체적으로, 오일 생합성이 유질화 동안 다중불포화와 경쟁한다고 예측되기 때문에, 유기체의 1종 이상의 아실트랜스퍼라제 (예컨대, PDAT 및/또는 DGAT1 및/또는 DGAT2) 활성을 저하시키거나 불활성화시켜서, 오일 생합성의 전반적인 비율은 저하시키면서 그와 동시에 지질 및 오일 분획으로 혼입되는 PUFA의 백분율 (전체 지방산에 대한 백분율)은 증가시킬 수 있다. 이러한 결과는 다중불포화가 보다 효율적으로 일어날 수 있기 때문이며, 또는 다른 말로 하면 특이적 DAG AT 활성을 하향 조절함으로써 오일 생합성과 다중불포화 사이의 기질 경쟁이 유질화 동안 다중포화가 유리하도록 저하될 수 있기 때문이다.

당업자는 최적 수준의 하향 조절을 이루는데 필요한 기술 및 이러한 억제 달성에 필요한 수단을 보유하고 있을 것이다. 예를 들어, 일부 바람직한 실시양태에서, 단일 DAG AT의 활성을 조작하는 것이 바람직할 수 있다 (예컨대, DGAT1 넉아웃을 생성하면서 PDAT와 DGAT2의 활성은 변경시키지 않음). 별법의 실시양태에서, 유질 유기체는 총 "n"종의 천연 DAG AT를 포함하고, 총 "n-1"종의 아실트랜스퍼라제 활성이 변형되어 오일 생합성의 속도를 저하시키는 반면에 나머지 아실트랜스퍼라제는 그의 야생형 활성을 유지한다. 그리고, 일부 상황에서는, 다중불포화 비율에 대한 오일 생합성의 최적 비율의 달성을 위해서 일부 바람직한 유질 유기체 중 모든 천연 DAG AT의 활성을 조작하는 것이 바람직할 수 있다.

유사한 방식으로, 본 발명자들은 이종 아실트랜스퍼라제의 발현 및 상응하는 천연 야로위아 리폴리티카 아실트랜스퍼라제의 넉아웃이 숙주 세포에서 생산되는 전반적 DHA를 유의하게 증가시킬 수 있다는 가설을 세웠다. 구체적으로, 이전에 제안되었던 바와 같이, C20 이상인 지방산에 특이성을 갖는 이종 GPAT, LPAAT, DGAT1, DGAT2, PDAT 및 LPCAT 아실트랜스퍼라제가 천연 효소보다 더 바람직할 수 있는데, 이는 야로위아 리폴리티카에서 천연적으로 생산된 PUFA가 18:2 지방산으로 제한되고 천연 효소가 더욱 장쇄의 지방산을 이용한 반응을 효율적으로 촉매할 수 없기 때문이다. 이러한 결론을 기초로 하여, 본 발명자들은 모르티에렐라 알피나에서 GPAT, LPAAT, DGAT1 및 DGAT2를 코딩하는 유전자를 동정하고 이들 유전자를 유전자 조작된 EPA-생산 야로위아 숙주에서 발현시켜서 PUFA 생합성을 증가시켰다 (본원 실시예 17 내지 실시예 20). 이후, 천연 및 이종 아실트랜스퍼라제 사이의 기질 경쟁을 저하시키는 수단으로서, 야로위아 리폴리티카 중의 여러가지 천연 아실트랜스퍼라제 (예컨대, DGAT1 및 DGAT2)의 활성을 감소시키거나 넉아웃시켰다. DHA를 생산하는 유전자 조작된 야로위아 숙주에서 유사한 결과가 예상된다.

당업자는 또한 DHA 생산에 영향을 주는 경로 및 포괄적인 조절자의 조작을 고려해야 한다. 예를 들어, 더욱 장쇄의 포화 및 불포화 지방산, 예컨대 팔미테이트 (16:0) 및 스테아르산 (18:0)의 전구체의 이용가능성을 증가시켜서 PUFA 생합성 경로로의 탄소 유동을 증가시키는 것이 유용하다. 상기한 장쇄의 포화 지방산의 합성은 $C_{14/16}$ 엘론가제의 활성에 의존적인 반면, 상기한 장쇄의 불포화 지방산의 합성은 $C_{16/18}$ 엘론가제의 활성에 의존적이다. 따라서, 천연 야로위아 리폴리티카 $C_{14/16}$ 엘론가제 (서열 97 및 98)의 과발현은 16:0 및 16:1 지방산의 생산을 실질적으로 증가시켰고 (대조군 균주에 비해 22% 증가), 유사하게 천연 야로위아 리폴리티카 $C_{16/18}$ 엘론가제 (서열 94 및 95)의 과발현은 18:0, 18:1, 18:2 및 18:3

지방산의 생산을 실질적으로 증가시켰으며 (대조군 균주에 비해 18% 증가), C_{16} 지방산의 축적을 저하시켰다 (대조군 균주에 비해 22% 감소). 물론, 본원에서 입증되고 이나가끼, 케이.(Inagaki, K.) 등의 연구 [Biosci. Biotech. Biochem. 66 (3):613-621 (2002)]에서 제안된 바와 같이, 본 발명의 일부 실시양태에서는 이종 $C_{16/18}$ 엘론가제 (예컨대, 라투스 노르베기쿠스 [진뱅크 관리 번호 AB071986, 본원의 서열 83 및 84] 및/또는 모르티에렐라 알피나 [서열 86 및 87]로부터 유래된 것)의 동시 발현이 유용할 수 있다. 따라서, 야로위아 리폴리티카 숙주 균주는 DHA 생합성을 위해서 최소한 (1) Δ6 데새투라제, $C_{18/20}$ 엘론가제, Δ5 데새투라제 및 Δ17 데새투라제 또는 Δ15 데새투라제 (또는 둘다), $C_{20/22}$ 엘론가제, Δ4 데새투라제; 또는 (2) Δ9 엘론가제, Δ8 데새투라제, Δ5 데새투라제 및 Δ17 데새투라제 또는 Δ15 데새투라제 (또는 둘다), $C_{20/22}$ 엘론가제, 및 Δ4 데새투라제를 발현하도록 조작되어야 하지만, 추가의 바람직한 실시양태에서는 숙주 균주가 Δ9 데새투라제, Δ12 데새투라제, $C_{14/16}$ 엘론가제 및/또는 $C_{16/18}$ 엘론가제 중 1종 이상을 추가로 포함한다.

또다른 바람직한 실시양태에서, 본 발명의 야로위아 리폴리티카에서 지방산 분해 및 TAG 분해에 영향을 주는 경로를 변형시켜서, 세포 내에서 아실-CoA 풀 또는 TAG 분획에 축적된 DHA의 분해를 최소화하였다. 이들 경로는 아실-CoA 옥시다제 및 리파제 유전자 각각에 의해 대표된다. 보다 구체적으로, 아실-CoA 옥시다제 (EC 1.3.3.6)는 각각의 분해 주기로 아세틸-CoA 분자 및 지방산 기질보다 2개의 탄소 원자가 더 짧은 지방산이 수득되는 퍼옥시좀 β-산화 반응을 촉매한다. 야로위아 리폴리티카에는 5종의 아실-CoA 옥시다제 동종효소가 존재하며, 진뱅크 관리 번호 AJ001299 내지 AJ001303 각각에 상응하는 POX1, POX2, POX3, POX4 및 POX5 유전자 (또한, Aco1, Aco2, Aco3, Aco4 및 Aco5 유전자라고도 공지되어 있음)에 의해 코딩된다 (또한, 상응하는 진뱅크 관리 번호 XP_504703, XP_505264, XP_503244, XP_504475 및 XP_502199 참조). 상기 동종효소 각각은 기질 특이성이 상이하다. 예를 들어, POX3 유전자는 단쇄 지방산에 대하여 활성인 아실-CoA 옥시다제를 코딩하는 반면, POX2 유전자는 더욱 장쇄의 지방산에 대하여 활성인 아실-CoA 옥시다제를 코딩한다 [Wang H.J., et al. J. Bacteriol., 181:5140-5148 (1999)]. 본 발명의 숙주 세포에서 퍼옥시좀 β-산화를 본원에서의 목적에 유리할 수 있는 방식으로 변형시키면, 상기 유전자 중 임의의 것의 활성이 저하되거나 제거될 수 있다고 여겨진다. 최종적으로, 혼동되는 것을 피하기 위해서, 본 발명자들은 상기한 아실-CoA 옥시다제를 POX 유전자라고 지칭할 것이지만, 이러한 용어는 공개적으로 입수가능한 몇몇 문헌에 따라 Aco 유전자 명명법과 혼용될 수 있다.

유사하게, 야로위아 리폴리티카에서 세포내 효소, 막-결합 효소 및 세포외 효소를 비롯한 여러 리파제 (EC 3.1.1.3)가 검출되었다 ([Choupina, A., et al. Curr. Genet. 35:297 (1999)], [Pignede, G., et al. J. Bacteriol. 182:2802-2810 (2000)]). 예를 들어, Lip1 (진뱅크 관리 번호 Z50020) 및 Lip3 (진뱅크 관리 번호 AJ249751)은 세포내 효소 또는 막-결합 효소인 반면, Lip2 (진뱅크 관리 번호 AJ012632)는 세포외 리파제를 코딩한다. 이들 리파제 각각은 이들 효소가 TAG 및 물이 DAG 및 지방산 음이온으로 직접 분해되는 반응을 촉매하기 때문에 파괴할 표적이었다.

추가의 별법의 실시양태에서, 야로위아 리폴리티카의 바람직한 숙주 균주에서 여러가지 포스포리파제의 활성이 조작될 수 있다. 포스포리파제는 막 지질의 생합성 및 분해에 중대한 역할을 한다. 보다 구체적으로, 용어 "포스포리파제"는 글리세로인지질 중의 1개 이상의 에스테르 연결부를 가수분해하는 능력을 공유하는 이질성 군의 효소를 지칭한다. 모든 포스포리파제가 인지질을 기질로서 표적화하지만, 각 효소는 특이적 에스테르 결합을 절단하는 능력을 보유한다. 따라서, 포스포리파제에 대한 명명법은 개개의 포스포리파제를 차별화하고, 인지질 분자 내에 표적화된 특이적 결합을 지시한다. 예를 들어, 포스포리파제 A_1 (PLA_1)은 글리세롤 부분의 sn-1 위치에서 지방 아실 에스테르 결합을 가수분해하지만, 포스포리파제 A_2 (PLA_2)는 이 분자의 sn-2 위치에서 지방산을 제거한다. PLA_1 (EC 3.1.1.32) 및 PLA_2 (EC 3.1.1.4)의 작용은 유리 지방산 및 2-아실 리소인지질 또는 1-아실 리소인지질 각각의 축적을 초래한다. 포스포리파제 C (PLC) (EC 3.1.4.3)는 인지질 주쇄 중의 포스포디에스테르 결합을 가수분해하여 1,2-DAG를 생성하고, 또한 포함된 구체적인 인지질 종에 따라 포스파티딜콜린, 포스파티딜에탄올아민 등을 생성한다 (예컨대, PLC_1은 1-포스파티딜-1D-마이오-이노시톨 4,5-비스포스페이트 + H_2O = 1D-마이오-이노시톨 1,4,5-트리포스페이트 + DAG의 반응을 담당함. ISC1은 이노시톨 스핑고인지질-특이적 포스포리파제 C를 코딩함 [Sawai, H., et al. J. Biol. Chem. 275, 39793-39798 (2000)]). 제2의 포스포디에스테르 결합은 포스포리파제 D (PLD) (EC 3.1.4.4)에 의해 절단되어 포스파티드산이 생성되고, 또한 마찬가지로 포함된 인지질 클래스에 따라 콜린 또는 에탄올아민이 생성된다. 포스포리파제 B (PLB)는 sn-1 및 sn-2 지방산 둘다를 제거할 수 있으며, 히드롤라제 (상기 효소는 2가지 지방산 인지질 [PLB 활성]과 리소인지질 [리소포스포리파제 활성] 둘다를 절단하여 지방산을 방출함) 및 리소포스포리파제-트랜스아실라제 활성 (상기 효소는 유리 지방산을 리소인지질에 전달하여 인지질을 생성할 수 있음)을 둘다 보유한다는 점에서 독특하다. 형질전환체 야로위아 숙주 세포의 전체 오일 분획에 축적되는 DHA의 농도를 증가시키기 위해서는 이들 포스포리파제 중 1종 이상을 과발현하는 것이 유용할 수 있다. 이러한 영향은 포스포리파제가 신장 또는 트리글리세리드로의 혼입을 위해 PC로부터의 아실기를 CoA 풀에 방출하기 때문에 관찰된 것이라는 가설을 세웠다.

또다른 별법의 실시양태에서, 야로위아 리폴리티카의 바람직한 숙주 균주에서, 포스파티딜콜린 (PC) 생합성을 담당하는 CDP-콜린 경로의 효소가 DHA 생합성을 전반적으로 증가시키기 위한 수단으로서 조작될 수도 있다. 상기 기술의 유용성은 본 명세서에서 디아실글리세롤 콜린포스포트랜스퍼라제 (EC 2.7.8.2)를 코딩하는 야로위아 리폴리티카 CPT1 유전자의 과발현에 의해 야로위아 리폴리티카의 유전자 조작된 균주에서 EPA 생합성이 증가되는 것으로 입증된다. 당업자는 PC 생합성 경로를 잘 알고 있을 것이고, 다른 적절한 후보 효소를 인지하고 있을 것이다.

상기한 바와 같은 생화학적 경로를 조작하는 방법은 당업자에게 공지되어 있지만, 천연 유전자의 활성을 저하시키거나 없애기 위한 몇가지 기술의 개요를 이하에서 간략하게 제시할 것이다. 이들 기술은 상기 논의한 바와 같은 천연 야로위아 리폴리티카 Δ12 데새투라제, GPAT, LPAAT, DGAT1, DGAT2, PDAT, LPCAT, 아실-CoA 옥시다제 2 (Aco2 또는 Pox2), 아실-CoA 옥시다제 3 (Aco3 또는 Pox3) 및/또는 리파제 유전자의 활성을 하향 조절하는데 유용하다.

당업자는 천연 유전자의 활성을 저하시키거나 없애기 위해 사용될 가장 적절한 기술을 확인하겠지만, 일반적으로는 특정 유전자의 내인성 활성은 예를 들어 1) 표적 유전자 전체 또는 일부의 삽입, 치환 및/또는 결실에 의한 유전자 파괴, 2) 안티센스 서열을 유전자의 전사 생성물로 전사하기 위한 카세트의 제공, 3) 특이적 유전자 활성이 천연적으로 거의 없거나 전혀 없는 [또는 그러하도록 돌연변이된] 숙주 세포의 사용, 4) 돌연변이화된 헤테로서브유닛 (즉, 2종 초과의 헤테로서브유닛을 포함하는 효소의 것)을 과발현시켜서, "우성 음성 효과"에 의해 효소의 활성을 저하시킴, 및 5) iRNA 기술의 이용에 의해 저하 또는 제거될 수 있다. 일부 경우에서, 원치않는 유전자 경로의 억제는 특이적 억제제 (예컨대, 미국 특허 제 4,778,630호에 기재된 것과 같은 데새투라제 억제제)의 사용을 통해 수행될 수도 있다.

유전자 파괴를 위해, 외래 DNA 단편 (전형적으로는 선별가능한 마커 유전자, 그러나 임의로는 발현시에 바람직한 표현형을 나타내는 키메라 유전자 또는 키메라 유전자 클러스터)을 파괴될 구조적 유전자에 삽입시켜서 그의 코딩 서열을 중단시키고 이에 의해 그 유전자를 기능적으로 불활성화시킨다. 파괴 카세트로 숙주 세포를 형질전환시키면, 기능적 천연 유전자가 기능적으로 파괴되지 않은 유전자와의 상동성 재조합에 의해 대체된다 (예컨대 문헌 [Hamilton et al. J. Bacteriol. 171:4617-4622 (1989)], [Balbas et al. Gene 136:211-213 (1993)], [Gueldener et al. Nucleic Acids Res. 24:2519-2524 (1996)] 및 [Smith et al. Methods Mol. Cell. Biol. 5:270-277(1996)] 참조).

안티센스 기술은 표적 유전자의 서열이 알려져 있는 경우에 유전자를 하향 조절하는 또다른 방법이다. 이를 수행하기 위해서, 원하는 유전자로부터의 핵산 절편을 클로닝하여 프로모터에 작동가능하게 연결시켜서 RNA의 안티센스 가닥이 전사되도록 한다. 이어서, 상기 구축물을 숙주 세포에 도입하고, RNA의 안티센스 가닥을 발현시킨다. 안티센스 RNA는 관심 단백질을 코딩하는 mRNA의 축적을 방해하여 유전자 발현을 억제한다. 당업자라면, 특정 유전자의 발현 저하를 위해서는 안티센스 기술의 이용에 대하여 특별히 고려해야 함을 알고 있을 것이다. 예를 들어, 안티센스 유전자의 발현 수준을 적당하게 하기 위해서는 당업자에게 공지된 상이한 조절 요소를 이용하는 여러 키메라 유전자의 사용이 필요할 수 있다.

표적화된 유전자 파괴 및 안티센스 기술은 공지 서열의 유전자를 하향 조절하는 효과적인 수단을 제공하지만, 서열-기재가 아닌 다른 덜 특이적 방법이 개발되어 왔다 (예컨대, UV 조사/화학 작용제 또는 위치전이성(transposable) 요소/트란스포손의 사용을 통한 돌연변이유발. 제WO 04/101757호 참조).

별법의 실시양태에서, 특정 유전자의 내인성 활성은 단백질의 발현을 조절하는 조절 서열의 조작에 의해 저하될 수 있다. 당업계에 공지된 바와 같이, 코딩 서열과 관련이 있는 조절 서열은 코딩 서열의 상류 (5' 비-코딩 서열), 내부, 또는 하류 (3' 비-코딩 서열)에 위치한 전사 및 번역 "조절" 뉴클레오티드 서열을 포함하며, 전사, RNA 가공 또는 안정성, 또는 관련된 코딩 서열의 번역에 영향을 미친다. 따라서, 특정 유전자의 조절 서열의 조작은 유전자의 프로모터, 번역 리더 서열, 인트론, 인핸서, 개시 조절 영역, 폴리아데닐화 인식 서열, RNA 가공 부위, 이펙터 결합 부위 및 스템-루프 구조의 조작을 지칭하는 것일 수 있다. 따라서, 예를 들어, DAG AT의 발현을 하향 조절하여 지질 및 오일 생합성의 비율 저하를 달성하기 위해서 DAG AT의 프로모터를 결실시키거나 파괴시킬 수 있다. 별법으로, DAG AT의 발현을 구동하는 천연 프로모터를, 프로모터 활성이 천연 프로모터에 비해 감소된 이종 프로모터로 치환될 수 있다. 조절 서열 조작에 유용한 방법은 당업자에게 공지되어 있다.

요약하면, 본원에서 교시된 내용을 이용하여, 형질전환체 유질 미생물 숙주는 DHA를 미생물 숙주 전체 지질의 약 5% 이상, 바람직하게는 DHA를 전체 지질의 약 10% 이상, 더욱 바람직하게는 DHA를 전체 지질의 약 15% 이상, 더욱 바람직하게는 DHA를 전체 지질의 약 20% 이상, 가장 바람직하게는 DHA를 전체 지질의 약 25 내지 30% 이상으로 생산할 것이다.

DHA 생산을 위한 발효 공정

형질전환된 미생물 숙주 세포는, 키메라 유전자 (예컨대, 데새투라제, 엘론가제, 아실트랜스퍼라제 등을 코딩함)의 발현을 최적화하고 DHA를 가장 많이 가장 경제적으로 수득하는 조건하에 성장시킨다. 일반적으로, 최적화될 수 있는 배지 조건으로는, 탄소 공급원의 유형과 양, 질소 공급원의 유형과 양, 탄소:질소 비율, 산소 수준, 성장 온도, pH, 생물체 생성기의 길이, 오일 축적기의 길이 및 세포 수확 시간 등이 있다. 야로위아 리폴리티카는 일반적으로 복합 배지 (예컨대, 효모 추출물-펩톤-덱스트로즈 브로쓰(broth) (YPD)) 또는 성장에 필요한 성분이 없어서 원하는 발현 카세트가 선별될 수 있도록 하는 규정된 최소 배지 (예컨대, 효모 질소 염기 (디프코 래버러토리즈(DIFCO Laboratories), 미국 미시간주 디트로이트 소재)) 중에서 성장한다.

본 발명에서의 발효 배지는 적합한 탄소 공급원을 함유해야 한다. 적합한 탄소 공급원으로는 단당류 (예컨대, 글루코스, 프럭토스), 이당류 (예컨대, 락토스, 수크로스), 올리고당류, 다당류 (예컨대, 전분, 셀룰로스 또는 이들의 혼합물), 당 알콜 (예컨대, 글리세롤) 또는 보완가능한 공급원료 (예컨대, 치즈 유장 투과물, 옥수수 침유(cornsteep liquor), 사탕무 당밀, 보리 맥아)로부터의 혼합물 등을 들 수 있으나 이에 제한되지 않는다. 추가로, 탄소 공급원은 알칸, 지방산, 지방산의 에스테르, 모노글리세리드, 디글리세리드, 트리글리세리드, 인지질 및 식물성 오일 (예컨대, 대두유) 및 동물 지방을 비롯한 지방산의 다양한 시판 공급원을 포함할 수 있다. 추가로, 탄소 공급원은 핵심적인 생화학적 중간체로의 대사 전환이 입증된 바 있는 1-탄소 공급원 (예컨대, 이산화탄소, 메탄올, 포름알데히드, 포르메이트 및 탄소-함유 아민)을 포함할 수 있다. 따라서, 본 발명에 이용되는 탄소 공급원은 광범위하게 다양한 탄소-함유 공급원을 포괄하는 것으로 고려된다. 상기 언급한 모든 탄소 공급원 및 이들의 혼합물이 본 발명에 적합하다고 예상되지만, 바람직한 탄소 공급원은 당 및/또는 지방산이다. 가장 바람직한 것은 글루코스 및/또는 10 내지 22개 탄소를 함유하는 지방산이다.

질소는 무기 (예컨대, $(NH_4)_2SO_4$) 또는 유기 (예컨대, 우레아 또는 글루타메이트) 공급원으로부터 공급될 수 있다. 발효 배지는 적절한 탄소 및 질소 공급원 뿐만이 아니라 적합한 미네랄, 염, 보조인자, 완충제, 비타민, 및 유질 효모의 성장 및 DHA 생산에 필요한 효소 경로의 촉진에 적합하다고 당업계에 공지된 기타 성분도 함유해야 한다. 특히, 지질 및 PUFA의 합성을 촉진시키는 여러 금속 이온 (예컨대, Mn^{+2}, Co^{+2}, Zn^{+2}, Mg^{+2})에 주목한다 [Nakahara, T. et al., Ind. Appl. Single Cell Oils, D. J. Kyle and R. Colin, eds. pp 61-97 (1992)].

본 발명에서 바람직한 성장 배지는 시판용으로 제조된 통상의 배지, 예컨대 효모 질소 염기 (디프코 래버러토리즈, 미국 미시간주 디트로이트 소재)이다. 다른 규정된 성장 배지 또는 합성 성장 배지가 사용될 수도 있고, 야로위아 리폴리티카의 성장에 적절한 배지는 미생물학 또는 발효 과학 분야의 당업자에게 공지되어 있다. 발효에 적합한 pH 범위는 전형적으로 약 pH 4.0 내지 pH 8.0이며, pH 5.5 내지 pH 7.0이 초기 성장 조건의 범위로서 바람직하다. 발효는 호기성 또는 혐기성 조건하에서 수행될 수 있고, 미호기성 조건이 바람직하다.

전형적으로, 유질 효모 세포 중 고수준의 PUFA의 축적은 2-단계 공정을 필요로하는데, 대사 상태가 성장 및 지방의 합성/저장 사이에서 "균형"을 이루어야 하기 때문이다. 따라서, 가장 바람직하게는, 야로위아 리폴리티카에서의 DHA 생산에 2-단계 발효 공정이 필요하다. 이러한 접근법은 제WO 2004/101757호에 다양한 적합한 발효 공정 디자인 (즉, 배치식, 공급형-배치식 및 연속식) 및 성장 동안의 고려사항으로서 기재되어 있다.

DHA의 정제 및 가공

DHA를 비롯한 PUFA는 숙주 미생물 중에서 유리 지방산 또는 아실글리세롤, 인지질, 술포지질 또는 당지질 등과 같은 에스테르화 형태로 존재할 수 있으며, 당업계에 공지된 각종 수단을 통해 상기 숙주 세포로부터 추출될 수 있다. 효모 지질의 추출 기술, 품질 분석 및 허용 기준에 대한 검토 중 하나는 문헌 [Z. Jacobs (Critical Reviews in Biotechnology, 12(5/6):463-491 (1992))]이다. 하류 가공에 대한 간략한 검토는 또한 문헌 [A. Singh and O. Ward (Adv. Appl. Microbiol., 45:271-312 (1997))]에서도 찾을 수 있다.

일반적으로, DHA 및 다른 PUFA의 정제 수단으로는 유기 용매를 이용한 추출, 초음파처리, 초임계 유체 추출 (예컨대, 이산화탄소 사용), 비누화 및 압착과 같은 물리적 수단, 또는 이들의 조합 등을 들 수 있다. 추가의 세부사항에 대하여는 제WO 2004/101757호의 교시 내용을 참조한다.

정련되고/되거나 정제된 DHA를 함유하는 오일은 수소화되어 다양한 용융 성질 및 조직을 갖는 지방이 생성될 수 있다. 스프레드, 제과 지방, 고체 버터, 마가린, 베이킹 쇼트닝 등을 비롯한 많은 가공 지방에는 실온에서의 고화도 정도가 다양할 것이 요구되며, 공급원 오일의 물리적 성질을 변경하여야 생산될 수 있다. 이것은 가장 통상적으로 촉매적 수소화에 의해 달성된다.

수소화는 니켈과 같은 촉매의 보조하에 수소가 불포화 지방산 이중 결합에 첨가되는 화학 반응이다. 수소화는 2가지 주요 효과를 갖는다. 우선, 불포화 지방산 함량의 감소로 인해 오일의 산화적 안정성이 증가한다. 다음으로, 지방산 변형이 융점을 증가시켜서 실온에서 반액체 또는 고체인 지방을 초래하기 때문에, 오일의 물리적 성질이 변화된다.

수소화 반응에 영향을 미치고 따라서 최종 생성물의 조성을 변경시키는 변수는 많이 있다. 압력, 온도, 촉매 유형 및 농도, 교반 및 반응기 디자인 등을 비롯한 작업 조건은 조절될 수 있는 더 중요한 파라미터이다. 선택적인 수소화 조건이 이용되어 더 높은 수준으로 불포화된 지방산, 우선적으로는 덜 불포화된 지방산을 수소화할 수 있다. 매우 적은 정도 또는 약간의 수소화는 액체 오일의 안정성을 증가시키는데 흔히 이용된다. 추가의 수소화는 액체 오일을 물리적으로 고체인 지방으로 전환시킨다. 수소화 정도는 특정 최종 생성물에 대해 디자인된 원하는 성능 및 용융 특징에 따라 달라진다. 베이킹 제품의 제조에 사용되는 액체 쇼트닝, 상업적으로 튀기고 굽는 작업에 사용되는 고체 지방 및 쇼트닝, 및 마가린 제조를 위한 베이스 스톡(base stock)이 수소화를 통해 달성될 수 있는 무수한 오일 및 지방 생성물 중에 속한다. 수소화 및 수소화 생성물에 대한 보다 상세한 설명은 문헌 [Patterson, H. B. W., Hydrogenation of Fats and Oils: Theory and Practice. The American Oil Chemists' Society, 1994]에서 찾을 수 있다.

식료품에 사용하기 위한, 야로위아 리폴리티카의 DHA-생산 균주

시장은 현재 ω-3 및/또는 ω-6 지방산 (특히 ARA, EPA 및 DHA)이 혼입된 매우 다양한 식품 및 사료 제품을 선호하고 있다. DHA를 포함하는 본 발명의 효모 미생물성 오일은 식품 및 사료 제품에서 이들 제제의 건강상의 이점을 부여하는 기능을 하는 것으로 고려된다.

본원에 기재한 효모 숙주에 의해 생산된 ω-3 및/또는 ω-6 지방산-함유 미생물성 오일은 식품 유사물, 육류 제품, 곡류 제품, 베이킹 식품, 스낵 식품, 및 유제품을 포함하지만 이에 제한되지 않는 각종 식품 및 사료 제품에서 사용하기에 적합할 것이다. 추가로, 본 발명의 미생물성 오일은 의료용 영양제, 식이 보조제, 유아용 조제식 및 또한 약제 생성물 등을 비롯한 의료용 식품에서 건강상의 이점을 부여하기 위한 제제에 사용될 수 있다. 식품 가공 및 식품 제제화 분야의 당업자는 소정량과 소정의 조성을 갖는 미생물성 오일이 식품 또는 사료 제품에 첨가될 수 있는 방법을 이해할 것이다. 이러한 양은 본원에서 "유효량"이라 지칭되며, 상기 생성물이 보충할 음식물인 식품 또는 사료 제품, 또는 의료용 식품 또는 의료용 영양제가 교정되거나 처치될 의학적 상태에 따라 달라질 것이다.

식품 유사물은 당업자에게 공지된 공정을 이용하여 제조될 수 있다. 육류 유사물, 치즈 유사물, 우유 유사물 등을 언급할 수 있다. 대두로부터 제조된 육류 유사물은 콩 단백질 또는 두부 및 다양한 종류의 육류를 모방하도록 함께 혼합된 다른 성분을 함유한다. 이들 육류 대안물은 냉동 식품, 통조림 식품 또는 건조 식품으로 판매된다. 통상적으로, 이들은 이것들이 대체할 식품과 동일한 방식으로 사용될 수 있다. 육류 유사물의 예로는 햄 유사물, 소시지 유사물, 베이컨 유사물 등이 있으나 이에 제한되지 않는다.

식품 유사물은 이들의 기능적 특성과 조성적 특성에 따라 대용물 또는 대체물로 분류될 수 있다. 예를 들어, 대용 치즈는 대체하도록 디자인된 치즈와 유사하기만 할 것이 요구된다. 그러나, 생성물이 치즈와 영양적으로 동등하고 치즈에 대한 최소 조성 요건을 대신하여 충족시키기만 한다면, 그 생성물은 통상적으로 대체용 치즈라고 불릴 수 있다. 따라서, 대체용 치즈는 종종 대용 치즈보다 단백질 수준이 더 높고 비타민과 미네랄이 강화되어 있다.

우유 유사물 또는 비-유제품 식품으로는 대용 우유 및 비-유제품 냉동 후식 (예컨대, 대두 및/또는 콩 단백질 생성물로부터 제조된 것) 등이 있으나 이에 제한되지 않는다.

육류 제품은 매우 다양한 생성물을 포함한다. 미국에서는 "육류"가 소, 돼지 및 양에서 얻은 "붉은 고기"를 포함한다. 붉은 고기뿐 아니라, 닭, 칠면조, 거위, 뿔닭, 오리 등을 비롯한 가금류 품목 및 어류와 조개류도 있다. 조리되고 가공된 육류 제품은 광범위하게 분류된다: 날 것, 보존처리되고 튀긴 것, 및 보존처리되고 요리된 것. 소시지 및 핫도그는 가공된 육류 제품의 예이다. 따라서, 본원에서 사용되는 용어 "육류 제품"은 가공된 육류 제품 등을 포함하지만 이에 제한되지 않는다.

시리얼 식품은 시리얼 곡물을 가공하여 유래된 식품이다. 시리얼 곡물은 식용 곡물 (종자)을 생성하는 목초류로부터의 임의의 식물을 포함한다. 가장 대중적인 곡물은 보리, 옥수수, 기장, 귀리, 퀴노아, 쌀, 호밀, 사탕수수, 라이밀, 밀 및 야생 벼이다. 시리얼 식품의 예로는 전립 곡물, 분쇄 곡물, 그릿(grits), 밀가루, 겨, 싹(germ), 아침식사용 시리얼, 압출식 식품, 파스타 등이 있으나 이에 제한되지 않는다.

베이킹 제품은 상기 언급한 임의의 시리얼 식품을 포함하며, 베이킹된 것이거나 베이킹과 유사한 방식으로 가공된 것, 즉 열을 가하여 건조 또는 경화시킨 것이다. 베이킹 제품의 예로는 빵, 케이크, 도넛, 바아(bars), 파스타, 빵가루, 베이킹 스낵, 미니-비스킷, 미니-크래커, 미니-쿠키, 및 미니-프레첼 등이 있으나 이에 제한되지 않는다. 상기 언급한 바와 같이, 본 발명의 오일이 성분으로 사용될 수 있다.

스낵 식품은 임의의 상기 또는 하기 식품을 포함한다.

튀긴 식품은 임의의 상기 또는 하기 튀긴 식품을 포함한다.

음료는 액체 또는 건조 분말 형태일 수 있다.

예를 들어, 과일 주스, (신선한 것, 냉동된 것, 캔에 들어 있거나 농축된 것), 향미료가 첨가되거나 첨가되지 않은 우유 드링크 등과 같은 비-탄산 음료가 언급될 수 있다. 성인 및 유아용 영양제는 당업계에 공지되어 있으며 시판되고 있다 (예컨대, 로스 프로덕츠 디비젼(Ross Products Division), 애보트 래버러토리즈(Abbott Laboratories)의 시밀락®, 엔슈어®, 제비티(Jevity®) 및 알리멘툼(Alimentum®)). 유아용 조제식은 유아와 영아에게 공급되는 액체 또는 재구성된 분말이다. 유아용 조제식은 모유를 대신한다. 유아용 조제식은 흔히 유아를 위한 유일한 영양 공급원이기 때문에, 유아용 조제식은 유아의 음식에서 특별한 역할을 한다. 유아에게는 수유가 여전히 가장 우수한 영양공급이지만, 유아용 조제식은 아기가 생존만 하게 하는 것이 아니라 성장하도록 하는데도 충분하다. 유아용 조제식은 모유에 점점 더 근접하고 있다.

유제품은 우유로부터 유래된 생성물이다. 우유 유사물 또는 비-유제품은 우유 이외의 공급원, 예를 들어 앞서 논의한 두유로부터 유래한다. 상기 생성물로는 전유(全乳), 탈지유, 발효유 제품, 예컨대 요구르트 또는 산미유(酸味乳), 크림, 버터, 연유, 탈수유(dehydrated milk), 커피 화이트너, 커피 크리머, 아이스크림, 치즈 등이 있으나 이에 제한되지 않는다.

본 발명의 DHA-함유 오일이 포함될 수 있는 추가의 식품으로는, 예를 들어 츄잉검, 제과류 및 프로스팅류(frostings), 젤라틴류 및 푸딩류, 경질 및 연질의 캔디, 잼 및 젤리, 과립화된 백설탕, 당 대체물, 달콤한 소스, 토핑류 및 시럽류, 및 건조 배합된 분말 혼합물 등이 있다.

건강 식품 및 약제

건강 식품은 건강상의 이점을 부여하는 임의의 식품이며, 기능성 식품, 의료용 식품, 의료용 영양제 및 식이 보조제 등이 있다. 추가로, 본 발명의 미생물성 오일은 표준 제약 조성물에 사용될 수 있다. 본 발명의 유전자 조작된 야로위아 리폴리티카 균주 또는 그로부터 생산된 DHA-포함 미생물성 오일은 상기 언급한 임의의 식품에 쉽게 혼입되어 예를 들어 기능적 식품 또는 의료용 식품을 생성할 수 있다. 예를 들어, DHA를 포함하는 보다 농축된 제제로는 인간 또는 인간 이외의 동물에서 식이 보조제로 사용될 수 있는 캡슐, 분말, 정제, 소프트겔, 겔캡, 액상 농축물 및 에멀젼 등이 있다.

식이 보조제에서의 용도

DHA를 포함하는 보다 농축된 제제로는 인간 또는 인간 이외의 동물에서 식이 보조제로 사용될 수 있는 캡슐, 분말, 정제, 소프트겔, 겔캡, 액상 농축물 및 에멀젼 등이 있다. 특히, 본 발명의 DHA-오일은 유아용 조제식 또는 아기 음식물과 같은 식이 보조제로 혼입시키는 데 특히 적합하다.

유아용 조제식은 유아와 영아에게 공급되는 액체 또는 재구성된 분말이다. 본원에서, "유아용 조제식"은 젖먹이 유아에서 모유를 대신할 수 있는 장용 영양 생성물로 정의되며, 전형적으로 수용액 중에서 원하는 백분율(%)의 탄수화물 및 단백질과 혼합된 원하는 백분율(%)의 지방으로 구성된다 (예컨대, 미국 특허 제4,670,285호 참조). 전 세계적인 조성 연구 및 전문가 그룹에 의해 규명된 수준을 기초로 할 때, 평균적인 모유는 전형적으로 총 지방산을 약 0.20% 내지 0.40%로 함유하며 (지방 칼로리의 약 50%로 추측됨), 일반적으로 DHA:ARA의 비율은 약 1:1 내지 1:2의 범위이다 (예컨대, 엔파밀 리필™(Enfamil LIPIL™) [미드 존슨 앤드 컴파니(Mead Johnson & Company)] 및 시밀락 어드밴스™(Similac Advance™) [로스 프로덕츠 디비젼, 애보트 래버러토리즈]의 제제 참조). 유아용 조제식은 흔히 유아를 위한 유일한 영양 공급원이기 때문에, 유아용 조제식은 유아의 음식에서 특별한 역할을 하며, 유아에게는 수유가 여전히 가장 우수한 영양공급이지만 유아용 조제식은 아기가 생존만 하게 하는 것이 아니라 성장하도록 하는데도 충분하다.

동물 사료에서의 용도

본원에서, 동물 사료는 인간 이외의 동물을 위한 사료로 사용되거나 사료에 혼합하기 위한 생성물을 총칭하여 정의된다. 그리고, 상기 언급된 바와 같이, 본 발명의 DHA-포함 오일은 다양한 동물 사료 중의 성분으로서 사용될 수 있다.

보다 구체적으로, 본 발명의 오일은 애완동물 사료, 반추 동물 및 가금류 사료, 및 양식용 사료 등을 포함하지만 이에 제한되지 않는 제품에 사용될 수 있다고 예측된다. 애완동물 사료는 애완동물 (예컨대, 개, 고양이, 새, 파충류, 설치류)에게 공급될 제품이고, 이들로는 상기한 시리얼 및 건강 식품 뿐만이 아니라 육류 및 육류 부산물, 콩 단백질 생성물, 목초 및 건초 제품 (예컨대, 자주개자리, 큰조아재비, 귀리 또는 브롬 그래스, 야채) 등이 있다. 반추 동물 및 가금류 사료는 예를 들어 칠면조, 닭, 소 및 돼지에게 공급될 제품이다. 상기한 애완동물 사료와 마찬가지로, 이들 제품은 앞서 언급한 시리얼 및 건강 식품, 콩 단백질 생성물, 육류 및 육류 부산물, 및 목초 및 건초 제품을 포함할 수 있다. 그리고, 양식 사료 (또는 "양식사료")는 신선한 물 또는 해양 물 중에서 수상 유기체 및/또는 동물을 증식, 배양 또는 사육하는 양식에 사용되는 제품이다.

대부분의 동물 사료 제제에는 ARA, EPA 및/또는 DHA를 고농도로 생산하는 본 발명의 유전자 조작된 야로위아 리폴리티카 균주를 포함시키는 것이 특히 유용할 것이라고 여겨진다. 필요한 ω-3 및/또는 ω-6 PUFA를 제공하는 것 이외에도, 상기 효모 그 자체가 동물의 전반적 건강 및 영양에 기여하고 제제의 맛을 증가시킬 수 있는 단백질 및 다른 영양분 (예컨대, 비타민, 미네랄, 핵산, 복합 탄수화물 등)의 유용한 공급원이다. 보다 구체적으로, 야로위아 리폴리티카 (ATCC #20362)에서 건조 세포 중량에 대한 백분율(%)로서의 대략적인 화학적 조성은 다음과 같다: 35% 단백질, 40% 지질, 10% 탄수화물, 5% 핵산, 5% 회분(ash) 및 5% 수분. 추가로, 탄수화물 분획 내에 β-글루칸은 대략 45.6 mg/g, 만나스는 대략 11.4 mg/g, 키틴은 대략 52.6 mg/g로 포함한다 (반면, 트레할로스는 소량 성분임 [대략 0.7 mg/g]).

상당한 문헌들에서 β-글루칸, 만나스 및 키틴의 면역 조정 효과가 조사된 바 있다. 키틴과 만나스는 둘다 유용한 면역자극제로서 유사하게 인식되지만, 박테리아 및 진균 세포벽의 주요 구성성분인 β-글루칸이 비-특이적 면역을 자극 (즉, "면역자극 효과")함으로써 양식 종, 애완 동물 및 농장 동물 및 인간의 건강을 개선시키는 수단이 가장 잘 연구되어 있다. 아주 간단하게 설명하면, 면역 반응의 전반적 향상은 β-글루칸의 사용으로 달성될 수 있는데, 이는 이들 β-1,3-D-폴리글루코스 분자가 백혈구 (예컨대, 대식세포, 호중구 및 단핵구)의 생성을 비-특이적 방식으로 자극함으로써 각종 병원성 항원 또는 환경적인 스트레서에 대한 민감도와 방어력을 증가시킬 수 있기 때문이다. 보다 구체적으로, 수많은 연구들은, β-글루칸이 바이러스, 박테리아, 진균 및 기생충 감염에 대한 보호력을 향상시키고, 항생제 및 백신과 병용될 때 보강 효과를 발휘하며, 창상 치유를 향상시키고, 유리 라디칼로 인한 손상을 호전시키고, 종양 퇴행을 향상시키고, 박테리아 내독소의 독성을 조정하며, 점막 면역을 강화시킨다는 것을 입증해 왔다 (문헌 [Raa, J. et al., Norwegian Beta Glucan Research, Clinical Applications of Natural Medicine. Immune: Depressions Dysfunction & Deficiency (1990)]에서 검토됨). 전통적인 동물 사육업 및 양식업 분야에 있어서 효모 β-글루칸, 만나스 및 키틴의 유용성을 문서화하고 있는 당업계 문헌의 예로는 [L.A. White et al. (J. Anim. Sci, 80:2619-2628 (2002)] (새끼 돼지를 위한 보충물), [K.S. Swanson et al. (J. Nutr. 132:980-989 (2002)] (개를 위한 보충물), [J. Ortuno et al. (Vet. Immunol. Immonopath. 85:41-50 (2002)] (도미(gilthead seabream)에게 투여된 온전한 사카로마이세스 세레비지애), [A. Rodriguez et al. (Fish Shell. Immuno. 16:241-249 (2004)] (도미에게 투여된 전장 무코르 써시넬로이데스), [M. Bagni et al. (Fish Shell. Immuno. 18:311-325 (2005) [(β-글루칸을 함유하는 효모 추출물을 함유하는, 농어(sea bass)를 위한 보충물), [J. Raa (In: Cruz - Suarez, L. E., Ricque-Marie, D., Tapia-Salazar, M., Olvera-Novoa, M.A. y Civera-Cerecedo, R., (Eds.). Avances en Nutricion Acuicola V. Memorias del V Simposium lnternacional de Nutricion Acuicola. 19-22 Noviembre, 2000. Merida, Yucatan, Mexico] (어류 및 조개류 사료에 면역-자극제를 사용하는 것에 대한 검토) 등이 있다.

야로위아 리폴리티카의 독특한 단백질:지질:탄수화물 조성 및 독특한 복합 탄수화물 프로파일 (대략 1:4:4.6 비율의 만난:β-글루칸:키틴을 포함함)을 기초로 하여, 본 발명의 유전적으로 유전자 조작된 효모 세포 (또는 그의 일부)는 동물 사료 제제에 유용한 첨가제라고 여겨진다 (예컨대, 온전한 [동결건조된] 효모 세포, 정제된 세포벽, 정제된 효모 탄수화물로서 사용되거나, 또는 다양한 다른 분획화 형태 중의 것으로 사용됨).

양식 산업과 관련하여, 다양한 어류 종에 대한 영양 요건의 이해가 높아지고 사료 제조에 있어서 기술적 진보가 이루어져서 양식 산업에서의 천연 사료를 보충하거나 대체하기 위해서 제조된 먹이 또는 인공 먹이 (제제화된 사료)를 개발하고 사용할 수 있게 되었다. 그러나, 일반적으로는 어류 양식용 사료에 포함된 다양한 영양분의 통상적인 비율은 (건조 먹이에 대한 백분율(%)로서 표현할 때) 32 내지 45%의 단백질, 4 내지 28%의 지방 (이 중 1 내지 2% 이상이 ω-3 및/또는 ω-6 PUFA임), 10 내지 30%의 탄수화물, 1.0 내지 2.5%의 미네랄 및 1.0 내지 2.5%의 비타민을 포함한다. 상기 제제에는 각종 다른 성분들이 임의로 첨가될 수 있다. 이들로는 (1) 살과 비늘의 색 각각을 향상시키기 위한 카로테노이드, 특히 연어류 및 관상용 "수족관" 어류를 위한 카로테노이드, (2) 펠렛에 안정성을 제공하고 영양분이 물로 침출되는 것을 저하시키기 위한 결합제 (예컨대, 소 심장, 전분, 셀룰로스, 펙틴, 젤라틴, 아라비아 고무, 로커스트 빈, 한천, 카르기닌 및 다른 알기네이트), (3) 어류 먹이의 저장 수명을 연장시키고 지방의 악취를 저하시키기 위한 보존제, 예컨대 항미생물제 및 항산화제

(예컨대, 비타민 E, 부틸화 히드록시아니솔, 부틸화 히드록시톨루엔, 에톡시퀸, 및 프로피온산, 벤조산 또는 소르브산의 나트륨염 및 칼륨염), (4) 사료의 맛과 그의 섭취를 향상시키기 위한 화학주성인자 및 향미료, 및 (5) 기타 먹이. 이들 다른 먹이는 먹이의 영양적 가치를 향상시키고 어류에 의한 섭취를 증가시키기 위해서 섬유 및 회분 (각각, 충전재 및 칼슘과 인의 공급원으로 사용하기 위함)과 같은 물질 및 식물성 물질 및/또는 어류 가루 또는 오징어 가루 (예컨대, 살아있거나, 냉동된 것이거나 또는 건조된 조류, 브라인 슈림프, 담륜충 또는 기타 동물성 플랑크톤) 등의 물질을 포함할 수 있다. 문헌 [Nutrient Requirements of Fish (National Research Council, National Academy: Washington D. C, 1993)]은 어류를 위한 필수 영양분 및 다양한 성분들의 영양분 함량에 대한 상세한 설명을 제공한다.

양식사료 제제의 제조에는 각종 인자를 고려할 것이 요구되는데, 완전한 먹이는 영양적으로 균형을 이루어야 하고, 맛이 좋아야 하며, 물에 안정적이어야 하고, 적당한 크기와 조직을 가져야 하기 때문이다. 양식사료의 영양분 조성과 관련하여서는 문헌 ([Handbook on Ingredients for Aquaculture Feeds (Hertrampf, J. W. and F. Piedad-Pascual. Kluwer Academic: Dordrecht, The Netherlands, 2000)] 및 [Standard Methods for the Nutrition and Feeding of Farmed Fish and Shrimp (Tacon, A. G. J. Argent Laboratories: Redmond, 1990)])을 참조한다. 일반적으로, 사료는 건조 (즉, 최종 수분 함량이 6 내지 10%), 반-습식 (즉, 35% 내지 40% 물 함량) 또는 습식 (즉, 50 내지 70% 물 함량) 상태가 되도록 제제화된다. 건조 사료는 건조 성분들이 단순히 섞여 있는 혼합물 (즉, "매쉬(mash)" 또는 "가루"), 압착형 펠렛, 부스러기 또는 과립, 및 플레이크를 포함한다. 어류의 먹이 공급 요건에 따라, 펠렛은 가라앉도록 제조될 수도 있고 또는 떠오르도록 제조될 수도 있다. 반-습식 사료 및 습식 사료는 단일 성분 또는 혼합 성분들 (예컨대, 잡어(trash fish) 또는 요리된 콩류)로 제조되고, 케이크나 볼로 성형될 수 있다.

여기서, 대부분의 양식용 사료에는 고농도의 DHA를 생산하는 본 발명의 유전자 조작된 야로위아 리폴리티카 균주를 포함시키는 것이 특히 유용할 것임은 명백하다. 필요한 ω-6 PUFA를 제공하는 것 이외에도, 상기 효모 그 자체는 제제의 맛을 증가시킬 수 있는 유용한 단백질 공급원이다. 별법의 실시양태에서, 본 발명의 야로위아 리폴리티카 균주에 의해 생산되는 오일은 상기 세포 덩어리로부터 추출 및 정제된 후에 양식용 사료 제제에 직접 도입될 수 있다.

바람직한 실시양태의 기재

본 발명은 유질 효모인 야로위아 리폴리티카가 전체 지질 분획 중 최대 5.6%로 DHA를 합성함을 입증한다. 도 5에 나타난 바와 같이, 다양한 유전자를 야생형 ATCC #20362 야로위아 리폴리티카에 통합하여 수많은 야로위아 리폴리티카 균주를 생성하였고, 각각의 형질전환체 균주는 PUFA (DHA 포함)를 상이한 양으로 생산할 수 있었다. 몇몇 대표적인 형질전환체 생물의 완전한 지질 프로파일을 하기 표 10에 나타내었다. 엘론가제는 다음과 같은 약어 명명법을 이용하여 표시한다: C_{18}EL1은 고 친화성 $C_{18/20}$ 엘론가제 (ELO1)를 지칭하고, C_{18}EL2는 $C_{18/20}$ 엘론가제 (ELO2)를 지칭하고, C_{16}EL은 $C_{16/18}$ 엘론가제를 지칭하며, C_{20}EL은 $C_{20/22}$ 엘론가제를 지칭한다. 지방산은 16:0, 16:1, 18:0, 18:1 (올레산), 18:2 (LA), GLA, DGLA, ARA, ETA, EPA, DPA 및 DHA로 표시하였고, 각각의 조성은 전체 지방산의 백분율(%)로 제시하였다. "지질 % dcw"는 건조 세포 중량으로 측정한, 세포 중의 지질의 백분율(%)을 나타낸다.

[표 10]

ω-6 Δ6 데새투라제/Δ6 엘론가제 경로를 발현하는 대표적인 야로위아 리폴리티카 균주의 지질 프로파일

균주	부가된 유전자 수										유전자	지방산 함량										지질
	Δ12	Δ6	C_{18} EL1	C_{18} EL2	Δ5	Δ17	C_{16} EL	Δ4	C_{20} EL		KOs	16:0	16:1	18:0	18:1	18:2	GLA	DGLA	ARA	ETA	EPA	DPA	DHA	% dcw
M4	1	1	1	1	--	--	--	--	--		*Ura*	15	4	2	5	27	35	8	0	0	0	--	--	--
EU1	1	1	1	1	2	3	--	--	--		*Aco3 Ura*	19	10.3	2.3	15.8	12	18.7	5.7	0.2	3	10.3	--	--	36
Y2072	2	1	2	1	4	3	1	--	--		*Aco3 Δ12*	7.6	4.1	2.2	16.8	13.9	27.8	3.7	1.7	2.2	15	--	--	--
Y2089	3	2	3	1	5	3	2	--	--		"	7.9	3.4	2.5	9.9	14.3	37.5	2.5	1.8	1.6	17.6	--	--	--
Y2098	3	2	3	1	5	3	2	--	--		"	8.4	1.3	2	7.2	11	24.6	4.3	1.2	2.4	25.6	--	--	--
Y3000	3	2	3	1	5	3	3	1	1		"	5.9	1.2	5.5	7.7	11.7	30.1	2.6	1.2	1.2	4.7	18.3	5.6	--

상기 표에서 보인 바와 같이, DHA 생산 균주는 하기 기재된 유전자 변형을 포함하였다 (이들에 대한 더 상세한 사항은 실시예에 제공되어 있음):

(1) FBA::F.Δ12::LIP2 및 TEF::F.Δ12::PEX16 키메라 유전자 내 푸사륨 모닐리포르메 Δ12 데새투라제의 2개 카피의 발현,

(2) FBAIN::M.D12::PEX20 키메라 유전자 내 모르티에렐라 이사벨리나 Δ12 데새투라제의 1개 카피의 발현,

(3) TEF::Δ6S::LIP1 및 FBAIN::Δ6S::LIP1 키메라 유전자 내 모르티에렐라 알피나 Δ6 데새투라제로부터 유래된 합성 Δ6 데새투라제 유전자 (야로위아 리폴리티카에서의 발현을 위해 코돈-최적화됨)의 2개 카피의 발현,

(4) FBAIN::MAΔ5S::PEX20 및 TEF::MAΔ5S::LIP1 키메라 유전자 내 모르티에렐라 알피나 Δ5 데새투라제의 2개 카피의 발현,

(5) TEF::H.D5S::PEX16 및 GPAT::H.D5S::PEX20 키메라 유전자 내 호모 사피엔스 Δ5 데새투라제로부터 유래된 합성 Δ5 데새투라제 유전자 (야로위아 리폴리티카에서의 발현을 위해 코돈-최적화됨)의 2개 카피의 발현,

(6) TEF::I.D5S::PEX20 키메라 유전자 내 이소크리시스 갈바나 Δ5 데새투라제로부터 유래된 합성 Δ5 데새투라제 유전자 (야로위아 리폴리티카에서의 발현을 위해 코돈-최적화됨)의 1개 카피의 발현,

(7) FBAIN::Δ17S::LIP2, TEF::Δ17S::PEX20 및 FBAINm::Δ17S::PEX16 키메라 유전자 내 사프롤레그니아 디클리나 Δ17 데새투라제로부터 유래된 합성 Δ17 데새투라제 유전자 (야로위아 리폴리티카에서의 발현을 위해 코돈-최적화됨)의 3개 카피의 발현,

(8) FBAIN::EL1S::PEX20, GPAT::EL1S::XPR 및 GPDIN::EL1S::LIP2 키메라 유전자 내 모르티에렐라 알피나 고친화도 엘론가제로부터 유래된 합성 고친화도 $C_{18/20}$ 엘론가제 유전자 (야로위아 리폴리티카에서의 발현을 위해 코돈-최적화됨)의 3개 카피의 발현,

(9) TEF::EL2S::XPR 키메라 유전자 내 트라우스토키트륨 아우레움 엘론가제로부터 유래된 합성 $C_{18/20}$ 엘론가제 유전자 (야로위아 리폴리티카에서의 발현을 위해 코돈-최적화됨)의 1개 카피의 발현,

(10) TEF::rELO2S::PEX20 키메라 유전자 내 라투스 노르베기쿠스 rELO 유전자로부터 유래된 합성 $C_{16/18}$ 엘론가제 유전자 (야로위아 리폴리티카에서의 발현을 위해 코돈-최적화됨)의 3개 카피의 발현,

(11) YAT1::Δ4S::Pex16 키메라 유전자 내 트라우스토키트륨 아우레움 Δ4 데새투라제 유전자로부터 유래된 합성 Δ4 데새투라제 유전자 (야로위아 리폴리티카에서의 발현을 위해 코돈-최적화됨)의 1개 카피의 발현,

(12) FBAIN::OtE2S::Oct 키메라 유전자 내 오스트레오코쿠스 타우리 $C_{20/22}$ 엘론가제 유전자로부터 유래된 합성 $C_{20/22}$ 엘론가제 유전자 (야로위아 리폴리티카에서의 발현을 위해 코돈-최적화됨)의 1개 카피의 발현,

(13) 오로티딘-5'-포스페이트 데카르복실라제를 코딩하는 천연 야로위아 리폴리티카 Ura3 유전자의 파괴,

(14) 이소프로필 말레이트 데히드로게나제를 코딩하는 천연 야로위아 리폴리티카 Leu2 유전자의 파괴,

(15) 아실-coA 옥시다제를 코딩하는 천연 야로위아 리폴리티카 Pox3 유전자의 파괴,

(16) Δ12 데새투라제를 코딩하는 천연 야로위아 리폴리티카 유전자의 파괴,

(17) 리파제 1을 코딩하는 천연 야로위아 리폴리티카 Lip1 유전자의 파괴, 및

(18) 아실-CoA 옥시다제를 코딩하는 천연 야로위아 리폴리티카 Pox2 유전자의 파괴.

본 발명자들은 야로위아 리폴리티카의 상기 특정 재조합 균주에서 5.6% DHA (여기서, 총 지방산에 대한 C22 지방산의 총 누적량(%)은 약 24%임)의 생산을 입증하였지만, 숙주 세포에서 DHA의 농도는 본 발명에 있어서 추가의 유전자 변형에 의해 극적으로 증가될 수 있는 것으로 고려된다. 추가로, 본원에서 기재한 교시 내용 및 결과를 기초로 하여, 유질 효모를 각종 ω-3 및/또는 ω-6 PUFA의 합성을 위한 생성 플랫폼으로서 사용하고, ω-6 Δ6 데새투라제/Δ6 엘론가제 경로 및/또는 ω-3 Δ6 데새투라제/Δ6 엘론가제 경로 및/또는 ω-6 Δ9 엘론가제/Δ8 데새투라제 경로 및/또는 ω-3 Δ9 엘론가제/Δ8 데새투라제 경로를 이용하여 생성되는 실행가능성 및 상업적 유용성을 인지할 것이라 예측된다.

실시예

본 발명은 하기 실시예에서 추가로 설명된다. 이들 실시예가 본 발명의 바람직한 실시양태를 나타내는 것이긴 하지만 이는 오직 예시하기 위한 것임을 이해해야 한다. 상기 논의 및 이들 실시예로부터, 당업자는 본 발명의 본질적 특성을 본 발명의 사상과 범위에서 벗어나지 않고 파악할 수 있을 것이고 본 발명에 각종 변화와 변형을 가하여 여러 용도 및 조건에 적합하게 할 수 있다.

통상의 방법

실시예에서 사용된 표준 재조합 DNA 및 분자 클로닝 기술은 당업계에 공지되어 있으며 하기 문헌에 기재되어 있다:

1.) Sambrook, J., Fritsch, E. F. and Maniatis, T. *Molecular Cloning: A Laboratory Manual*; Cold Spring Harbor Laboratory: Cold Spring Harbor, NY (1989) (Maniatis), 2.) T. J. Silhavy, M. L. Bennan, and L. W. Enquist, Experiments with Gene Fusions; Cold Spring Harbor Laboratory: Cold Spring Harbor, NY (1984), 및 3.) Ausubel, F. M. et al., Current Protocols in Molecular Biology, published by Greene Publishing Assoc. and Wiley-Interscience (1987).

미생물 배양물의 유지 및 성장에 적합한 재료 및 방법은 당업계에 공지되어 있다. 하기 예에서 사용하기에 적합한 기술은 문헌 [Manual of Methods for General Bacteriology (Phillipp Gerhardt, R. G. E. Murray, Ralph N. Costilow, Eugene W. Nester, Willis A. Wood, Noel R. Krieg and G. Briggs Phillips, Eds), American Society for Microbiology: Washington, D. C. (1994)] 또는 문헌 [Thomas D. Brock in Biotechnology: A Textbook of Industrial Microbiology, 2nd ed., Sinauer Associates: Sunderland, MA (1989)] 등에서 찾을 수 있다. 달리 명시하지 않는 한, 미생물 세포의 성장 및 유지에 사용되는 모든 시약, 제한 효소 및 재료는 알드리치 케미칼스(Aldrich Chemicals) (미국 위스콘신주 밀워키 소재), 디프코 래버러토리즈 (미국 미시간주 디트로이트 소재), GIBCO/BRL (미국 메릴랜드주 가이터스버그 소재) 또는 시그마 케미칼 컴퍼니(Sigma Chemical Company) (미국 미주리주 세인트 루이스 소재)로부터 얻었다.

이. 콜라이 (XL1-Blue) 감응성(competent) 세포는 스트라타진 컴퍼니(Stratagene Company) (미국 캘리포니아주 샌 디에고 소재)로부터 구입하였다. 이. 콜라이 균주는 전형적으로 37℃의 루리아 베르타니(Luria Bertani, LB) 플레이트에서 성장시켰다.

통상의 분자 클로닝은 표준 방법 [Sambrook et al., 상기 문헌]에 따라 수행하였다. 올리고뉴클레오티드는 시그마-제노시스(Sigma-Genosys) (미국 텍사스주 스프링 소재)가 합성한 것이었다. 개개의 PCR 증폭 반응은 총 부피를 50 ㎕로 하여 수행하였고, 달리 명시하지 않는 한은 PCR 완충제 (10 mM KCl, 10 mM $(NH_4)_2SO_4$, 20 mM Tris-HCl (pH 8.75), 2 mM $MgSO_4$, 0.1% 트리톤 X(Triton X)-100 함유), 100 ㎍/mL BSA (최종 농도), 200 μM씩의 각 데옥시리보뉴클레오티드 트리포스페이트, 각각의 프라이머 10 pmol 및 Pfu DNA 폴리머라제 (스트라타진, 미국 캘리포니아주 샌 디에고 소재) 1 ㎕를 포함하였다. 부위-지정 돌연변이유발은 스트라타진의 퀵체인지(QuickChange)™ 부위-지정 돌연변이유발 키트를 제조사의 지시에 따라 사용하여 수행하였다. PCR 또는 부위-지정 돌연변이유발이 서브클로닝에 이용된 경우에는 상기 구축물에 대하여 서열결정을 수행하여 서열에 오류가 없음을 확인하였다. PCR 생성물을 프로메가(Promega) (미국 위스콘신주 매디슨 소재)의 pGEM-T-이지(pGEM-T-easy) 벡터에 클로닝하였다.

DNA 서열은 벡터 및 삽입물-특이적 프라이머의 조합물을 이용한 염료 종결자 기술 (미국 특허 제5,366,86호, 유럽 특허 제272,007호)을 이용하는 ABI 오토매틱 시퀀서(ABI Automatic sequencer)에서 생성하였다. 서열 교정(editing)은 시퀀서 (진 코드스 코포레이션, 미국 미시건주 앤 아버 소재)에서 수행하였다. 모든 서열은 양 방향으로 2배 이상의 범위를 포함한다. 유전자 서열의 비교는 DNASTAR 소프트웨어 (DNA 스타, 인크.(DNA Star, Inc.))를 이용하여 수행하였다. 별법으로, 유전자 서열의 조작은 제네틱스 컴퓨터 그룹 인크.(Genetics Computer Group Inc.) (위스콘신 패키지 버전 9.0, 제네틱스 컴퓨터 그룹 (GCG), 미국 위스콘신주 매디슨 소재)에서 구한 프로그램 스위트를 이용하여 수행하였다. GCG 프로그램 "Pileup"은 갭 생성 디폴트 값을 12로 하고 갭 신장 디폴트 값을 4로 하여 사용하였다. GCG "Gap" 또는 "Bestfit" 프로그램은 디폴트 갭 생성 패널티를 50으로 하고 디폴트 갭 신장 패널티를 3으로 하여 사용하였다. 달리 언급하지 않는다면, 모든 다른 경우에서도 상기한 GCG 프로그램 디폴트 파라미터를 사용하였다.

BLAST (Basic Local Alignment Search Tool, 문헌 ([Altschul, S. F., et al., J. Mol. Biol. 215:403-410 (1993)] 및 [Nucleic Acids Res. 25:3389-3402 (1997)])) 검색을 수행하여, BLAST "nr" 데이타베이스 내의 서열 (3-차원 구조 브룩하벤 프로테인 데이타 뱅크(Brookhaven Protein Data Bank), SWISS-PROT 단백질 서열 데이타베이스, EMBL 및 DDBJ 데이타베이스로부터 유래된 서열인 모든 비-풍부 진뱅크 CDS 번역부를 포함함)과 유사성을 갖는 단리된 서열을 동

정하였다. 서열은 모든 리딩 프레임에서 번역되었고, NCBI에서 제공된 BLASTX 알고리즘 [Gish, W. and States, D. J. Nature Genetics 3:266-272 (1993)]을 이용하여 상기 "nr" 데이타베이스 내의 모든 공개적으로 입수가능한 단백질 서열과의 유사성에 대해 비교하였다.

의문(query) 서열과 최대의 유사성을 갖는 서열을 정리한 BLAST 비교 결과는 동일성(%), 유사성(%) 및 기대값으로 보고된다. "동일성(%)"은 2종 단백질 사이에서 동일한 아미노산의 백분율(%)로 정의된다. "유사성(%)"은 2종 단백질 사이에서 동일하거나 보존된 아미노산의 백분율(%)로 정의된다. "기대값"은 매치수를 주어진 스코어로 표시하여 상기 매치의 통계적 유의성을 추정하는데, 해당 크기의 데이타베이스 검색시에 아주 우연하게 기대되는 값이다.

약어의 의미는 다음과 같다: "sec"은 초를 의미하고, "min"은 분을 의미하며, "h"는 시간을 의미하고, "d"는 일(日)수를 의미하고, "㎕"는 마이크로리터를 의미하고, "mL"은 밀리리터를 의미하고, "L"은 리터를 의미하고, "μM"은 마이크로몰농도를 의미하고, "mM"은 밀리몰농도를 의미하고, "M"은 몰농도를 의미하고, "mmol"은 밀리몰수를 의미하고, "μmol"은 마이크로몰수를 의미하고, "g"은 그램을 의미하고, "㎍"은 마이크로그램을 의미하고, "ng"은 나노그램을 의미하고, "U"는 유닛을 의미하고, "bp"는 염기쌍을 의미하며, "kB"는 킬로베이스를 의미한다.

야로위아 리폴리티카의 형질전환 및 배양

야로위아 리폴리티카 균주 ATCC #20362, #76982 및 #90812은 아메리칸 타입 컬쳐 콜렉션 (미국 메릴랜드주 록크빌 소재)로부터 구입하였다. 야로위아 리폴리티카 균주를 일반적으로 28℃의 YPD 아가 (1% 효모 추출물, 2% 박토펩톤, 2% 글루코스, 2% 한천)에서 성장시켰다. 별법으로, "SD" 배지는 아미노산 없이 암모늄 술페이트가 들어 있는 0.67% 효모 질소 염기 및 2% 글루코스를 포함한다.

야로위아 리폴리티카의 형질전환은 달리 언급하지 않는 한은 문헌 [Chen, D. C. et al. (Appl. Microbiol Biotechnol. 48 (2):232-235 (1997)]의 방법에 따라 수행하였다. 간략하게 설명하면, 야로위아를 YPD 플레이트에 스트리킹(streaking)하여 30℃에서 대략 18시간 동안 성장시켰다. 세포를 백금이로 여러번 크게 떠서 플레이트로부터 긁어 내고, 50% PEG (평균 MW 3350) 2.25 mL, 2 M Li 아세테이트 (pH 6.0) 0.125 mL, 2 M DTT 0.125 mL 및 전단된 연어 정자 DNA 50 ㎍를 함유하는 형질전환 완충제 1 mL 중에 재현탁시켰다. 이어서, 선형화된 플라스미드 DNA 대략 500 ng을 재현탁된 세포 100 ㎕ 중에서 인큐베이션하고, 15분 간격으로 볼텍싱 혼합하면서 39℃에서 1시간 동안 유지시켰다. 세포를 선별 배지 플레이트에 플레이팅하고, 30℃에서 2일 내지 3일 동안 유지시켰다.

형질전환체 선별을 위해, 일반적으로 SD 배지 또는 최소 배지 ("MM")를 사용하였다. MM의 조성은 다음과 같다: 암모늄 술페이트 또는 아미노산이 없는 0.17% 효모 질소 염기 (디프코 래버러토리즈, 미국 미시간주 디트로이트 소재), 2% 글루코스, 0.1% 프롤린, pH 6.1). 적절하다면, 아데닌, 루이신, 리신 및/또는 우라실의 보충제를 최종 농도 0.01%로 첨가하였다 (이로써 "MMA", "MMLe", "MMLy" 및 "MMU" 선별 배지가 생성됨. 각각 20 g/L 한천으로 제조됨).

별법으로, 형질전환체를 암모늄 술페이트 또는 아미노산이 없는 0.17% 효모 질소 염기 (디프코 래버러토리즈), 2% 글루코스, 0.1% 프롤린, 75 mg/L 우라실, 75 mg/L 유리딘, 900 mg/L FOA (자이모 리써치 코포레이션(Zymo Research Corp.), 미국 캘리포니아주 오렌지 소재) 및 20 g/L 한천을 포함하는 5-플루오로오로트산 ("FOA"; 또한, 5-플루오로우라실-6-카르복실산 1수화물) 선별 배지에서 선별하였다.

최종적으로, 유질성 조건을 촉진하기 위해 디자인된 "2-단계 성장 조건"을 위해서, 고농도 글루코스 배지 ("HGM")를 다음과 같이 제조하였다: 14 g/L KH_2PO_4, 4 g/L K_2HPO_4, 2 g/L $MgSO_4 \cdot 7H_2O$, 80 g/L 글루코스 (pH 6.5). 균주는, 하기 프로토콜에 따라 "2-단계 성장 조건"하에서 배양하였다: 먼저, 세포를 30℃에서 액체 MM 중에서 250 rpm/분으로 48시간 동안 3벌로 진탕 성장시켰다. 세포를 원심분리로 수집하고, 액체 상등액을 추출하였다. 펠렛화된 세포를 HGM 중에 재현탁하고 72시간 또는 96시간 동안 30℃에서 250 rpm/분으로 진탕 성장시켰다. 세포를 다시 원심분리로 수집하고, 액체 상등액을 추출하였다.

"2-단계 성장 조건"을 위해 사용된 변형 배지는 "SD + AA" 배지였으며, 이것은 아미노산은 없지만 암모늄 술페이트가 함유된 6.7 g 효모 질소 염기, 20 g 글루코스, 및 1× 아미노산 혼합물 (20 mg/mL 아데닌 술페이트, 20 mg/mL 우라실, 20 mg/mL L-트립토판, 20 mg/mL L-히스티딘-HCL, 20 mg/mL L-아르기닌-HCL, 20 mg/mL L-메티오닌, 30 mg/mL L-티로신, 30 mg/mL L-루이신, 30 mg/mL L-이소루이신, 30 mg/mL L-리신-HCl, 50 mg/mL L-페닐알라닌, 100 mg/mL L-글루탐산, 100 mg/mL L-아스파르트산, 150 mg/mL L-발린, 200 mg/mL L-트레오닌 및 400 mg/mL L-세린)로 구성된 것이었다.

야로위아 리폴리티카의 지방산 분석

지방산 분석을 위해, 원심분리로 세포를 수집하고, 문헌 [Bligh, E. G. & Dyer, W. J. (Can. J. Biochem. Physiol. 37:911-917 (1959)]에 기재된 바와 같이 하여 지질을 추출하였다. 지질 추출물을 나트륨 메톡시드와 에스테르교환반응시켜서 지방산 메틸 에스테르를 제조하고 [Roughan, G., and Nishida I. Arch Biochem Biophys. 276(1):38-46 (1990)], 그 후에는 30-m × 0.25 mm (i.d.) HP-INNOWAX (휴렛-팩커드(Hewlett-Packard)) 컬럼이 장착된 휴렛-팩커드 6890 GC로 분석하였다. 오븐 온도는 170℃ (25분 동안 고정)에서부터 3.5℃/분씩 185℃로 하였다.

직접적인 염기 에스테르교환반응을 위해서, 야로위아 배양물 (3 mL)을 수확하여 증류수 중에서 1회 세척하고 스피드-백(Speed-Vac)에서 진공하에 5 내지 10분 동안 건조시켰다. 상기 샘플에 나트륨 메톡시드 (1%, 100 ㎕)를 첨가한 후에 상기 샘플을 볼텍싱하고 20분 동안 흔들어 주었다. 1 M NaCl 3 방울 및 헥산 400 ㎕를 첨가한 후, 샘플을 볼텍싱하고 회전시켰다. 상층을 취하여 상기한 바와 같이 GC로 분석하였다.

실시예 1

야로위아 리폴리티카에서의 고농도 발현을 위한 프로모터의 동정

TEF, GPD, GPDIN, GPM, GPAT, FBA, FBAIN 및 YAT1 프로모터의 프로모터 활성을 조사하는 비교 연구를, 각 프로모터 및 리포터 유전자로서 β-글루쿠로니다제 (GUS)를 코딩하는 이. 콜라이 유전자를 포함하는 구축물을 합성하여 수행하였다 [Jefferson, R.A. Nature. 14(342):837-838 (1989)]. 이어서, 조직화학적 및 형광측정 검정 [Jefferson, R. A. Plant Mol. Biol. Reporter 5:387-405 (1987)] 및/또는 mRNA 정량화를 위한 실시간 PCR로 GUS 활성을 측정하였다.

키메라 프로모터::GUS::XPR 유전자를 포함하는 플라스미드의 구축

플라스미드 pY5-30 (도 6A, 서열 154)은 야로위아 자가 복제 서열 (ARS18), ColE1 플라스미드 복제 기점, 이. 콜라이에서의 선별을 위한 앰피실린-내성 유전자 (Amp^R), 야로위아에서의 선별을 위한 야로위아 LEU2 유전자, 및 키메라 TEF::GUS::XPR 유전자를 함유하였다. 상기 플라스미드를 기초로 하여, TEF 프로모터가 각종 다른 천연 야로위아 리폴리티카 프로모터로 대체된 일련의 플라스미드를 제작하였다.

하기 표 11에 나타낸 프라이머를 사용하고 주형으로서는 게놈 야로위아 리폴리티카 DNA 또는 pGEM-T-이지 벡터 (프로메가, 미국 위스콘신주 매디슨 소재)에 클로닝된 적절한 영역의 DNA를 함유하는 게놈 DNA의 단편을 사용하여 추정적 프로모터 영역을 PCR로 증폭시켰다:

[표 11]

키메라 프로모터::GUS::XPR 유전자를 포함하는 플라스미드의 구축				
프로모터	프라이머	유전자에 대한 위치	RE 부위	플라스미드 명칭
GPD	YL211, YL212 (서열 225 및 226)	gdp 유전자의 -968 bp 내지 'ATG' 번역 개시 부위 (서열 210)	SalI 및 NcoI	pYZGDG
GPDIN	YL376, YL377 (서열 227 및 228)	gdp 유전자 부근의 -973 bp 내지 +201 bp (이로써 위치 +49 bp 내지 +194 bp에 위치한 146 bp 인트론을 포함함) (서열 211)	PstI/NcoI (프로모터용), 및 PstI/SalI (벡터용)	pDMW222
GPM	YL203, YL204 (서열 229 및 230)	gmp 유전자의 -875 bp 내지 'ATG' 번역 개시 부위 (서열 212)	NcoI 및 SalI	pYZGMG
GPAT	GPAT-5-1, GPAT-5-2 (서열 231 및 232)	gpat 유전자의 -1130 bp 내지 'ATG' 번역 개시 부위 (서열 216)	SalI 및 NcoI	pYGPAT-GUS

FBA	ODMW314, YL341 (서열 233 및 234)	fba 유전자 부근의 -1001 bp 내지 -1 bp (서열 213)	NcoI 및 SalI	pDMW212
FBAIN	ODMW320, ODMW341 (서열 235 및 236)	fba 유전자 부근의 -804 bp 내지 +169 bp (이로써 +62 bp 내지 +165 bp에 위치한 102 bp 인트론을 포함함) (서열 214)	NcoI 및 SalI	pDMW214
YAT1	27203-F, 27203-R (서열 237 및 238)	yat1 유전자 부근의 -778 bp 내지 -1 bp (서열 217)	HindIII 및 SalI, 또한 NcoI 및 HindIII	pYAT-GUS
주: 'ATG' 번역 개시 코돈의 'A' 뉴클레오티드를 +1로 함.				

GPD, GPDIN, GPM, FBA 및 FBAIN에 대한 개개의 PCR 증폭 반응은 총 부피를 50 ㎕로 하여 상기 '통상의 방법' 항목에 기재한 바와 같이 수행하였다. 열순환기 조건은 95℃에서 1분, 56℃에서 30초 및 72℃에서 1분의 35 주기 후에 72℃에서 10분의 최종 신장으로 설정하였다.

GPAT 프로모터에 대한 PCR 증폭은 예비혼합된 2× PCR 용액 (다까라 바이오 인크.(TaKaRa Bio Inc.), 일본 520-2193 시가 오쯔 소재)의 1:1 희석물을 이용하여 총 부피를 50 ㎕로 하여 수행하였다. 최종 조성물은 25 mM TAPS (pH 9.3), 50 mM KCl, 2 mM $MgCl_2$, 1 mM 2-머캅토에탄올, 200 μM씩의 각 데옥시리보뉴클레오티드 트리포스페이트, 10 pmol씩의 각 프라이머, 50 ng 주형 및 1.25 U 다까라 Ex Taq™ DNA 폴리머라제 (다까라 미루스 바이오(Takara Mirus Bio), 미국 위스콘신주 매디슨 소재)를 함유하였다. 열순환기 조건은 94℃에서 2.5분, 55℃에서 30초 및 72℃에서 2.5분의 30 주기 후에 72℃에서 6분의 최종 신장으로 설정하였다.

YAT1 프로모터에 대한 PCR 증폭은 GPAT에 대하여 상기한 것과 유사한 조성물 중에서 수행하였다. 반응 혼합물을 우선 94℃로 150초 동안 가열하였다. 증폭은 94℃에서 30초, 55℃에서 30초 및 72℃에서 1분의 30 주기 후에 72℃에서 7분의 최종 신장으로 하여 수행하였다.

각 PCR 생성물을 퀴아젠(Qiagen) PCR 정제 키트로 정제한 후에 제한 효소를 사용하여 소화시키고 (표준 조건을 이용하여 상기한 표에 따라 수행함), 소화 생성물을 1% (w/v) 아가로스 중에서의 겔 전기영동으로 정제하였다. 이어서, 소화 PCR 생성물 (YAT1로부터의 생성물은 제외)을 유사하게 소화시킨 pY5-30 벡터로 라이게이션하였다. 이어서, 각 반응물의 라이게이션된 DNA를 사용하여 이. 콜라이 Top10, 이. 콜라이 DH10B 또는 이. 콜라이 DH5α을 개별적으로 형질전환시켰다. 앰피실린 (100 ㎍/mL)을 함유하는 LB 아가에서 형질전환체를 선별하였다.

YAT1은 pY5-30으로의 클로닝 이전에 추가의 조작이 필요하였다. 구체적으로, YAT1 PCR 생성물을 HindIII 및 SalI으로 소화시키면 약 600 bp 단편이 생성되었고, NcoI 및 HindIII으로 소화시키면 약 200 bp 단편이 생성되었다. 두가지 생성물 모두를 단리하여 정제하였다. 이어서, 플라스미드 pYGPAT-GUS를 SalI 및 NcoI으로 소화시키고, 약 9.5 kB 단편을 단리 및 정제하였다. 상기 3가지 DNA 단편을 함께 라이게이션하여 pYAT-GUS를 생성하였다.

각 형질전환 반응물의 플라스미드 DNA를 분석하여 예측된 플라스미드의 존재를 확인하였다. 이들 플라스미드를 다음과 같이 칭하였다: pYZGDG (GPD::GUS::XPR 키메라 유전자를 포함함), pDMW222 (GPDIN::GUS::XPR 키메라 유전자를 포함함), pYZGMG (GPM::GUS::XPR 키메라 유전자를 포함함), pYGPAT-GUS (GPAT::GUS::XPR 키메라 유전자를 포함함), pDMW212 (FBA::GUS::XPR 키메라 유전자를 포함함), pDMW214 (FBAIN::GUS::XPR 키메라 유전자를 포함함) 및 pYAT-GUS (YAT1::GUS::XPR 키메라 유전자를 포함함).

상기 '통상의 방법' 항목에 기재한 바와 같이, 상기 각 플라스미드 및 추가의 플라스미드 pY5-30 (TEF::GUS::XPR 키메라 유전자를 포함함)를 개별적으로 야로위아 리폴리티카에 형질전환시켰다. 야로위아 리폴리티카 숙주는 야로위아 리폴리티카 ATCC #76982 또는 야로위아 리폴리티카 ATCC #20362, 균주 Y2034 (하기함 [실시예 13], Δ6 데새투라제/Δ6 엘론가제 경로를 통해 ARA를 10%로 생산할 수 있음)였다. 형질전환된 모든 세포를 루이신이 없는 최소 배지 플레이트에 플레이팅하여 30℃에서 2 내지 3일 동안 유지시켰다.

GUS 발현의 조직화학적 분석에 의한 야로위아 프로모터들의 비교 분석

플라스미드 pY5-30, pYZGDG, pYZGMG, pDMW212 및 pDMW214를 함유하는 야로위아 리폴리티카 ATCC #76982 균주들을 MM 3 mL 중의 단일 콜로니로부터 30℃에서 OD_{600} 약 1.0으로 성장시켰다. 이어서, 세포 100 ㎕를 원심분리로 수집하여, 조직화학적 염색 완충제 100 ㎕ 중에 재현탁하고 30℃에서 인큐베이션하였다. 염색 완충제는 5-브로모-4-클로로-3-인돌릴 글루쿠로나이드 (X-Gluc) 5 mg을 디메틸 포름아미드 50 ㎕ 중에 용해한 후에 50 mM $NaPO_4$ (pH 7.0) 5 mL를 가하여 제조하였다. 조직화학 염색 결과 (도 6B)는 구축물 pY5-30 중의 TEF 프로모터, 구축물 pYZGDG 중의 GPD 프로모터, 구축물 pYZGMG 중의 GPM 프로모터, 구축물 pDMW212 중의 FBA 프로모터, 및 구축물 pDMW214 중의 FBAIN 프로모터 모두가 활성임을 보여준다. FBA 프로모터와 FBAIN 프로모터 둘다가 모든 다른 프로모터보다 훨씬 강력하며, FBAIN 프로모터가 가장 강력한 프로모터 활성을 보유하는 것으로 여겨졌다.

별도의 실험에서, 플라스미드 pY5-30, pYGPAT-GUS, pYAT-GUS 및 pDMW214를 함유하는 야로위아 리폴리티카 Y2034 균주들을 SD 배지 5 mL 중의 단일 콜로니로부터 30℃에서 24시간 동안 OD_{600} 약 8.0으로 성장시켰다. 이어서, 세포 1 mL를 원심분리로 수집하였다. 나머지 배양물을 원심분리하여 HGM으로 2회 세척하고, 각각 5 mL씩의 HGM 중에 재현탁하고, 30℃에서 추가로 성장시켰다. 24시간 및 120시간 후에, 각 배양물 약 0.25 mL씩을 원심분리하여 세포를 수집하였다. 세포 샘플을 조직화학적 염색 완충제 (상기) 100 ㎕ 중에 개별적으로 재현탁하였다. 각각에 자이몰라제 (Zymolase) 2OT (1 mg/mL 5 ㎕; ICN 바이오메디칼스(ICN Biomedicals), 미국 캘리포니아주 코스타 메사 소재)를 가하고, 혼합물을 30℃에서 인큐베이션하였다.

조직화학적 염색 결과는, SD 배지 중에서 24시간 동안 성장시킨 경우, 구축물 pYAT-GUS 중의 YAT1 프로모터가 활성인 것처럼 구축물 pYGPAT-GUS 중의 GPAT 프로모터가 활성임을 보여준다 (도 6C, "SD 배지 중 24시간"). 비교하면, GPAT 프로모터가 TEF 프로모터보다는 훨씬 더 강력하고 FBAIN 프로모터에 비하여는 활성이 낮다고 여겨졌다. 마찬가지로, 세포를 SD 배지 중에서 24시간 동안 성장시킨 경우, YAT1 프로모터는 TEF 프로모터보다는 더 강력하였으나 FBAIN 프로모터 및 GPAT 프로모터보다는 유의하게 더 약한 것으로 여겨졌다. 그러나 더 흥미롭게도, HGM 중에서 24시간 동안 성장시킨 세포 중에서 YAT1 프로모터는 GPAT 프로모터보다 더 강력하고 FBAIN 프로모터에 필적하는 것으로 여겨졌다 (도 6C, "HG 배지 중 24시간"). 이것은 HGM 중에서 120시간이 지난 후에도 마찬가지였다 (도 6C, "HG 배지 중 120시간"). 따라서, YAT1 프로모터는 질소 제한으로 인해 유질 성장 조건을 촉진시키는 배지인 HGM 중에서 유도되는 것으로 여겨졌다.

<u>GUS 발현의 형광측정 검정에 의한 야로위아 프로모터의 비교 분석</u>

또한, GUS 활성을, 상응하는 기질 β-글루쿠로나이드로부터의 4-메틸룸벨리페론 (4-MU) 생성을 형광 측정함으로써 검정하였다 [Jefferson, R. A. Plant Mol. Biol. Reporter 5:387-405 (1987)].

플라스미드 pY5-30, pYZGDG, pYZGMG, pDMW212 및 pDMW214를 함유하는 야로위아 리폴리티카 ATCC #76982의 균주들을 3 mL MM (상기한 바와 같음) 중의 단일 콜로니로부터 30℃에서 OD_{600} 약 1.0으로 성장시켰다. 이어서, 3 mL 배양물 각각을 50 mL MM을 함유하는 500 mL 플라스크에 첨가하고 30℃의 진탕 인큐베이터 중에서 약 24시간 동안 성장시켰다. 세포를 원심분리로 수집하여 프로메가 세포 용해 완충제 중에 재현탁하고 Bio101 바이오풀버라이저 (BioPulverizer) 시스템 (비스타(Vista), 미국 캘리포니아주 소재)을 사용하여 용해시켰다. 원심분리 후, 상등액을 취하여 얼음상에 보관하였다.

유사하게, 플라스미드 pY5-30, pYAT-GUS, pYGPAT-GUS 및 pDMW214 구축물 각각을 함유하는 야로위아 리폴리티카 균주 Y2034를 10 mL SD 배지 중의 단일 콜로니로부터 30℃에서 OD_{600} 약 5.0으로 48시간 동안 성장시켰다. 각 배양물 2 mL를 하기하는 GUS 활성 검정을 위해 수집하였고, 각 배양물 5 mL는 HGM으로 변경했다.

구체적으로, 5 mL 분취액으로부터의 세포를 원심분리로 수집하고, HGM 5 mL로 1회 세척하여 HGM 중에 재현탁하였다. 이어서, HGM 중의 배양물을 30℃의 인큐베이터에서 24시간 동안 진탕 성장시켰다. GUS 활성 검정을 위해 각 HGM 배양물 2 mL씩을 수집하였고, 나머지 배양물은 추가의 96시간 동안 더 성장시킨 후에 검정을 위해서 각 배양물을 2 mL씩 더 수집하였다.

SD 배지 중 2 mL씩의 배양물 샘플을 0.5× 세포 배양물 용해 시약 (프로메가) 1 mL 중에 재현탁하였다. 재현탁된 세포를 고무 O-링이 장착된 2.0 mL 스크류 캡 튜브 중에서 유리 비드 (0.5 mm 직경) 0.6 mL와 혼합하였다. 이어서, 상기 세포를

바이오스펙 미니 비드비터(Biospec mini beadbeater) (미국 오클라호마주 바르틀레스빌 소재)에서 90초 동안 가장 높은 셋팅으로 균질화했다. 균질화 혼합물을 에펜도르프(Eppendorf) 원심분리기에서 2분 동안 14,000 rpm으로 원심분리하여 세포 파편 및 비드를 제거하였다. 상등액을 이용하여 GUS 검정 및 단백질 결정을 수행하였다.

각각의 형광측정 검정을 위해, 추출물 100 ㎕를 GUS 검정 완충제 (추출 완충제 중 2 mM 4-메틸룸벨리페릴-β-D-글루쿠로나이드 ("MUG")) 700 ㎕에 첨가하거나, 또는 추출물 200 ㎕를 GUS 검정 완충제 800 ㎕에 첨가하였다. 상기 혼합물을 37℃로 유지시켰다. 0분, 30분 및 60분 시점에서 분취액을 100 ㎕씩 취하여 중지 완충제 (1 M Na_2CO_3) 900 ㎕에 첨가하였다. 각 시점에 여기 파장은 360 nm로 하고 방출 파장은 455 nm로 하여 설정된 사이토플루어 시리즈 4000 플루오레슨스 멀티-웰 플레이트 리더(CytoFluor Series 4000 Fluorescence Multi-Well Plate Reader) (퍼셉티브 바이오시스템즈(PerSeptive Biosystems), 미국 매사추세츠주 프레이밍햄 소재)를 이용하여 판독하였다. 각 샘플의 총 단백질 농도를, 추출물 10 ㎕와 바이오라드 브래드포드(BioRad Bradford) 시약 200 ㎕를 사용하거나 또는 추출물 20 ㎕와 바이오라드 브래드포드 시약 980 ㎕를 사용하여 결정하였다 [Bradford, M. M. Anal. Biochem. 72:248-254 (1976)]. GUS 활성은 단백질 1 mg 당 1분 당의 4-MU 나노몰(nmol)로 표시하였다.

야로위아 리폴리티카 ATCC #76982 균주에서 TEF, GPD, GPM, FBA 및 FBAIN 프로모터를 비교하도록 디자인된 이들 형광측정 검정의 결과를 도 7A에 나타내었다. 구체적으로, FBA 프로모터는 야로위아 리폴리티카에서 GPD 프로모터보다 2.2배 더 강력하였다. 추가로, FBAIN 프로모터의 GUS 활성은 GPD 프로모터보다 약 6.6배 더 강력하였다.

야로위아 리폴리티카 균주 Y2034에서 TEF, GPAT, YAT1 및 FBAIN 프로모터를 비교하도록 디자인된 이들 형광측정 검정의 결과를 하기 표에 나타내었다:

[표 12]

다양한 성장 조건하에서 TEF, FBAIN. YAT1 및 GPAT 프로모터의 활성 비교				
배양 조건	프로모터			
	TEF	FBAIN	YAT1	GPAT
48시간, SD	0.401	43.333	0.536	5.252
24 시간, HGM	0.942	30.694	19.154	2.969
120시간, HGM	0.466	17.200	13.400	3.050

세포 추출물의 GUS 활성을 기준으로 YAT1 프로모터의 활성을 정량한 상기 데이타에 기초할 때, YAT1 프로모터의 활성은 세포를 SD 배지에서 HGM으로 변경하여 24시간 동안 성장시킬 경우에 약 37배 증가하였다. HGM 중에서 120시간이 지난 후, 상기 활성은 다소 감소하였으나 SD 배지 중에서의 활성에 비해 여전히 25배 더 높았다. 반대로, FBAIN 프로모터 및 GPAT 프로모터의 활성은 SD 배지에서 HGM으로 변경하여 24시간 동안 성장시킬 경우에 각각 30% 및 40% 감소하였다. TEF 프로모터의 활성은 HGM 중에서 24시간이 지난 후에 2.3배 증가하였다. 따라서, YAT1 프로모터는 유질 조건하에서 유도가능하다.

<u>GUS 발현의 정량적 PCR 분석에 의한 야로위아 프로모터의 비교 분석</u>

TEF, GPD, GPDIN, FBA 및 FBAIN 프로모터의 전사 활성을 pY5-30, pYZGDG, pDMW222, pDMW212 및 pDMW214 구축물을 함유하는 야로위아 리폴리티카 중에서 정량적 PCR 분석으로 측정하였다. 이를 위해서는 RNA의 단리 및 실시간 RT-PCR이 필요했다.

보다 구체적으로, pY5-30, pYZGDG, pDMW222, pDMW212 및 pDMW214를 함유하는 야로위아 리폴리티카 ATCC # 76982 균주를 25 mL 에를렌마이어(Erlenmeyer) 플라스크 중의 MM 6 mL에서 16시간 동안 30℃에서 단일 콜로니로부터 성장시켰다. 이어서, 6 mL 출발 배양물 각각을 HGM 140 mL를 함유하는 개개의 500 mL 플라스크에 첨가하고 30℃에서 4일 동안 인큐베이션시켰다. 각각 24시간씩의 간격을 두고, 각 플라스크로부터 배양물을 1 mL씩 취하여 광학 밀도를 측정하였고, 27 mL를 취하여 형광측정 GUS 검정 (상기한 바와 같음)에 사용하였으며, RNA 단리를 위해 1.5 mL씩의 2개 분취액을 취하였다. RNA 단리를 위한 배양물은 원심분리하여 세포 펠렛을 생성하였다.

변형된 퀴아젠 RNeasy 미니 프로토콜(RNeasy mini protocol) (퀴아젠, 미국 캘리포니아주 샌 디에고 소재)에 따라, 야로위아 균주로부터 RNA를 단리하였다. 간략하게 설명하면, 각 샘플마다 각 시점에서 퀴아젠 완충제 RLT 340 ㎕를 사용하여 2개의 세포 펠렛 각각을 재현탁했다. 2개 튜브 각각으로부터의 완충제 RLT/세포 현탁액 혼합물을 비드 비팅(beating) 튜브 (Bio101, 미국 캘리포니아주 샌 디에고 소재)에서 합하였다. 0.5 mL 유리 비드 약 500 ㎕를 상기 튜브에 넣고, 설정치 5에서 2분 동안의 비드 비팅으로 세포를 파괴하였다 (바이오풀버라이저, Bio101 컴파니, 미국 캘리포니아주 샌 디에고 소재). 이어서, 파괴된 세포를14,000 rpm에서 1분 동안 원심분리하여 펠렛화하고, 상등액 350 ㎕를 새로운 마이크로원심분리 튜브에 옮겼다. 각각의 균질화된 용해물에 에탄올 (70%, 350 ㎕)을 첨가하였다. 완만한 혼합 후, 전체 샘플을 2 mL 수집 튜브 중의 RNeasy 미니 컬럼에 첨가하였다. 상기 샘플을 15초 동안 10,000 rpm에서 원심분리하였다. 완충제 RW1 (350 ㎕)을 RNeasy 미니 컬럼에 첨가하고, 상기 컬럼을 15초 동안 10,000 rpm에서 원심분리하여 세포를 세척하였다. 용출액은 버렸다. 퀴아젠의 DNase1 스톡(stock) 용액 (10 ㎕)을 완충제 RDD 70 ㎕에 첨가하고 완만하게 혼합하였다. 상기 전체 DNase 용액을 RNeasy 미니 컬럼에 첨가하고 실온에서 15분 동안 인큐베이션시켰다. 인큐베이션 단계 후에는 완충제 RW1 350 ㎕를 상기 미니 컬럼에 첨가하고, 컬럼을 15초 동안 10,000 rpm에서 원심분리하였다. 상기 컬럼을 완충제 RW1 700 ㎕로 2회 세척하였다. RNase-무함유 물 (50 ㎕)을 컬럼에 첨가하였다. 컬럼을 1분 동안 10,000 rpm에서 원심분리하여 RNA를 용출시켰다.

2-단계 RT-PCR 프로토콜을 이용하였고, 이때 전체 야로위아 RNA를 우선 cDNA로 전환시킨 후에 cDNA를 실시간 PCR로 분석하였다. cDNA로의 전환은 어플라이드 바이오시스템즈(Applied Biosystems)의 하이 캐퍼서티 cDNA 아르키브 (High Capacity cDNA Archive) 키트 (PN#4322171; 미국 캘리포니아주 포스터 시티 소재) 및 미디아테크, 인크. (MediaTech, Inc.)의 분자생물학 등급 물 (PN# 46-000-Con; 미국 플로리다주 홀리 힐 소재)을 이용하여 수행하였다. 야로위아로부터 얻은 전체 RNA (100 ng)를 RT 완충제 10 ㎕, 25× dNTP 4 ㎕, 10× 랜덤 헥사머(Random Hexamer) 프라이머 10 ㎕, 멀티스크라이브 리버스 트랜스크립타제(Multiscribe Reverse Transcriptase) 5 ㎕ 및 RNase 인히비터 (Inhibitor) 0.005 ㎕와 혼합하고 물을 사용하여 총 반응 부피를 100 ㎕로 맞추어 cDNA로 전환시켰다. 상기 반응물을 열순환기 중에서 10분 동안 25℃에서 인큐베이션한 후에 2시간 동안 37℃에서 인큐베이션했다. 실시간 분석 이전에는 cDNA를 -20℃에서 보관하였다.

실시간 분석은 어플라이드 바이오시스템즈의 SYBR 그린 PCR 마스터 믹스(SYBR Green PCR Master Mix) (PN# 4309155)로 수행하였다. 역전사 반응물 (2 ㎕)을 2× SYBR PCR 믹스 10 ㎕, URA (즉, 프라이머 YL-URA-16F 및 YL-URA-78R [서열 239 및 서열 240]) 또는 GUS (즉, 프라이머 GUS-767F 및 GUS-891R [서열 241 및 서열 242])를 위한 100 μM의 전방향 및 역방향 프라이머 0.2 ㎕ 및 물 7.2 ㎕에 첨가하였다. 상기 반응물을 ABI 7900 시퀀스 디텍션 시스템(Sequence Ditection System) 기기에서 10분 동안 95℃로 열주기(thermocycle) 적용시킨 후, 95℃에서 5초 및 60℃에서 1분으로 하여 40회 주기를 실시하였다. 각각의 주기 중 60℃ 신장 동안의 실시간 형광 데이타를 수집하였다.

상대적 정량화는 유저 블러틴(User Bulletin) #2 ("유전자 발현의 상대적 정량화", 어플라이드 바이오시스템즈, 2001년 10월 업데이트)에 따라 ΔΔCT 방법으로 수행했다. 상기 URA 유전자를 이용하여 GUS 발현을 표준화하였다. 표준물 유전자로서의 URA 용도를 확인하기 위해서 GUS 및 URA의 PCR 효율을 비교하였고, 이들은 각각 1.04 및 0.99인 것으로 밝혀졌다 (여기서, 1.00은 100% 효율에 해당함). PCR 효율이 둘다 100%에 가까웠기 때문에, 발현 정량화를 위한 ΔΔCT 방법의 이용과 같이 GUS 발현을 위한 표준물로서 URA의 용도가 확인되었다. 표준화된 양은 ΔCT로 표시한다.

각각의 상이한 균주 (즉, pYZGDG, pDMW222, pDMW212 및 pDMW214 구축물을 함유하는 야로위아 리폴리티카 ATCC #76982 균주) 중의 GUS mRNA를 pY5-30 (TEF::GUS)를 갖는 균주의 mRNA 수준에 대하여 정량화하였다. 따라서, 발현의 상대적 정량화는 TEF::GUS를 함유하는 균주의 mRNA 수준을 참조용 샘플로서 사용하여 계산하였다. GPD::GUS, GPDIN::GUS, FBA::GUS 및 FBAIN::GUS에 대한 표준화 값을 TEF::GUS 기준물의 표준화 값과 비교하였다. 이 양을 ΔΔCT라고 지칭한다. 이어서, ΔΔCT 값을 식 $2^{-\Delta\Delta CT}$를 이용하여 절대값으로 전환시켰다. 이 값은, 키메라 GPD::GUS, GPDIN::GUS, FBA::GUS 및 FBAIN::GUS 유전자를 포함하는 균주 중에서의 GUS mRNA 수준을 키메라 TEF::GUS 유전자를 포함하는 경우에서의 값과 비교한 증가분 (배)으로 표시한다. 이 방법을 이용하면, TEF 프로모터의 활성을 GPD, GPDIN, FBA 및 FBAIN 프로모터와 비교하는 것이 가능하였다.

각각의 GUS 키메라 유전자에 대한 mRNA의 상대적 정량화 결과는 도 7B에 나타내었다. 보다 구체적으로, 상기 검정법은 HGM에서 24시간이 지난 후에는 FBA 및 FBAIN 프로모터의 전사 활성이 TEF 프로모터보다 각각 약 3.3배 및 6배 더 강력함을 보여주었다. 유사하게, GPD 및 GPDIN 프로모터의 전사 활성은 TEF 프로모터보다 각각 약 2배 및 4.4배 더 강력

하였다. FBA::GUS, FBAIN::GUS, GPD::GUS 및 GPDIN::GUS 유전자 융합체의 전사 활성은 4일의 실험 기간에 걸쳐 감소했지만, FBAIN 및 GPDIN 프로모터의 전사 활성은 실험 마지막 날에 TEF 프로모터보다 여전히 약 3배 및 2.6배 더 강력했다.

실시예 2

야로위아 리폴리티카에서의 유전자 전사를 증가시키는데 유용한

인핸서의 동정

본 실시예의 연구는, FBAIN 및 GPDIN의 강력한 프로모터 활성 (상기 활성은 각각 FBA 및 GPD 프로모터의 활성보다 높음) 및 각 프로모터 영역 내 인트론 동정 결과를 기초로 하여, 각각의 인트론 내에 인핸서가 존재하는지 여부를 결정하기 위한 것이었다.

구체적으로, GPM::FBAIN 프로모터 융합체 및 GPM::GPDIN 프로모터 융합체로 구성된 2종의 키메라 프로모터를 생성하여, GUS 리포터 유전자의 발현을 구동하였다. 키메라 프로모터 ("성분 1" 및 "성분 2"로 구성됨)를 하기 표 13에 기재한다:

[표 13]

키메라 프로모터::GUS::XPR 유전자 내에 키메라 프로모터를 포함하는 플라스미드의 구축			
키메라 프로모터	성분 1	성분 2	플라스미드 명칭
GPM::FBAIN (서열 219)	GPM의 -1 bp 내지 -843 bp 영역	FBAIN의 +1 bp 내지 +171 bp. 여기서의 인트론은 위치 +62 bp 내지 +165 bp에 위치함.	pDMW224
GPM::GPDIN (서열 220)	GPM의 -1 bp 내지 -843 bp 영역	GPDIN의 +1 bp 내지 +198 bp. 여기서의 인트론은 위치 +49 bp 내지 +194 bp에 위치함.	pDMW225

상기 키메라 프로모터들은 이들 각각이 플라스미드 pDMW224 및 pDMW225 내에서 GUS 리포터 유전자의 발현을 구동하도록 위치시켰다.

pDMW224 및 pDMW225를 포함하는 야로위아 리폴리티카 균주 중에서의 GUS 활성을 조직화학적 검정 결과 (실시예 1에 기재한 바와 같음)에 기초하여 pY5-30, pYZGDG, pYZGMG 및 pDMW214 구축물을 포함하는 야로위아 리폴리티카 균주 중에서의 GUS 활성과 비교함으로써, GPM::FBAIN 프로모터 및 GPM::GPDIN 프로모터의 활성을 TEF, FBAIN, GPDIN 및 GPM 프로모터의 활성과 비교하였다. 앞서 측정된 바와 같이, FBAIN 프로모터가 가장 강력한 프로모터였다. 그러나, 키메라 GPM::FBAIN 프로모터 및 키메라 GPM::GPDIN 프로모터는 둘다 GPM 프로모터보다 훨씬 더 강력하였고, GPDIN 프로모터와 활성면에서 동등한 것으로 여겨졌다. 따라서, 이것은 GPDIN 프로모터와 FBAIN 프로모터 둘다에 인핸서가 존재함을 확인시켜 주었다.

당업자는 GPDIN 인트론 또는 FBAIN 인트론을 사용한 유사한 키메라 프로모터를 쉽게 구축할 수 있을 것이다.

실시예 3

숄포닐우레아 선별

야로위아의 유전적 개선은 적합한 비-항생제 선별가능한 형질전환 마커의 결여로 인해 방해를 받았다. 본 실시예는 술포닐우레아 내성을 기초로 하는, 야로위아 리폴리티카에 대한 우성의 비-항생제 마커의 개발을 기재하며, 이것은 반수체, 이배체, 이수체 또는 이형접합체일 수 있는 다른 산업상의 효모 균주에도 적용될 수 있다.

이론 및 초기 감수성 스크리닝

아세토히드록시산 신타제 (AHAS)는 분지쇄 아미노산의 생합성 경로에 공통적인 첫번째 효소이다. 이것은 술포닐우레아 및 이미다졸리논 제초제의 표적이다. 이와 같이, 미생물과 식물 모두에서 술포닐 우레아 제초제 내성이 보고된 바 있다. 예를 들어, 사카로마이세스 세레비지애에서는 AHAS에서의 단일 W586L 돌연변이가 술포닐우레아 제초제에 내성을 부여한다 ([Falco, S. C, et al., Dev. Ind. Microbiol. 30:187-194 (1989)], [Duggleby, R.G., et. al. Eur. J. Biochem. 270:2895 (2003)]).

야생형 AHAS 야로위아 리폴리티카 효소 (진뱅크 관리 번호 XP_501277) 및 사카로마이세스 세레비지애 효소 (진뱅크 관리 번호 P07342)의 아미노산 서열을 정렬했을 때, 사카로마이세스 세레비지애 효소의 위치 586에서의 Trp 아미노산 잔기는 야로위아 리폴리티카 효소의 위치 497에서의 Trp 아미노산 잔기와 동등하였다. 따라서, 야생형 세포 자체가 술포닐우레아에 감수성이 있다면 야로위아 리폴리티카 효소 중의 W497L 돌연변이가 술포닐우레아 제초제 내성을 부여할 것이라는 가설을 세웠다. 당업자에게 공지된 방법을 이용하여, 최소 배지 중 100 ㎍/mL 농도의 술포닐우레아 (클로리무론 에틸)가 야생형 야로위아 리폴리티카 균주 ATCC #20362 및 ATCC #90812의 성장을 억제하는데 충분한 것으로 결정되었다.

돌연변이체 W497L AHAS 유전자의 합성

W497L 돌연변이를 함유하는 야로위아 리폴리티카 AHAS 유전자 (서열 243)를 게놈 DNA로부터 2-단계 반응으로 생성하였다. 우선, Pfu 울트라(Pfu Ultra)™ 하이-피델러티(High-Fidelity) DNA 폴리머라제 (스트라타진, 카탈로그 # 600380) 및 프라이머 410 및 411 [서열 244 및 245]를 사용하여 AHAS 유전자의 5' 부분을 게놈 DNA로부터 증폭시켰고, 상기 유전자의 3' 부분은 프라이머 412 및 413 [서열 246 및 247]를 사용하여 유사하게 증폭시켰다. 2쌍의 프라이머를, W497L 돌연변이 (여기서의 돌연변이는 'CT'의 'TG'로의 변화임)를 함유하는 영역이 중복되도록 중복시켰다.

정확한 크기의 5' 및 3' PCR 생성물을 겔 정제하여, 2차 PCR의 주형으로 사용하였고, 이때 프라이머 414 및 415 (서열 248 및 249) 및 2가지 1차 PCR 반응에서 얻은 생성물들의 혼합물을 사용하여, 전체 돌연변이체 유전자가 증폭되었다. 이 돌연변이체 유전자는 고유의 천연 프로모터 및 종결자 서열을 보유하였다. 정확한 크기의 2차 PCR 생성물을 겔 정제하여 인-퓨전(in-fusion) 기술로 플라스미드 pY35 [키메라 TEF::푸사륨 모닐리포르메 Δ12 데새투라제 (Fm2) 유전자, 이. 콜라이 복제 기점, 박테리아 앰피실린 내성 유전자, 야로위아 Leu 2 유전자 및 야로위아 자가 복제 서열 (ARS)을 함유함. 더욱 상세한 사항에 대하여는 제WO 2005/047485호 참조]의 벡터 주쇄에 클로닝한 후, 효소 SalI/BsiWI으로 소화시켰다. 인-퓨전 반응 혼합물로 TOP10 감응성 세포 (인비트로젠(Invitrogen), 카탈로그 #C4040-10)를 형질전환시켰다. 1일 후에 LB/Amp 플레이트상에서 선별하고, 여덟개 (8개)의 콜로니를 DNA 미니프렙(miniprep)으로 분석하였다. 제한 소화를 통해 7개의 클론이 제대로 된 것임을 확인하였다. 이들 중 술포닐우레아 내성 유전자 및 또한 LEU 유전자를 함유하는 하나를 "pY57" (또는 "pY57.YI.AHAS.w4971", 도 3B)라고 명명하였다.

표준 리튬 아세테이트 방법에 의해, 야생형 야로위아 리폴리티카 균주 ATCC #90812 및 ATCC #20362를 pY57 및 '빈 (empty)' LEU로 형질전환시켰다. DNA를 포함하지 않는 ('No-DNA') 형질전환 대조군도 사용하였다. 형질전환체를 MM 또는 MM + 술포닐우레아 (SU; 100 ㎍/mL) 아가 플레이트에 플레이팅하고, 4일 동안 성장시킨 후에 콜로니의 존재 또는 부재를 평가하였다.

[표 14]

야로위아 리폴리티카에서의 AHAS 선별				
플라스미드	ATCC #90812		ATCC #20362	
	MM	MM + SU (100 ㎍/mL)	MM	MM + SU (100 ㎍/mL)
pY57	콜로니 생성	콜로니 생성	콜로니 생성	콜로니 생성
Leu 벡터 대조군	콜로니 생성	콜로니 생성 없음	콜로니 생성	콜로니 생성 없음
No DNA 대조군	콜로니 생성 없음	콜로니 생성 없음	콜로니 생성 없음	콜로니 생성 없음

상기 나타낸 결과를 기초로 할 때, AHAS W497L은 야로위아 리폴리티카 ATCC #90812와 ATCC #20362 모두에서 양효한 비-항생제 선별 마커였다. 이어서, 본 발명자들은 술포닐우레아를 150 ㎍/mL의 농도로 사용하였다. 상기 신규 마커는 외래 유전자에는 의존적이 아니지만 돌연변이체 천연 유전자에는 의존적이며, 영양요구성도 필요 없고 영양요구성을 야기하지도 않기 때문에, 야로위아 리폴리티카의 형질전환에 유리하다. 제초제는 인간과 동물에게 비독성이다.

이러한 선별 방법은, 돌연변이체 AHAS 효소가 본원에 기재한 것과 유사한 방식으로 제작된 경우에는 반수체, 이배체, 이수체 또는 이형접합체일 수 있는 다른 산업상의 효모 균주에 일반적으로 적용될 수 있다고 예측된다.

실시예 4

야로위아 리폴리티카 중의 코돈-최적화 Δ4 데새투라제 유전자의 합성 및 기능적 발현

트라우스토키트륨 아우레움의 Δ4 데새투라제 유전자 (진뱅크 관리 번호 AAN75707)의 코돈 사용방식을 WO 2004/101753에 기재된 바와 유사한 방식으로 야로위아 리폴리티카에서의 발현에 최적화시켰다. 구체적으로, 야로위아 코돈 사용방식 패턴, ATG 번역 개시 코돈 주변의 컨센서스 서열, 및 RNA 안정성의 일반 규칙 [Guhaniyogi, G. and J. Brewer, Gene 265(1-2):11-23 (2001)]에 따라, 코돈-최적화 Δ4 데새투라제 유전자는 트라우스토키트륨 아우레움 유전자 (서열 104)의 DNA 서열을 기초로 하여 디자인하였다 (서열 106 및 107). 번역 개시 부위의 변형 외에, 1545 bp 코딩 영역 중 170 bp가 변형되었고 (11%), 166개 코돈이 최적화되었다. 번역 개시 부위를 변형시키기 위해, 서열 105 (야생형)의 제2 아미노산 (T)는 코돈-최적화 Δ4 데새투라제 유전자 (서열 107)에 포함시키지 않았다.

코돈-최적화 Δ4 데세투라제 유전자의 시험관내 합성

코돈-최적화 Δ4 데새투라제 유전자를 하기와 같이 합성하였다. 먼저, 16쌍의 올리고뉴클레오티드를 디자인하여 트라우스토키트륨 아우레움 Δ4 데새투라제 유전자의 코돈-최적화 코딩 영역의 전장을 신장시켰다 (예컨대, 서열 250 내지 281에 상응하는 D4-1A, D4-1B, D4- 2A1 D4-2B, D4-3A, D4-3B, D4-4A, D4-4B, D4-5A, D4-5B, D4-6A, D4-6B, D4-7A, D4-7B, D4-8A, D4-8B, D4-9A, D4-9B, D4-10A, D4-10B, D4-11A, D4-11B, D4-12A, D4-12B, D4-13A, D4-13B, D4-14A, D4-14B, D4-15A, D4-15B, D4-16A 및 D4-16B). 센스 (A) 및 안티센스 (B) 올리고뉴클레오티드의 각 쌍은 각 5'-말단에서의 4 bp 오버행 외에는 상보적이었다. 추가로, 프라이머 D4-1F (서열 282)은 NcoI 부위를 도입하였고, 프라이머 D4-4R (서열 283) 및 D4-5F (서열 284)은 BamHI 부위를 도입하였고, 프라이머 D4-8R (서열 285) 및 D4-9F (서열 286)는 274번 아미노산을 Phe에서 Leu로 변화시킨 HindIII 부위를 도입하였고, 프라이머 D4-12R (서열 287) 및 D4-13 (서열 288)은 ApaLI 부위를 도입하였고, 프라이머 D4-16R (서열 289)은 차후의 서브클로닝을 위해 NotI 부위를 도입하였다. 이후, 프라이머 D4-8R 및 D4-9F에 의해 도입된 274번 위치의 아미노산 변화는 이하의 전체 유전자 어셈블리를 이용해서 교정하였다.

각각의 올리고뉴클레오티드 (100 ng)는 37℃에서 1시간 동안 50 mM Tris-HCl (pH 7.5), 10 mM $MgCl_2$, 10 mM DTT, 0.5 mM 스페르미딘, 0.5 mM ATP 및 10U T4 폴리뉴클레오티드 키나제를 함유한 부피 20 ㎕에서 인산화하였다. 센스 및 안티센스 올리고뉴클레오티드의 각 쌍을 혼합하고, 하기 파라미터를 이용하는 열순환기에서 어닐링시켰다: 95℃ (2분), 85℃ (2분), 65℃ (15분), 37℃ (15분), 24℃ (15분) 및 4℃ (15분). 따라서, D4-1A (서열 250)를 D4-1B (서열 251)에 어닐링시켜 이중-가닥 생성물 "D4-1AB"를 생산하였다. 유사하게, D4-2A (서열 252)를 D4-2B (서열 253)에 어닐링시켜 이중-가닥 생성물 "D4-2AB" 등을 생산하였다.

어닐링된 이중-가닥 올리고뉴클레오티드의 4개의 개별적 풀을 하기에 나타내는 바와 같이 이어서 함께 라이게이션하였다: 풀 1 (D4-1AB, D4-2AB, D4-3AB 및 D4-4AB 포함); 풀 2 (D4-5AB, D4-6AB, D4-7AB 및 D4-8AB 포함); 풀 3 (D4-9AB, D4-10AB, D4-11AB 및 D4-12AB 포함); 및 풀 4 (D4-13AB, D4-14AB, D4-15AB 및 D4- 16AB 포함). 어닐링된 올리고뉴클레오티드의 각 풀을 10U T4 DNA 리가제를 갖는 부피 20 ㎕에서 혼합하고, 라이게이션 반응물을 밤새 16℃에서 인큐베이션하였다.

각 라이게이션 반응의 생성물을 이어서 주형으로서 사용하여 디자인된 DNA 단편을 PCR로 증폭시켰다. 구체적으로, 라이게이션된 "풀 1" 혼합물 (즉, D4-1AB, D4-2AB, D4-3AB 및 D4-4AB)을 주형으로 사용하고, 올리고뉴클레오티드 D4-1F 및 D4-4R (서열 282 및 283)을 프라이머로 사용하여, 코돈-최적화 Δ4 데새투라제 유전자의 제1 부분을 PCR로 증폭시켰다. PCR 증폭은 총 부피를 50 ㎕로 하여 상기 '통상의 방법' 항목에 기재한 바와 같이 수행하였다. 증폭을 하기와 같이 수행하였다: 95℃에서 3분 동안 초기 변성, 이어서 95℃에서 1분 동안, 56℃에서 30초, 72℃에서 40초 동안의 35 주기. 72℃에서 10분 동안의 최종 신장 주기를 수행하고, 이어서 4℃에서의 반응 종결을 수행하였다. 433 bp PCR 단편을 pGEM-T 이지 벡터 (프로메가) 내로 서브클로닝하여 pT4(1-4)을 생성시켰다.

라이게이션된 "풀 2" 혼합물 (즉, D4-5AB, D4-6AB, D4-7AB 및 D4-8AB)을 주형으로 사용하고, 올리고뉴클레오티드 D4-5F 및 D4-8R (서열 284 및 285)을 프라이머로서 사용하여, 코돈-최적화 Δ4 데새투라제 유전자의 제2 부분을 PCR로 유사하게 증폭시키고, pGEM-T-이지 벡터 내로 클로닝시켜 pT4(5-8)를 생성시켰다.

라이게이션된 "풀 3" 혼합물 (즉, D4-9AB, D4-10AB, D4-11AB 및 D4-12AB)을 주형으로 사용하고, 올리고뉴클레오티드 D4-9F 및 D4-12R (서열 286 및 287)을 프라이머로서 사용하여, 코돈-최적화 Δ4 데새투라제 유전자의 제3 부분을 PCR로 유사하게 증폭시키고, pGEM-T-이지 벡터 내로 클로닝시켜 pT4(9-12)를 생성시켰다.

라이게이션된 "풀 4" 혼합물 (즉, D4-13AB, D4-14AB, D4-15AB 및 D4-16AB)을 주형으로 사용하고, 올리고뉴클레오티드 D4-13F 및 D4-16R (서열 288 및 289)을 프라이머로서 사용하여, 코돈-최적화 Δ4 데새투라제 유전자의 제4 부분을 PCR로 유사하게 증폭시키고, pGEM-T-이지 벡터 내로 클로닝시켜 pT4(13-16)를 생성시켰다.

이. 콜라이를 따로 pT4(1-4), pT4(5-8), pT4(9-12) 및 pT4(13-16)로 형질전환시키고, 플라스미드 DNA를 앰피실린-내성 형질전환체로부터 단리시켰다. 플라스미드 DNA를 정제하고, 적절한 제한 엔도뉴클레아제로 소화시켜 pT4(1-4)의 433 bp NcoI/BamHI 단편, p4(5-8)의 383 bp BamHI/HindIII 단편, p4(9-12)의 436 bp HindIII/ApaLI 단편, 및 p4(13-16)의 381 bp ApaLI/NotI 단편을 유리시켰다. 이어서 이들 4개 단편을 합하고, Nco1/Not1 소화 pZUF17 (서열 162; 도 9B)와 함께 정해진 방향으로 라이게이션시켜 pZUF4 (서열 163)을 생성시켰다.

클로닝 편의를 이유로, pZUF4에서 합성 Δ4 데새투라제 유전자 ("D4S")의 274번 아미노산을 처음부터 Phe에서 Leu로 변화시켰다. 주형으로서 pZUF4를 사용하고 올리고뉴클레오티드 YL251 및 YL252를 프라이머 (서열 290 및 291)로서 사용해서 부위-지정 돌연변이유발법에 의해 274번 아미노산인 Leu를 Phe로 교정하였다. 따라서, 생성된 플라스미드는 Δ4 데새투라제 유전자 (서열 106)의 정확한 합성 아미노산 서열을 함유했으며, pZUF4S (서열 164; 도 9C)로 표시했다.

야로위아 리폴리티카 중의 코돈-최적화 Δ4 데새투라제 유전자의 발현

키메라 FBAIN::D4S::Pex20 유전자를 포함한 구축물 pZUF4S인 자가-복제 플라스미드를 야로위아 리폴리티카 Y20362U (FOA 내성에 대해 선별함으로써 생성된, ATCC 번호 20362의 자발적 Ura- 돌연변이체) 내로 상기 '통상의 방법' 항목에 기재한 바와 같이 형질전환시켰다. 형질전환체 세포를 MM 선별 배지 플레이트에 플레이팅하고, 30 ℃에서 2 내지 3일 동안 유지하였다. MM 플레이트에서 성장한 형질전환체 (3)을 골라내서 신선한 MM 플레이트에 재스트리킹하였다. 일단 성장했으면, 이들 균주를 30 ℃에서 10 ㎍ DPA가 공급된 3 mL 액체 MM에 개별적으로 접종시키고, 250 rpm/분으로 2일 동안 진탕시켰다. 세포를 원심분리로 수집하여 지질을 추출하고 에스테르교환반응으로 지방산 메틸 에스테르를 제조한 후에 휴렛-팩커드 6890 GC로 분석하였다.

GC 결과는 이들 3개의 형질전환체에서 DHA가 약 2%로 생산됨을 보여주었다. 코돈-최적화 유전자의 "기질 전환율(%)"은 약 20%인 것으로 측정되었다.

실시예 5

야로위아 리폴리티카 중의 코돈-최적화 $C_{20/22}$ 엘론가제 유전자의 합성 및 기능적 발현

오스트레오코쿠스 타우리의 $C_{20/22}$ 엘론가제 유전자 (진뱅크 관리 번호 AY591336; "OtElo2")의 코돈 사용방식을 WO 2004/101753 및 실시예 4 (상기 문헌)에 기재된 바와 유사한 방식으로 야로위아 리폴리티카에서의 발현에 최적화시켰다.

구체적으로, 야로위아 코돈 사용방식 패턴 (WO 2004/101753), 'ATG' 번역 개시 코돈 주변의 컨센서스 서열, 및 RNA 안정성의 일반 규칙 [Guhaniyogi, G. and J. Brewer, Gene 265(1-2):11-23 (2001)]에 따라, 코돈-최적화 OtElo2 엘론가제 유전자 ("OtE2S"로 표시됨, 서열 102)는 오스트레오코쿠스 타우리의 공개된 서열 (진뱅크 관리 번호 AY591336, 서열 100)을 기초로 하여 디자인하였다. 번역 개시 부위의 변형 외에, 903 bp 코딩 영역 중 160 bp가 변형되었고 (17.7%), 147개 코돈이 최적화되었다 (49%). 코돈-최적화 ORF (서열 102 및 103) 내의 번역 개시 코돈 주변에 NcoI 부위를 부가하기 위해 제2 아미노산을 'S'에서 'A'로 변화시킨 것을 제외하면, 코돈-최적화 유전자에서의 어떤 변형도 코딩된 단백질의 아미노산 서열 (서열 101)을 변화시키지 않았다. 디자인된 OtE2S 유전자는 뉴저지주 피스카타웨이 소재의 젠스크립트 코포레이션 (GenScript Corporation)사에 의해 합성되었으며, 이를 pUC57 (진뱅크 관리 번호 Y14837)에 클로닝하여 pOtE2S를 생성시켰다.

야로위아 리폴리티카 중의 코돈-최적화 OtE2S 유전자의 발현

pOtE2S의 NcoI/NotI 단편을 단리하고, Nco1/Not1 소화 pZUF17 (서열 162; 도 9B)와 함께 라이게이션하여 pFOE2S (도 9D)를 생성시켰다. 키메라 FBAIN::OtE2S::Pex20 유전자를 포함한 구축물 pFOE2S인 자가-복제 플라스미드를 야로위아 리폴리티카 균주 Y20362U (FOA 스크리닝 하에 선별된, ATCC 번호 20362의 자발적 Ura- 돌연변이체) 내로 상기 '통상의 방법' 항목에 기재한 바와 같이 형질전환시켰다. 형질전환체 세포를 MM 선별 배지 플레이트에 플레이팅하고, 30 ℃에서 2 내지 3일 동안 유지하였다. MM 플레이트에서 성장한 pFOE2S 보유 Y20362U의 형질전환체 (3)을 골라내서 신선한 MM 플레이트에 재스트리킹하였다. 일단 성장했으면, 이들 균주를 30 ℃에서 10 ㎍ EPA가 공급된 3 mL 액체 MM에 개별적으로 접종시키고, 250 rpm/분으로 2일 동안 진탕시켰다. 세포를 원심분리로 수집하여 지질을 추출하고 에스테르교환반응으로 지방산 메틸 에스테르를 제조한 후에 휴렛-팩커드 6890 GC로 분석하였다.

GC 결과는 이들 3개의 형질전환체에서 DPA가 전체 지질의 약 1% 내지 1.3%로 존재하며 EPA가 약 0.5%로 존재함을 보여주었다. 이들 형질전환체에서 코돈-최적화 OtE2S 유전자의 EPA에서 DPA로의 "기질 전환율(%)"은 67%인 것으로 측정되었다.

실시예 6

ω-6 Δ6 데새투라제/Δ6 엘론가제 경로: DHA를 전체 지질의 5% 초과로 생산하는 균주 Y3000의 생성

본 실시예는 야로위아 리폴리티카 ATCC #20362로부터 유래되어 DHA를 전체 지질에 대하여 5.6%로 생산할 수 있는 균주 Y3000의 구축을 기재한다 (도 5). 이 균주는 ω-6 Δ6 데새투라제/Δ6 엘론가제 경로를 발현하도록 유전자 조작되었다.

균주 Y3000의 개발에는 균주 M4 (DGLA를 8%로 생성함), 균주 Y2047 (ARA를 11%로 생산함), 균주 Y2048 (EPA를 11%로 생산함), 균주 Y2060 (EPA를 13%로 생산함), 균주 Y2072 (EPA를 15%로 생산함), 균주 Y2072U3 (EPA를 16%로 생산함), 균주 Y2098 (EPA를 22%로 생산함), 및 균주 Y2098U (EPA를 21%로 생산함)의 구축이 요구되었다.

DGLA를 전체 지질의 약 8%로 생산하는 M4 균주의 생성

4개의 키메라 유전자 (Δ12 데새투라제, Δ6 데새투라제 및 2개의 $C_{18/20}$ 엘론가제 포함)를 야생형 야로위아 균주 ATCC #20362의 Ura3 유전자좌에 통합함으로써 DGLA가 생성될 수 있도록 하는 구축물 pKUNF12T6E (도 8A, 서열 156)를 제조하였다. pKUNF12T6E 플라스미드는 하기 성분을 함유하였다:

[표 15]

플라스미드 pKUNF12T6E (서열 156)의 기재	
서열 156 내의 RE 부위 및 뉴클레오티드	단편 및 키메라 유전자 성분의 기재
AscI/BsiWI (9420-8629)	야로위아 Ura3 유전자의 784 bp 5' 부분 (진뱅크 관리 번호 AJ306421)
SphI/PacI (12128-1)	야로위아 Ura3 유전자의 516 bp 3' 부분 (진뱅크 관리 번호 AJ306421)

SwaI/BsiWI (6380-8629)	· FBAIN: FBAIN 프로모터 (서열 214) · EL1S: 코돈-최적화 엘론가제 1 유전자 (서열 24) (모르티에렐라 알피나 (진뱅크 관리 번호 AX464731)로부터 유래됨) · Pex20: 야로위아 Pex20 유전자 (진뱅크 관리 번호 AF054613)로부터의 Pex20 종결자 서열 을 포함하는 FBAIN::EL1S::Pex20
BglII/SwaI (4221-6380)	· TEF: TEF 프로모터 (진뱅크 관리 번호 AF054508) · Δ6S: 코돈-최적화 Δ6 데새투라제 유전자 (서열 3) (모르티에렐라 알피나 (진뱅크 관리 번호 AF465281)로부터 유래됨) . Lip1: 야로위아 Lip1 유전자 (진뱅크 관리 번호 Z50020)로부터의 Lip1 종결자 서열 을 포함하는 TEF::Δ6S::Lip1
PmeI/ClaI (4207-1459)	· FBA:FBA 프로모터 (서열 213) · F.Δ12:푸사륨 모닐리포르메 Δ12 데새투라제 유전자 (서열 32) · Lip2: 야로위아 Lip2 유전자 (진뱅크 관리 번호 AJ012632)로부터의 Lip2 종결자 서열 을 포함하는 FBA::F.Δ12::Lip2
ClaI/PacI (1459-1)	· TEF: TEF 프로모터 (진뱅크 관리 번호 AF054508) · EL2S: 코돈-최적화 엘론가제 유전자 (서열 27) (트라우스토키트륨 아우레움 (미국 특허 제6,677,145호)으로부터 유래됨) · XPR: 야로위아 Xpr 유전자 (진뱅크 관리 번호 M17741)의 3' 영역의 약 100 bp 를 포함하는 TEF::EL2S::XPR

pKUNF12T6E 플라스미드를 AscI/SphI으로 소화시킨 후에, 이것을 상기 '통상의 방법'에 따라 사용하여 야생형 야로위아 리폴리티카 ATCC #20362를 형질전환시켰다. 형질전환 세포를 FOA 선별 배지 플레이트에 플레이팅하고, 30℃에서 2 내지 3일 동안 유지시켰다. FOA 내성 콜로니를 골라내서 MM 및 MMU 선별 플레이트에 스트리킹했다. MMU 플레이트에서는 성장할 수 있으나 MM 플레이트에서는 성장할 수 없는 콜로니를 Ura- 균주로서 선별하였다. 이어서, Ura- 균주의 단일 콜로니를 30℃에서 액체 MMU에 접종하고 250 rpm/분으로 2일 동안 진탕시켰다. 세포를 원심분리로 수집하여 지질을 추출하고 에스테르교환반응으로 지방산 메틸 에스테르를 제조한 후에 휴렛-팩커드 6890 GC로 분석하였다.

GC 분석은 DGLA가 pKUNF12T6E의 4개 키메라 유전자를 함유하는 형질전환체 중에는 존재하지만 야생형 야로위아 대조군 균주 중에는 존재하지 않음을 보여주었다. 선별된 32개의 Ura- 균주 대개는 DGLA를 전체 지질의 약 6%로 생산하였다. DGLA를 전체 지질의 약 8%로 생산하는 균주는 2종이었다 (즉, 균주 M4 및 13-8).

ARA를 전체 지질의 약 10%로 생산하는 Y2047 균주의 생성

3개의 Δ5 키메라 유전자를 야로위아 균주 M4의 Leu2 유전자에 통합하여 구축물 pDMW271 (도 8B, 서열 157)를 제조하였다. 플라스미드 pDMW271은 표 16에 나타낸 하기 성분을 함유하였다:

[표 16]

플라스미드 pDMW271 (서열 157)의 기재	
서열 157 내의 RE 부위 및 뉴클레오티드	단편 및 키메라 유전자 성분의 기재
AscI/BsiWI (5520-6315)	야로위아 Leu2 유전자 (진뱅크 관리 번호 AF260230)의 788 bp 5' 부분
SphI/PacI (2820-2109)	야로위아 Leu2 유전자 (진뱅크 관리 번호 AF260230)의 703 bp 3' 부분

SwaI/BsiWI (8960-6315)	· FBAIN: FBAIN 프로모터 (서열 214) · MAΔ5: 모르티에렐라 알피나 Δ5 데새투라제 유전자 (서열 6) (진뱅크 관리 번호 AF067654) · Pex20: 야로위아 Pex20 유전자 (진뱅크 관리 번호 AF054613)의 Pex20 종결자 서열 을 포함하는 FBAIN::MAΔ5::Pex20
SwaI/ClaI (8960-11055)	· TEF: TEF 프로모터 (진뱅크 관리 번호 AF054508) · MAΔ5: 서열 6 (상기 참조) · Lip1: 야로위아 Lip1 유전자 (진뱅크 관리 번호 Z50020)의 Lip1 종결자 서열 을 포함하는 TEF::MAΔ5::Lip1
PmeI/ClaI (12690-11055)	야로위아 Ura3 유전자 (진뱅크 관리 번호 AJ306421)
ClaI/PacI (1-2109)	· TEF: TEF 프로모터 (진뱅크 관리 번호 AF054508) · HΔ5S: 코돈-최적화 Δ5 데새투라제 유전자 (서열 13) (호모 사피엔스 (진뱅크 관리 번호 NP_037534)로부터 유래됨) · Pex16: 야로위아 Pex16 유전자의 (진뱅크 관리 번호 U75433)의 Pex16 종결자 서열 을 포함하는 TEF::HΔ5S::Pex16

플라스미드 pDMW271을 AscI/SphI로 소화시킨 후에, 이것을 상기 '통상의 방법'에 따라 사용하여 균주 M4를 형질전환시켰다. 형질전환 후, 세포를 MMLe 플레이트에 플레이팅하고, 30℃에서 2 내지 3일 동안 유지시켰다. MMLe 플레이트에서 성장된 개개의 콜로니를 골라내서 MM 및 MMLe 플레이트에 스트리킹했다. MMLe 플레이트에서는 성장할 수 있으나 MM 플레이트에서는 성장할 수 없는 콜로니를 $Leu2^-$ 균주로서 선별하였다. 이어서, $Leu2^-$ 균주의 단일 콜로니를 30℃에서 액체 MMLe 배지에 접종하고 250 rpm/분으로 2일 동안 진탕시켰다. 세포를 원심분리로 수집하여 지질을 추출하고 에스테르교환반응으로 지방산 메틸 에스테르를 제조한 후에 휴렛-팩커드 6890 GC로 분석하였다.

GC 분석은 ARA가 pDMW232 형질전환체 중에는 존재하지만 모(母) M4 균주 중에는 존재하지 않음을 보여주었다. 구체적으로, pDMW271을 함유하는 것으로 선별된 48개의 $Leu2^-$ 형질전환체 중에서 유전자 조작된 야로위아에서 ARA를 전체 지질의 5% 미만으로 생산하는 균주는 35개였고, ARA를 6 내지 8%로 생산하는 균주는 12개였으며, ARA를 약 11%로 생산하는 균주는 1개였다. ARA를 11%로 생산하는 균주를 "Y2047"로 명명하였다.

<u>EPA를 전체 지질의 약 11%로 생산하는 Y2048 균주의 생성</u>

3개의 합성 Δ17 데새투라제 키메라 유전자를 Y2047 균주의 아실-CoA 옥시다제 3 유전자에 통합하여 구축물 pZP3L37 (도 8C, 서열 158)을 제조하였다. 플라스미드 pZP3L37은 하기 성분을 함유하였다:

[표 17]

플라스미드 pZP3L37 (서열 158)의 기재	
서열 158 내의 RE 부위 및 뉴클레오티드	단편 및 키메라 유전자 성분의 기재
AscI/BsiWI (6813-6043)	야로위아 Pox3 유전자 (진뱅크 관리 번호 AJ001301)의 763 bp 5' 부분
SphI/PacI (9521-10345)	야로위아 Pox3 유전자 (진뱅크 관리 번호 AJ001301)의 818 bp 3' 부분

ClaI/BsiWI (4233-6043)	· TEF: TEF 프로모터 (진뱅크 관리 번호 AF054508) · Δ17S: 코돈-최적화 Δ17 데새투라제 유전자 (서열 21) (사프롤레그니아 디클리나로부터 유래됨) (미국 특허 출원 제2003/0196217 A1호) · Pex20: 야로위아 Pex20 유전자 (진뱅크 관리 번호 AF054613)의 Pex20 종결자 서열 을 포함하는 TEF::Δ17S::Pex20
ClaI/PmeI (4233-1811)	· FBAIN: FBAIN 프로모터 (서열 214) · Δ17S:서열 21 (상기 참조) · Lip2: 야로위아 Lip2 유전자 (진뱅크 관리 번호 AJ012632)의 Lip2 종결자 서열 을 포함하는 FBAIN::Δ17S::Lip2
PmeI/SwaI (1811-1)	야로위아 Leu2 유전자 (진뱅크 관리 번호 AF260230)
PacI/SwaI (10345-1)	· FBAINm: FBAINm 프로모터 (서열 215) · Δ17S:서열 21 (상기 참조) · Pex16: 야로위아 Pex16 유전자 (진뱅크 관리 번호 U75433)의 Pex16 종결자 서열 을 포함하는 FBAINm::Δ17S::Pex16

플라스미드 pZP3L37을 AscI/SphI으로 소화시킨 후에, 이것을 상기 '통상의 방법'에 따라 사용하여 균주 Y2047을 형질전환시켰다. 형질전환 후, 세포를 MM 플레이트에 플레이팅하고, 30℃에서 2 내지 3일 동안 유지시켰다. MM 플레이트에서 성장한 총 96개의 형질전환체를 골라내서 신선한 MM 플레이트에 재스트리킹했다. 일단 성장했으면, 이들 균주를 개별적으로 30℃에서 액체 MM에 접종하고 250 rpm/분으로 2일 동안 진탕시켰다. 세포를 원심분리로 수집하여 지질을 추출하고 에스테르교환반응으로 지방산 메틸 에스테르를 제조한 후에 휴렛-팩커드 6890 GC로 분석하였다.

GC 분석은 EPA가 pZP3L37을 함유하는 대부분의 형질전환체 중에는 존재하지만 모 Y2047 균주 중에는 존재하지 않음을 보여주었다. pZP3L37을 함유하는 것으로 선별된 96개의 형질전환체 중에서, 유전자 조작된 야로위아에서 EPA를 전체 지질의 2% 미만으로 생산하는 균주는 20개였고, EPA를 2 내지 3%로 생산하는 균주는 23개였고, EPA를 3 내지 4%로 생산하는 균주는 5개였으며, EPA를 약 6%로 생산하는 균주는 2개 (즉, 균주 번호 71 및 균주 번호 94)였다. (EPA를 6%로 생산하는) 균주 번호 71은 상기 '통상의 방법' 항목에 기재한 바와 같이 "2-단계 성장 조건" (즉, 48시간 MM, 72시간 HGM)을 이용하여 추가로 분석하였다. GC 분석은 상기 균주 번호 71이 EPA를 전체 지질의 약 11%로 생산함을 보여주었다. 이 균주를 "Y2048"로 칭하였다.

<u>EPA를 전체 지질의 약 13%로 생산하는, Ura- 표현형을 갖는 Y2060 균주의 생성</u>

균주 Y2048에서 Ura3 유전자를 파괴하기 위해, TEF::rELO2S::Pex20 키메라 유전자를 균주 Y2048의 Ura3 유전자에 통합시켜 구축물 pZKUT16 (도 8D; 서열 159)를 생성시켰다. rELO2S는 16:0을 18:0로 연장시키는 래트 간 효소 (즉, $C_{16/18}$ 엘론가제)를 코딩하는 코돈-최적화 rELO 유전자이다. 플라스미드 pZKUT16은 하기 성분을 함유하였다:

[표 18]

플라스미드 pZKUT16 (서열 159)의 기재	
서열 159 내의 RE 부위 및 뉴클레오티드	단편 및 키메라 유전자 성분의 기재
BsiWI/PacI (1-721)	야로위아 Ura3 유전자 (진뱅크 관리 번호 AJ306421)의 721 bp 5' 부분
SalI/ClaI (3565-4289)	야로위아 Ura3 유전자 (진뱅크 관리 번호 AJ306421)의 724 bp 3' 부분

ClaI/BsiWI (4289-1)	· TEF: TEF 프로모터 (진뱅크 관리 번호 AF054508) · rEL02S: 코돈-최적화 rEL02 엘론가제 유전자 (서열 85) (래트 (진뱅크 관리 번호 AB071986)로부터 유래함) · Pex20: 야로위아 Pex20 유전자 (진뱅크 관리 번호 AF054613)의 Pex20 종결자 서열 을 포함하는 TEF::rEL02S::Pex20

구체적으로, 플라스미드 pZKUT16을 SalI/PacI으로 소화시킨 후에, 이것을 상기 '통상의 방법'에 따라 사용하여 균주 Y2048을 형질전환시켰다. 형질전환 후, 세포를 MM + 5-FOA 선별 플레이트에 플레이팅하고, 30℃에서 2 내지 3일 동안 유지시켰다.

MM + 5-FOA 플레이트에서 성장한 총 40개의 형질전환체를 골라내서 MM 플레이트 및 MM + 5-FOA 플레이트에 따로 재스트리킹했다. MM + 5-FOA 플레이트에서 성장할 수 있지만 MM 플레이트에서는 성장할 수 없는 균주를 Ura- 균주로서 선별했다. 이들 40개의 Ura- 균주를 개별적으로 액체 MMU에 접종하고, 30℃에서 250 rpm/분으로 2일 동안 진탕 성장시켰다. 세포를 원심분리로 수집하여 지질을 추출하고 에스테르교환반응으로 지방산 메틸 에스테르를 제조한 후에 휴렛-팩커드 6890 GC로 분석하였다.

GC 분석은 MMU 배지에서 2일간 성장시킨 후 EPA를 전체 지질의 5% 미만으로 생산하는 균주는 14개였고, EPA를 5 내지 5.9%로 생산하는 균주는 9개였으며, EPA를 6 내지 6.9%로 생산하는 균주는 15개였고, EPA를 7 내지 8%로 생산하는 균주는 7개였음을 보여주었다. EPA를 7 내지 8%로 생산하는 균주는 상기 '통상의 방법' 항목에 기재한 바와 같은 2-단계 성장 조건 (즉, 48시간 MM, 96시간 HGM)을 이용하여 추가로 분석하였다. GC 분석은 상기 균주 모두가 EPA를 10% 초과로 생산하며 이들 중 하나는 EPA를 전체 지질의 약 13%로 생산함을 보여주었다. 이 균주를 균주 "Y2060"이라고 칭하였다.

EPA를 전체 지질의 약 15%로 생산하는 Y2072 균주의 생성

3개의 키메라 유전자 ($C_{18/20}$ 엘론가제, Δ12 데새투라제 및 Δ5 데새투라제를 포함) 및 Ura3 유전자의 클러스터를 균주 Y2060의 천연 야로위아 Δ12 데새투라제 유전자 부위에 통합시킴으로써 구축물 pKO2UM25E (도 8E; 서열 160)를 생성시켰다. 플라스미드 pKO2UM25E는 하기 성분을 함유하였다:

[표 19]

플라스미드 pKO2UM25E (서열 160)의 기재	
서열 160 내의 RE 부위 및 뉴클레오티드	단편 및 키메라 유전자 성분의 기재
HindIII/AscI (1-728)	야로위아 Δ12 데새투라제 유전자 (서열 28)의 728 bp 5' 부분
SphI/EcoRI (3436-3992)	야로위아 Δ12 데새투라제 유전자 (서열 28)의 556 bp 3' 부분
BsiWI/HindIII (10437-1)	· GPAT: GPAT 프로모터 (서열 216) · EL1S: 코돈-최적화 엘론가제 1 유전자 (서열 24) (모르티에렐라 알피나 (진뱅크 관리 번호 AX464731)로부터 유래됨) · XPR: 야로위아 Xpr 유전자 (진뱅크 관리 번호 M17741)의 3' 영역의 약 100 bp 를 포함하는 GPAT::EL1S::XPR
BglII/BsiWI (7920-10437)	· FBAIN: FBAIN 프로모터 (서열 214) · M.Δ12: 모르티에렐라 이사벨리나 Δ12 데새투라제 유전자 (진뱅크 관리 번호 AF417245; 서열 30) · Pex20: 야로위아 Pex20 유전자 (진뱅크 관리 번호 AF054613)의 Pex20 종결자 서열 을 포함하는 FBAIN::M.Δ12::Pex20

SalI/PacI (6046-7544)	야로위아 Ura3 유전자 (진뱅크 관리 번호 AJ306421)
EcoRI/SalI (3992-6046)	· TEF: TEF 프로모터 (진뱅크 관리 번호 AF054508) · I.Δ5S: 코돈-최적화 Δ5 데새투라제 유전자 (서열 10) (이소크리시스 갈바나 (WO 2002/081668)로부터 유래됨) · Pex20: 야로위아 Pex20 유전자 (진뱅크 관리 번호 AF054613)의 Pex20 종결자 서열 을 포함하는 TEF::I.Δ5S::Pex20

구체적으로, 플라스미드 pKO2UM25E를 SphI/AscI로 소화시킨 후에, 이것을 상기 '통상의 방법'에 따라 사용하여 Y2060을 형질전환시켰다. 형질전환 후, 세포를 MM 플레이트에 플레이팅하고, 30℃에서 2 내지 3일 동안 유지시켰다.

MM 플레이트에서 성장한 총 63개의 형질전환체를 골라내서 신선한 MM 플레이트에 재스트리킹했다. 일단 성장했으면, 이들 균주를 개별적으로 30℃에서 액체 MM에 접종하고 250 rpm/분으로 2일 동안 진탕 배양했다. 세포를 원심분리로 수집하여 지질을 추출하고 에스테르교환반응으로 지방산 메틸 에스테르를 제조한 후에 휴렛-팩커드 6890 GC로 분석하였다.

GC 분석은, MMU 배지 중에서 1일 동안 성장시킨 후에는 pKO2UM25E를 함유하는 거의 모든 형질전환체에 EPA가 존재함을 보여주었다. 선별된 63개의 형질전환체 중에서, EPA를 6 내지 8.9%로 생산하는 균주는 26개였으며, EPA를 9% 초과로 생산하는 균주는 46개였다. EPA를 9% 초과로 생산하는 균주는 상기 '통상의 방법'에 기재된 바와 같은 2-단계 성장 조건 (즉, 48시간 MM, 96시간 HGM)을 이용하여 추가의 분석을 위해 선별하였다. GC 분석은 선별된 46개의 균주 중 45개가 EPA를 11 내지 14.5%로 생산하지만 배양물 번호 2는 2-단계 성장 후 EPA를 전체 지질의 15.1%로 생산함을 보여주었다. 이 균주 (즉, 2번 균주)를 균주 "Y2072"라고 칭하였다.

<u>EPA를 전체 지질의 약 15 내지 16%로 생산하는, Ura- 표현형을 갖는 Y2072U3 및 Y2072U4 균주의 생성</u>

구축물 pZKUT16 (도 8D, 서열 159; 상기 참조)을 사용해서 TEF::rELO2S::Pex20 키메라 유전자를 균주 Y2072의 Ura3 유전자에 통합시켰다. 구체적으로, 상기 '통상의 방법'에 따라 SalI/PacI-소화 플라스미드 pZKUT16을 사용해서 균주 Y2072를 형질전환시켰다. 형질전환 후, 세포를 MM + 5-FOA 선별 플레이트에 플레이팅하고, 30 ℃에서 3 내지 4일 동안 유지하였다.

MM + 5-FOA 플레이트에서 성장한 총 24개의 형질전환체를 골라내서 MM 플레이트 및 MM + 5-FOA 플레이트에 따로 재스트리킹했다. MM + 5-FOA 플레이트에서 성장할 수 있지만 MM 플레이트에서는 성장할 수 없는 균주를 Ura- 균주로서 선별했다. 이들 24개의 Ura- 균주를 개별적으로 30℃에서 액체 MMU에 접종하고, 250 rpm/분으로 2일 동안 진탕 배양시켰다. 세포를 원심분리로 수집하여 지질을 추출하고 에스테르교환반응으로 지방산 메틸 에스테르를 제조한 후에 휴렛-팩커드 6890 GC로 분석하였다.

GC 분석은 MMU 배지에서 2일간 성장시킨 후 EPA를 전체 지질의 8.9% 미만으로 생산하는 균주는 14개였고, EPA를 9 내지 9.9%로 생산하는 균주는 8개였으며, EPA를 10.1%로 생산하는 균주는 1개 (즉, 12번 균주)였음을 보여주었다. 균주 번호 12 (10.1% EPA) 및 균주 번호 11 (9.6% EPA)은 2-단계 성장 절차 (즉, 48시간 MM, 96시간 HGM)을 이용하여 추가로 분석하였다. GC 분석은 균주 번호 12가 EPA를 약 15%로 생산했음을 보여주었으며, 이 균주를 "Y2072U3"으로 칭하였다. 이와는 달리, 균주 번호 11은 EPA를 약 16%로 생산하였으며, 이 균주를 균주 "Y2072U4"로 칭하였다.

<u>EPA를 전체 지질의 23 내지 28%로 생산하는 Y2096, Y2097, Y2098, Y2105 및 Y2106 균주의 생성</u>

4개의 키메라 유전자 ($C_{18/20}$ 엘론가제, Δ6 데새투라제, Δ5 데새투라제 및 Δ12 데새투라제를 포함) 및 Ura3 유전자의 클러스터를 균주 Y2072U3의 야로위아 리파제 1 유전자 부위에 통합시킴으로써 구축물 pDMW303 (도 9A; 서열 161)를 생성시켰다. 플라스미드 pDMW303은 하기 성분을 함유하였다:

[표 20]

플라스미드 pDMW303 (서열 161)의 기재	
서열 161 내의 RE 부위 및 뉴클레오티드	단편 및 키메라 유전자 성분의 기재
BsiWI/AscI (1-819)	야로위아 리파제 1 유전자 (진뱅크 관리 번호 Z50020)의 819 bp 5' 부분
SphI/PacI (35278-4297)	야로위아 리파제 1 유전자 (진뱅크 관리 번호 Z50020)의 769 bp 3' 부분
SwaI/BsiWI (13300-1)	· GPAT: GPAT 프로모터 (서열 216) · HΔ5S: 코돈-최적화 Δ5 데새투라제 유전자 (서열 13) (호모 사피엔스 (진뱅크 관리 번호 NP_037534)로부터 유래됨) · Pex20: 야로위아 Pex20 유전자 (진뱅크 관리 번호 AF054613)의 Pex20 종결자 서열 을 포함하는 GPAT::HΔ5S::Pex20
BglII/SwaI (10602-13300)	· FBAIN: FBAIN 프로모터 (서열 214) · Δ6S: 코돈-최적화 Δ6 데새투라제 유전자 (서열 3) (모르티에렐라 알피나 (진뱅크 관리 번호 AF465281)로부터 유래됨) . Lip1: 야로위아 Lip1 유전자 (진뱅크 관리 번호 Z50020)로부터의 Lip1 종결자 서열 을 포함하는 FBAIN::Δ6S::Lip1
ClaI/PmeI (8081-10558)	· GPDIN: GPDIN 프로모터 (서열 211) · EL1S: 코돈-최적화 엘론가제 1 유전자 (서열 24) (모르티에렐라 알피나 (진뱅크 관리 번호 AX464731)로부터 유래됨) · Lip2: 야로위아 리파제 2 유전자 (진뱅크 관리 번호 AJ012632)의 Lip2 종결자 를 포함하는 GPDIN::EL1S::Lip2
EcoRI/ClaI (6453-8081)	야로위아 Ura3 유전자 (진뱅크 관리 번호 AJ306421)
PacI/EcoRI (4297-6453)	· TEF: TEF 프로모터 (진뱅크 관리 번호 AF054508) · F.Δ12: 푸사륨 모닐리포르메 Δ12 데새투라제 유전자 (서열 32) · Pex16: 야로위아 Pex16 유전자 (진뱅크 관리 번호 U75433)의 Pex16 종결자 를 포함하는 TEF::F.Δ12::Pex16

구체적으로, SphI/AscI-소화 플라스미드를 상기 '통상의 방법'에 따라 사용하여 균주 Y2072U3을 형질전환시켰다. 형질전환 후, 세포를 MM 플레이트에 플레이팅하고, 30℃에서 3 내지 4일 동안 유지시켰다.

MM 플레이트에서 성장한 총 48개의 형질전환체를 골라내서 신선한 MM 플레이트에 재스트리킹했다. 일단 성장했으면, 이들 균주를 개별적으로 30℃에서 액체 MM에 접종하고 250 rpm/분으로 2일 동안 진탕 성장시켰다. 세포를 원심분리로 수집하여 지질을 추출하고 에스테르교환반응으로 지방산 메틸 에스테르를 제조한 후에 휴렛-팩커드 6890 GC로 분석하였다.

GC 분석은, MM 배지 중에서 2일 동안 성장시킨 후에는 pDMW303을 함유하는 Y2072U3의 거의 모든 형질전환체에서 EPA가 생산됨을 보여주었다. 선별된 48개의 형질전환체 중에서, EPA를 전체 지질의 13.9% 미만으로 생산하는 균주는 35개였고, EPA를 14 내지 16.9%로 생산하는 균주는 8개였으며, EPA를 17 내지 18.3%로 생산하는 균주는 4개였다.

(MM에서 2일 후) EPA를 전체 지질의 14% 초과로 생산하는 균주는 2-단계 성장 절차 (즉, 48시간 MM, 96시간 HGM)를 이용해서 추가의 분석을 위해 선별하였다. GC 분석은 상기 12개의 균주 모두가 EPA를 전체 지질의 18% 초과로 생산함을 보여주었다. 이들 중에서, 균주 번호 6 (균주 "Y2096"으로 표시함)은 EPA를 전체 지질의 약 24%로 생산했고, 균주 번호 43 (균주 "Y2097"로 표시함)은 EPA를 약 22.3%로 생산했고, 균주 번호 45 (균주 "Y2098"로 표시함)는 EPA를 약 22.4%로 생산했고, 균주 번호 47 (균주 "Y2099"로 표시함)은 EPA를 약 22.6%로 생산했고, 균주 번호 5 (균주 "Y2015"로 표시함)는 EPA를 약 23.3%로 생산했으며, 균주 번호 48 (균주 "Y2106"으로 표시함)은 EPA를 약 23%로 생산했다.

균주 Y2096 중의 EPA 함량 및 오일 양을 이하의 변형된 2-단계 성장 절차를 이용해서 추가로 분석하였다. 균주 Y2096은 단일 콜로니로부터 SD+AA 배지 3 mL에서 30℃ 및 250 rpm/분으로 진탕 배양시켰다. 성장 24시간 후에, SD + AA 배지 32 mL를 함유한 에를렌마이어 플라스크에 시발 배양액 3 mL를 첨가하였다. 30℃ 및 250 rpm/분의 진탕에서 48시간 추가 성장시킨 후에, 세포를 펠렛화하고, 상등액을 제거하였다. 펠렛을 250 mL 플라스크 중의 HGM 35 mL에 재현탁하였다. 배양액 35 mL를 30℃에서 인큐베이션하고, 250 rpm/분으로 추가 4일 동안 진탕 성장시켰다. 배양액의 분취액 (1 mL)을 GC 분석에 사용하고, 배양액 30 mL를 세포 건중량의 측정에 사용하였다. (내부 대조군으로서 사용하기 위한) C15:0 40 ㎍을 에스테르교환반응을 위한 나트륨 메톡시드에 첨가한 것을 제외하면 GC 분석을 상기 '통상의 방법' 항목에 기재한 바와 같이 수행하였다. 세포 건중량은 배양액 30 mL로부터의 H_2O-세척 세포 펠렛을 동결건조시킴으로써 측정하였다.

GC 분석은 Y2096이 전체 지질의 약 28.1% EPA와 약 20.8% 오일/세포 건중량을 생산하였음을 보여주었다. 균주 Y2096은 야생형 야로위아 리폴리티카 ATCC #20362에 대해 하기 유전자형을 보유하였다:

POX3-, LIP1-, Y.Δ12-, FBA::F.Δ12::Lip2, TEF::F. Δ12::Pex16, FBAIN::MΔ12::Pex20, TEF::Δ6S::Lip1, FBAIN::Δ6S::Lip1, FBAIN::E1S::Pex20, GPAT::E1S::Oct, GPDIN::E1S::Lip2, TEF::E2S::Xpr, FBAIN::MAΔ5::Pex20, TEF::MAΔ5::Lip1, TEF::HΔ5S::Pex16, TEF::IΔ5S::Pex20, GPAT::ID5S::Pex20, FBAIN::Δ17S::Lip2, FBAINm::Δ17S::Pex16, TEF::Δ17S::Pex16 and *2X TEF::rELO2S::Pex20.*

EPA를 전체 지질의 약 21%로 생산하는, Ura- 표현형을 갖는 Y2098U 균주의 생성

상기 '통상의 방법'에 따라 SalI/PacI-소화 구축물 pZKUT16 (서열 159)를 사용해서 TEF::rELO2S::Pex20 키메라 유전자를 균주 Y2098의 Ura3 유전자에 통합시켰다. 형질전환 후, 세포를 MM + 5-FOA 선별 플레이트에 플레이팅하고, 30℃에서 2 내지 3일 동안 유지하였다.

MM + 5-FOA 플레이트에서 성장한 총 48개의 형질전환체를 골라내서 MM 플레이트 및 MM + 5-FOA 플레이트에 따로 재스트리킹했다. MM + 5-FOA 플레이트에서 성장할 수 있지만 MM 플레이트에서는 성장할 수 없는 균주를 Ura- 균주로서 선별했다. 이들 48개의 균주 모두를 상기 '통상의 방법'에 기재된 바와 같은 2-단계 성장 조건 (즉, 48시간 MMU, 96시간 HGM)을 이용해서 분석하였다. 세포를 원심분리로 수집하여 지질을 추출하고 에스테르교환반응으로 지방산 메틸 에스테르를 제조한 후에 휴렛-팩커드 6890 GC로 분석하였다.

GC 분석은 2-단계 성장 후 pZKUT16을 갖는 모든 형질전환체에서 EPA가 12 내지 20%로 존재했음을 보여주었다. 한가지 균주 (즉, 번호 33 균주)는 EPA를 약 21%로 생산했으며, 균주 "Y2098U"로 표시했다.

DHA를 전체 지질의 5.6% 초과로 생산하는 Y3000 균주의 생성

구축물 pZP2FOEN4 (도 9E; 서열 165)를 사용해서 2개의 키메라 유전자 (합성 $C_{20/22}$ "OtE2S" 엘론가제 및 합성 Δ4 데새투라제 "D4S"를 포함)의 클러스터를 균주 Y2098U의 Pox2 유전자 부위에 통합시켰다. 플라스미드 pZP2FOEN4는 하기 성분을 함유하였다:

[표 21]

플라스미드 pZP2FOEN4 (서열 165)의 기재	
서열 165 내의 RE 부위 및 뉴클레오티드	단편 및 키메라 유전자 성분의 기재
BsiWI/AscI (6152-6962)	야로위아 Aco2 유전자 (진뱅크 관리 번호 AJ001300)의 810 bp 5' 부분
SphI/EcoRI (9670-10325)	야로위아 Aco2 유전자 (진뱅크 관리 번호 AJ001300)의 655 bp 3' 부분

EcoRI/PacI (10640-2648)	· YAT 프로모터: YAT1 프로모터 (서열 217) · Δ4S: 코돈-최적화 Δ4 데새투라제 유전자 (서열 106) (트라우스토키트륨 아우레움 (진뱅크 관리 번호 AAN75707) 으로부터 유래됨) · Pex16: 야로위아 Pex16 유전자 (진뱅크 관리 번호 U75433)의 Pex16 종결자 서열 을 포함하는 YAT::Δ4S::Pex16
EcoRV를 갖는 EcoRI/SwaI (10618-8345)	· FBAIN: FBAIN 프로모터 (서열 214) · OtE2S: 코돈-최적화 OtE2S 유전자 (서열 102) (오스트레오코쿠스 타우리 (진뱅크 관리 번호 AY591336) 로부터 유래됨) · OCT: 야로위아 OCT 유전자 (진뱅크 관리 번호 X69988)의 OCT 종결자 서열 을 포함하는 FBAIN::OtE2S::Oct
SpeI를 갖는 SpeI/XbaI (12582-11095)	야로위아 Ura3 유전자 (진뱅크 관리 번호 AJ306421)

플라스미드 pZP2FOEN4를 SphI/AscI으로 소화시킨 후에, 이것을 상기 '통상의 방법'에 따라 사용하여 Y2098U 균주를 형질전환시켰다. 형질전환 후, 세포를 MM 선별 배지 플레이트에 플레이팅하고, 30℃에서 2 내지 3일 동안 유지시켰다. MM 플레이트에서 성장한 총 24개의 형질전환체를 골라내서 신선한 MM 플레이트에 재스트리킹했다. 일단 성장했으면, 이들 균주를 개별적으로 30℃에서 액체 배지 3 mL에 접종하고 250 rpm/분으로 2일 동안 진탕시켰다. 세포를 원심분리로 수집하여 지질을 추출하고 에스테르교환반응으로 지방산 메틸 에스테르를 제조한 후에 휴렛-팩커드 6890 GC로 분석하였다.

GC 분석은, 24개의 형질전환체들 중 12개에서 DPA가 전체 지질의 약 3.6 내지 5.3%로 존재하며 DHA가 0.4 내지 1%로 존재함을 보여주었다. DPA 및 DHA를 생산하는 12개의 균주는 상기 2-단계 성장 조건 (즉, 48시간 MM, 96시간 HGM)을 이용하여 추가의 분석을 위해 선별하였다. 이어서 세포를 원심분리로 수집하여 지질을 추출하고 에스테르교환반응으로 지방산 메틸 에스테르를 제조한 후에 휴렛-팩커드 6890 GC로 분석하였다.

GC 분석은 이들 12개의 형질전환체들에서 DPA가 15.6 내지 20%로 생산되며 DHA가 2.9 내지 5.6%로 생산됨을 보여주었다. DHA를 5.6%로 생산하는 균주를 균주 "Y3000"으로 칭했다.

실시예 7

Ura- 유전자형을 가지며 LA를 전체 지질의 45%로 생산하는 중간체 균주 Y2031의 생성

플라스미드 pKUNT2 (도 10A)의 TEF::Y.Δ12:Pex20 키메라 유전자를 야생형 야로위아 균주 ATCC #20362의 Ura3 유전자의 유전자좌에 통합시켜 Ura-유전자형을 구성함으로써 균주 Y2031을 생성시켰다.

구체적으로, 플라스미드 pKUNT2는 하기 성분을 함유하였다:

[표 22]

플라스미드 pKUNT2 (서열 116)의 기재	
서열 166 내의 RE 부위 및 뉴클레오티드	단편 및 키메라 유전자 성분의 기재
AscI/BsiWI (3225-3015)	야로위아 Ura3 유전자 (진뱅크 관리 번호 AJ306421)의 784 bp 5' 부분

SphI/PacI (5933-13)	야로위아 Ura3 유전자 (진뱅크 관리 번호 AJ306421)의 516 bp 3' 부분
EcoRI/BsiWI (6380-8629)	· TEF: TEF 프로모터 (진뱅크 관리 번호 AF054508) · Y.Δ12: 야로위아 Δ12 데새투라제 유전자 (서열 28) · Pex20: 야로위아 Pex20 유전자 (진뱅크 관리 번호 AF054613)의 Pex20 종결자 서열 을 포함하는 TEF::Y.Δ12::Pex20

pKUNT2 플라스미드를 AscI/SphI으로 소화시킨 후, 이것을 상기 '통상의 방법'에 따라 사용하여 야생형 야로위아 리폴리티카 ATCC #20362를 형질전환시켰다. 형질전환체 세포를 FOA 선별 배지 플레이트에 플레이팅하고 30℃에서 2 내지 3일 동안 유지시켰다. FOA 내성 콜로니를 골라내서 MM 및 MMU 선별 플레이트에 스트리킹했다. MMU 플레이트에서는 성장할 수 있으나 MM 플레이트에서는 성장할 수 없는 콜로니를 Ura- 균주로서 선별하였다. 이어서 Ura- 균주의 단일 콜로니 (5)를 30℃에서 액체 MMU 에 접종시키고 250 rpm/분으로 2일 동안 진탕시켰다. 세포를 원심분리로 수집하여 지질을 추출하고 에스테르교환반응으로 지방산 메틸 에스테르를 제조한 후에 휴렛-팩커드 6890 GC로 분석하였다.

GC 분석은 야생형 ATCC #20362에서 LA가 약 20%로 존재한 것에 비해 2개의 Ura- 균주 (즉, 균주 번호 2 및 3)에서 LA가 약 45%로 존재했음을 보여주었다. 형질전환체 균주 번호 2를 균주 "Y2031"로 칭했다.

실시예 8

야로위아 리폴리티카 중의 코돈-최적화 Δ9 엘론가제 유전자의 합성 및 기능적 발현

이소크리시스 갈바나의 Δ9 엘론가제 유전자 (진뱅크 관리 번호 AF390174)의 코돈 사용방식을 WO 2004/101753 및 실시예 4에 기재된 바와 유사한 방식으로 야로위아 리폴리티카에서의 발현에 최적화시켰다. 구체적으로, 야로위아 코돈 사용방식 패턴, ATG 번역 개시 코돈 주변의 컨센서스 서열, 및 RNA 안정성의 일반 규칙 [Guhaniyogi, G. and J. Brewer, Gene 265(1-2):11-23 (2001)]에 따라, 코돈-최적화 Δ9 엘론가제 유전자는 이소크리시스 갈바나 유전자 (서열 69)의 DNA 서열을 기초로 하여 디자인하였다 (서열 71). 번역 개시 부위의 변형 외에, 792 bp 코딩 영역 중 126 bp가 변형되었고, 123개 코돈이 최적화되었다. 코돈-최적화 유전자의 변형의 어느 것도 코딩된 단백질의 아미노산 서열을 변화시키지는 않았다 (진뱅크 관리 번호 AF390174, 서열 70).

야로위아에 대한 코돈-최적화 Δ9 엘론가제 유전자의 시험관내 합성

코돈-최적화 Δ9 엘론가제 유전자를 하기와 같이 합성하였다. 먼저, 8쌍의 올리고뉴클레오티드를 디자인하여 이소크리시스 갈바나 Δ9 엘론가제 유전자의 코돈-최적화 코딩 영역의 전장을 신장시켰다 (예컨대, 서열 292 내지 307에 상응하는 IL3-1A, IL3-1B, IL3-2A, IL3-2B, IL3-3A, IL3-3B, IL3-4A, IL3-4B, IL3-5A, IL3-5B, IL3-6A, IL3-6B, IL3-7A, IL3-7B, IL3-8A 및 IL3-8B). 센스 (A) 및 안티센스 (B) 올리고뉴클레오티드의 각 쌍은 각 5'-말단에서의 4 bp 오버행 외에는 상보적이었다. 추가로, 프라이머 IL3-1F, IL3-4R, IL3-5F 및 IL3-8R (서열 308-311)은 또한 차후의 서브클로닝을 위해 NcoI, PstI, PstI 및 Not1 제한 부위를 각각 도입하였다.

각각의 올리고뉴클레오티드 (100 ng)는 실시예 4에 기재된 바와 같이 인산화하였으며, 이어서 센스 및 안티센스 올리고뉴클레오티드의 각 쌍을 혼합하고, 함께 어닐링시켰다 [예를 들면, IL3-1A (서열 292)를 IL3-1B (서열 293)에 어닐링시켜 이중-가닥 생성물 "IL3-1AB"를 생산하였으며, IL3-2A (서열 294)를 IL3-2B (서열 295)에 어닐링시켜 이중-가닥 생성물 "IL3-2AB"을 생산함, 등.]

이어서 어닐링된 이중-가닥 올리고뉴클레오티드의 2개의 개별적 풀을 하기에 나타내는 바와 같이 함께 라이게이션하였다: 풀 1 (IL3-1AB, IL3-2AB, IL3-3AB 및 IL3-4AB 포함); 및 풀 2 (IL3-5AB, IL3-6AB, IL3-7AB 및 IL3-8AB 포함). 어닐링된 올리고뉴클레오티드의 각 풀을 10U T4 DNA 리가제를 갖는 부피 20 ㎕에서 혼합하고, 라이게이션 반응물을 밤새 16℃에서 인큐베이션하였다.

이어서 각 라이게이션 반응의 생성물을 주형으로 사용하여, 디자인된 DNA 단편을 PCR로 증폭시켰다. 구체적으로, 라이게이션된 "풀 1" 혼합물 (즉, IL3-1AB, IL3-2AB, IL3-3AB 및 IL3-4AB)을 주형으로 사용하고, 올리고뉴클레오티드 IL3-1F 및 IL3-4R (서열 308 및 309)을 프라이머로 사용하여, 코돈-최적화 Δ9 엘론가제 유전자의 제1 부분을 PCR로 증폭시켰다. PCR 증폭은 총 부피를 50 ㎕로 하여 실시예 4에 기재한 바와 같이 수행하였다. 417 bp PCR 단편을 pGEM-T 이지 벡터 (프로메가) 내로 서브클로닝하여 pT9(1-4)을 생성시켰다.

라이게이션된 "풀 2" 혼합물 (즉, IL3-5AB, IL3-6AB, IL3-7AB 및 IL3-8AB)을 주형으로 사용하고, 올리고뉴클레오티드 IL3-5F 및 IL3-8R (서열 310 및 311)을 프라이머로서 사용하여, 코돈-최적화 Δ9 엘론가제 유전자의 제2 부분을 PCR로 유사하게 증폭시키고, pGEM-T-이지 벡터 내로 클로닝시켜 pT9(5-8)을 생성시켰다.

이. 콜라이를 따로 pT9(1-4) 및 pT9(5-8)로 형질전환시키고, 플라스미드 DNA를 앰피실린-내성 형질전환체로부터 단리시켰다. 플라스미드 DNA를 정제하고, 적절한 제한 엔도뉴클레아제로 소화시켜 pT9(1-4)의 417 bp NcoI/PstI 단편 (서열 312) 및 pT9(5-8)의 377 bp PstI/Not1 단편 (서열 313)을 유리시켰다. 이어서 이들 2개 단편을 합하고, Nco1/Not1 소화 pZUF17 (서열 162; 도 9B)와 함께 정해진 방향으로 라이게이션시켜 pDMW237 (도 10B, 서열 167)을 생성시켰다. pDMW237 중의 생성된 합성 Δ9 엘론가제 유전자 ("IgD9e")의 DNA 서열은 야로위아에 대해 원래 디자인된 코돈-최적화 유전자 (즉, 서열 71)와 정확하게 동일하였다.

야로위아 리폴리티카 중의 코돈-최적화 Δ9 엘론가제 유전자의 발현

키메라 FBAIN::IgD9e::Pex20 유전자를 함유한 구축물 pDMW237 (도 10B)인 자가-복제 플라스미드를 야로위아 리폴리티카 Y2031 균주 (실시예 7) 내로 상기 '통상의 방법' 항목에 기재한 바와 같이 형질전환시켰다. pDMW237을 갖는 Y2031의 3개 형질전환체를 개별적으로 MM 배지에서 2일 동안 성장시키고, 세포를 원심분리로 수집하고 지질을 추출하고 에스테르교환반응으로 지방산 메틸 에스테르를 제조한 후에 휴렛-팩커드 6890 GC로 분석하였다.

GC 결과는 pDMW237을 갖는 이들 형질전환체에서 약 7.1%, 7.3% 및 7.4% EDA가 각각 생산되었음을 보여주었다. 이들 데이타는 합성 코돈-최적화 IgD9e가 C18:2를 EDA로 전환시킬 수 있음을 입증하였다. 코돈-최적화 유전자의 "기질 전환율(%)"은 약 13%인 것으로 측정되었다.

실시예 9

야로위아 리폴리티카에서의 코돈-최적화 Δ8 데새투라제 유전자의 합성

유글레나 그라실리스의 Δ8 데새투라제 유전자 (진뱅크 관리 번호 AAD45877)의 코돈 사용방식을 WO 2004/101753 및 실시예 4 및 8 (상기 문헌)에 기재된 바와 유사한 방식으로 야로위아 리폴리티카에서의 발현에 최적화시켰다. 3개 상이한 코돈-최적화 유전자 (즉, "D8S-1", "D8S-2" 및 "D8S-3")의 합성에도 불구하고, 이들 유전자의 어느 것도 EDA를 DGLA로 탈포화시킬 수 없었다. 따라서, 이전에 공개된 Δ8 데새투라제 서열은 정확하지 않으며, Δ8 데새투라제를 유글레나 그라실리스로부터 직접 단리한 후에 mRNA 단리, cDNA 합성 및 PCR을 행하는 것이 필요하다고 가정하였다. 이로 인해 2개의 유사한 서열을 본원에서 Eg5 (서열 77 및 78) 및 Eg12 (서열 79 및 80)로 동정하였다.

각 유전자 서열의 기능적 분석은 유전자를 사카로마이세스 세레비지애 효모 발현 벡터 내에 클로닝하고, 기질 공급 시험을 행하여 수행하였다. Eg5 및 Eg12 모두는 데새투라제 EDA 및 ETrA가 DGLA 및 ETA를 각각 생산할 수 있게 하였지만, Eg5는 Eg12보다 유의하게 큰 활성을 가졌다.

Eg5의 확인된 Δ8 데새투라제 활성에 기초하여, 상기 서열은 야로위아 리폴리티카에서의 발현을 위해 코돈-최적화되었고, 그에 의해 "D8SF" (서열 81 및 82)로 지칭되는 합성 기능적 코돈-최적화 Δ8 데새투라제를 합성하였다.

코돈-최적화 Δ8 데새투라제 유전자의 예비적 시험관내 합성

코돈-최적화 Δ8 데새투라제 유전자 ("D8S-1"이라고 칭함, 서열 75)는 유글레나 그라실리스의 공개된 서열 (서열 72 및 73)을 기초로 하여 야로위아 코돈 사용방식 패턴 (WO 2004/101753), 'ATG' 번역 개시 코돈 주변의 컨센서스 서열, 및 RNA 안정성의 일반 규칙 [Guhaniyogi, G. and J. Brewer, Gene 265(1-2): 11-23 (2001)]에 따라 디자인하였다. 번역

개시 부위의 변형 외에, 1260 bp 코딩 영역 중 200 bp가 변형되었다 (15.9%). 코돈-최적화 유전자의 변형은 번역 개시 코돈 주변에 NcoI 부위를 첨가하는 'K'에서 'E'로의 제2 아미노산 변화를 제외하면 코딩된 단백질의 아미노산 서열 (서열 73)을 변화시키지 않았다.

구체적으로, 코돈-최적화 Δ8 데새투라제 유전자를 하기와 같이 합성하였다. 먼저, 13쌍의 올리고뉴클레오티드를 디자인하여 유글레나 그라실리스 Δ8 데새투라제 유전자의 코돈-최적화 코딩 영역의 전장을 신장시켰다 (예컨대, 서열 314 내지 339에 상응하는 D8-1A, D8-1B, D8-2A, D8-2B, D8-3A, D8-3B, D8-4A, D8-4B, D8-5A, D8-5B, D8-6A, D8-6B, D8-7A, D8-7B, D8-8A, D8-8B, D8-9A, D8-9B, D8-10A, D8-10B, D8-11A, D8-11B, D8-12A, D8-12B, D8-13A 및 D8-13B). 센스 (A) 및 안티센스 (B) 올리고뉴클레오티드의 각 쌍은 각 5'-말단에서의 4 bp 오버행 외에는 상보적이었다. 추가로, 프라이머 D8-1A, D8-3B, D8-7A, D8-9B 및 D8-13B (서열 314, 319, 326, 331 및 339)는 또한 차후의 서브클로닝을 위해 NcoI, BglII, XhoI, SacI 및 Not1 제한 부위를 각각 도입하였다.

각각의 올리고뉴클레오티드 (100 ng)를 실시예 4에 기재된 바와 같이 인산화하고, 이이서 센스 및 안티센스 올리고뉴클레오티드의 각 쌍을 혼합하고, 함께 어닐링시켰다 [예를 들어, D8-1A (서열 314)를 D8-1B (서열 315)에 어닐링시켜 이중-가닥 생성물 "D8-1AB"를 생산함, D8-2A (서열 316)를 D8-2B (서열 317)에 어닐링시켜 이중-가닥 생성물 "D8-2AB"를 생산함, 등].

어닐링된 이중-가닥 올리고뉴클레오티드의 4개의 개별적 풀을 하기에 나타내는 바와 같이 이어서 함께 라이게이션하였다: 풀 1 (D8-1AB, D8-2AB 및 D8-3AB 포함); 풀 2 (D8-4AB, D8-5AB 및 D8-6AB 포함); 풀 3 (D8-7AB, D8-8AB, 및 D8-9AB 포함); 및 풀 4 (D8-10AB, D8-11AB, D8-12AB 및 D8-13AB 포함). 어닐링된 올리고뉴클레오티드의 각 풀을 10U T4 DNA 리가제를 갖는 부피 20 ㎕에서 혼합하고, 라이게이션 반응물을 16℃에서 밤새 인큐베이션하였다.

각 라이게이션 반응의 생성물을 이어서 주형으로서 사용하여 디자인된 DNA 단편을 PCR로 증폭시켰다. 구체적으로, 라이게이션된 "풀 1" 혼합물 (즉, D8-1AB, D8-2AB 및 D8-3AB)을 주형으로 사용하고, 올리고뉴클레오티드 D8-1F 및 D8-3R (서열 340 및 341)을 프라이머로 사용하여, 코돈-최적화 Δ8 데새투라제 유전자의 제1 부분을 PCR로 증폭시켰다. PCR 증폭은 총 부피를 50 ㎕로 하여 실시예 4에 기재한 바와 같이 수행하였다. 309 bp PCR 단편을 pGEM-T 이지 벡터 (프로메가) 내로 서브클로닝하여 pT8(1-3)을 생성시켰다.

라이게이션된 "풀 2" 혼합물 (즉, D8-4AB, D8-5AB 및 D8-6AB)을 주형으로 사용하고, 올리고뉴클레오티드 D8-4F 및 D8-6R (서열 342 및 343)을 프라이머로서 사용하여, 코돈-최적화 Δ8 데새투라제 유전자의 제2 부분을 PCR로 유사하게 증폭시키고, pGEM-T-이지 벡터 내로 클로닝시켜 pT8(4-6)을 생성시켰다. 라이게이션된 "풀 3" 혼합물 (즉, D8-7AB, D8-8AB 및 D8-9AB)을 주형으로 사용하고, 올리고뉴클레오티드 D8-7F 및 D8-9R (서열 344 및 345)을 프라이머로서 사용하여, 코돈-최적화 Δ8 데새투라제 유전자의 제3 부분을 PCR로 유사하게 증폭시키고, pGEM-T-이지 벡터 내로 클로닝시켜 pT8(7-9)을 생성시켰다. 최종적으로, 라이게이션된 "풀 4" 혼합물 (즉, D8-10AB, D8-11AB, D8-12AB 및 D8-13AB)을 주형으로 사용하고, 올리고뉴클레오티드 D8-10F 및 D8-13R (서열 346 및 347)을 프라이머로서 사용하여, 코돈-최적화 Δ8 데새투라제 유전자의 제4 부분을 PCR로 유사하게 증폭시키고, pGEM-T-이지 벡터 내로 클로닝시켜 pT8(10-13)을 생성시켰다.

이. 콜라이를 따로 pT8(1-3), pT8(4-6), pT8(7-9) 및 pT8(10-13)으로 형질전환시키고, 플라스미드 DNA를 앰피실린-내성 형질전환체로부터 단리시켰다. 플라스미드 DNA를 정제하고, 적절한 제한 엔도뉴클레아제로 소화시켜 pT8(1-3)의 309 bp NcoI/BglII 단편 (서열 348), pT8(4-6)의 321 bp BglII/XhoI 단편 (서열 349), pT8(7-9)의 264 bp XhoI/SacI 단편 (서열 350) 및 pT8(10-13)의 369 bp Sac1/Not1 단편 (서열 351)을 유리시켰다. 이어서 이들 단편을 합하고, Nco1/Not1 소화 pY54PC (서열 168; WO2004/101757)와 함께 정해진 방향으로 라이게이션시켜 pDMW240 (도 10C)을 생성시켰다. 이는 pDMW240 중의 합성 Δ8 데새투라제 유전자 ("D8S-1", 서열 75)를 생성시켰다.

유글레나 그라실리스의 공개된 Δ8 데새투라제 아미노산 서열 (서열 73)에 비해, 번역 개시 코돈 주변에 NcoI 부위를 첨가하기 위해 D8S-1의 제2 아미노산을 'K'에서 'E'로 변화시켰다. 공개된 유글레나 그라실리스 Δ8 데새투라제 서열 (서열 73)과 정확하게 동일한 아미노산 서열을 갖는 또다른 형태의 합성된 유전자는 pDMW240을 주형으로서 사용하고 올리고뉴클레오티드 ODMW390 및 ODMW391 (서열 352 및 353)을 프라이머로서 사용하여 시험관내 돌연변이유발 (스트라타진, 미국 캘리포니아주 샌 디에고 소재)에 의해 구축하였다. 생성된 플라스미드를 pDMW255라고 칭하였다. pDMW255 중의 합성 Δ8 데새투라제 유전자를 "D8S-2"라고 칭하였으며, 아미노산 서열은 서열 73에 도시한 서열과 정확하게 동일하였다.

비기능적 코돈-최적화 Δ8 데새투라제 유전자

야로위아 리폴리티카 균주 ATCC #76982(Leu-)를 pDMW240 (도 10C) 및 pDMW255로 각각 상기 '통상의 방법' 항목에 기재한 바와 같이 형질전환시켰다. 재조합 구축물 함유 효모를 EDA [20:2(11,14)]를 보충한 MM에서 성장시켰다. 구체적으로, pDMW240 (D8S-1 함유) 또는 pDMW255 (D8S-2 함유)를 함유한 형질전환체 야로위아 리폴리티카의 단일 콜로니를 MM 3 mL 중에서 30℃에서 OD_{600} 약 1.0까지 성장시켰다. 이어서 기질 공급을 위해 세포 100 ㎕를, EDA 기질 10 ㎍을 함유하는 MM 3 mL 중에서 30℃에서 약 24시간 동안 계대배양하였다. 세포를 원심분리로 수집하고 지질을 추출하고 에스테르교환반응으로 지방산 메틸 에스테르를 제조한 후에 휴렛-팩커드 6890 GC로 분석하였다.

상기 형질전환체 중 어느 것도 EDA로부터 DGLA를 생산하지 않았으며, 따라서 D8S-1 및 D8S-2가 기능하지 않았고, EDA를 탈포화시킬 수 없었다. 키메라 D8S-1::XPR 및 D8S-2::XPR 유전자를 서열 354 및 355에 각각 나타낸다.

진뱅크 (관리 번호 AAD45877 [서열 73])에 기탁되고 WO 00/34439 또는 문헌 [Wallis et al., Archives of Biochem. Biophys, 365:307-316 (1999)]에 기재한 Δ8 데새투라제의 단백질 서열 (본원에서 서열 74)들 사이에서 3개의 아미노산 차이가 발견되었다. 구체적으로, 3개 아미노산은 진뱅크 관리 번호 AAD45877에서 누락된 것으로 나타났다. pDMW255를 주형으로 사용하고 ODMW392 및 ODMW393 (서열 356 및 357)을 프라이머로서 사용하여, 시험관내 돌연변이유발 (스트라타진, 미국 캘리포니아주 샌 디에고 소재)에 의해 9 bp를 합성 D8S-2 유전자 내에 부가함으로써 WO 00/34439 및 문헌 [Wallis et al. (상기 문헌)]에 기재한 서열 (서열 74)과 동일한 단백질을 생산하였다. 생성된 플라스미드를 pDMW261이라고 칭하였다. pDMW261 중의 합성 Δ8 데새투라제 유전자를 "D8S-3" (서열 76)이라고 칭하였다. pDMW261 구축물을 야로위아 내로 형질전환시킨 후에, EDA를 사용하는 유사한 공급 실험을 상기 기재한 바와 같이 수행하였다. D8S-3의 경우 EDA에서 DGLA로의 탈포화가 관찰되지 않았다.

유글레나 그라실리스 Δ8 데새투라제 유전자의 단리

유글레나 그라실리스를 미시간 주립대학교(미국 미시건주 이스트 랜싱 소재)에 재직하는 리차드 트리에머(Richard Triemer) 박사의 실험실에서 얻었다. 활성적 성장 배양물 10 mL로부터, 분취액 1 mL를 500 mL 유리병 중의 유글레나 그라실리스 (Eg) 배지 250 mL에 옮겼다. Eg 배지는 물 970 mL 중에 나트륨 아세테이트 1 g, 소 추출물 (카탈로그 #U126-01, 디프코 래버러토리즈, 미국 미시간주 디트로이트 소재) 1 g, 박토(Bacto)® 트립톤 (카탈로그 #0123-17-3, 디프코 래버러토리즈) 2 g 및 박토® 효모 추출물 (카탈로그 #0127-17-9, 디프코 래버러토리즈) 2 g을 합하여 제조하였다. 여과 멸균 후에, 토양수 상등액 (카탈로그 #15-3790, 캐롤라이나 바이올로지컬 서플라이 컴파니(Carolina Biological Supply Company), 미국 노쓰캐롤라이나 벌링톤 소재) 30 mL를 무균적으로 첨가하여 최종 Eg 배지를 제조하였다. 유글레나 그라실리스 배양물을 23℃에서 교반 없이 2주 동안 광 주기 16시간, 암 주기 8시간으로 성장시켰다.

2주 후에, 배양물 10 mL을 지질 분석을 위해 들어내고, 1,800 x g에서 5분 동안 원심분리하였다. 펠렛을 물로 1회 세척하고, 재-원심분리하였다. 생성된 펠렛을 5분 동안 진공하에 건조시키고, 트리메틸술포늄 히드록시드 (TMSH) 100 ㎕에 재현탁하고, 실온에서 15분 동안 진탕 인큐베이션하였다. 이 후에, 헥산 0.5 mL를 첨가하고, 바이알을 15분 동안 실온에서 진탕 인큐베이션하였다. 지방산 메틸 에스테르 (헥산층으로부터 5 ㎕ 주사됨)를 분리하고, 오메가왁스 320 융합 실리카 모세관 컬럼 (카탈로그 #24152, 수펠코 인크.(Supelco Inc.))을 장착한 휴렛-팩커드 6890 가스 크로마토그래프로 정량하였다. 오븐 온도를 220℃에서 2.7분 동안 유지하고, 20℃/분으로 240℃까지 증가시키고, 이어서 추가 2.3분 동안 유지하는 프로그램대로 진행시켰다. 운반 기체를 왓트만(Whatman) 발생기로 제공하였다. 체류 시간을 상업적으로 입수가능한 표준의 메틸 에스테르 (카탈로그 #U-99-A, 누-체크 프렙, 인크.(Nu-Chek Prep, Inc.))와 비교하고, 생성된 크로마토그램을 도 11에 나타낸다.

잔존 2주 배양물 (240 mL)을 1,800 x g에서 10분 동안 원심분리하여 펠렛화하고, 물로 1회 세척하고, 재-원심분리하였다. 총 RNA를 생성된 펠렛으로부터 RNA STAT-60™ 시약 (텔-테스트, 인크.(TEL-TEST, Inc.), 미국 택사스주 프랜즈우드 소재)으로 제공된 제조자의 프로토콜에 따라 (물 0.5 mL에 RNA를 용해한 시약 5 mL 사용) 추출하였다. 이러한 방법으로, 총 RNA (2 mg/mL) 1 mg을 펠렛으로부터 수득하였다. mRNA를 제공된 제조자의 프로토콜에 따라 mRNA 정제 키트 (아머샴 바이오사이언스즈(Amersham Biosciences), 미국 뉴저지주 피스카타웨이 소재)로 총 RNA 1 mg으로부터 단리하였다. 이런 방식으로, mRNA 85 ㎍을 수득하였다.

제조자의 프로토콜에 따라 제공된 올리고(dT) 프라이머와 함께 cDNA 합성용 수퍼스크립트™ 초이스 (Superscript™ Choice) 시스템 (인비트로젠™ 라이프 테크놀로지즈(Invitrogen™ Life Technologies), 미국 캘리포니아주 칼스배드 소재)을 사용하여 cDNA를 mRNA 765 ng으로부터 합성하였다. 합성된 cDNA를 물 20 ㎕에 용해하였다.

하기 조건을 이용하여 올리고뉴클레오티드 프라이머 Eg5-1 및 Eg3-3 (서열 358 및 359)로 유글레나 그라실리스 Δ8 데새투라제를 cDNA로부터 증폭시켰다. 구체적으로, cDNA (1 ㎕)를 50 pmol Eg5-1, 50 pmol Eg5-1, PCR 뉴클레오티드 믹스 (10 mM, 프로메가, 미국 위스콘신주 매디슨 소재) 1 ㎕, 10X PCR 완충제 (인비트로젠) 5 ㎕, $MgCl_2$ (50 mM, 인비트로젠) 1.5 ㎕, Taq 폴리머라제 (인비트로젠) 0.5 ㎕ 및 최종 50 ㎕가 되는 양의 물과 합하였다. 반응 조건은 3분 동안 94℃, 이어서 45초 동안 94℃, 45초 동안 55℃ 및 1분 동안 72℃의 35 주기였다. PCR을 72℃에서 7분 동안, 이어서 4℃에서 유지하여 마쳤다. PCR 반응물을 5 ㎕에 대해 아가로스 겔 전기영동으로 분석하고, 약 1.3 kB의 분자 중량을 갖는 DNA 밴드를 관찰하였다. 잔존 45 ㎕의 생성물을 아가로스 겔 전기영동으로 분리하고, DNA 밴드는 제조자의 프로토콜에 따라 자이모클린™ 겔 DNA 리커버리 키트(Zymoclean™ Gel DNA Recovery Kit)(자이모 리써치, 미국 캘리포니아주 오렌지 소재)를 사용하여 정제하였다. 생성된 DNA를 제조자의 프로토콜에 따라 pGEM®-T 이지 벡터 (프로메가) 내로 클로닝하였다. 다중 클론을 T7, M13-28Rev, Eg3-2 및 Eg5-2 (각각 서열 360 내지 363)로 서열결정하였다.

따라서, 단지 몇개 bp가 상이한 Eg5 (서열 77) 및 Eg12 (서열 79)인 2개 클래스의 DNA 서열을 얻었다. Eg5 및 Eg12의 번역은 단지 1개 아미노산이 상이한 단백질 서열인 서열 78 및 80을 각각 생성하였다. 따라서, Eg5에 대한 DNA 및 단백질 서열을 서열 77 및 서열 78로 각각 나타내고; Eg12에 대한 DNA 및 단백질 서열을 서열 79 및 서열 80으로 각각 나타낸다.

단리된 유글레나 그라실리스 Δ8 데새투라제 서열 대 공개된 유글레나 그라실리스 Δ8 데새투라제 서열의 비교

진뱅크 관리 번호 AAD45877 (gi: 5639724, 본원에서 서열 73)로부터의 단백질 서열 및 문헌 [Wallis et al., Archives of Biochem. Biophys., 365:307-316 (1999); WO 00/34439][본원에서 서열 74]의 공개된 단백질 서열과 서열 78 (Eg5) 및 서열 80 (Eg12)으로 나타내는 단백질 서열의 정렬을 도 12에 나타낸다. 모든 4개 서열 중에 보존된 아미노산을 별표 (*)로 나타낸다. 서열의 정렬을 최대화하기 위해 대시(-)를 프로그램에 의해 사용하였다. 추정적 시토크롬 b_5 도메인에 밑줄을 쳤다. 추정적 His 박스를 볼드체로 나타낸다. 동일성(%) 계산은 Eg5 Δ8 데새투라제 단백질 서열이 서열 73에 대해 95.5% 및 서열 74에 대해 96.2% 동일함을 나타내며, 여기서 "동일성(%)"을 2개 단백질 간의 동일한 아미노산의 비율(%)로서 정의한다. 서열 정렬 및 동일성(%) 계산은 LASERGENE 바이오인포매틱스 컴퓨팅 스위트(bioinformatics computing suite) (드나스타 인크., 미국 위스콘신주 매디슨 소재)의 메갈린 프로그램으로 수행하였다. 서열의 다중 정렬은 디폴트 파라미터 (GAP PENALTY=10, GAP LENGTH PENALTY=10)를 갖는 클러스탈 정렬 방법 [Higgins and Sharp, CABIOS. 5:151-153 (1989)]으로 수행하였다. 클러스탈 방법을 사용하는 쌍쌍 정렬에 대한 디폴트 파라미터는 KTUPLE 1, GAP PENALTY=3, WINDOW=5 및 DIAGONALS SAVED=5이었다. 다양한 유글레나 그라실리스 Δ8 데새투라제 서열들 간의 차이에 대한 더 완전한 분석을 위해서는 동시 계류 중인 미국 특허 출원 제11/166993호를 참조한다.

사카로마이세스 세레비지애 중의 유글레나 그라실리스 Δ8 데새투라제 서열의 기능적 분석

효모 에피솜 플라스미드 (YEp)-유형 벡터 pRS425 [Christianson et al., Gene, 110:119-22 (1992)]는 사카로마이세스 세레비지애 2μ 내인성 플라스미드로부터의 서열, LEU2 선별가능한 마커 및 다기능적 파지미드 pBluescript II SK + 의 주쇄에 기초한 서열을 함유한다. 사카로마이세스 세레비지애의 강력한 구성적 글리세르알데히드-3-포스페이트 데히드로게나제 (GPD) 프로모터를 pRS425의 SacII와 SpeI 부위 사이에 문헌 [Jia et al., Physiological Genomics, 3:83-92 (2000)]에 기재된 것과 동일한 방식으로 클로닝하여 pGPD-425를 생산하였다. NotI 부위를 pGPD-425의 BamHI 부위 내에 도입하여 (따라서, BamHI 부위에 의해 플랭킹된 NotI 부위를 생산함) 플라스미드 pY-75를 생산하였다. Eg5 (서열 77) 및 Eg12 (서열 79)을 NotI 소화에 의해 상기 기재한 pGEM®-T 이지 벡터로부터 유리시키고, pY-75의 NotI 부위 내로 클로닝하여 pY89-5 (ATCC #PTA-6048로 기탁됨) 및 pY89-12를 각각 생산하였다. 이런 방식으로, Δ8 데새투라제 (즉, Eg5 [서열 77] 및 Eg12 [서열 79])를 사카로마이세스 세레비지애에서의 발현을 위해 강력한 구성적 프로모터 뒤에 클로닝하였다. pY89-5의 지도를 도 10D에 나타낸다.

플라스미드 pY89-5, pY89-12 및 pY-75를 사카로마이세스 세레비지애 BY4741 (ATCC #201388) 내로 표준 리튬 아세테이트 형질전환 절차로 형질전환시켰다. 형질전환체는 CSM-leu (큐바이오젠(Qbiogene), 미국 캘리포니아주 칼스배드 소재)를 보충한 DOBA 배지에서 선별하였다. 각 플레이트로부터의 형질전환체를 CSM-leu (큐바이오젠)으로 보충한 DOB 배지 2 mL에 접종하고, 1일 동안 30℃에서 성장시킨 후에, 0.5 mL를 1 mM로 EDA 또는 EtrA를 보충한 동일한 배지로 옮겼다. 이들을 30℃에서 250 rpm으로 밤새 인큐베이션하고, 펠렛을 원심분리로 수득하고, 진공하에 건조시켰다. 펠렛을 TMSH 50 ㎕로 에스테르교환하고, 상기 '통상의 방법' 항목에 기재한 바와 같이 GC로 분석하였다. pY-75 (즉, 클론 75-1 및 75-2) 및 pY89-5 (즉, 클론 5-6-1 및 5-6-2)에 대한 2개 클론을 각각 분석하면서, 두 가지 독립적 형질전환으로부터의 pY89-12에 대한 2개 세트의 클론 (즉, 클론 12-8-1, 12-8-2, 12-9-1 및 12-9-2)을 분석하였다.

EDA 공급 클론의 GC 분석에 의해 얻은 지질 프로파일을 표 23에 나타내고; EtrA 공급 클론의 GC 분석에 의해 얻은 지질 프로파일을 표 24에 나타냈다. 지방산을 16:0 (팔미테이트), 16:1 (팔미트올레산), 18:0, 18:1 (올레산), 20:2 [EDA], 20:3 (8,11,14) [DGLA], 20:3 (11,14,17) [ETrA] 및 20:4 (8,11,14,17) [ETA]로서 확인하고; 각각의 조성을 전체 지방산의 백분율(%)로 나타낸다.

[표 23]

유글레나 그라실리스 Δ8 데새투라제를 과발현하는 형질전환체

사카로마이세스 세레비지애의 지질 분석 : EDA 기질 공급

클론	16:0	16:1	18:0	18:1	20:2	20:3 (8,11,14)	% 20:2 전환됨
75-1 (대조군)	14	32	5	38	10	0	0
75-2 (대조군)	14	31	5	41	9	0	0
5-6-1 (Eg5)	14	32	6	40	6	2	24
5-6-2 (Eg5)	14	30	6	41	7	2	19
12-8-1 (Eg12)	14	30	6	41	9	1	7
12-8-2 (Eg12)	14	32	5	41	8	1	8
12-9-1 (Eg12)	14	31	5	40	9	1	8
12-9-2 (Eg12)	14	32	5	41	8	1	7

[표 24]

유글레나 그라실리스 Δ8 데새투라제를 과발현하는 형질전환체

사카로마이세스 세레비지애의 지질 분석 : ETrA 기질 공급

클론	16:0	16:1	18:0	18:1	20:3 (11,14,17)	20:4 (8,11,14,17)	% 20:3 전환됨
75-1 (대조군)	12	25	5	33	24	0	0
75-2 (대조군)	12	24	5	36	22	1	5
5-6-1 (Eg5)	13	25	6	34	15	7	32
5-6-2 (Eg5)	13	24	6	34	17	6	27
12-8-1 (Eg12)	12	24	5	34	22	2	8
12-8-2 (Eg12)	12	25	5	35	20	2	9
12-9-1 (Eg12)	12	24	5	34	22	2	9
12-9-2 (Eg12)	12	25	6	35	20	2	9

표 23 및 24의 데이타는 클로닝된 유글레나 Δ8 데새투라제가 EDA 및 EtrA를 탈포화시킬 수 있었음을 보여주었다. 서열 80에 나타낸 서열은 서열 78에 나타낸 서열에 비해 1개 아미노산 변화를 가지고, 저하된 Δ8 데새투라제 활성을 가졌다.

표 24 중의 클론 75-2에 의해 생성되는 소량의 20:4(8,11,14,17)는 20:4(8,11,14,17)에 대한 표준과 약간 다른 체류 시간을 가졌다. 이 피크는 그 실험에서 야생형 효모에 의해 생성되는 소량의 상이한 지방산일 가능성이 더 크다.

야로위아 리폴리티카에 대해 최적화된 Δ8 데새투라제 유전자 코돈의 추가 개질

pDMW261 중의 합성 D8S-3 유전자의 아미노산 서열을 기능적 유글레나 Δ8 데새투라제의 아미노산 서열 (서열 77 및 78)에 따라 교정하였다. pDMW261을 주형으로 사용하고, 올리고뉴클레오티드 ODMW404 (서열 364) 및 D8-13R (서열 347)을 사용하여, 합성 D8S-3 데새투라제 유전자를 코딩하는 DNA 단편을 증폭시켰다. 생성된 PCR 단편을 Bio101 진클린(Geneclean) 키트로 정제하고, 이어서 Kpn1 및 Not1로 소화시켰다 (프라이머 ODMW404가 KpnI 부위를 도입하고, 프라이머 D8-13R은 NotI 부위를 도입함). Kpn1/Not1 단편 (서열 365)을 Kpn1/Not1 소화 pKUNFmKF2 (도 10E, 서열 169) 내로 클로닝하여 pDMW277 (도 13A)을 생산하였다.

증폭시키고, D8S-3 유전자의 5' 말단을 교정하기 위해 디자인한 올리고뉴클레오티드 YL521 및 YL522 (서열 366 및 367)를 pDMW277이 주형으로서 사용되는 또다른 PCR 반응에서 프라이머로서 사용하였다. 프라이머는 PCR 단편의 5' 및 3' 말단에 Nco1 부위 및 BglII 부위를 각각 도입하였다. 318 bp PCR 생성물을 Bio101 진클린 키트로 정제하고, 이어서 Nco1 및 BglII로 소화시켰다. 소화 단편을 pDMW277로부터의 954 bp BglII/NotI 단편과 함께 사용하여 pZF5T-PPC (도 13B, 서열 170)를 교환하여 pDMW287을 형성시켰다. 합성 D8S-3 유전자의 5' 말단을 교정하는 것 외에, 이 클로닝 반응은 또한 합성 Δ8 데새투라제 유전자를 야로위아 리폴리티카 FBAIN 프로모터 (서열 214)의 조절 하에 두었다.

부위-지정 돌연변이유발 반응의 최종 시리즈 중의 제1 반응은 이어서 pDMW287 상에서 수행하였다. 프라이머의 제1 세트인 YL525 및 YL526 (서열 368 및 369)은 pDMW287 중의 합성 D8S-3 유전자의 아미노산 F에서 S (위치 #50)로의 교정을 위해 디자인하였다. 이 돌연변이유발 반응으로부터 생성된 플라스미드는 이어서 프라이머 YL527 및 YL528 (서열 370 및 371)을 사용하는 다음 부위-지정 돌연변이유발 반응에 주형이 되었다. 이들 프라이머는 아미노산 D8S-3 유전자의 아미노산 F에서 S (위치 #67)로의 교정을 위해 디자인하고, 플라스미드 pDMW287/YL527을 생성하였다.

유전자의 제2 쿼터 내에 서열 교정을 완료하기 위해, 하기 반응을 유전자의 제1 쿼터 상의 돌연변이와 동시에 수행하였다. pDMW287을 주형으로서 사용하고, 올리고뉴클레오티드 YL529 및 YL530 (서열 372 및 373)을 프라이머로서 사용하여, 시험관내 돌연변이유발 반응을 수행하여 합성 D8S-3 유전자의 아미노산 C에서 W (위치 #177)로의 교정을 수행하였다. 이 돌연변이유발 반응의 생성물 (즉, pDMW287/Y529)을 프라이머 YL531 및 YL532 (서열 374 및 375)를 사용하는 하기 반응에 주형으로서 사용하여 아미노산 P에서 L (위치 #213)로 교정하였다. 이 반응의 생성물을 pDMW287/YL529-31이라고 칭하였다.

유전자의 제1 및 제2 쿼터 상의 돌연변이와 동시에, 상기 반응을 유전자의 3' 말단 상에서 유사하게 수행하였다. 각각의 차후 돌연변이유발 반응은 이전 반응으로부터의 플라스미드 생성물을 사용하였다. 프라이머 YL533 및 YL534 (서열 376 및 377)를 pDMW287 상에 사용하여 아미노산 C에서 S (위치 #244)로 교정함으로써 pDMW287/YL533을 생성하였다. 프라이머 YL535 및 YL536 (서열 378 및 379)을 사용하여 pDMW287/YL533의 합성 D8S-3 유전자 중의 아미노산 A를 T (위치 #280)로 교정하여 pDMW287/YL533-5를 형성시켰다. 마지막으로, pDMW287/YL533-5를 주형으로서 사용하고 YL537 및 YL538 (서열 380 및 381)을 프라이머로 사용하여 합성 D8S-3 유전자 중의 위치 #333에서의 아미노산 P를 S로 교정하였다. 생성된 플라스미드를 pDMW287/YL533-5-7이라고 명명하였다.

pDMW287/YL529-31의 BglII/XhoI 단편 및 pDMW287/YL533-5-7의 XhoI/NotI 단편을 사용하여 pDMW287/YL257의 BglII/NotI 단편을 변화시켜 완전히 교정된 합성 Δ8 데새투라제 유전자 (D8SF라고 칭하고, 서열 81에 나타냄)를 함유한 pDMW287F (도 13C)을 생산하였다. 서열 82는 개시 메티오닌 이후의 추가의 발린을 제외하면 서열 78에 나타낸 서열과 본질적으로 동일한 서열 81의 뉴클레오티드 2 내지 1270을 코딩하는 아미노산 서열을 나타낸다.

실시예 10

야로위아 리폴리티카 중의 코돈-최적화 Δ9 엘론가제 유전자 및

코돈-최적화 Δ8 데새투라제의 기능적 발현

본 실시예는 실시예 8 및 9로부터의 코돈-최적화 Δ9 엘론가제 및 코돈-최적화된 Δ8 데새투라제를 동시 발현하기 위해 형질전환된 야로위아 리폴리티카 중의 DGLA 생합성 및 축적을 기재한다. 이에 의해 본 실험은 유전자 활성 및 야로위아 리폴리티카의 Δ9 엘론가제/Δ8 데새투라제 경로 발현 능력을 모두 확인하였다.

구체적으로, 구축물 pDMW287F (실시예 9)의 키메라 FBAIN::D8SF::Pex16 유전자를 포함하는 ClaI/PacI 단편을 pDMW237 (실시예 8)의 ClaI/PacI 부위 내에 삽입시켜 구축물 pDMW297 (도 13D; 서열 123)을 생성하였다.

플라스미드 pDMW297은 하기 성분을 함유하였다:

[표 25]

플라스미드 pDMW297 (서열 171)의 기재

서열 171 중의 RE 부위 및 뉴클레오티드	단편 및 키메라 유전자 성분의 기재
EcoRI/ClaI (9053-10448)	ARS18 서열 (진뱅크 관리 번호 A17608)
ClaI/PacI (1-2590)	· FBAIN: FBAIN 프로모터 (서열 214) · D8SF: 코돈-최적화 Δ8 데새투라제 유전자 (서열 181) (유글레나 그라실리스 (진뱅크 관리 번호 AF139720)로부터 유래됨) · Pex16: 야로위아 Pex16 유전자 (진뱅크 관리 번호 U75433)의 Pex16 종결자 서열 을 포함하는 FBAIN::D8SF::Pex16
PacI/SalI (2590-4082)	야로위아 Ura3 유전자 (진뱅크 관리 번호 AJ306421)
SalI/BsiWI (4082-6257)	· FBAIN: FBAIN 프로모터 (서열 214) · IgD9e: 코돈-최적화 Δ9 엘론가제 유전자 (서열 71) (이소크리시스 갈바나 (진뱅크 관리 번호 390174) 로부터 유래됨) · Pex20: 야로위아 Pex20 유전자 (진뱅크 관리 번호 AF054613)의 Pex20 종결자 서열 을 포함하는 FBAIN::IgD9e::Pex20

이어서 구축물 pDMW297을 균주 Y2031 (실시예 13)의 형질전환에 대해 통상의 방법에 따라 사용하였다. 형질전환체 세포를 MM 선별 배지 플레이트에 플레이팅하고, 30℃에서 2 내지 3일 동안 유지시켰다. MM 플레이트에서 성장한 총 8개 형질전환체를 골라내서 신선한 MM 플레이트에 재스트리킹했다. 일단 성장했으면, 이들 균주를 개별적으로 30℃에서 액체 MM에 접종하고 250 rpm/분으로 2일 동안 진탕시켰다. 세포를 원심분리로 수집하고 지질을 추출하고 에스테르교환반응으로 지방산 메틸 에스테르를 제조한 후에 휴렛-팩커드 6890 GC로 분석하였다.

GC 분석은 DGLA가 분석된 모든 형질전환체에서 생산됨을 보여주었다. 1개 균주는 DGLA를 약 3.2% 생산하고, 4개 균주는 DGLA를 4.3 내지 4.5% 생산하고, 2개 균주는 DGLA를 5.5 내지 5.8% 생산하고, 1개 균주는 DGLA를 6.4% (본원에서 균주 "Y0489"라고 칭함) 생산하였다. 균주 Y0489 중의 코돈-최적화 D8SF 유전자의 "기질 전환율(%)"은 75%인 것으로 측정되었다.

실시예 11

ω-6 Δ9 엘론가제/Δ8 데새투라제 경로:

야로위아 리폴리티카에서 DHA를 생산하는 균주의 생성 (예언적)

본 실시예는 야로위아 리폴리티카 ATCC #20362로부터 유래되어 ω-6 Δ9 엘론가제/Δ8 데새투라제 경로의 발현을 통해 DHA를 생산하도록 유전자 조작된 균주의 구축을 기재하고 있으며, 따라서 이들 DHA-생산 균주의 완전한 지질 프로파일의 분석은 최종 DHA-함유 오일에서 GLA 동시-합성이 없음을 나타내었다.

이들 DHA-생산 균주의 개발에는 첫째로 야로위아 리폴리티카 ATCC #20362로부터 유래되어 EPA를 전체 지질에 대하여 9%로 생산할 수 있는 균주 Y2201 및 Y2203의 생성이 요구되었다 (도 5). 균주 Y2201 및 Y2203에는 균주 Y2152 및 Y2153 (DGLA를 ~3.5%로 생산함), 균주 Y2173 (DGLA를 14%로 생산함) 및 균주 Y2189 (EPA를 5%로 생산함)의 구축이 요구되었다.

후속적으로, Y2201 및 Y2203은 Lys-균주이기 때문에, 당업자에게 익히 공지된 방법을 사용하여 pZP2FOEN4 (서열 165)의 뉴클레오티드 위치 12582-11095를 스패닝하는 야로위아 Ura3 유전자 (진뱅크 관리 번호 AJ306421)를 야로위아 Lys5 유전자 (진뱅크 관리 번호 M34929)로 교환하여 플라스미드 pZP2FOEN4-Lys를 생성하는 것이 필요하였다. 이러한 조작 후에, 균주 Y2201 및 Y2203은 상기 실시예 6에 기재된 바와 같이 플라스미드 pZP2FOEN4-Lys로 형질전환되어, ω-6 Δ9 엘론가제/Δ8 데새투라제 경로를 사용하여 DHA를 생산하는 형질전환체 균주가 생성되었다.

DGLA를 전체 지질의 약 ~3.5%로 생산하는 균주 Y2152 및 Y2153의 생성

구축물 pZP2C16M899 (도 14A, 서열 172)를 사용하여 4개의 키메라 유전자 (2개의 Δ9 엘론가제, 합성 $C_{16/18}$ 지방산 엘론가제 및 Δ8 데새투라제 포함), 뿐만 아니라 단일 아미노산 돌연변이를 함유한 야로위아 AHAS 유전자 (아세토히드록시-산 신타제)의 클러스터를 통합시켰다. 야로위아 중의 돌연변이된 AHAS 효소는 양성 스크리닝 마커로서 사용되는 술포닐우레아 내성을 부여하였다. 플라스미드 pZP2C16M899는 야로위아 균주 ATCC #20362의 Pox2 유전자 부위 내로 통합되도록 디자인하고, 따라서 하기 성분을 함유하였다:

[표 26]

플라스미드 pZP2C16M899 (서열 172)의 기재	
서열 172 중의 RE 부위 및 뉴클레오티드	단편 및 키메라 유전자 성분의 기재
BsiWI/AscI (6152-6962)	야로위아 Aco2 유전자 (진뱅크 관리 번호 AJ001300)의 810 bp 5' 부분
SphI/EcoRI (9670-10325)	야로위아 Aco2 유전자 (진뱅크 관리 번호 AJ001300)의 655 bp 3' 부분
BsiWI/PmeI와 함께 EcoRV (929-3195)	· GPM/FBAIN: GPM::FBAIN 키메라 프로모터 (서열 219) · rEL02S: 코돈-최적화된 rEL02 엘론가제 유전자 (서열 85) (래트 (진뱅크 관리 번호 AB071986)에서 유래됨) · OCT: 야로위아 OCT 유전자 (진뱅크 관리 번호 X69988)의 OCT 종결자 서열 을 포함하는 GPM/FBA인트론::rEL02S::Oct
BsiWI/EcoRI (929-14447, 역방향)	· GPAT: GPAT 프로모터 (서열 216) · IgD9e: 코돈-최적화된 Δ9 엘론가제 유전자 (서열 71) (이소크리시스 갈바나에서 유래됨) · Pex20: 야로위아 Pex20 유전자 (진뱅크 관리 번호 AF054613)의 Pex20 종결자 서열 을 포함하는 GPAT::IgD9e::Pex20
EcoRI/SwaI (14447 -12912)	· TEF: TEF 프로모터 (진뱅크 관리 번호 AF054508) · IgD9e: 서열 71 (상기 문헌) · Lip1: 야로위아 Lip1 유전자 (진뱅크 관리 번호 Z50020)의 Lip1 종결자 서열 을 포함하는 TEF::IgD9e::Lip1
SwaI/PacI (12912-10325)	· FBAIN: FBAIN 프로모터 (서열 214) · D8SF: 코돈-최적화된 Δ8 데새투라제 유전자 (서열 81) (유글레나 그라실리스 (진뱅크 관리 번호 AF139720)에서 유래됨) · Pex16: 야로위아 Pex16 유전자 (진뱅크 관리 번호 U75433)의 Pex16 종결자 서열 을 포함하는 FBAIN::D8SF::Pex16
PmeI와 함께 EcoRV/BsiWI (3195-6152)	W497L 돌연변이를 포함하는 야로위아 리폴리티카 AHAS 유전자 (서열 243)

플라스미드 pZP2C16M899를 SphI/AscI로 소화시킨 후에, 이것을 사용하여 ATCC #20362를 통상의 방법에 따라 형질전환시켰다. 형질전환 후, 세포를 술포닐우레아 150 mg을 함유한 MM 플레이트에 플레이팅하고, 30℃에서 2 내지 3일 동안 유지시켰다. 술포닐우레아 내성 콜로니를 골라내고 술포닐우레아 함유 MM 선별 플레이트 상에 스트리킹하였다. 총 96개 형질전환체를 이어서 술포닐우레아를 함유한 액체 MM에 30℃에서 접종시키고, 250 rpm/분에서 2일 동안 진탕시켰다. 세포를 원심분리로 수집하고 지질을 추출하고 에스테르교환반응으로 지방산 메틸 에스테르를 제조한 후에 휴렛-팩커드 6890 GC로 분석하였다.

GC 분석은 pZP2C16M899의 4개 키메라 유전자를 함유한 형질전환체에서 DGLA의 존재를 보였지만, 야생형 야로위아 대조군 균주에서는 보이지 않았다. 선별된 96개 균주 중 대부분은 전체 지질의 2% 미만으로 DGLA를 생산하였다. DGLA를 전체 지질의 2 내지 2.9%로 생산하는 균주가 28개였다. DGLA를 전체 지질의 3.5%로 생산하는 균주가 2개였다. 균주 #65 및 #73을 본원에서 균주 "Y2152" 및 "Y2153"이라고 각각 칭하였다.

DGLA를 전체 지질의 약 14 내지 16%로 생산하는 균주 Y2173 및 Y2175의 생성

구축물 pDMW314 (도 14B, 서열 173)을 사용하여 4개의 키메라 유전자 (2개의 Δ9 엘론가제, Δ8 데새투라제 및 Δ12 데새투라제 포함)의 클러스터를 야로위아 균주 Y2152 및 Y2153의 Ura3 유전자 부위 내로 통합하여 DGLA의 생성을 증대시켰다. 플라스미드 pDMW314는 하기 성분을 함유하였다:

[표 27]

플라스미드 pDMW314 (서열 173)의 기재	
서열 173 중의 RE 부위 및 뉴클레오티드	단편 및 키메라 유전자 성분의 기재
AscI/BsiWI (10066-9275)	야로위아 Ura3 유전자 (진뱅크 관리 번호 AJ306421)의 784 bp 5' 부분
SphI/PacI (12774-1)	야로위아 Ura3 유전자 (진뱅크 관리 번호 AJ306421)의 516 bp 3' 부분
SwaI/BsiWI (6582-9275)	· FBAIN: FBAIN 프로모터 (서열 214) · F.Δ12: 푸사륨 모닐리포르메 Δ12 데새투라제 유전자 (서열 32) · Pex20: 야로위아 Pex20 유전자 (진뱅크 관리 번호 AF054613)로부터의 Pex20 종결자 서열 을 포함하는 FBAIN::F.D12S::Pex20
ClaI/EcoRI (6199-4123)	GPAT::IgD9e::Pex20: pZP2C16M899 (상기 참조)에 기재된 바와 같음
EcoRI/SwaI (4123-2588)	TEF::IgD9E::Lip1: pZP2C16M899 (상기 참조)에 기재된 바와 같음
SwaI/PacI (2588-1)	FBAIN::D8SF::Pex16: pZP2C16M899 (상기 참조)에 기재된 바와 같음

플라스미드 pDMW314를 AscI/SphI로 소화시킨 후에, 이를 야로위아 리폴리티카 균주 Y2152 및 Y2153의 형질전환을 위해 통상의 방법에 따라 사용하였다. 형질전환체 세포를 FOA 선별 배지 플레이트 상에 플레이팅하고, 30℃에서 2 내지 3일 동안 유지시켰다. FOA 내성 콜로니를 골라내고 MM 및 MMU 선별 플레이트 상에 스트리킹하였다. MMU 플레이트에는 성장할 수 있지만, MM 플레이트에서는 성장할 수 없는 콜로니를 Ura^- 균주로서 선별하였다. Ura^- 균주의 단일 콜로니를 이어서 액체 MMU 내에 30℃에서 접종하고, 250 rpm/분에서 2일 동안 진탕시켰다. 세포를 원심분리로 수집하고 지질을 추출하고 에스테르교환반응으로 지방산 메틸 에스테르를 제조한 후에 휴렛-팩커드 6890 GC로 분석하였다.

GC 분석은 pDMW314의 4개 키메라 유전자를 함유한 거의 모든 형질전환체에서 DGLA의 증가된 생성을 보여주었다. pDMW314를 갖는 Y2152의 선별된 Ura^- 균주 48개 중 대부분은 DGLA를 전체 지질의 약 6 내지 8%로 생산하였다. DGLA를 전체 지질의 약 13.9%로 생산하는 균주는 1개였다 (즉, #47은 본원에서 "Y2173"이라고 칭함).

유사하게, pDMW314를 갖는 Y2153의 선별된 24개 Ura^- 균주 중 대부분은 DGLA를 전체 지질의 약 6 내지 8%로 생산하였다. DGLA를 전체 지질의 약 16.3% 및 17.2%로 생산하는 균주는 2개였다 (즉, #6 및 #11은 본원에서 균주 "Y2175" 및 "Y2176"이라고 칭함).

EPA를 전체 지질의 약 4.8%로 생산하는 균주 Y2189의 생성

구축물 pDMW325 (도 14C, 서열 174)를 사용하여 4개의 키메라 유전자 (2개의 Δ5 데새투라제 및 2개의 Δ17 데새투라제 포함)의 클러스터를 야로위아 Y2173 균주의 Leu2 유전자 부위 내로 통합하여 EPA를 생성할 수 있었다. 플라스미드 pDMW325는 하기 성분을 함유하였다:

[표 28]

플라스미드 pDMW325 (서열 174)의 기재	
서열 174 중의 RE 부위 및 뉴클레오티드	단편 및 키메라 유전자 성분의 기재
AscI/BsiWI (4837-5632)	야로위아 Leu2 유전자 (진뱅크 관리 번호 AF260230)의 788 bp 5' 부분
SphI/PacI (2137-1426)	야로위아 Leu2 유전자 (진뱅크 관리 번호 AF260230)의 703 bp 3' 부분
SwaI와 함께 PmeI/BsiWI (8277-5632)	· FBAIN: FBAIN 프로모터 (서열 214) · MAΔ5: 모르티에렐라 알피나 Δ5 데새투라제 유전자 (서열 6) (진뱅크 관리 번호 AF067654) · Pex20: 야로위아 Pex20 유전자 (진뱅크 관리 번호 AF054613)의 Pex20 종결자 서열 을 포함하는 FBAIN::MAΔ5::Pex20
EcoRI/SwaI와 함께 PmeI (10876-8278)	· GPM/FBAIN: GPM::FBAIN 키메라 프로모터 (서열 219) · I.Δ5S: 코돈-최적화된 Δ5 데새투라제 유전자 (서열 10) (이소크리시스 갈바나 (WO 2002/081668)에서 유래됨) · OCT: 야로위아 OCT 유전자 (진뱅크 관리 번호 X69988)의 OCT 종결자 서열 을 포함하는 GPM/FBAIN::I.Δ5S::Oct
EcoRI/PmeI (10876-12497)	야로위아 Ura3 유전자 (진뱅크 관리 번호 AJ306421)
PmeI/ClaI (12497-14651)	· YAT: YAT1 프로모터 (서열 217) · Δ17S: 코돈-최적화된 Δ17 데새투라제 유전자 (서열 21) (사프롤레그니아 디클리나에서 유래됨) · Lip2: 야로위아 리파제2 유전자 (진뱅크 관리 번호 AJ012632)의 Lip2 종결자 을 포함하는 YAT::D17S::Lip2
ClaI/PacI (14651-1426)	· GPD: GPD 프로모터 (서열 210) · Δ17S: 서열 21 (상기 참조) · Pex16: 야로위아 Pex16 유전자 (진뱅크 관리 번호 U75433)의 Pex16 종결자 서열 을 포함하는 GPD::D17S::Pex16

플라스미드 pDMW325를 AscI/SphI로 소화시킨 후에, 이를 사용하여 균주 Y2173을 통상의 방법에 따라 형질전환시켰다. 형질전환 후, 세포를 MMLe 플레이트 상에 플레이팅하고, 30℃에서 2 내지 3일 동안 유지시켰다. 각각의 형질전환으로부터 MMLe 플레이트 상에 성장한 개개의 콜로니를 골라내고 MM 및 MMLe 플레이트 상에 스트리킹하였다. MMLe 플레이트 상에 성장할 수 있지만 MM 플레이트 상에 성장할 수 없는 이들 콜로니를 $Leu2^-$ 균주로서 선별하였다. $Leu2^-$ 균주의 단일 콜로니를 이어서 액체 MMLe 배지에 30℃에서 접종하고, 250 rpm/분에서 2일 동안 진탕시켰다. 세포를 원심분리로 수집하고 지질을 추출하고 에스테르교환반응으로 지방산 메틸 에스테르를 제조한 후에 휴렛-팩커드 6890 GC로 분석하였다.

GC 분석은 pDMW325 형질전환체에 EPA가 존재하지만, 모 Y2173 균주에는 존재하지 않음을 보여주었다. 구체적으로, pDMW325를 갖는 Y2173의 48개 선별된 $Leu2^-$ 형질전환체 중에, 대부분의 균주가 EPA를 전체 지질의 3% 미만으로 생산하였다. EPA를 각각 전체 지질의 약 4.8% 및 3.4%로 생산하는 균주가 2개 (즉, #21 및 #46, 본원에서 "Y2189" 및 "Y2190"이라고 칭함) 있었다.

EPA를 전체 지질의 약 9%로 생산하는 균주 Y2201 및 Y2203의 생성

구축물 pZKSL5598 (도 14D, 서열 175)을 사용하여 4개의 키메라 유전자 (Δ9 엘론가제, Δ8 데새투라제 및 2개의 Δ5 데새투라제 포함)의 클러스터를 야로위아 Y2189 균주의 Lys5 유전자 (진뱅크 관리 번호 M34929) 부위 내에 통합하여 EPA의 생성을 증대시켰다. 플라스미드 pZKSL5598은 하기 성분을 함유하였다:

[표 29]

플라스미드 pZKSL5598 (서열 175)의 기재	
서열 175 중의 RE 부위 및 뉴클레오티드	단편 및 키메라 유전자 성분의 기재
AscI/BsiWI (10409-9573)	야로위아 Lys5 유전자 (진뱅크 관리 번호 M34929)의 794 bp 5' 부분
SphI/PacI (13804-13117)	야로위아 Lys5 유전자 (진뱅크 관리 번호 M34929)의 687 bp 3' 부분
BsiWI/SwaI (7150-9573)	· NT: YAT1 프로모터 (서열 217) · I.Δ5S: 코돈-최적화된 Δ5 데새투라제 유전자 (서열 10) (이소크리시스 갈바나 (WO 2002/081668)에서 유래됨) · Lip1: 야로위아 Lip1 유전자 (진뱅크 관리 번호 Z50020)의 Lip1 종결자 서열 을 포함하는 NT::I.D5S::Lip1
SalI/BsiWI (4537-7150)	· GPAT: GPAT 프로모터 (서열 216) · MAΔ5: 모르티에렐라 알피나 Δ5 데새투라제 유전자 (서열 6) (진뱅크 관리 번호 AF067654) · Pex20: 야로위아 Pex20 유전자 (진뱅크 관리 번호 AF054613)로부터의 Pex20 종결자 서열 을 포함하는 GPAT::MAΔ5::Pex20
SwaI/PmeI (2381-348)	· FBAINm: FBAINm 프로모터 (서열 215) · IgD9e: 코돈-최적화된 Δ9 엘론가제 유전자 (서열 71) (이소크리시스 갈바나에서 유래됨) · OCT: 야로위아 OCT 유전자 (진뱅크 관리 번호 X69988)의 OCT 종결자 서열 을 포함하는 FBAINm::IgD9e::OCT
ClaI/PacI (1-13804)	· GPD: GPD 프로모터 (서열 210) · D8SF: 코돈-최적화된 Δ8 데새투라제 유전자 (서열 81) (유글레나 그라실리스 (진뱅크 관리 번호 AF139720)에서 유래됨) · Pex16: 야로위아 Pex16 유전자 (진뱅크 관리 번호 U75433)의 Pex16 종결자 서열 을 포함하는 GPD::D8SF::Pex16
SalII/PmeI (4537 -2417)	야로위아 Leu2 유전자 (진뱅크 관리 번호 AF260230)

플라스미드 pZKSL5598을 AscI/SphI로 소화시킨 후에, 이를 사용하여 균주 Y2189을 통상의 방법에 따라 형질전환시켰다. 형질전환 후, 세포를 MMLys 플레이트 상에 플레이팅하고, 30℃에서 2 내지 3일 동안 유지시켰다. 각각의 형질전환으로부터 MMLys 플레이트 상에 성장한 개개의 콜로니를 골라내고 MM 및 MMLys 플레이트 상에 스트리킹하였다. 플레이트 상에 성장할 수 있지만 MM 플레이트 상에서는 성장할 수 없는 이들 콜로니를 Lys^- 균주로서 선별하였다. Lys^- 균주의 단일 콜로니를 이어서 액체 MMLys 배지에 30℃에서 접종하고, 250 rpm/분에서 2일 동안 진탕시켰다. 세포를 원심분리로 수집하고 지질을 추출하고 에스테르교환반응으로 지방산 메틸 에스테르를 제조한 후에 휴렛-팩커드 6890 GC로 분석하였다.

GC 분석은 pZKSL5598 형질전환체에서 EPA의 증가된 생성을 나타냈다. pZKSL5598을 갖는 Y2189의 96개 선별된 Lys^- 형질전환체 중에서, 대부분의 균주는 EPA를 전체 지질의 4 내지 8%로 생산하였다. EPA를 각각 전체 지질의 약 9% 및 8.7%로 생산하는 균주가 2개 있었다 (즉, #34 및 #77, 본원에서 "Y2201" 및 "Y2203"이라고 칭함).

실시예 12

ω-3 Δ9 엘론가제/Δ8 데새투라제 경로:

야로위아 리폴리티카에서 DHA를 생산하는 균주의 생성 (예언적)

본 실시예는 야로위아 리폴리티카 ATCC #20362로부터 유래되어 ω-3 Δ9 엘론가제/Δ8 데새투라제 경로의 발현을 통해 DHA를 생산하도록 유전자 조작된 균주의 구축을 기재하고 있으며, 따라서 이들 DHA-생산 균주의 완전한 지질 프로파일의 분석은 최종 DHA-함유 오일에서 GLA 동시-합성이 없음을 나타내었다.

이들 DHA-생산 균주의 발달에는 첫째로 야로위아 리폴리티카 ATCC #20362로부터 유래되어 EPA를 전체 지질에 대하여 1.3%로 생산할 수 있는 균주 L116의 생성이 요구되었다 (도 5). 균주 L116에는 균주 L98 (ALA를 생산함), 균주 L103 (증가된 ALA를 생산함) 및 균주 L115 (ETA를 약 4%로 생산함)의 구축이 요구되었다. 추가로, 균주 L116은 오직 (또는 강력한) ω-3 특이성을 갖는 것으로 본원에서 특성화된, 다니오 레리오 (진뱅크 관리 번호 BC068224)로부터 유래된 신규 이작용성 Δ5/Δ6 데새투라제의 합성 및 발현이 요구되었다.

후속적으로, L116 Ura^- 균주는 pKUT16을 갖는 L116에서 Ura3 유전자의 파괴에 의해 생성된 다음 상기 실시예 6에 기재된 바와 같이 플라스미드 pZP2FOEN4로 형질전환되어, ω-3 Δ9 엘론가제/Δ8 데새투라제 경로를 사용하여 DHA를 생산하는 형질전환체 균주가 생성되었다.

재순환가능 선별을 위한 LoxP::Ura3/HPT::LoxP 통합 구축물 및 Cre-SU 복제 플라스미드의 생성

Δ15 데새투라제의 다중 카피를 야로위아 리폴리티카 내로 도입하기 위해 재순환가능 선별 마커 및 부위-특이적 재조합 시스템 (즉, Cre/Lox)에 기반한 전략을 사용하였다. 간략하게 설명하면, 표적 유전자 (즉, 푸사륨 모닐리포르메 Δ15 데새투라제 [서열 51])를 통합 구축물에서 LoxP 부위에 의해 플랭킹되는 선별 마커 (예컨대, Ura3 및 하이그로마이신 포스포트랜스퍼라제 [HPT])에 인접하게 클로닝하였다. 형질전환 및 형질전환체의 선별 후에, 선별 마커는 술포닐우레아 내성 (SU) 유전자 및 Cre 레콤비나제 유전자를 갖는 복제 플라스미드의 도입에 의해 염색체로부터 제거하였다. 선별 마커의 손실 후에, Cre 플라스미드를 치유하였다. 따라서, 치유된 균주는 또다른 형질전환에 사용가능하였다.

보다 구체적으로, 플라스미드 pY72 (도 15A, 서열 176)는 1개 카피의 푸사륨 모닐리포르메 Δ15 데새투라제 및 Ura3/HPT 선별 마커 (LoxP 부위에 의해 플랭킹됨)를 포함하는 통합 구축물이었다. 구축물 pY72는 하기 성분을 함유하였다:

[표 30]

플라스미드 pY72 (서열 176)의 기재	
서열 176 중의 RE 부위 및 뉴클레오티드	단편 및 키메라 유전자 성분의 기재
6763-7643	야로위아 Lip1 유전자 (진뱅크 관리 번호 Z50020)의 881 bp 5' 부분
9422-10184	야로위아 Lip1 유전자 (진뱅크 관리 번호 Z50020)의 763 bp 3' 부분
SwaI/SbfI (16-2522)	· FBAIN: FBAIN 프로모터 (서열 214) · FmD15: 푸사륨 모닐리포르메 Δ15 데새투라제 유전자 (서열 51) . Lip2: 야로위아 Lip2 유전자 (진뱅크 관리 번호 AJ012632)로부터의 Lip2 종결자 서열 을 포함하는 FBAIN::FmD15:Lip2
2531-2564	LoxP 서열 (서열 382)
2566-4184	야로위아 Ura3 유전자 (진뱅크 관리 번호 AJ306421)

4198-5861	· TEF: TEF 프로모터 (진뱅크 관리 번호 AF054508) · HPT: 히그로마이신 내성을 부여하는 에스케리키아 콜라이 히그로마이신 포스포트랜스퍼라제 코딩 영역 [Kaster, K.R., et al., Nucleic Acids Res. 11 :6895-6911 (1983)] · XPR: 야로위아 Xpr 유전자 (진뱅크 관리 번호 M17741)의 3' 영역의 ~100 bp 를 포함하는 TEF::HPT::XPR
5862-5895	LoxP 서열 (서열 382)

유사하게, 플라스미드 pY80 (도 15B, 서열 177)을 사용하여 2개 카피의 푸사륨 모닐리포르메 Δ15 데새투라제 및 Ura3/HPT 선별 마커 (LoxP 부위에 의해 플랭킹됨)를 포함하는 통합 구축물을 생성하였다. 프라이머 436 및 437 (서열 383 및 384)을 사용하며, PCR을 사용하여 8878 bp 플라스미드 pY34 (WO 2005/047480)로부터의 GPD::Fm1::XPR2를 포함하는 PacI/FseI 단편을 증폭하였다. 이 PacI/FseI 단편을 PacI/FseI-소화 벡터 pY72 내로 인-퓨전 클로닝 (클론테크 래버러토리즈, 인크., 미국 캘리포니아주 마운틴 뷰 소재)에 의해 클로닝하고, XL-2 울트라 감응성 세포 (BRL, 미국 메릴랜드주 베쎄스다 소재) 내로 형질전환시켰다. 10개 양성 형질전환체를 PacI/FseI 소화 후에 미니프렙 분석에 의해 동정하고, 단지 클론 #3 및 #4가 올바른 것이었다. 올바른 클론 중 하나를 "pY80"이라 칭하였다. 따라서, 구축물 pY80은 하기 성분을 함유하였다:

[표 31]

플라스미드 pY80 (서열 177)의 기재	
서열 177 중의 RE 부위 및 뉴클레오티드	**단편 및 키메라 유전자 성분의 기재**
PacI/FseI (4-2375)	· GPD: GPD 프로모터 (서열 210) · FmD15: 푸사륨 모닐리포르메 Δ15 데새투라제 유전자 (서열 51) · XPR: 야로위아 Xpr 유전자 (진뱅크 관리 번호 M17741)의 3' 영역의 ~100 bp 을 포함하는 GPD::FmD15:XPR
FseI/SbfI 2385-4891	FBAIN::FmD15:Lip2: pY72 (상기 참조)에 기재된 바와 같음
4900-4933	LoxP 서열 (서열 382)
4935-6533	야로위아 Ura3 유전자 (진뱅크 관리 번호 AJ306421)
6567-8230	TEF::HPT::XPR: pY72 (상기 참조)에 기재된 바와 같음
8231-8264	LoxP 서열 (서열 382)
8271-9079	야로위아 Lip1 유전자 (진뱅크 관리 번호 250020)의 809 bp 5' 부분
11791-12553	야로위아 Lip1 유전자 (진뱅크 관리 번호 250020)의 763 bp 3' 부분

구축물 pY79 (도 15C, 서열 178)는 술포닐우레아 내성 (SU) 유전자 (즉, AHAS) 및 TEF::Cre 레콤비나제 유전자를 갖는 복제 플라스미드이었다. 구체적으로, 구축물 pY79는 하기 성분을 함유하였다:

[표 32]

플라스미드 pY79 (서열 178)의 기재	
RE 부위 및 뉴클레오티드 서열 178	**단편 및 키메라 유전자 성분의 기재**
4329-7315	W497L 돌연변이체 (서열 243)를 포함한 야로위아 리폴리티카 AHAS 유전자

7362-1	· TEF: TEF 프로모터 (진뱅크 관리 번호 AF054508) · Cre: 리컴비나제 단백질에 대한 장내세균 파지 P1 Cre 유전자 (진뱅크 관리 번호 X03453) · XPR: 야로위아 Xpr 유전자의 3' 영역의 ~100 bp (진뱅크 관리 번호 M17741) 을 포함하는 TEF::Cre::XPR

ALA를 생산하는 균주 L98의 생성

플라스미드 pY72 (서열 176)를 AscI/SphI로 소화시킨 후에, 이것으로 표준 리튬 아세테이트 방법을 사용하여 야생형 야로위아 리폴리티카 ATCC #20362를 형질전환시켰다. 형질전환 후, 세포를 YPD+ 하이그로마이신 (250 μg/mL) 플레이트 상에 플레이팅하였다. 2일 후에, 20개 형질전환체를 골라내고 신선한 YPD+ 하이그로마이신 (250 μg/mL) 플레이트 상에 스트리킹하고, 30℃에서 밤새 인큐베이션하였다. 세포를 원심분리로 수집하고 지질을 추출하고 에스테르교환반응으로 지방산 메틸 에스테르를 제조한 후에 휴렛-팩커드 6890 GC로 분석하였다.

GC 분석은 pY72 형질전환체에서 ALA가 존재하지만, 야생형 야로위아 대조군 균주에서는 존재하지 않았음을 보여주었다. 최고 클론은 ALA를 전체 지질의 약 27%로 생산하고, 80%의 기질 전환율을 나타냈다.

ATCC #20362/pY72 형질전환체를 pY79 (서열 178, 술포닐우레아 (SU) 내성 마커를 가짐)로 형질전환시키고, 형질전환체를 3일 동안 MM + SU (150 μg/mL) 플레이트 상에서 선별함으로써 pY72 중에서 LoxP 부위에 의해 플랭킹된 Ura3/HPT 마커를 게놈으로부터 제거하였다. SU-내성 (SU^R) 형질전환체를 신선한 MM + SU (150 μg/mL) 플레이트 상에 1일 동안 재스트리킹하고, 이어서 YPD+ 하이그로마이신 (250 μg/mL) 플레이트 상에 레플리카-플레이팅하였다. 모든 클론 (클론 #1 제외)은 하이그로마이신 (Hyg^s)에 감수성이어서 이는 HPT 내성 유전자가 Cre 레콤비나제에 의해 성공적으로 제거되었음을 나타낸다.

플라스미드 pY79는 세포를 YPD에서 선별 없이 30℃에서 밤새 성장시킴으로써 Hyg^s 클론 #6 및 #14로부터 치유되었다. 배양물 (0.1 mL)을 YPD 1 mL에 희석하고, 이를 사용하여 가장 높은 희석이 20,000배가 되도록 계단 희석하였다. 각각의 희석을 이어서 새로운 YPD 플레이트 상에 플레이팅하고, 30℃에서 밤새 인큐베이션하였다. 플레이트를 MM + SU (150 μg/mL) 플레이트 상에 레플리카-플레이팅하였다. 모든 클론은 SU-감수성 (SU^S)이어서, 이는 이들이 pY79를 성공적으로 치유하였음을 나타낸다. 클론 #6-1을 추가의 형질전환에 대해 사용하였다.

구체적으로, 상기 방법을 사용하여 플라스미드 pY80 (서열 177)을 AscI/SphI로 소화시킨 후에, 이를 사용하여 균주 #6-1을 형질전환시켰다. YPD+ 하이그로마이신 (250 μg/mL) 플레이트 상의 선별, 전체 지질의 GC 분석, 플라스미드 pY79 (서열 178)로의 형질전환, SU^R 및 Hyg^s 클론의 동정, 및 플라스미드 pY79의 치유 후에, 균주 #1을 동정하였다. 이에 의해 이 균주는 3개 카피의 FmΔ15를 가지고, LA에서 ALA로의 96.1%의 기질 전환율을 가졌다.

균주 #1을 pY80 및 이어서 pY79로 상기 기재한 바와 같이 형질전환시켰다. 이는 불충분한 기질 (즉, LA)의 결과로서 5개 카피의 FmΔ15를 갖는 균주 L98을 생성하였지만, 이 균주 중의 Δ15 탈포화는 균주 #1 (3개 카피의 FmΔ15를 가짐)에 상대적으로 유의하게 개선되지 않았다.

증가된 ALA를 생산하는 균주 L103의 생성

플라스미드 pY86 (도 15D, 서열 179)을 1개 카피의 푸사륨 모닐리포르메 Δ12 데새투라제 및 Ura3/HPT 선별 마커 (LoxP 부위에 의해 플랭킹됨)를 포함하는 통합 구축물이었다. 구체적으로, pY86은 하기 성분을 함유하였다:

[표 33]

플라스미드 pY86 (서열 179)의 기재

서열 179 중의 RE 부위 및 뉴클레오티드	단편 및 키메라 유전자 성분의 기재
3399-4207	야로위아 Lip1 유전자의 809 bp 5' 부분 (진뱅크 관리 번호 Z50020)
6919-7681	야로위아 Lip1 유전자의 763 bp 3' 부분 (진뱅크 관리 번호 Z50020)
28-61	LoxP 서열 (서열 382)
63-1681	야로위아 Ura3 유전자 (진뱅크 관리 번호 AJ306421)
1695-3358	· TEF: TEF 프로모터 (진뱅크 관리 번호 AF054508) · HPT: 히그로마이신 내성을 부여하는 에스케리키아 콜라이 히그로마이신 포스포트랜스퍼라제 코딩 영역 [Kaster, K.R., et al., Nucleic Acids Res. 11 :6895-6911 (1983)] · XPR: 야로위아 Xpr 유전자 (진뱅크 관리 번호 M17741)의 3' 영역의 ~100 bp 를 포함하는 TEF::HPT::XPR
3359-3392	LoxP 서열 (서열 382)
PacI/FseI (7690-7)	· FBAIN: FBAIN 프로모터 (서열 214) · FmD12: 모르티에렐라 이자벨리나 Δ12 데새투라제 유전자 (서열 32) · Lip2: 야로위아 Lip2 유전자 (진뱅크 관리 번호 AJ012632)의 Lip2 종결자 서열 을 포함하는 FBAIN::FmD12::Lip2

상기 방법으로, 플라스미드 pY86을 AscI/SphI로 소화시킨 후에, 이것으로 균주 L98을 형질전환시켰다. YPD+ 하이그로마이신 (250 ㎍/mL) 플레이트 상의 선별, 전체 지질의 GC 분석, 플라스미드 pY79 (서열 178)를 사용하는 형질전환 및 SU^R 및 Hyg^S 클론의 동정 후에, 균주 L103을 동정하였다. 이에 의해 이 균주는 5개 카피의 FmΔ15, 1개 카피의 FmΔ12를 가졌고, Ura3-이었다. 균주 L98에 상대적으로, 균주 L103 중의 18:1의 양 (전체 지방산의 백분율로서)은 42%에서 약 10%로 저하되었고, 균주 L103 중의 18:2의 양 (전체 지방산의 백분율로서)은 2%에서 약 10%로 증가하고, 균주 L103 중의 ALA의 양 (전체 지방산의 백분율로서)은 22%에서 47%로 증가하였다.

<u>ETA를 전체 지질의 약 4%로 생산하는 균주 L115의 생성</u>

플라스미드 pY94 (도 16A, 서열 180)는 1개 카피의 Δ8 데새투라제, 1개 카피의 Δ9 엘론가제, 및 Ura3 선별 마커 (LoxP 부위에 의해 플랭킹됨)를 포함하는 통합 구축물이었다. 이 플라스미드는 하기 성분을 함유하였다:

[표 34]

플라스미드 pY94 (서열 180)의 기재	
서열 180 중의 RE 부위 및 뉴클레오티드	단편 및 키메라 유전자 성분의 기재
PacI/SwaI (1-2587)	· FBAIN: FBAIN 프로모터 (서열 214) · D8: 코돈-최적화된 Δ8 데새투라제 유전자 (서열 81) (유글레나 그라실리스 (진뱅크 관리 번호 AF139720)에서 유래됨) · Pex16: 야로위아 Pex16 유전자 (진뱅크 관리 번호 U75433)의 Pex16 종결자 서열 을 포함하는 FBAIN::D8:Pex16
2592-4684	· GPAT: GPAT 프로모터 (서열 216) · D9E: 코돈-최적화된 Δ9 엘론가제 유전자 (서열 71) (이소크리시스 갈바나에서 유래됨) · Lip1: 야로위아 Lip1 유전자 (진뱅크 관리 번호 Z50020)의 Lip1 종결자 서열 을 포함하는 GPAT::D9E::Lip1
4714-4747	LoxP 서열 (서열 382)
4761-6378	야로위아 Ura3 유전자 (진뱅크 관리 번호 AJ306421)

6380-6413	LoxP 서열 (서열 382)
6470-7253	야로위아 Ura3 유전자 (진뱅크 관리 번호 AJ306421)의 784 bp 5' 부분
9965-10480	야로위아 Ura3 유전자 (진뱅크 관리 번호 AJ306421)의 516 bp 3' 부분

표준 리튬 아세테이트 방법을 이용하여 플라스미드 pY94로 균주 L103을 형질전환시켰다. 형질전환 후, 상기 세포를 MM 플레이트에 플레이팅하고 3일 동안 유지시켰다. 이어서, 22개의 콜로니를 골라내고 신선한 MM 플레이트에 스트리킹하여 30℃에서 밤새 성장시켰다. 세포를 원심분리로 수집하고 지질을 추출하고 에스테르교환반응으로 지방산 메틸 에스테르를 제조한 후에 휴렛-팩커드 6890 GC로 분석하였다. 클론 #8 (본원에서는 균주 L104라고 칭함)은 가장 높은 Δ9 엘론가제 및 Δ8 데새투라제 기질 전환율(%)을 가졌다.

균주 L104의 대수증식기 세포를 pY79 (서열 178) 1 ㎕ (약 0.5 ㎍/㎕)로 형질전환시키고, 형질전환체를 4일 동안 MMU + SU (100 ㎍/mL) 플레이트에서 선별함으로써 게놈으로부터 pY94에서 LoxP 부위에 의해 플랭킹된 Ura3 마커를 잘라냈다. 12개 SU^R 형질전환체를 신선한 MM 및 MMU 플레이트에서 2일 동안 재스트리킹하였다. 모든 클론 (1개는 제외)은 URA 영양요구성 (즉, Ura^S)이므로, 이는 Ura3 내성 유전자를 Cre 레콤비나제에 의해 성공적으로 잘라냈음을 나타낸다.

세포의 1/3 백금이로부터 MMU 중에 1:10,000 내지 1:50,000 희석함으로써 플라스미드 pY79를 1개 URA 영양요구균으로부터 치유하였다. 희석물 (100 ㎕/플레이트)을 YPD 플레이트 상에 플레이팅하고, 30℃에서 2일 동안 인큐베이션하였다. 8개 콜로니를 YPD 플레이트로부터 골라내고, MMU 플레이트 및 MMU + SU 플레이트 상에 스트리킹하고, 30℃에서 24 시간 동안 인큐베이션하였다. 모든 클론은 SU-감수성 (SU^S)이어서, 이는 이들이 pY79를 성공적으로 치유하였음을 나타낸다. 이들 중 하나를 L111이라 칭하고, 이에 의해 5개 카피의 FmΔ15, 1개 카피의 FmΔ12, 1개 카피의 Δ8 데새투라제, 1개 카피의 Δ9 엘론가제를 가지고, Ura3-이었다.

균주 L115 (5개 카피의 FmΔ15, 1개 카피의 FmΔ12, 2개 카피의 Δ8 데새투라제, 2개 카피의 Δ9 엘론가제를 보유하고 Ura3-을 특징으로 함)는 균주 L111을 pY94 (서열 180)로 상기 방법을 이용하여 형질전환시킴으로써 생성하였다. GC 분석은 균주 L115가 ETA를 전체 지질 (완전한 지질 프로파일, 하기 참조)의 약 4%로 생산하였음을 보여주었다.

EPA를 전체 지질의 약 1.3%로 생산하는 균주 L116의 생성

진뱅크 관리 번호 AF309556 [Hastings et al., PNAS 98(25): 14304-14309 (2001)]로서 동정된 다니오 레리오 데새투라제는 사카로마이세스 세레비지애에서 하기 이작용성 Δ6 및 Δ5 데새투라제 활성을 나타내는 것으로 보고되었다: (1) ω-6 기질에 비해 ω-3 기질의 별개의 성호성; 및 (2) Δ5 데새투라제 활성에 상대적으로 유의하게 더 높은 Δ6 데새투라제 활성.

본 발명자들은 진뱅크 관리 번호 BC068224를 진뱅크 관리 번호 AF309556의 상동체 (이들은 단지 ORF의 위치 984에서의 1 bp (T) 결실 (무효화 돌연변이를 생성함) 및 위치 1171에서의 1 bp 변화 ('G'에서 'A'로) ('V에서 'M'으로 아미노산을 변화시킴)가 상이함)로서 동정하였다.

V1171M 돌연변이 외에 진뱅크 관리 번호 AF309556 (본원에서 "Drd6/d5(V)"로서 동정됨, 서열 15)에 동일하게 돌연변이체 단백질을 이어서 생성하였다 (본원에서 "Drd6/d5(M)"로서 동정함, 서열 18). 구체적으로, 2개 중복 단편을 먼저 진뱅크 관리 번호 BC068224 cDNA 파지미드로부터 프라이머 쌍들 475 및 477 (서열 385 및 386) 및 478 및 476 (서열 387 및 388) [여기서, 프라이머 477 및 478은 "누락 T"를 가짐]으로 증폭하였다. 이어서, 전체 Drd6/d5(M) ORF는 프라이머 475 및 476 및 주형으로서의 2개 중복 단편을 사용하여 증폭시켰다. ORF는 하기 성분을 함유한 복제 플라스미드 중에 두고, 본원에서 플라스미드 "pY91M" (도 16B)로서 동정하였다:

[표 35]

플라스미드 pY91M (서열 181)의 기재	
서열 181 중의 RE 부위 및 뉴클레오티드	단편 및 키메라 유전자 성분의 기재

2866-4170	ARS18 서열 (진뱅크 관리 번호 A17608)
4216-5703	야로위아 Ura3 유전자 (진뱅크 관리 번호 AJ306421)
SalI/BsiwI (5705-8423)	· FBAIN: FBAIN 프로모터 (서열 214) · DrD6: Drd6/d5(M) 유전자 (서열 17) (다니오 레리오 Δ5/Δ6 데새투라제 (진뱅크 관리 번호 BC068224) 에서 유래됨) · Pex20: 야로위아 Pex20 유전자 (진뱅크 관리 번호 AF054613)의 Pex20 종결자 서열 을 포함하는 FBAIN::DrD6:Pex20

플라스미드 pY91V는 플라스미드 pY91M으로부터 부위-특이적 돌연변이유발에 의해 QuikChange® II 부위-지정 돌연변이유발 키트, (스트라타진, 카탈로그 #200523) 및 프라이머 505 및 506 (서열 389 및 390)을 사용하여 생성하였다. pY91V는 상기 기재한 M에서 V로 아미노산 돌연변이시키는 단일 bp 변화 외에 pY91M과 동일하였다.

플라스미드 pY91M 및 pY91V, 뿐만 아니라 대조군로서 제공된 빈 벡터를 균주 L115의 대수증식기 세포 내로 각각 표준 리튬 아세테이트 방법으로 형질전환시켰다. 형질전환 후, 세포를 MM 플레이트에 플레이팅하고, 3일 동안 유지하였다. 콜로니를 이어서 골라내고, 신선한 MM 플레이트 상에 성장시키고, 30℃에서 밤새 성장시켰다. 각 클론의 세포의 1/3 백금이를 MM 3 mL에 접종하고, 30℃의 진탕기에서 24시간 동안 성장시켰다. 별법으로, 세포를 24 시간 동안 MM에서 성장시키고, 이어서 3일 동안 HGM에서 배양하였다. 모든 세포를 수확하고, 그의 지방산 조성을 이전에 기재한 바와 같이 GC에 의해 분석하였다.

빈 벡터 (대조군), pY91M 및 pY91V로 형질전환시킨 균주 L115 (FmΔ15, FmΔ12, Δ8 데새투라제 및 Δ9 엘론가제 키메라 유전자의 결과로서 ω-3 Δ9 엘론가제/Δ8 데새투라제 경로를 발현함)의 완전한 지질 프로파일을 하기 표 36에 나타낸다. 지방산을 16:0, 16:1, 17:1, 18:0, 18:1 (올레산), 18:2 (LA), GLA, 20:2 (EDA), DGLA, ARA, ALA, STA, 20:3 (ETrA), ETA 및 EPA로서 확인하고; 각각의 조성을 전체 지방산(%)으로서 나타낸다. 3개의 별도 실험을 수행하였으며, 이를 컬럼 "실험 번호"에 실험 번호 1, 2 및 3으로 나타냈다. 추가로, 각각의 균주에 대한 Δ6 및 Δ5 기질 전환율(%)은 ω-6 및 ω-3 기질을 모두 사용하는 활성에 대해 보고한다 (표 37).

[표 36]

pY91 M 및 pY91V로 형질전환시킨 야로위아 리폴리티카 균주 L115의 지질 프로파일

실험 번호	균주	시간/배지	16:0%	16:1%	18:0%	18:1%	18:2%	GLA%	20:2%	DGLA %	ARA%	ALA%	STA%	20:3%	ETA%	EPA%
1	L115 + pY91M (클론 11)	1D MM	16	10	2	8	11	0.0	0	1	0.0	40	3.2	1.1	4.2	1.3
1	L115 + 대조군	1D MM	18	9	5	18	12	0.0	0	1	0.0	31	0.0	0.9	3.9	0.0
2	L115 + pY91M (클론 11)	1D MM/ 3D HGM	14	11	6	26	15	0.5	1	2	0.0	18	2.4	1.0	2.9	0.6
2	L115 + 대조군	1D MM/ 3D HGM	13	11	6	26	15	0.2	1	2	0.0	20	0.0	1.5	4.1	0.2
3	L115 + pY91V (클론 10)	1D MM	17	8	6	20	15	0.0	0	1	0.0	27	0.0	0.9	3.9	0.0
3	L115 + pY91M (클론 11)	1D MM	17	9	3	11	11	0.0	0	0	0.0	38	2.6	1.1	4.4	1.2
3	L115 + c 대조군	1D MM	17	8	6	21	13	0.0	0	1	0.0	28	0.0	1.1	4.0	0.0

* 클론 #11로서 동정된 L115/pY91M 형질전환체를 야로위아 리폴리티카 균주 "L116"이라고 칭하였다.

[표 37]

Drd6/d5M 및 Drd6/d5V에 의한 기질 전환율(%)

실험 번호	균주	ω-3/ω-6	Δ6 (ω-6)	Δ6 (ω-3)	Δ5 (ω-6)	Δ5 (ω-3)
1	L115 + pY91M (클론 11)	4.2	0	7	0	23
1	L115 + 대조군	2.8	0	0	0	0
2	L115 + pY91M (클론 11)	1.3	3	12	0	16
2	L115 + 대조군	1.5	1	0	0	5
3	L115 + pY91V (클론 10)	2.0	0	0	0	0
3	L115 + pY91M (클론 11)	4.4	0	7		21
3	L115 + 대조군	2.5	0	0	0	0

상기 결과에서 입증되는 바와 같이, 야로위아 리폴리티카 (즉, 균주 L115 + pY91M)에서의 Drd6/d5(M)의 발현은 Δ6 데새투라제 활성 (즉, ALA에서 STA로)보다 Δ5 데새투라제 활성 (즉, ETA에서 EPA로)에 대한 더 높은 기질 전환율(%) 및 Δ6 및 Δ5 데새투라제 활성 모두에 대해 매우 높은 ω-3 기질 선호성을 갖는, Δ6 및 Δ5 데새투라제 활성을 모두 갖는 이작용성 효소를 생산하였다. 예상외로, Drd6/d5(V) (즉, 균주 L115 + pY91V)는 ω-6 기질에 대해 Δ6 또는 Δ5 활성을 나타

내지 않은 반면, Drd6/d5(M)은 ω-6 기질에 대해 Δ5 활성을 결실하였다. 따라서, Drd6/d5(M)은 Drd6/d5(V)와 상이한 특징을 가졌다. Drd6/d5(V)와 공개된 구축물과의 활성에서의 차이는 단백질이 발현되는 상이한 숙주 유기체 및/또는 기질의 기원 (즉, 기질 공급 [Hastings et al., 상기 문헌] 또는 기질 생합성 [본원에서 입증됨])과 관련이 있는 것으로 여겨진다.

Drd6/d5(M) 및 Drd6/d5(V)의 기질 특이성의 더 나은 이해를 위해, FBAIN::Drd6/d5(M)::Pex20 및 FBAIN::Drd6/d5(V)::Pex20 키메라 유전자를 야로위아 복제 플라스미드 내로 LEU 선별로 전달하고, 이에 의해 플라스미드 pY102(M) 및 pY102(V)를 각각 생성하였다. 이들 플라스미드를 이어서 균주 Q-d12D인, Δ12 데새투라제 넉아웃 (WO 2004/104167)을 포함하는 야로위아 리폴리티카 균주 내로 형질전환시켰다. 형질전환체를 1일 동안 MM에서 0.5 mM LA, ALA, ETrA [20:3 (11,14,17)], EDA, DGLA 또는 ETA의 존재하에 성장시키고, 기질 전환율(%)을 시험하였다. 결과를 하기 표 38에 나타낸다:

[표 38]

형질전환체 야로위아 균주 Q-d12D 중의
Drd6/d5(M) 및 Drd6/d5(V)에 의한 기질 전환율(%)

플라스미드	지방산	기질 전환율(%)					
		Δ6 (ω-6)	Δ6 (ω-3)	Δ8 (ω-6)	Δ8 (ω-3)	Δ5 (ω-6)	Δ5 (ω-3)
pY102(M)	LA	**17**	--	--	--	--	--
pY102(V)	LA	**4**	--	--	--	--	--
pY102(M)	ALA	--	**24**	--	--	--	--
pY102(V)	ALA	--	**6**	--	--	--	--
pY102(M)	EDA	17	--	**0**	--	--	--
pY102(V)	EDA	0	--	**0**	--	--	--
pY102(M)	ETrA	--	30	--	**13**	--	--
pY102(V)	ETrA	--	9	--	**0**	--	--
pY102(M)	DGLA	--	--	--	--	**12**	--
pY102(V)	DGLA	--	--	--	--	**0**	--
pY102(M)	ETA	--	--	--	--	--	**34**
pY102(V)	ETA	--	--	--	--	--	**0**

상기 결과는 신규 Drd6/d5(M) 데새투라제가 (공개된 Drd6/d5(V) 데새투라제에 비해) (1) 시험된 모든 기질에 대해 더 높은 기질 전환율(%); (2) ω-6 지방산에 비해 ω-3 지방산에 대해 더 높은 선택성 [ω-3 또는 ω-6 기질를 갖는 Drd6/d5(V)에는 Δ5 활성이 없었지만]; 및 (3) 예측되지 않은 Δ8 데새투라제 활성을 가졌음을 보여주었다.

Q-d12D 형질전환체 대 L115 형질전환체 간의 기질 전환율(%)에서의 차이는 기질 공급의 결과로 여겨졌다. Drd6/d5가 아실-CoA 기질 상에 작용하는 것으로 보고되었기 때문에, 데새투라제 활성은 지방산 공급 또는 야로위아 숙주에 의한 새로운 합성의 결과로서 상이할 수 있다. 또다른 예측되지 않은 관찰은 Drd6/d5(M)가 ETrA [20:3 (11, 14, 17)]를 ETA로 전환시키지만, EDA [20:2 (11, 14)]를 DGLA로 전환시키지 않았는데, 즉 단백질이 ω-3 기질에 대해서만 Δ8 데새투라제 활성을 가졌다는 것이었다.

이 신규 Drd6/d5(M) 데새투라제는 야로위아 리폴리티카에서 발현되는 경우에 경로 유전자 조작에 대한 특이한 이점을 제공할 수 있다는 특징을 가지는 것으로 고려된다.

실시예 13

EPA를 전체 지질의 14%로 생산하는, 중간체 균주 Y2067U의 생성

본 실시예는 야로위아 리폴리티카 ATCC #20362로부터 유래되어 EPA를 전체 지질에 대하여 14%로 생산할 수 있는 균주 Y2067U의 구축을 기재한다 (도 5). 이 균주는 ω-6 Δ6 데새투라제/Δ6 엘론가제 경로를 발현하도록 유전자 조작되었

다. 이 EPA 생산 균주에서 모르티에렐라 알피나 LPAAT2, DGAT1 및 DGAT2 및 야로위아 리폴리티카 CPT1 유전자 과발현이 미치는 효과를 실시예 17, 18, 19 및 24 각각 (하기 참조)에 기재한 바와 같은 TAG 함량 및/또는 조성의 분석을 기초로 조사하였다.

균주 Y2067U의 개발에는 균주 M4 (DGLA를 8%로 생성하고, 실시예 6에 기재됨), 균주 Y2034 (ARA를 10%로 생산함), 균주 E (EPA를 10%로 생산함), 균주 EU (EPA를 10%로 생산함) 및 균주 Y2067 (EPA를 15%로 생산함)의 구축이 요구되었다.

ARA를 전체 지질의 약 10%로 생산하는 Y2034 균주의 생성

2개의 Δ5 키메라 유전자를 야로위아 균주 M4의 Leu2 유전자에 통합하여 구축물 pDMW232 (도 16C, 서열 182)를 제조하였다. 플라스미드 pDMW232는 표 39에 나타낸 하기 성분을 함유하였다:

[표 39]

플라스미드 pDMW232 (서열 182)의 기재	
서열 182 내의 RE 부위 및 뉴클레오티드	단편 및 키메라 유전자 성분의 기재
AscI/BsiWI (5550-4755)	야로위아 Leu2 유전자 (진뱅크 관리 번호 AF260230)의 788 bp 5' 부분
SphI/PacI (8258-8967)	야로위아 Leu2 유전자 (진뱅크 관리 번호 AF260230)의 703 bp 3' 부분
SwaI/BsiWI (2114-4755)	· FBAIN: FBAIN 프로모터 (서열 214) · MAΔ5: 모르티에렐라 알피나 Δ5 데새투라제 유전자 (서열 6) (진뱅크 관리 번호 AF067654) · Pex20: 야로위아 Pex20 유전자 (진뱅크 관리 번호 AF054613)의 Pex20 종결자 서열 을 포함하는 FBAIN::MAΔ5::Pex20
SwaI/ClaI (2114-17)	· TEF: TEF 프로모터 (진뱅크 관리 번호 AF054508) · MAΔ5: 서열 6 (상기함) · Lip1: 야로위아 Lip1 유전자 (진뱅크 관리 번호 Z50020)의 Lip1 종결자 서열 을 포함하는 TEF::MAΔ5::Lip1
PmeI/ClaI (5550-4755)	야로위아 Ura3 유전자 (진뱅크 관리 번호 AJ306421)

플라스미드 pDMW232를 AscI/SphI으로 소화시킨 후에, 이것을 상기 '통상의 방법'에 따라 사용하여 균주 M4를 형질전환시켰다. 형질전환 후, 세포를 MMLe 플레이트에 플레이팅하고, 30℃에서 2 내지 3일 동안 유지시켰다. 각각의 형질전환으로부터 MMLe 플레이트에서 성장한 개개의 콜로니를 골라내고 MM 및 MMLe 플레이트에 스트리킹했다. MMLe 플레이트에서는 성장할 수 있으나 MM 플레이트에서는 성장할 수 없는 콜로니를 $Leu2^-$ 균주로서 선별하였다. 이어서, $Leu2^-$ 균주의 단일 콜로니를 30℃의 액체 MMLe 배지에 접종하고 250 rpm/분으로 2일 동안 진탕시켰다. 세포를 원심분리로 수집하여 지질을 추출하고 에스테르교환반응으로 지방산 메틸 에스테르를 제조한 후에 휴렛-팩커드 6890 GC로 분석하였다.

GC 분석은 ARA가 pDMW232 형질전환체 중에는 존재하지만 모(母) M4 균주 중에는 존재하지 않음을 보여주었다. 구체적으로, pDMW232를 함유하는 것으로 선별된 48개의 $Leu2^-$ 형질전환체 중에서 유전자 조작된 야로위아에서 ARA를 전체 지질의 5% 미만으로 생산하는 균주는 34개였고, ARA를 6 내지 8%로 생산하는 균주는 11개였으며, ARA를 약 10%로 생산하는 균주는 3개였다. ARA를 10%로 생산하는 균주 중 하나를 "Y2034"로 명명하였다.

EPA를 전체 지질의 약 10%로 생산하는 E 균주의 생성

플라스미드 pZP3L37 (실시예 6)을 AscI/SphI으로 소화시킨 후에, 이것을 상기 '통상의 방법'에 따라 사용하여 균주 Y2034를 형질전환시켰다. 형질전환 후, 세포를 MM 플레이트에 플레이팅하고, 30℃에서 2 내지 3일 동안 유지시켰다. MM 플레이트에서 성장한 총 48개의 형질전환체를 골라내고 신선한 MM 플레이트에 재스트리킹했다. 일단 성장했으면, 이들 균주를 개별적으로 30℃의 액체 MM에 접종하고 250 rpm/분으로 2일 동안 진탕시켰다. 세포를 원심분리로 수집하여 지질을 추출하고 에스테르교환반응으로 지방산 메틸 에스테르를 제조한 후에 휴렛-팩커드 6890 GC로 분석하였다.

GC 분석은 EPA가 pZP3L37을 함유하는 대개의 형질전환체 중에는 존재하지만 모 균주 (즉, Y2034) 중에는 존재하지 않음을 보여주었다. pZP3L37을 함유하는 것으로 선별된 48개의 형질전환체 중에서, 유전자 조작된 야로위아에서 EPA를 전체 지질의 2% 미만으로 생산하는 균주는 18개였고, EPA를 2 내지 3%로 생산하는 균주는 14개였으며, EPA를 약 7%로 생산하는 균주는 1개였다.

EPA를 7%로 생산하는 균주는 상기 '통상의 방법' 항목에 기재한 바와 같이 "2-단계 성장 조건" (즉, 48시간 MM, 72시간 HGM)을 이용하여 배양한 후에 추가로 분석하였다. GC 분석은 상기 유전자 조작된 균주가 상기 2-단계 성장 후에 EPA를 전체 지질의 약 10%로 생산함을 보여주었다. 상기 균주를 "E" 균주라고 칭하였다.

EPA를 전체 지질의 약 10%로 생산하는, Ura^- 표현형을 갖는 EU 균주의 생성

5-FOA 내성인 균주 E의 돌연변이체 세포를 동정하여 균주 EU (Ura^-)를 제조하였다. 구체적으로, 일군의 야로위아 E 균주 세포를 YPD 배지 3 mL에 접종하고, 30℃에서 250 rpm/분으로 24시간 동안 진탕 성장시켰다. 배양물을 YPD으로 OD_{600} 0.4까지 희석한 후에 추가의 4시간 동안 인큐베이션했다. 상기 배양물을 MM + FOA 플레이트에 플레이팅 (플레이트 1개 당 100 ㎕)하고, 30℃에서 2 내지 3일 동안 유지시켰다. 총 16개의 FOA 내성 콜로니를 골라내고 MM 및 MM + FOA 선별 플레이트에 스트리킹했다. 이들 중 10개의 콜로니가 FOA 선별 플레이트에서는 성장하였으나 MM 플레이트에서는 성장하지 못했고, 이들을 잠재적인 Ura^- 균주로서 선별하였다.

이들 균주 중 1개를, 키메라 GPD::푸사륨 모닐리포르메 Δ15::XPR2 유전자 및 선별 마커로서의 Ura3 유전자를 포함하는 pY37/F15를 사용한 형질전환의 숙주로 사용하였다 (도 16D, 서열 183). MM 플레이트에서 3일 동안의 선별 후, 상기 플레이트에서는 수백개의 콜로니가 성장하였고, 플라스미드를 보유하지 않는 형질전환 대조군에서는 콜로니 성장이 없었다. 이 실험으로 5-FOA 내성 숙주 균주가 Ura^-인 것으로 확인되었으며, 이 균주를 균주 "EU"라고 칭하였다.

이어서, EU 균주의 단일 콜로니를 0.1 g/L 유리딘을 추가로 함유하는 액체 MMU에 접종하고, 30℃에서 250 rpm/분으로 2일 동안 진탕 배양하였다. 세포를 원심분리로 수집하여 지질을 추출하고 에스테르교환반응으로 지방산 메틸 에스테르를 제조한 후에 휴렛-팩커드 6890 GC로 분석하였다. GC 분석은 EU 균주가 EPA를 전체 지질의 약 10%로 생산함을 보여주었다.

EPA를 전체 지질의 약 15%로 생산하는 Y2067 균주의 생성

2개의 키메라 유전자 (이종 Δ12 데새투라제 및 $C_{18/20}$ 엘론가제를 포함함) 및 Ura3 유전자를 함유하는 클러스터를 균주 EU의 천연 야로위아 Δ12 데새투라제 유전자에 통합하여 플라스미드 pKO2UF2PE (도 17A, 서열 184)를 제조하였다. 플라스미드 pKO2UF2PE는 하기 성분을 함유하였다:

[표 40]

플라스미드 pKO2UF2PE (서열 184)의 기재	
서열 184 내의 RE 부위 및 뉴클레오티드	단편 및 키메라 유전자 성분의 기재
AscI/BsiWI (3382-2645)	야로위아 Δ12 데새투라제 유전자 (서열 28)의 730 bp 5' 부분
SphI/EcoRI (6090-6646)	야로위아 Δ12 데새투라제 유전자 (서열 28)의 556 bp 3' 부분

SwaI/BsiWI (1-2645)	·FBAINm: FBAINm 프로모터 (서열 215) ·F.Δ12: 푸사륨 모닐리포르메 Δ12 데새투라제 유전자 (서열 32) ·Pex20: 야로위아 Pex20 유전자 (진뱅크 관리 번호 AF054613)의 Pex20 종결자 서열 을 포함하는 FBAINm::F.Δ12DS::Pex20
SwaI/PmeI (1-8525)	·GPAT: GPAT 프로모터 (서열 216) ·EL1S: 코돈-최적화된 엘론가제 1 유전자 (서열 24) (모르티에렐라 알피나 (진뱅크 관리 번호 AX464731)로부터 유래됨) ·OCT: 야로위아 OCT 유전자 (진뱅크 관리 번호 X69988)의 OCT 종결자 서열 을 포함하는 GPAT::EL1S::OCT
EcoRI/PacI (6646-8163)	야로위아 Ura3 유전자 (진뱅크 관리 번호 AJ306421)

플라스미드 pKO2UF2PE를 AscI/SphI으로 소화시킨 후에, 이것을 상기 '통상의 방법'에 따라 사용하여 균주 EU를 형질전환시켰다 (균주 EU는 형질전환 완충제 중에 현탁시키기 전에 YPD 플레이트에 스트리킹하여 대략 36시간 [*vs.* 18시간] 동안 성장시켰음). 형질전환 후, 세포를 MM 플레이트에 플레이팅하고, 30℃에서 2 내지 3일 동안 유지시켰다. MM 플레이트에서 성장한 총 72개의 형질전환체를 골라내고 신선한 MM 플레이트에 따로 재스트리킹했다. 일단 성장했으면, 이들 균주를 개별적으로 30℃의 액체 MM에 접종하고 250 rpm/분으로 2일 동안 진탕시켰다. 세포를 원심분리로 수집하여 지질을 추출하고 에스테르교환반응으로 지방산 메틸 에스테르를 제조한 후에 휴렛-팩커드 6890 GC로 분석하였다.

GC 분석은 pKO2UF2PE를 함유하는 거의 모든 형질전환체 중에는 EPA가 존재함을 보여주었다. 보다 구체적으로, 72개의 선별된 형질전환체 중에서 유전자 조작된 야로위아에서 EPA를 전체 지질의 8 내지 9.9%로 생산하는 균주는 17개였고, EPA를 10 내지 10.9%로 생산하는 균주는 27개였으며, EPA를 11 내지 11.9%로 생산하는 균주는 16개였고, EPA를 12 내지 12.7%로 생산하는 균주는 7개였다. EPA를 12.7%로 생산하는 균주는 상기 '통상의 방법' 항목에 기재한 바와 같은 2-단계 성장 조건 (즉, 48시간 MM, 72시간 HGM)을 이용하여 추가로 분석하였다. GC 분석은 상기 유전자 조작된 균주가 상기 2-단계 성장 후에 EPA를 전체 지질의 약 15%로 생산함을 보여주었다. 상기 균주를 균주 "Y2067"이라고 칭하였다.

<u>EPA를 전체 지질의 약 14%로 생산하는, Ura⁻ 표현형을 갖는 Y2067U 균주의 생성</u>

플라스미드 pZKUT16 (실시예 6)을 SalI/PacI로 소화시킨 후에, 이것을 상기 '통상의 방법'에 따라 사용하여 Y2067 균주를 형질전환시켰다. 형질전환 후, 세포를 MM + 5-FOA 선별 플레이트에 플레이팅하고, 30℃에서 2 내지 3일 동안 유지시켰다.

MM + 5-FOA 플레이트에서 성장한 총 24개의 형질전환체를 골라내고 MM 플레이트 및 MM + 5-FOA 플레이트에 따로 재스트리킹했다. MM + 5-FOA 플레이트에서는 성장할 수 있으나 MM 플레이트에서는 성장할 수 없는 균주를 Ura^- 균주로서 선별하였다. 총 10개의 Ura^- 균주를 개별적으로 30℃의 액체 MMU 배지에 접종하고 250 rpm/분으로 1일 동안 진탕 성장시켰다. 세포를 원심분리로 수집하여 지질을 추출하고 에스테르교환반응으로 지방산 메틸 에스테르를 제조한 후에 휴렛-팩커드 6890 GC로 분석하였다.

GC 분석은, MMU 배지 중에서 1일 동안 성장시킨 후에는 pZKUT16을 함유하는 모든 형질전환체에 EPA가 5 내지 7%로 존재함을 보여주었다. EPA를 6.2%로 생산하는 균주는 2-단계 성장 조건 (즉, 48시간 MM, 96시간 HGM)을 이용하여 추가로 분석하였다. GC 분석은 상기 유전자 조작된 균주가 EPA를 전체 지질의 약 14%로 생산함을 보여주었다. 상기 균주를 균주 "Y2067U"라고 칭하였다.

야생형 야로위아 리폴리티카 ATCC #20362에 대한 이 균주의 최종 유전자형은 다음과 같았다:

Ura3-, Pox3-, Y.Δ12-,
FBA::F.Δ12::Lip2, FBAINm::F. Δ12::Pex20, TEF::Δ6S::Lip1,
FBAIN::E1S::Pex20, GPAT::E1S::Oct, TEF::E2S::Xpr, FBAIN::Δ5::Pex20,
TEF::Δ5::Lip1, FBAIN::Δ17S::Lip2, FBAINm::Δ17S::Pex16, TEF::Δ17S 및
TEF::rELO2S::Pex20.

실시예 14

EPA를 전체 지질의 16%로 생산하는, 중간체 균주 Y2107U1의 생성

본 실시예는 야로위아 리폴리티카 ATCC #20362로부터 유래되어 EPA를 전체 지질에 대하여 상당한 농도로 생산할 수 있는 균주 Y2107U1의 구축을 기재한다 (도 5). 모르티에렐라 알피나 GPAT 유전자 과발현의 영향을 실시예 20에 기재된 바와 같이 TAG 함량 및/또는 조성의 분석에 기초하여 이 EPA 생산 균주에서 조사하였다 (하기 참조).

본원에서 균주 Y2107U1 (EPA를 16%로 생산하고, Ura^- 표현형을 보유함)의 발달에는 균주 M4 (DGLA를 8%로 생산함, 실시예 6에 기재됨), 균주 Y2047 (ARA를 11%로 생산함, 실시예 6에 기재됨), 균주 Y2048 (EPA를 11%로 생산함, 실시예 6에 기재됨), 균주 Y2060 (EPA를 13%로 생산함, 실시예 6에 기재됨), 균주 Y2072 (EPA를 15%로 생산함, 실시예 6에 기재됨), 균주 Y2072U1 (EPA를 14%로 생산함) 및 Y2089 (EPA를 18%로 생산함)의 구축이 요구되었다.

EPA를 전체 지질의 약 14%로 생산하는, Ura^- 표현형을 갖는 Y2072U1 균주의 생성

GPAT::I.Δ5S::Pex20 키메라 유전자를 Y2072 균주 (실시예 6)의 Ura3 유전자에 통합하여 구축물 pZKUGPI5S (도 17B, 서열 187)를 제조하였다. 보다 구체적으로, 플라스미드 pZKUGPI5S는 하기 성분을 함유하였다:

[표 41]

플라스미드 pZKUGPI5S (서열 187)의 기재	
서열 187 내의 RE 부위 및 뉴클레오티드	단편 및 키메라 유전자 성분의 기재
BsiWI/PacI (318-1038)	야로위아 Ura3 유전자 (진뱅크 관리 번호 AJ306421)의 721 bp 5' 부분
SalI/ClaI (3882-4606)	야로위아 Ura3 유전자 (진뱅크 관리 번호 AJ306421)의 724 bp 3' 부분
ClaI/BsiWI (4606-318)	· GPAT: GPAT 프로모터 (서열 216) · I.Δ5S: 코돈-최적화된 Δ5 데새투라제 유전자 (서열 10) (이소크리시스 갈바나로부터 유래됨) (제WO 2002/081668호) · Pex20: 야로위아 Pex20 유전자 (진뱅크 관리 번호 AF054613)의 Pex20 종결자 서열 을 포함하는 GPAT::I.Δ5S::Pex20

플라스미드 pZKUGPI5S를 SalI/PacI으로 소화시킨 후, 이것을 상기 '통상의 방법'에 따라 사용하여 균주 Y2072를 형질전환시켰다. 형질전환 후에는 세포를 MM + 5-FOA 선별 플레이트에 플레이팅하고 30℃에서 3 내지 4일 동안 유지시켰다.

MM + 5-FOA 플레이트에서 성장한 총 24개의 형질전환체를 골라내고 MM 플레이트 및 MM + 5-FOA 플레이트에 따로 재스트리킹했다. MM + 5-FOA 플레이트에서는 성장할 수 있으나 MM 플레이트에서는 성장할 수 없는 균주를 Ura^- 균주로서 선별하였다. 이들 24개의 Ura^- 균주 각각을 개별적으로 액체 MMU에 접종하고 30℃에서 250 rpm/분으로 2일 동안 진탕 성장시켰다. 세포를 원심분리로 수집하여 지질을 추출하고 에스테르교환반응으로 지방산 메틸 에스테르를 제조한 후에 휴렛-팩커드 6890 GC로 분석하였다.

GC 분석은, MMU 중에서 2일 동안 성장시킨 후에는 EPA를 전체 지질의 7.3 내지 8.9%로 생산하는 균주가 8개이고, EPA를 9 내지 9.9%로 생산하는 균주가 14개이고, EPA를 10.5%로 생산하는 균주가 1개 (즉, #1)이며, EPA를 10.7%로 생산하는 균주가 1개 (즉, #23)임을 보여주었다. 균주 #1 및 균주 #23을 2-단계 성장 조건 (즉, 48시간 MM, 96시간 HGM)을 이용하여 추가로 분석하였다. GC 분석은 이들 2개의 균주가 상기 2-단계 성장 후에 EPA를 전체 지질의 약 14%로 생산함을 보여주었다. 균주 #1은 균주 "Y2072U1"이라고 칭하였다.

EPA를 전체 지질의 약 18%로 생산하는 Y2089 균주의 생성

4개의 키메라 유전자 ($C_{16/18}$ 엘론가제, $C_{18/20}$ 엘론가제, Δ6 데새투라제 및 Δ12 데새투라제를 포함함) 및 Ura3 유전자의 클러스터를 Y2072U1 균주의 야로위아 리파제 1 유전자 부위에 통합하여 구축물 pDMW302T16 (도 17C, 서열 186)을 제조하였다. 플라스미드 pDMW302T16은 하기 성분을 함유하였다:

[표 42]

플라스미드 pDMW302T16 (서열 186)의 기재	
서열 186 내의 RE 부위 및 뉴클레오티드	단편 및 키메라 유전자 성분의 기재
BsiWI/AscI (1-817)	야로위아 리파제 1 유전자 (진뱅크 관리 번호 Z50020)의 817 bp 5' 부분
SphI/PacI (3525-4294)	야로위아 리파제 1 유전자 (진뱅크 관리 번호 Z50020)의 769 bp 3' 부분
EcoRI/BsiWI (13328-1)	· TEF: TEF 프로모터 (진뱅크 관리 번호 AF054508) · rEL02S: 코돈-최적화된 rEL02 엘론가제 유전자 (서열 85) (래트 (진뱅크 관리 번호 AB071986)로부터 유래됨) · Pex20: 야로위아 Pex20 유전자 (진뱅크 관리 번호 AF054613)의 Pex20 종결자 서열 을 포함하는 TEF::rEL02S::Pex20
BglII/EcoRI (10599-13306)	· FBAIN: FBAIN 프로모터 (서열 214) · Δ6S: 코돈-최적화된 Δ6 데새투라제 유전자 (서열 3) (모르티에렐라 알피나 (진뱅크 관리 번호 AF465281)로부터 유래됨) · Lip1: 야로위아 Lip1 유전자 (진뱅크 관리 번호 Z50020)의 Lip1 종결자 서열 을 포함하는 FBAIN::D6S::Lip1
ClaI/PmeI (8078-10555)	· GPDIN: GPDIN 프로모터 (서열 211) · EL1S: 코돈-최적화된 엘론가제 1 유전자 (서열 24) (모르티에렐라 알피나 (진뱅크 관리 번호 AX464731)로부터 유래됨) · Lip2: 야로위아 리파제 2 유전자 (진뱅크 관리 번호 AJ012632)의 Lip2 종결자 를 포함하는 GPDIN::EL1S::Lip2
EcoRI/ClaI (6450-8078)	야로위아 Ura3 유전자 (진뱅크 관리 번호 AJ306421)
PacI/EcoRI (4294-6450)	· TEF: TEF 프로모터 (진뱅크 관리 번호 AF054508) · F.Δ12: 푸사륨 모닐리포르메 Δ12 데새투라제 유전자 (서열 32) · Pex16: 야로위아 Pex16 유전자 (진뱅크 관리 번호 U75433)의 Pex16 종결자 를 포함하는 TEF::F.Δ12::Pex16

플라스미드 pDMW302T16을 SphI/AscI으로 소화시킨 후에, 이것을 상기 '통상의 방법'에 따라 사용하여 균주 Y2072U1을 형질전환시켰다. 형질전환 후, 세포를 MM 플레이트에 플레이팅하고, 30℃에서 3 내지 4일 동안 유지시켰다.

MM 플레이트에서 성장한 총 48개의 형질전환체를 골라내고 신선한 MM 플레이트에 재스트리킹했다. 일단 성장했으면, 이들 균주를 개별적으로 30℃의 액체 MM에 접종하고 250 rpm/분으로 2일 동안 진탕 성장시켰다. 세포를 원심분리로 수집하여 지질을 추출하고 에스테르교환반응으로 지방산 메틸 에스테르를 제조한 후에 휴렛-팩커드 6890 GC로 분석하였다.

GC 분석은, MM 배지 중에서 2일 동안 성장시킨 후에는 pDMW302T16을 함유하는 Y2072U1의 거의 모든 형질전환체에서 EPA가 생산됨을 보여주었다. 선별된 48개의 형질전환체 중에서, EPA를 10% 미만으로 생산하는 균주는 27개였고, EPA를 10 내지 12.9%로 생산하는 균주는 14개였으며, EPA를 13 내지 13.9%로 생산하는 균주는 5개였다. 균주 #34 (EPA를 13.9%로 생산함)를 선별하여 2-단계 성장 절차 (즉, 48시간 MM, 96시간 HGM)를 이용한 추가의 분석을 수행하였다. GC 분석은 균주 #34가 EPA를 전체 지질의 약 18%로 생산함을 보여주었다. 균주 #34를 균주 "Y2089"라고 칭하였다.

야생형 야로위아 리폴리티카 ATCC #20362에 대한 균주 Y2089의 유전자형은 다음과 같았다:

Pox3-, LIP1-, Y.Δ12-,
FBA::F.Δ12::Lip2, TEF::F. Δ12::Pex16, FBAIN::MΔ12::Pex20,
TEF::Δ6S::Lip1, FBAIN::Δ6S::Lip1, FBAIN::E1S::Pex20, GPAT::E1S::Oct,
GPDIN::E1S::Lip2, TEF::E2S::Xpr, FBAIN::MAΔ5::Pex20,
TEF::MAΔ5::Lip1, TEF::HΔ5S::Pex16, TEF::IΔ5S::Pex20,
GPAT::IΔ5S::Pex20, FBAIN::Δ17S::Lip2, FBAINm::Δ17S::Pex16,
TEF::Δ17S::Pex16 및 *2X TEF::rELO2S::Pex20.*

<u>EPA를 전체 지질의 약 16%로 생산하는, Ura- 표현형을 갖는 Y2107U1 균주의 생성</u>

GPAT::EL1S::Pex20 키메라 유전자를 균주 Y2089의 Ura3 유전자 내로 통합하여 구축물 pZKUGPE1S (서열 187)를 제조하였다. 보다 구체적으로, 플라스미드 pZKUGPE1S는 하기 성분을 함유하였다:

[표 43]

플라스미드 pZKUGPE1S (서열 187)의 기재	
서열 187 중의 RE 부위 및 뉴클레오티드	단편 및 키메라 유전자 성분의 기재
BsiWI/PacI (318-1038)	야로위아 Ura3 유전자의 721 bp 5' 부분 (진뱅크 관리 번호 AJ306421)
SalI/ClaI (3882-4606)	야로위아 Ura3 유전자의 724 bp 3' 부분 (진뱅크 관리 번호 AJ306421)
ClaI/BsiWI (4606-318)	· GPAT:GPAT 프로모터 (서열 216) · EL1S: 코돈-최적화된 엘론가제 1 유전자 (서열 24) (모르티에렐라 알피나 (진뱅크 관리 번호 AX464731)에서 유래됨) · Pex20: 야로위아 Pex20 유전자 (진뱅크 관리 번호 AF054613)로부터의 Pex20 종결자 서열 을 포함하는 GPAT::EL1S::Pex20

플라스미드 pZKUGPE1S를 PstI/PacI로 소화시키고, 이어서 이를 사용하여 통상의 방법에 따라 균주 Y2089를 형질전환시켰다. 형질전환 후, 세포를 MM + 5-FOA 선별 플레이트 상에 플레이팅하고, 30℃에서 3 내지 4일 동안 유지시켰다.

MM + 5-FOA 플레이트 상에 성장한 총 8개 형질전환체를 골라내고 MM 플레이트 및 MM + 5-FOA 플레이트 상에 따로 재스트리킹하였다. MM + 5-FOA 플레이트에서는 성장할 수 있으나 MM 플레이트에서는 성장할 수 없는 이들 균주를

Ura⁻ 균주로서 선별하였다. 각각의 이들 8 Ura⁻ 균주를 액체 MMU에 개별적으로 접종시키고, 30℃에서 250 rpm/분으로 2일 동안 진탕하여 성장시켰다. 세포를 원심분리로 수집하고 지질을 추출하고 에스테르교환반응으로 지방산 메틸 에스테르를 제조한 후에 휴렛-팩커드 6890 GC로 분석하였다.

GC 분석은 2일 동안 MMU에서 성장시킨 후에는 EPA를 전체 지질의 6.6 내지 8.7%로 생산하는 균주가 6개였고, EPA (즉, #4 및 #5)를 9.4 내지 10%로 생산하는 균주가 2개였음을 보여주었다. 균주 #4 및 #5는 2-단계 성장 조건 (즉, 48시간 MM, 72시간 HGM)을 이용하여 추가로 분석하였다. GC 분석은 이들 2개의 균주가 EPA를 전체 지질의 약 16%로 상기 2-단계 성장 후에 생산하였음을 보여주었다. 균주 #4를 균주 "Y2107U1"이라고 칭하고, 균주 #5를 균주 "Y2107U2"라고 칭하였다.

실시예 15

EPA를 전체 지질의 9 내지 12%로 생산하는, 중간체 균주 MU의 생성

본 실시예는 야로위아 리폴리티카 ATCC #20362로부터 유래되어 EPA를 전체 지질에 대하여 상당한 농도로 생산할 수 있는 균주 MU의 구축을 기재한다 (도 5). 다양한 천연 야로위아 리폴리티카 아실트랜스퍼라제 넉아웃의 영향을 실시예 27에 기재된 바와 같이 TAG 함량 및/또는 조성의 분석에 기초하여 이 EPA 생산 균주에서 조사하였다 (하기 참조).

균주 MU (본원에서 EPA를 9 내지 12%로 생산함)의 발달에는 균주 M4 (DGLA를 8%로 생산함, 실시예 6에 기재됨), 균주 Y2034 (ARA를 10%로 생산함, 실시예 13에 기재됨), 균주 E (EPA를 10%로 생산함, 실시예 13에 기재됨), 균주 EU (EPA를 10%로 생산함, 실시예 13에 기재됨) 및 균주 M26 (EPA를 14%로 생산함)의 구축이 요구되었다.

EPA를 전체 지질의 약 14%로 생산하는 M26 균주의 생성

구축물 pKO2UM26E (서열 188, 도 17D)를 사용하여 3개의 키메라 유전자 ($C_{18/20}$ 엘론가제, Δ6 데새투라제 및 Δ12 데새투라제 포함) 및 Ura3 유전자의 클러스터를 EU 균주의 야로위아 Δ12 데새투라제 유전자 부위 내로 통합하였다 (실시예 13). 플라스미드 pKO2UM26E는 하기 성분을 함유하였다:

[표 44]

	플라스미드 pKO2UM26E (서열 188)의 기재
서열 188 중의 RE 부위 및 뉴클레오티드	단편 및 키메라 유전자 성분의 기재
HindIII/AscI (1-728)	야로위아 Δ12 데새투라제 유전자 (서열 28)의 728 bp 5' 부분
SphI/EcoRI (3436-3992)	야로위아 Δ12 데새투라제 유전자 (서열 28)의 556 bp 3' 부분
BsiWI/HindIII (11095-1)	·GPAT: GPAT 프로모터 (서열 216) ·EL1S: 코돈-최적화된 엘론가제 1 유전자 (서열 24) (모르티에렐라 알피나 (진뱅크 관리 번호 AX464731)에서 유래됨) ·XPR: 야로위아 Xpr 유전자 (진뱅크 관리 번호 M17741)의 3' 영역의 ~100 bp 를 포함하는 GPAT::EL1S::XPR
BglII/BsiWI (8578-11095)	·FBAIN: FBAIN 프로모터 (서열 214) ·M.Δ12: 모르티에렐라 이자벨리나 Δ12 데새투라제 유전자 (진뱅크 관리 번호 AF417245; 서열 30) ·Pex20: 야로위아 Pex20 유전자 (진뱅크 관리 번호 AF054613)의 Pex20 종결자 서열 을 포함하는 FBAIN::M.Δ12::Pex20
SalI/PacI (6704 내지 8202)	야로위아 Ura3 유전자 (진뱅크 관리 번호 AJ306421)

EcoRI/SalI (3992-6704)	· TEF: TEF 프로모터 (진뱅크 관리 번호 AF054508) · M.Δ6B: 모르티에렐라 알피나 Δ6 데새투라제 유전자 "B" (진뱅크 관리 번호 AB070555; 서열 4) · Pex20: 야로위아 Pex20 유전자 (진뱅크 관리 번호 AF054613)의 Pex20 종결자 서열 을 포함하는 FBAIN::M.Δ6B::Pex20

플라스미드 pKO2UM26E를 SphI/AscI로 소화시킨 후에, 이것을 사용하여 EU 균주 (실시예 13)를 통상의 방법에 따라 형질전환시켰다. 형질전환 후, 세포를 MM 플레이트에 플레이팅하고, 30℃에서 2 내지 3일 동안 유지시켰다.

MM 플레이트에서 성장한 총 48개의 형질전환체를 골라내고 신선한 MM 플레이트에 재스트리킹했다. 일단 성장했으면, 이들 균주를 개별적으로 30℃의 액체 MM에 접종하고 250 rpm/분에서 1일 동안 진탕하여 성장시켰다. 세포를 원심분리로 수집하고 지질을 추출하고 에스테르교환반응으로 지방산 메틸 에스테르를 제조한 후에 휴렛-팩커드 6890 GC로 분석하였다.

GC 분석은 EPA가 1일 동안 MM 배지에서 성장시킨 후에는 pKO2UM26E를 갖는 거의 모든 형질전환체에서 생산되었음을 보였다. 선별된 48개의 형질전환체 중에서, 유전자 조작된 야로위아에서 EPA를 전체 지질의 4% 미만으로 생산하는 균주가 5개였고, EPA를 4 내지 5.9%로 생산하는 균주가 23개였고, EPA를 6 내지 6.9%로 생산하는 균주가 9개였고, EPA를 7 내지 8.2%로 생산하는 균주가 11개였다. EPA를 8.2%로 생산하는 균주를 2-단계 성장 절차 (즉, 48시간 MM, 96시간 HGM)를 이용한 추가의 분석을 위해 선별하였다. GC 분석은 상기 유전자 조작된 균주가 EPA를 전체 지질의 약 14%로 생산하였음을 보였다. 상기 균주를 균주 "M26"이라고 칭하였다.

야생형 야로위아 리폴리티카 ATCC #20362에 대한 M26 균주의 유전자형은 다음과 같았다:

Pox3-, Y.Δ12-,
FBA::F.Δ12::Lip2, FBAIN::MΔ12::Pex20, TEF::Δ6S::Lip1,
FBAIN::Δ6B::Pex20, FBAIN::E1S::Pex20, GPAT::E1S:: Xpr,
TEF::E2S::Xpr, FBAIN::MAΔ5::Pex20, TEF::MAΔ5::Lip1,
TEF::HΔ5S::Pex16, FBAIN::Δ17S::Lip2, FBAINm::Δ17S::Pex16,
TEF::Δ17S::Pex16 및 *TEF::rELO2S::Pex20.*

EPA를 전체 지질의 약 14%로 생산하는 MU 균주의 생성

균주 MU는 균주 M26의 Ura 영양요구균이었다. 이 균주는 균주 M26을 PacI 및 HincII로 소화시킨 플라스미드 pZKUM (서열 189) 5 ㎍으로 형질전환시킴으로써 제조하였다. 형질전환을 동결-EZ 효모 형질전환 키트 (Frozen-EZ Yeast Transformation kit; 자이모 리서치 코포레이션, 미국 캘리포니아 오렌지 소재)로 수행하고, 형질전환체는 하기 배지를 갖는 아가 플레이트 상에 형질전환된 세포 믹스 100 ㎕를 플레이팅함으로써 선별하였다: 6.7 g/L 효모 질소 염기 (디프코 래버러토리즈, 미국 미시간주 디트로이트 소재), 20 g/L 덱스트로즈, 50 mg/L 우라실 및 800 mg/L FOA. 7일 후에, MM 및 MMU 아가 플레이트 상에 플레이팅된 작은 콜로니들이 도출되었다. 모두가 URA 영양요구이었다. 균주 중 하나를 "MU"라고 칭하였다.

실시예 16

모르티에렐라 알피나 게놈 DNA 및 cDNA의 구축

본 실시예는 모르티에렐라 알피나 (ATCC #16266)로부터 게놈 DNA 및 cDNA의 제조를 기재한다. 이는 실시예 17, 18, 19, 20 및 21에 각각 기재하는 바와 같이 모르티에렐라 알피나 LPAAT2, DGAT1, DGAT2, GPAT 및 ELO3을 단리할 수 있다.

모르티에렐라 알피나로부터의 게놈 DNA의 제조

게놈 DNA를 모르티에렐라 알피나 (ATCC #16266)로부터 QiaPrep 스핀 미니프렙 키트(Spin Miniprep Kit) (퀴아젠, 카탈로그 #627106)로 단리하였다. YPD 아가 플레이트 (2% 박토-효모 추출물, 3% 박토-펩톤, 2% 글루코스, 2.5% 박토-아가) 상에 성장한 세포를 긁어 내어 키트 완충제 P1 1.2 mL에 재현탁하였다. 재현탁한 세포를 2개 2.0 mL 스크루 캡 튜브에 두었으며, 이들 각각은 0.6 mL 유리 비드 (0.5 mm 직경)를 함유하였다. 세포를 HOMOGENIZE 셋팅에서의 바이오스펙(Biospec) (미국 오클라호마주 바르틀레스빌 소재) 미니 비드 비터(beater) 상에서 2분 동안 균질화하였다. 튜브를 이어서 14,000 rpm에서의 에펜도르프 미세원침에서 2분 동안 원심분리하였다. 상등액 (0.75 mL)을 3개의 1.5 mL 미세원침관 내로 전달하였다. 키트 완충제 P2의 동일 부피를 각각의 튜브에 첨가하였다. 3회 뒤집어서 튜브를 혼합한 후에, 완충제 N3 0.35 mL를 각각의 튜브에 첨가하였다. 각각의 튜브의 함유물을 총 5회 뒤집어서 다시 혼합하였다. 혼합물을 14,000 rpm에서의 에펜도르프 미세원침에서 5분 동안 원심분리하였다. 각 튜브로부터의 상등액을 개별적으로 3개의 개별 키트 스핀 컬럼 내로 옮겼다. 컬럼에 이어서 하기 단계를 수행하였다: 원심분리 (14,000 rpm에서 1분), 완충제 PE로 1회 세척, 원심분리 (14,000 rpm에서 1분), 및 이어서 최종 원심분리 (14,000 rpm에서 1분). 완충제 EB (50 ㎕)를 각 컬럼에 첨가하고, 1분 동안 정치하였다. 게놈 DNA를 이어서 14,000 rpm에서 1분 동안 원심분리함으로써 용출시켰다.

<u>모르티에렐라 알피나로부터의 cDNA의 제조</u>

모르티에렐라 알피나의 cDNA를 제조자의 프로토콜에 따라 BD-클론테크 크리에이터 스마트® cDNA 라이브러리 키트 (캐나다 온타리오주 미시소거 소재)로 제조하였다.

구체적으로, 모르티에렐라 알피나 균주 ATCC #16266을 YPD 배지 (2% 박토-효모 추출물, 3% 박토르(Bactor)-펩톤, 2% 글루코스) 60 mL에 3일 동안 23℃에서 성장시켰다. 세포를 3750 rpm에서의 벡크만(Beckman) GH3.8 로터에서 10분 동안 원심분리하여 펠렛화하고, 트리졸(Trizole) 시약 (인비트로젠) 6 X 0.6 mL에 재현탁하였다. 재현탁한 세포를 6개의 2 mL 스크루 캡 튜브에 옮겼으며, 이들 각각은 0.5 mm 유리 비드 0.6 mL를 함유하였다. 세포를 HOMOGENIZE 셋팅에서의 바이오스펙 미니 비드 비터 상에서 2분 동안 균질화하였다. 튜브를 잠시 원심분리하여 비드를 가라앉혔다. 상기 액체를 4개의 신선한 1.5 mL 미세원침관에 옮기고, 클로로포름/이소아밀 알콜 (24:1) 0.2 mL를 각 튜브에 첨가하였다. 튜브를 손으로 1분 동안 진탕하고, 3분 동안 정치하였다. 튜브를 이어서 14,000 rpm에서 10분 동안 4℃에서 원심분리하였다. 상부 층을 4개 새로운 튜브에 옮겼다. 이소프로필 알콜 (0.5 mL)을 각 튜브에 첨가하였다. 튜브를 실온에서 15분 동안 인큐베이션한 후에, 14,000 rpm 및 4℃에서 10분 동안 원심분리하였다. 펠렛을 RNase-무함유 물로 제조한 각 75% 에탄올 1 mL로 세척하고, 공기 건조시켰다. 총 RNA 샘플을 이어서 물 500 ㎕에 재용해하고, 1:50 희석 RNA 샘플을 사용하여 A260 nm로 RNA 양을 측정하였다. 총 3.14 mg의 RNA가 수득되었다.

이 총 RNA 샘플을 제조자의 프로토콜에 따라 퀴아젠 RNeasy 총 RNA 미디(Midi) 키트로 추가 정제하였다. 따라서, 총 RNA 샘플을 2 mL로 희석하고, β-머캅토에탄올 80 ㎕ 및 100% 에탄올 5.6 mL를 갖는 완충제 RLT 8 mL와 혼합하였다. 샘플을 4개 부분으로 나누고, 4개 RNeasy 미디드(midid) 컬럼 상에 적하하였다. 컬럼을 이어서 5분 동안 4500Xg에서 원심분리하였다. 컬럼을 세척하기 위해, 완충제 RPE 2 mL를 적하하고, 컬럼을 2분 동안 4500Xg에서 원심분리하였다. 원심분리 시간을 5분으로 연장한 것 외에는 세척 단계를 1회 반복하였다. 총 RNA는 RNase 무함유 물 250 ㎕를 각 컬럼에 가하고, 1분 동안 대기하고, 4500Xg 3분 동안 원심분리하여 용출시켰다.

이어서, 폴리A(+)RNA를 아머샴 바이오사이언스즈 mRNA 정제 키트의 프로토콜에 따라 상기 총 RNA 샘플로부터 단리하였다. 간략하게 설명하면, 2개 올리고-dT-셀룰로스 컬럼을 사용하였다. 컬럼을 각 고염 완충제 1 mL로 2회 세척하였다. 이전의 단계로부터의 총 RNA 샘플을 총 부피 2 mL로 희석하고, 10 mM Tris/HCl, pH 8.0, 1 mM EDTA로 조정하였다. 샘플을 65℃에서 5분 동안 가열하고, 이어서 얼음상에 두었다. 샘플 완충제 (0.4 mL)를 첨가하고, 샘플을 이어서 2개 올리고-dT-셀룰로스 컬럼 상에 중력 공급 하에 적하하였다. 컬럼을 350Xg에서 2분 동안 원심분리하고, 각 고염 완충제 0.25 mL로 2회 세척하였으며, 각 횟수 후에 350Xg에서 2분 동안 원심분리하였다. 컬럼을 저염 완충제로 추가로 3회 세척한 후에, 동일한 일상적 원심분리를 수행하였다. 폴리(A)+ RNA는 컬럼을 65℃로 예열한 각 용출 완충제 0.25 mL로 4회 세척한 후에, 동일한 원심분리 절차를 수행함으로써 용출하였다. 전체 정제 과정을 1회 반복하였다. 정제된 폴리(A)+ RNA를 30.4 ng/㎕의 농도로 얻었다.

cDNA를 BD-클론테크 및 폴리A(+) RNA 샘플 0.1 ㎍로 특정지어진 LD-PCR 방법으로 생성하였다. 구체적으로, 제1 가닥 cDNA 합성을 위해, 폴리(A)+ RNA 샘플 3 ㎕를 SMART IV 올리고 뉴클레오티드 (서열 391) 1 ㎕ 및 CDSIII/3' PCR 프라이머 (서열 392) 1 ㎕와 혼합하였다. 혼합물을 72℃에서 2분 동안 가열하고, 얼음상에서 2분 동안 냉각하였다. 상기 튜브에 제1 가닥 완충제 2 ㎕, 20 mM DTT 1 ㎕, 10 mM dNTP 믹스 1 ㎕ 및 파워스크립트(Powerscript) 역전사효소 1 ㎕를 첨가하였다. 혼합물을 42℃에서 1시간 동안 인큐베이션하고, 얼음상에서 냉각하였다.

제1 가닥 cDNA 합성 혼합물을 PCR 반응에 대해 주형으로서 사용하였다. 구체적으로, 반응 혼합물은 제1 가닥 cDNA 혼합물 2 ㎕, 5'-PCR 프라이머 (서열 393) 2 ㎕, CDSIII/3'-PCR 프라이머 (서열 392) 2 ㎕, 물 80 ㎕, 10X 어드밴티지 (Advantage) 2 PCR 완충제 10 ㎕, 5OX dNTP 믹스 2 ㎕ 및 5OX 어드밴티지 2개 폴리머라제 믹스 2 ㎕를 함유하였다. 열순환기 조건은 GenAmp 9600 기기상에서 95℃에서 20초 후에 95℃에서 5초 및 68℃에서 6분의 14 내지 20 주기로 설정하였다. PCR 생성물을 아가로스 겔 전기영동 및 에티듐 브로마이드 염색으로 정량하였다.

상기 PCR 생성물 (cDNA) 75 ㎕를 키트로 공급한 20 ㎍/㎕ 프로테이나제 K 3 ㎕와 혼합하였다. 혼합물을 45℃에서 20분 동안 인큐베이션하고, 이어서 물 75 ㎕를 첨가하고, 혼합물을 페놀:클로로포름:이소아밀 알콜 혼합물 (25:24:1) 150 ㎕로 추출하였다. 수성 상을 클로로포름:이소아밀 알콜 (25:1) 150 ㎕로 추가 추출하였다. 수성 상을 이어서 3M 나트륨 아세테이트 15 ㎕, 20 ㎍/㎕ 글리코겐 2 ㎕ 및 100% 에탄올 400 ㎕와 혼합하였다. 혼합물을 실온에서 20분 동안 14000 rpm에서의 미세원침에서 즉시 원심분리하였다. 펠렛을 80% 에탄올 150 ㎕로 1회 세척하고, 공기 건조시키고, 물 79 ㎕에 용해하였다.

용해된 cDNA를 이어서 SfiI (cDNA 79 ㎕를 1OX SfiI 완충제 10 ㎕, SfiI 효소 10 ㎕ 및 100X BSA 1 ㎕와 혼합하고, 혼합물을 50℃에서 2시간 동안 인큐베이션함)로 소화하였다. 크실렌 시아놀 염료 (1%, 2 ㎕)를 첨가하였다. 혼합물을 제조자 절차에 정확하게 따라 키트로 제공한 크로마 스핀(Chroma Spin)-400 컬럼 상에서 이어서 분획하였다. 컬럼으로부터 수집한 분획을 아가로스 겔 전기영동으로 분석하였다. cDNA를 함유한 제1의 3개 분획을 모으고, cDNA를 에탄올로 침전시켰다. 침전된 cDNA를 물 7 ㎕로 재용해하고, 키트-공급 pDNR-LIB 내로 라이게이션하였다.

라이브러리 서열분석

라이게이션 생성물을 사용하여 이. 콜라이 XL-1 블루(Blue) 전기천공법 감응성 세포 (스트라타진)를 형질전환시켰다. 추측된 총 2 x 10^6 콜로니를 수득하였다. cDNA 라이브러리의 서열분석을 M13 전진 프라이머 (서열 394)로 아젠코트 바이오사이언스 코포레이션(Agencourt Bioscience Corporation)(미국 매사추세츠주 비벌리 소재)에 의해 수행하였다.

실시예 17

모르티에렐라 알피나 LPAAT2 발현은 PUFA(%)를 증가시킨다.

본 실시예는 모르티에렐라 알피나 LPAAT2 (서열 110 및 111)를 동시 발현하도록 형질전환된 야로위아 리폴리티카 균주 Y2067U (실시예 13)에서의 증가된 EPA 생합성 및 축적을 기재한다. Δ6 데새투라제/Δ6 엘론가제 경로 또는 Δ9 엘론가제/Δ8 데새투라제 경로를 통해 DHA를 생산하도록 유전자 조작된 야로위아 리폴리티카 숙주 균주는 모르티에렐라 알피나 LPAAT2가 그에서 유사하게 동시 발현되는 경우에 (예컨대, 균주 Y3000에서) 증가된 DHA 생합성 및 축적을 입증할 수 있었음을 고려한다.

모르티에렐라 알피나 LPAAT2 ORF를 하기와 같이 클로닝하였다. 프라이머 MLPAT-F 및 MLPAT-R (서열 395 및 396)을 사용하여 모르티에렐라 알피나 (실시예 16)의 cDNA로부터 LPAAT2 ORF를 PCR에 의해 증폭시켰다. 반응 혼합물은 cDNA 1 ㎕, 각 프라이머 1 ㎕, 물 22 ㎕ 및 ExTaq 프리믹스 2X Taq PCR 용액 (다까라 바이오 인크., 일본 520-2193 시가현 오쯔 소재) 25 ㎕를 함유하였다. 증폭을 하기와 같이 수행하였다: 94℃에서 150초 동안 초기 변성, 이어서 94℃에서 30초 동안 변성, 55℃에서 30초 동안 어닐링 및 72℃에서 90초 동안 신장의 30주기. 72℃에서 10분 최종 신장 주기를 수행한 후에, 반응을 4℃에서 종결하였다. ~950 bp DNA 단편을 PCR 반응물로부터 수득하였다. 이를 퀴아젠 (미국 캘리포니아주 발렌시아 소재) PCR 정제 키트로 제조자의 프로토콜에 따라 정제하였다. 정제된 PCR 생성물을 NcoI 및 NotI로 소화시키고, NcoI-NotI 절단 pZUF17 벡터 (서열 162; 도 9B) 내로 클로닝하여 유전자가 야로위아 리폴리티카에서의 발현을 위한 자가-복제 벡터 중에서 야로위아 리폴리티카 FBAIN 프로모터 및 PEX20-3' 종결자 영역의 조절 하에 놓이게 하였다. 정확한 형질전환체를 미니프렙 DNA의 제한 분석에 의해 확인하고, 생성된 플라스미드를 "pMLPAT-17" (서열 190)이라고 칭하였다.

모르티에렐라 알피나 LPAAT2를 야로위아 리폴리티카의 게놈 내로 통합하기 위해, 플라스미드 pM LPAT-Int를 생성하였다. 프라이머 LPAT-Re-5-1 및 LPAT-Re-5-2 (서열 397 및 398)를 사용하여 야로위아 리폴리티카 LPAAT1 (서열 112)의 AUG 바로 상류에 1103 bp 단편의 야로위아 리폴리티카 게놈을 함유한 1129 bp DNA 단편인 YLPAT-5' (서열 399)를 증폭하였다. 반응 혼합물은 야로위아 리폴리티카 게놈 DNA 1 ㎕, 각 프라이머 1 ㎕, 물 22 ㎕ 및 ExTaq 프리믹스 2X Taq PCR 용액 (다까라) 25 ㎕를 함유하였다. 증폭을 상기 기재한 바와 같이 수행하였다. ~1130 bp DNA 단편을 PCR

반응물로부터 수득하였다. 이를 퀴아젠 PCR 정제 키트로 제조자의 프로토콜에 따라 정제하였다. 정제된 PCR 생성물을 SalI 및 ClaI으로 소화시키고, SalI-ClaI 절단 pBluescript SK (-) 벡터 내로 클로닝하여 플라스미드 "pYLPAT-5"를 생성하였다.

프라이머 LPAT-Re-3-1 및 LPAT-Re-3-2 (서열 400 및 401)를 이어서 사용하여 야로위아 리폴리티카 LPAAT1의 정지 코돈 직후에 903 bp 단편의 야로위아 리폴리티카 게놈을 함유한 938 bp 단편 YLPAT-3' (서열 402)을 상기와 동일한 조건으로 증폭하였다. 정제된 PCR 생성물을 ClaI 및 XhoI으로 소화시키고, ClaI-XhoI 소화 pYLPAT-5'내로 클로닝하였다. 정확한 형질전환체를 미니프렙 분석으로 확인하고, 생성된 플라스미드를 "pYLPAT-5'-3"'라고 칭하였다.

pMLPAT-17 (서열 190)을 이어서 ClaI 및 NotI로 소화시키고, 야로위아 리폴리티카 URA3 유전자, 야로위아 리폴리티카 FBAIN 프로모터 및 모르티에렐라 알피나 LPAAT2 유전자를 함유한 ~3.5 kb 단편을 퀴아젠 QiaexII 겔 정제 키트로 제조자의 프로토콜에 따라 단리하였다. 이 단편을 ClaI-NotI 소화 pYLPAT-5'-3' 내로 클로닝하였다. 정확한 형질전환체를 미니프렙 및 제한 분석에 의해 확인하였다. 생성된 플라스미드를 "pMLPAT-Int" (서열 191)라고 명명하였다.

"대조군" 벡터 pZUF-MOD-1 (서열 192, 도 18A)을 하기와 같이 제조하였다. 첫째, 프라이머 pzuf-mod1 및 pzuf-mod2 (서열 403 및 404)를 사용하며, pDNR-LIB (클론테크, 캘리포니아주 팔로 알토 소재)를 주형로서 사용하여 252 bp "스터퍼(stuffer)" DNA 단편을 증폭시켰다. 증폭시킨 단편을 퀴아젠 QiaQuick PCR 정제 키트로 정제하고, 표준 조건을 사용하여 NcoI 및 NotI로 소화시키고, 이어서 QiaQuick PCR 정제 키트로 다시 정제하였다. 이 단편을 유사하게 소화시킨 NcoI-/NotI-절단 pZUF17 벡터 (서열 162; 도 9B) 내로 라이게이션하고, 생성된 라이게이션 혼합물을 사용하여 이. 콜라이 Top10 세포 (인비트로젠)를 형질전환시켰다. 플라스미드 DNA를 4개의 생성된 콜로니로부터 퀴아젠 QiaPrep 스핀 미니프렙 키트로 정제하였다. 정제된 플라스미드를 NcoI 및 NotI로 소화시켜 ~250 bp 단편의 존재를 확인하였다. 생성된 플라스미드를 "pZUF-MOD-1" (서열 192)이라고 명명하였다.

야로위아 리폴리티카 균주 Y2067U (실시예 13으로부터, EPA를 전체 지질의 14%로 생산함)를 플라스미드 pMLPAT-17, 플라스미드 pZUF-MOD-1 (대조군) 및 SpeI/XbaI 소화 플라스미드 pMLPAT-Int로 개별적으로 통상의 방법에 따라 형질전환시켰다. 형질전환체를 2일 동안 아미노산 보충 합성 MM에서, 이어서 4일 동안 HGM에서 성장시켰다. pZUF-MOD-1를 함유한 2개의 형질전환체, pMLPAT-17을 함유한 2개의 형질전환체, 및 pMLPAT-Int를 갖는 2개의 형질전환체 (게놈 내에 통합됨)의 지방산 프로파일을 (상기 '통상의 방법' 항목에 기재한 바와 같이) GC 분석에 기초하여 하기 표에 나타낸다. 지방산을 18:0, 18:1 (올레산), 18:2 (LA), GLA, DGLA, ARA, ETA 및 EPA로서 확인하고; 각각의 조성을 전체 지방산(%)으로서 나타낸다.

[표 45]

모르티에렐라 알피나 LPAAT2를 과발현하도록 유전자 조작된 야로위아 균주 Y2067U 중의 지질 조성

균주	총 지방산							
	18:0	18:1	18:2	GLA	DGLA	ARA	ETA	EPA
Y2067U + pZUF-MOD-1 #1	1.1	4.7	10.9	19.4	6.3	0.9	3.9	13.8
Y2067U + pZUF-MOD-1 #2	0.9	4.4	9.5	19.3	6.6	0.9	4.0	14.1
Y2067U + pMLPAT-17 #1	1.0	4.4	9.8	18.6	5.9	0.8	3.4	15.5
Y2067U + pMLPAT-17 #2	0.7	3.5	8.4	16.7	6.2	1.0	2.9	16.0
Y2067U + pMLPAT-Int #1	1.9	4.9	13.9	21.1	4.8	1.1	2.7	16.6
Y2067U + pMLPAT-Int #2	1.7	4.2	12.1	21.3	5.2	1.2	2.9	17.3

상기 입증한 바와 같이, pMLPAT-17로부터의 모르티에렐라 알피나 LPAAT2의 발현은 EPA(%)를 "대조군" 균주에서의 ~14%에서 15.5 내지 16%로 증가시켰다. 16.6 내지 17.3%로의 EPA의 추가 증가는 모르티에렐라 알피나 LPAAT2가 pMLPAT-Int를 사용하여 게놈 내로 통합되는 경우에 달성되었다. 추가 증가는 천연 야로위아 리폴리티카 LPAAT1 (서열 112 및 113) 및/또는 LPAAT2 (서열 115 및 116)가 예를 들어 균주 Y2067U + pMLPAT-Int에서 넉아웃 되는 경우에 예측될 것이다.

실시예 18

모르티에렐라 알피나 DGAT1 발현은 PUFA(%)를 증가시킨다.

본 실시예는 모르티에렐라 알피나 DGAT1 cDNA (서열 124)를 동시 발현하도록 형질전환된 야로위아 리폴리티카 균주 Y2067U (실시예 13)에서의 증가된 EPA 생합성 및 축적을 기재한다. Δ6 데새투라제/Δ6 엘론가제 경로 또는 Δ9 엘론가제/Δ8 데새투라제 경로를 통해 DHA를 생산하도록 유전자 조작된 야로위아 리폴리티카 숙주 균주는 모르티에렐라 알피나 DGAT1이 그에서 유사하게 동시 발현되는 경우에 (예컨대, 균주 Y3000에서) 증가된 DHA 생합성 및 축적을 입증할 수 있었음을 고려한다.

모르티에렐라 알피나 DGAT1 ORF를 하기와 같이 클로닝하였다. 첫째, cDNA의 클로닝을 돕기 위해, DGAT1의 제2 코돈의 서열을 'ACA'에서 'GCA'로 바꾸어 트레오닌에서 알라닌으로 아미노산을 변화시켰다. 모르티에렐라 알피나 DGAT1 ORF의 완전한 코딩 영역을 프라이머 MACAT-F1 및 MACAT-R (서열 405 및 406)로 증폭시켜 이를 달성하였다. 구체적으로, PCR 반응 혼합물은 프라이머 MACAT-F1 및 MACAT-R 각 20 μM 용액 1 ㎕, 모르티에렐라 알피나 cDNA (상기 참조, 실시예 16) 1 ㎕, 물 22 ㎕ 및 ExTaq 프리믹스 2X Taq PCR 용액 (다까라 바이오 인크, 일본 520-2193 시가현 오쯔 소재) 25 ㎕를 함유하였다. 증폭을 하기와 같이 수행하였다: 94℃에서 150초 동안 초기 변성, 이어서 94℃에서 30초 동안 변성, 55℃에서 30초 동안 어닐링 및 72℃에서 90초 동안 신장의 30주기. 72℃에서 10분 동안의 최종 신장 주기를 수행한 후에, 반응을 4℃에서 종결하였다. ~1600 bp DNA 단편을 PCR 반응으로부터 수득하였다. 이를 퀴아젠 PCR 정제 키트로 제조자의 프로토콜에 따라 정제하였다.

모르티에렐라 알피나 DGAT1 ORF를 NcoI-NotI 소화 플라스미드 pZUF17 (서열 162; 도 9B) 내로 삽입하여 ORF를 FBAIN 프로모터 및 PEX20-3' 종결자 영역의 조절 하에 클로닝하였다. 그러나, DGAT1 ORF가 내부 NcoI 부위를 함유하였기 때문에, 클로닝을 위해 2개의 개별 제한 효소 소화를 수행하는 것이 요구되었다. 첫째, 정제된 PCR 생성물 ~2 ㎍을 BamHI 및 NcoI로 소화시켰다. 반응 혼합물은 총 부피 60 ㎕ 중에 20 U의 각 효소 (프로메가) 및 제한 완충제 D 6 ㎕를 함유하였다. 혼합물을 2시간 동안 37℃에서 인큐베이션하였다. ~320 bp 단편을 아가로스 겔 전기영동으로 분리하고, 퀴아젠 QiaexII 겔 정제 키트로 정제하였다. 달리, NcoI를 NotI로 대체한 것 외에 상기와 동일한 반응 조건을 이용하여 정제된 PCR 생성물 ~2 ㎍을 BamHI 및 NotI로 소화시켰다. ~1280 bp 단편을 상기와 같이 단리하고, 정제하였다. 최종적으로, pZUF17 ~3 ㎍을 NcoI 및 NotI로 소화시키고, 상기와 같이 정제하하여, ~7 kB 단편을 생성하였다.

~7 kB NcoI/NotI pZUF17 단편, ~320 bp NcoI/BamHI DGAT1 단편 및 ~1280 bp BamHI/NotI DGAT1 단편을 함께 실온에서 밤새 인큐베이션하는 3-방식 라이게이션으로 라이게이션하였다. 라이게이션 혼합물은 총 부피 20 ㎕ 중에 7 kB 단편 100 ng, 및 320 bp 및 1280 bp 단편 각각 200 ng, 리가제 완충제 2 ㎕, 및 2 U T4 DNA 리가제 (프로메가)를 함유하였다. 라이게이션 생성물을 사용하여 이. 콜라이 Top10 화학 감응성 세포 (인비트로젠)를 제조자의 프로토콜에 따라 형질전환시켰다.

형질전환으로부터의 개개 콜로니 (총 12개)를 사용하여 미니프렙 분석을 위해 배양물에 접종하였다. 제한 지도분석 및 서열분석은 12개 중 5개 콜로니가 원하는 플라스미드를 가졌음을 보여주었으며, 이를 "pMDGAT1-17" (도 18B, 서열 193)이라고 명명하였다.

야로위아 리폴리티카 균주 Y2067U (실시예 13으부터)를 pMDGAT1-17 및 pZUF-MOD-1 (상기 참조, 실시예 17)로 통상의 방법에 따라 각각 형질전환시켰다. 형질전환체를 2일 동안 아미노산 보충 합성 MM에서, 이어서 4일 동안 HGM에서 성장시켰다. pMDGAT1-17을 함유한 2개의 형질전환체 및 pZUF-MOD-1을 함유한 2개의 형질전환체의 지방산 프로파일을 (상기 '통상의 방법' 항목에 기재한 바와 같이) GC 분석에 기초하여 하기 표 46에 나타낸다. 지방산을 18:0, 18:1 (올레산), 18:2 (LA), GLA, DGLA, ARA, ETA 및 EPA로서 확인하고; 각각의 조성을 총 지방산(%)으로서 나타낸다.

[표 46]

모르티에렐라 알피나 DGAT1을 과발현하도록
유전자 조작된 야로위아 균주 Y2067U 중의 지질 조성

균주	총 지방산							
	18:0	18:1	18:2	GLA	DGLA	ARA	ETA	EPA
Y2067U + pZUF-MOD-1 #1	1.31	6.92	12.03	23.11	5.72	1.05	3.80	13.20
Y2067U + pZUF-MOD-1 #2	1.39	6.83	12.15	21.99	5.83	1.07	3.82	13.47
Y2067U + pMDGAT1-17 #1	0.89	7.13	10.87	24.88	5.82	1.19	3.97	14.09
Y2067U + pMDGAT1-17 #2	0.86	7.20	10.25	22.42	6.35	1.26	4.38	15.07

상기 입증한 바와 같이, 플라스미드 pMDGAT1-17로부터의 모르티에렐라 알피나 DGAT1의 발현은 EPA(%)를 "대조군" 균주에서의 ~13.3%에서 ~14.1% ("Y2067U + pMDGAT1-17 #1") 및 ~15.1% ("Y2067U + pMDGAT1-17 #2")로 각각 증가시켰다. EPA에서의 추가 증가는 천연 야로위아 리폴리티카 DGAT1 (서열 122 및 123)이 예를 들어 균주 Y2067U + pMDGAT1-17에서 넉아웃 되는 경우에 예측될 것이다.

실시예 19

모르티에렐라 알피나 DGAT2 발현은 PUFA(%)를 증가시킨다.

본 실시예는 모르티에렐라 알피나 DGAT2 cDNA (서열 136)를 동시 발현하도록 형질전환된 야로위아 리폴리티카 균주 Y2067U (실시예 13)에서의 증가된 EPA 생합성 및 축적을 기재한다. Δ6 데새투라제/Δ6 엘론가제 경로 또는 Δ9 엘론가제/Δ8 데새투라제 경로를 통해 DHA를 생산하도록 유전자 조작된 야로위아 리폴리티카 숙주 균주는 모르티에렐라 알피나 DGAT2가 그에서 유사하게 동시 발현되는 경우에 (예컨대, 균주 Y3000에서) 증가된 DHA 생합성 및 축적을 입증할 수 있었음을 고려한다.

모르티에렐라 알피나 DGAT2 ORF를 하기와 같이 플라스미드 pZUF17 내로 클로닝하였다. 첫째, ORF를 모르티에렐라 알피나 cDNA으로부터의 프라이머 MDGAT-F 및 MDGAT-R1 (서열 407 및 408) (상기 참조, 실시예 16)로 PCR-증폭시켰다. 예측된 1015 bp 단편을 단리하고, 정제하고, NcoI 및 NotI로 소화시키고, NcoI-NotI 절단 pZUF17 벡터 (서열 162; 도 9B) 내로 클로닝하여 유전자가 야로위아 리폴리티카 FBAIN 프로모터 및 PEX20-3' 종결자 영역의 조절 하에 놓이게 하였다. 정확한 형질전환체를 미니프렙 DNA의 제한 분석에 의해 확인하고, 생성된 플라스미드를 "pMDGAT2-17" (서열 194)이라고 칭하였다.

야로위아 리폴리티카 균주 Y2067U (실시예 13으부터)를 pMDGAT2-17 및 pZUF-MOD-1 (상기 참조, 실시예 17)로 통상의 방법에 따라 각각 형질전환시켰다. 형질전환체를 2일 동안 아미노산 보충 합성 MM에서, 이어서 4일 동안 HGM에서 성장시켰다. pMDGAT2-17을 함유한 2개의 형질전환체 및 pZUF-MOD-1을 함유한 2개의 형질전환체의 지방산 프로파일을 (상기 '통상의 방법' 항목에 기재한 바와 같이) GC 분석에 기초하여 하기에 나타낸다. 지방산을 18:0, 18:1 (올레산), 18:2 (LA), GLA, DGLA, ARA, ETA 및 EPA로서 확인하고; 각각의 조성을 총 지방산(%)으로서 나타낸다.

[표 47]

모르티에렐라 알피나 DGAT2를 과발현하도록
유전자 조작된 야로위아 균주 Y2067U 중의 지질 조성

균주	총 지방산							
	18:0	18:1	18:2	GLA	DGLA	ARA	ETA	EPA
Y2067U + pZUF-MOD-1 #1	1.31	6.92	12.03	23.11	5.72	1.05	3.80	13.20
Y2067U + pZUF-MOD-1 #2	1.39	6.83	12.15	21.99	5.83	1.07	3.82	13.47
Y2067U + pMDGAT2-17 #1	0.00	7.47	10.77	25.30	5.70	1.43	3.45	15.12
Y2067U + pMDGAT2-17 #2	1.45	7.79	9.96	25.16	6.06	1.25	3.99	15.37

플라스미드 pMDGAT2-17로부터의 모르티에렐라 알피나 DGAT2의 발현은 EPA(%)를 "대조군" 균주에서의 ~13.3%에서 ~15.25% ("Y2067U + pMDGAT2-17")로 증가시켰다. EPA에서의 추가 증가는 천연 야로위아 리폴리티카 DGAT2 (서열 130 내지 135)가 예를 들어 균주 Y2067U + pMDGAT2-17에서 넉아웃 되는 경우에 예측될 것이다.

실시예 20

모르티에렐라 알피나 GPAT는 PUFA(%)를 증가시킨다.

본 실시예는 모르티에렐라 알피나 GPAT ORF (서열 138)를 동시 발현하도록 형질전환된 야로위아 리폴리티카 균주 Y2107U1 (실시예 14)에서의 증가된 DGLA 생합성 및 축적 (및 저하된 18:1의 양)을 기재한다. Δ6 데새투라제/Δ6 엘론가제 경로 또는 Δ9 엘론가제/Δ8 데새투라제 경로를 통해 DHA를 생산하도록 유전자 조작된 야로위아 리폴리티카 숙주 균주는 모르티에렐라 알피나 GPAT가 그에서 유사하게 동시 발현되는 경우에 (예컨대, 균주 Y3000에서) 증가된 DHA 생합성 및 축적을 입증할 수 있었음을 고려한다.

추정 PCR 프라이머를 사용하는 모르티에렐라 알피나 GPAT의 동정

아스페르길루스 니둘란스 (진뱅크 관리 번호 EAA62242) 및 뉴로스포라 크라싸 (진뱅크 관리 번호 XP_325840)로부터의 GPAT의 서열을 기초로 하여, 하기 프라이머를 추정 PCR을 위해 디자인하였다:

MGPAT-N1 (서열 409) CCNCAYGCNAAYCARTTYGT

MGPAT-NR5 (서열 410) TTCCANGTNGCCATNTCRTC

[주의: 서열 409 및 410에 사용된 핵산 다의성 코드는 하기와 같음: R= A/G; Y=C/T; 및 N=A/C/T/G.]

PCR 증폭은 퍼킨 엘머(Perkin Elmer) GeneAmp 9600 PCR 머신에서 다까라 ExTaq 프리믹스 Taq 폴리머라제 (다까라 바이오 인크., 일본 시가현 오쯔 소재)를 사용하여 수행하였다. 증폭을 하기와 같이 수행하였다: 94℃에서 30초 동안 변성, 55℃에서 30초 동안 어닐링 및 72℃에서 90초 동안 신장의 30주기, 이어서 72℃에서 7분 동안의 최종 신장 주기.

~1.2 kB 단편을 수득하였다 (서열 140). 이 단편을 퀴아젠 QiaQuick PCR 정제 키트로 정제하고, TOPO® 클로닝 벡터 pCR2.1-TOPO (인비트로젠) 내로 클로닝하고, 서열결정하였다. 생성된 서열은 번역되는 경우에 BLAST 프로그램 분석에 기초하여 공지된 GPAT에 상동성을 가졌다.

1212 bp cDNA 단편의 서열에 기초하여, 모르티에렐라 알피나 GPAT의 5' 및 3' 말단 영역을 PCR 증폭 및 게놈 워킹 기술로 클로닝하였다. 이는 모르티에렐라 알피나 GPAT (서열 141)의 -1050 bp 내지 + 2885 bp 영역에 상응하는 콘티그를 어셈블리할 수 있다. 이 콘티그는 GPAT의 전체 코딩 영역 및 4개의 인트론 (서열 145, 146, 147 및 148)을 포함한다.

구체적으로, 실시예 16에 기재된 모르티에렐라 알피나 cDNA 샘플 (1 ㎕)을 GPAT의 3'-말단의 증폭을 위해 주형으로 사용하였다. MGPAT-5N1 (서열 411) 및 CDSIII/3' (서열 392)를 프라이머로서 사용하였다. PCR 증폭은 퍼킨 엘머 GeneAmp 9600 PCR 머신에서 다까라 ExTaq 프리믹스 Taq 폴리머라제 (다까라 바이오 인크., 일본 시가현 오쯔 소재)를 사용하여 수행하였다. 증폭을 하기와 같이 수행하였다: 94℃에서 30초 동안 변성, 55℃에서 30초 동안 어닐링 및 72℃에서 120초 동안 신장의 30주기, 이어서 72℃에서 7분 동안의 최종 신장 주기.

PCR 생성물을 1:10으로 희석하고, 희석된 PCR 생성물 1 ㎕는 MGPAT-5N2 (서열 412) 및 CDSIII/3'를 프라이머로서 사용하는 제2회 증폭에 주형으로서 사용하였다. 조건은 상기와 정확하게 동일하였다. 제2회 PCR 생성물을 다시 1:10으로 희석하고, 희석된 PCR 생성물 1 ㎕는 MGPAT-5N3 (서열 413) 및 CDSIII/3'를 프라이머로서 사용하는 제3회 PCR의 주형으로서 사용하였다. PCR 조건은 다시 동일하였다.

~1 kB 단편을 제3회 PCR에서 생성하였다. 이 단편을 퀴아젠 PCR 정제 키트로 정제하고, 서열 분석을 위해 pCR2.1-TOPO 벡터 내로 클로닝하였다. 서열 분석으로부터의 결과는 이 965 bp 단편 (서열 142)이 GPAT 유전자의 3'-말단에 상응함을 보여주었다.

클론테크 유니버셜 게놈워커™ 키트를 사용하여 모르티에렐라 알피나 GPAT의 5'-말단 영역에 상응하는 게놈 DNA의 한 조각을 수득하였다. 간략하게 설명하면, 각 모르티에렐라 알피나 게놈 DNA 2.5 ㎍을 DraI, EcoRV, PvuII 또는 StuI로 개별적으로 소화시키고, 소화된 DNA 샘플을 퀴아젠 Qiaquick PCR 정제 키트로 정제하고, 각 키트 완충제 EB 30 ㎕로 용출시키고, 정제된 샘플을 이어서 하기 나타내는 게놈 워커 어댑터 (서열 414 [상부 가닥] 및 415 [하부 가닥])로 라이게이션하였다:

```
5′-GTAATACGACTCACTATAGGGCACGCGTGGTCGACGGCCCGGGCTGGT-3′
                                    3′-H2N-CCCGACCA-5′
```

각각의 라이게이션 반응 혼합물은 25 μM 게놈 워커 어댑터 1.9 ㎕, 10X 라이게이션 완충제 1.6 ㎕, T4 DNA 리가제 0.5 ㎕ 및 정제된 소화 게놈 DNA 샘플 4 ㎕를 함유하였다. 반응 혼합물을 16℃에서 밤새 인큐베이션하였다. 반응을 70℃에서 5분 동안 인큐베이션에 의해 종결하였다. 이어서, 10 mM TrisHCl 72 ㎕, 1 mM EDTA, pH 7.4 완충제를 각 라이게이션 반응 믹스에 첨가하였다.

4개 개별 PCR 반응을 수행하고, 이들 각각은 4개의 라이게이션 혼합물 중 하나를 주형으로서 사용하였다. PCR 반응 혼합물은 라이게이션 혼합물 1 ㎕, 20 μM MGPAT-5-1A (서열 416) 0.5 ㎕, 10 μM 키트 프라이머 AP1 (서열 417) 1 ㎕, 물 22.5 ㎕, 및 ExTaq 프리믹스 Taq 2X PCR 용액 (다까라) 25 ㎕를 함유하였다. PCR 반응은 하기 조건을 사용하는 32주기 동안 수행하였다: 94℃에서 30초 동안 변성, 55℃에서 30초 동안 어닐링 및 72℃에서 180초 동안 신장. 72℃에서 7분 동안의 최종 신장 주기를 수행한 후에, 반응을 4℃에서 종결하였다.

각각의 PCR 반응의 생성물을 1:50으로 개별적으로 희석하고, 제2회 PCR의 주형으로서 사용하였다. 각각의 반응 혼합물은 주형으로서의 희석된 PCR 생성물 중 하나 1 ㎕, 20 μM MGPAT-3N1 (서열 418) 0.5 ㎕, 10 μM 키트 프라이머 AP2 (서열 419) 21 ㎕, 물 22.5 ㎕ 및 ExTaq 프리믹스 Taq 2X PCR 용액 (다까라) 25 ㎕를 함유하였다. PCR 반응을 상기한 바와 동일한 열순환기 조건으로 32주기 동안 수행하였다.

DNA 단편을 제2회 PCR로부터 수득하였다. 이 단편을 정제하고, pCR2.1-TOPO 내로 클로닝하고, 서열결정하였다. 서열 분석은 1908 bp 단편 (서열 143)이 모르티에렐라 알피나 GPAT 유전자의 5'-말단이었음을 보여주었다.

유사하게, 966 bp 단편 (서열 144)은 프라이머 MGPAT-5N1을 제1회 PCR에 대한 유전자 특이적 프라이머로서 사용하고, 프라이머 MGPAT-5N2를 제2회 PCR에 대한 유전자 특이적 프라이머로서 사용하는 것을 제외하고는 상기한 2개 회수의 게놈 워킹에 의해 수득하였다. 이 단편을 또한 정제하고, pCR2.1-TOPO 내로 클로닝하고, 서열결정하였다. 서열 분석은 이들이 GPAT 유전자의 부분을 함유함을 보여주었으나, 단편은 유전자의 말단까지 연장하기에 충분히 길지 않았다. 3' cDNA 서열 (서열 142)과의 비교는 ORF의 최종 171 bp가 함유되지 않았음을 보여주었다.

<u>모르티에렐라 알피나로부터의 전장 GPAT 서열의 어셈블리</u>

완전한 GPAT 유전자 (GPAT 번역 개시 'ATG' 코돈의 상류로 1050개 염기를 연장하고, GPAT 종결 코돈 너머로 22개 염기를 연장하는 영역을 포함함)를 함유한 3935 bp 서열 (서열 141)을 원래 부분적 cDNA 단편 (서열 140), 3' cDNA 단편 (서열 142), 내부 게놈 단편 (서열 144), 및 5' 게놈 단편 (서열 143)(상기함)(도 19에 도시함)의 서열로부터 어셈블리하였다. 이 영역에는 2151 bp GPAT ORF가 포함된다. 'ATG'에서 정지 코돈 TAG'로의 모르티에렐라 알피나 GPAT ORF의 완전한 뉴클레오티드 서열은 서열 138 (서열 141의 염기 1050 내지 2863에 상응함, 4개 인트론 (즉, 인트론 1 [서열 145], 서열 141의 염기 1195 내지 1469에 상응함; 인트론 2 [서열 146], 서열 141의 염기 1585 내지 1839에 상응함; 인트론 3 [서열 147], 서열 141의 염기 2795 내지 2877에 상응함 및 인트론 4 [서열 148], 서열 141의 염기 2940 내지 3038에 상응함)을 제외함)을 제공한다. 번역된 아미노산 서열 (서열 139)은 수많은 진균, 식물 및 동물 GPAT와 상동성을 보였다.

보다 구체적으로, 서열의 동일성을 BLAST (베이직 로컬 얼라인먼트 서치 툴(Basic Local Alignment Search Tool); [Altschul, S. F., et al., J. Mol. Biol. 215:403-410 (1993)]) 검색을 수행하여 측정하였다. 본원에서 서열 139라고 기재한 아미노산 단편은 우스틸라고 마이디스(*Ustilago maydis*)의 추정적 GPAT의 단백질 서열 (진뱅크 관리 번호 EAK84237)과 47% 동일성 및 65% 유사성을 가지며, 기대값이 1e-152이고; 추가로, 서열 139는 아스페르길루스 푸미가투스의 GPAT (진뱅크 관리 번호 EAL20089)와 47% 동일성 및 62% 유사성을 가지며, 기대값이 1e-142이었다.

<u>FBAIN::MGPAT::PEX20-3' 키메라 유전자를 포함하는 플라스미드 pMGPAT-17의 구축</u>

모르티에렐라 알피나 GPAT ORF를 하기와 같이 클로닝하였다. 프라이머 MGPAT-cDNA-5 및 MGPAT-cDNA-R (서열 420 및 421)을 사용하여 GPAT ORF를 모르티에렐라 알피나의 cDNA로부터 PCR에 의해 증폭시켰다. 반응 혼합물은 cDNA 1 ㎕, 각 프라이머 1 ㎕, 물 22 ㎕ 및 ExTaq 프리믹스 2X Taq PCR 용액 (다까라 바이오 인크., 일본 520-2193 시가현 오쯔 소재) 25 ㎕를 함유하였다. 증폭을 하기와 같이 수행하였다: 94℃에서 150초 동안 초기 변성, 이어서 94℃에서 30초 동안 변성, 55℃에서 30초 동안 어닐링 및 72℃에서 120초 동안 신장의 30주기. 72℃에서 10분 동안의 최종 신장 주기를 수행한 후에, 반응을 4℃에서 종결하였다. ~2.2 kB DNA 단편을 PCR 반응으로부터 수득하였다. 이를 퀴아젠 PCR 정제 키트로 제조자의 프로토콜에 따라 정제하였다.

정제된 PCR 생성물을 BamHI 및 EcoRI로 소화시키고, ~470 bp 단편을 겔 아가로스 전기영동에 의해 단리하고, 퀴아젠 겔 정제 키트로 정제하였다. 별법으로, PCR 생성물을 또한 EcoRI 및 NotI로 절단하고, 1.69 kB 단편을 상기와 같이 단리하고, 정제하였다. 2개의 단편을 BamHI 및 NotI 절단 pZUF-MOD-1 벡터 (서열 192; 도 18A) 내로 라이게이션하여 유전자가 야로위아 리폴리티카에서의 발현을 위한 자가-복제 벡터 중에서 야로위아 리폴리티카 FBAIN 프로모터 및 PEX20-3' 종결자 영역의 조절 하에 놓이게 하였다. 정확한 형질전환체를 미니프렙 DNA의 제한 분석에 의해 확인하고, 생성된 플라스미드를 "pMGPAT-17" (서열 195; 도 18C)이라고 칭하였다.

모르티에렐라 알피나 GPAT를 과발현하는 형질전환체 야로위아 리폴리티카 중의 지질 조성의 분석

야로위아 리폴리티카 균주 Y2107U1 (실시예 14로부터)을 플라스미드 pMGPAT-17 및 플라스미드 pZUF-MOD-1 (상기 참조, 실시예 17)로 통상의 방법에 따라 각각 형질전환시켰다. 형질전환체를 2일 동안 아미노산 보충 합성 MM에서, 이어서 4일 동안 HGM에서 성장시켰다. pZUF-MOD-1을 함유한 2개의 형질전환체 및 pMGPAT-17을 함유한 4개의 형질전환체의 지방산 프로파일을 (상기 '통상의 방법' 항목에 기재한 바와 같이) GC 분석에 기초하여 하기 표에 나타낸다. 지방산을 18:0, 18:1 (올레산), 18:2 (LA), GLA, DGLA, ARA, ETA 및 EPA로서 확인하고; 각각의 조성을 총 지방산(%)으로서 나타낸다.

[표 48]

모르티에렐라 알피나 GPAT를 과발현하도록 유전자 조작된 야로위아 균주 Y2107U1 중의 지질 조성

균주	총 지방산							
	18:0	18:1	18:2	GLA	DGLA	ARA	ETA	EPA
Y2107U1 + pZUF-MOD-1 #1	2.8	22.7	9.8	28.5	2.7	1.7	0.4	17.4
Y2107U1 + pZUF-MOD-1 #2	2.5	23.4	10.3	28.7	2.5	1.5	0.3	16.8
Y2107U1+ pMGPAT-17 #1	3.2	14.8	11.7	29.8	5.6	2.0	0.3	18.4
Y2107U1 + pMGPAT-17 #2	2.9	16.3	11.7	28.3	6.1	1.8	0.4	16.9
Y2107U1 + pMGPAT-17 #3	2.1	14.3	10.8	27.5	7.2	1.4	0.4	17.4
Y2107U1 + pMGPAT-17 #4	2.7	15.7	11.5	29.1	6.3	1.7	0.4	17.3

상기 입증한 바와 같이, pMGPAT-17로부터의 모르티에렐라 알피나 GPAT의 발현은 DGLA(%)를 "대조군" 균주에서의 ~2.5%에서 6.5%로 증가시켰다. 18:1의 수준은 ~23%에서 ~16%로 저하되었다. DGLA (또는 임의의 다른 하류 PUFA)에서의 추가 증가는 천연 야로위아 리폴리티카 GPAT가 pMGPAT-17을 발현하는 형질전환체 균주에서 넉아웃 되는 경우에 예측될 것이다.

실시예 21

모르티에렐라 알피나 지방산 엘론가제 "ELO3"은 PUFA(%)를 증가시킨다.

본 실시예는 모르티에렐라 알피나 $C_{16/18}$ 지방산 엘론가제 ("ELO3", 서열 86)를 동시 발현하도록 형질전환된 야로위아 리폴리티카 균주 Y2031 (실시예 7)에서 대조군 균주에 상대적으로 35% 증가의 C18 지방산 (18:0, 18:1, 18:2 및 GLA) 및 31% 저하의 C16 지방산을 기재한다. ELO3 (증가된 발현을 위해 임의로 코돈-최적화될 수 있음)이 원하는 PUFA, 즉

DHA의 생성을 증가시키기 위한 수단으로서 유전자 조작된 Δ6 데새투라제/Δ6 엘론가제 경로 또는 Δ9 엘론가제/Δ8 데새투라제 경로 내로 탄소 플럭스를 가할 수 있음을 고려한다. 예를 들어, 이 $C_{16/18}$ 지방산 엘론가제를 포함하는 키메라 유전자는 예를 들어, 균주 Y3000 내로 쉽게 도입될 수 있었다.

모르티에렐라 알피나 $C_{16/18}$ 지방산 엘론가제의 서열 확인

모르티에렐라 알피나 지방산 엘론가제의 부분을 코딩하는 cDNA 단편 (서열 88)을 9,984개 모르티에렐라 알피나 cDNA 서열 (실시예 16) 중에서 동정하였다. 이 단편은 수많은 지방산 엘론가제에 유의한 상동성을 가지고, 따라서 엘론가제로서 임시적으로 동정되었다.

서열 88이 가장 유사성을 갖는 서열을 요약한 BLAST 비교의 결과를 동일성(%), 유사성(%), 및 기대값에 따라 보고한다. 구체적으로, 서열 88의 번역된 아미노산 서열은 칸디다 알비칸스 SC5314 (진뱅크 관리 번호 EAL04510.1, 이는 사카로마이세스 세레비지애 EUR4, FEN1 및 ELO1과 유사한 3개의 잠재적인 지방산 엘론가제 유전자 중 하나로 본원에 주석을 달음)로부터의 잠재적인 지방산 엘론가제의 단백질 서열과 32% 동일성 및 46% 유사성을 가지며, 기대값이 4e-13이었다. 추가로, 서열 88은 사카로마이세스 세레비지애 (진뱅크 관리 번호 NC_001142, 염색체 X의 염기 67849 내지 68781)로부터의 ELO1과 35% 동일성 및 53% 유사성을 가졌다. 불포화 C12-C16 지방 아실-CoA에서 C16-C18 지방산으로의 카르복시-말단 신장을 촉매하는 배지-쇄 아실 엘론가제로서 사카로마이세스 세레비지애 ELO1을 기재한다.

상기 보고한 상동성을 기초하여, 서열 88의 야로위아 리폴리티카 유전자 생성물을 본원에서 "엘론가제 3" 또는 "ELO3"으로서 칭하였다.

부분적 지방산 엘론가제 cDNA 서열 (서열 88)의 분석은 5' 및 3'-말단이 모두 불완전하였음을 나타냈다. 모르티에렐라 알피나 ELO3의 누락된 3' 영역을 얻기 위해, 클론테크 유니버셜 게놈워커™ 키트를 사용하였다 (실시예 20에 기재함). 구체적으로, MA Elong 3'1 (서열 422) 및 AP1을 프라이머 (즉, 프라이머 MGPAT-5-1A 및 AP1 대신에)로서 사용한 것 외에는 4개의 라이게이션 혼합물의 동일한 세트를 제1회의 PCR에 이전에 기재한 동일한 성분 및 조건으로 사용하였다. 제2회의 PCR은 MA Elong 3'2 (서열 423) 및 AP2를 프라이머로서 사용하였다. 1042 bp DNA 단편을 제2회의 PCR로부터 수득하였다 (서열 89). 이 단편을 정제하고, pCR2.1-TOPO 내로 클로닝하고, 서열결정하였다. 서열 분석은 단편이 유전자의 'TAA' 정지 코돈의 ~640 bp 하류를 포함하는 ELO3의 3'-말단을 함유하였음을 보여주었다.

클론테크 3'-말단 RACE (상기 참조)에 사용된 4개의 라이게이션 혼합물의 동일한 세트를 또한 사용하여 모르티에렐라 알피나 ELO3의 5'-말단 영역을 수득하였다. 구체적으로, MA Elong 5'1 (서열 424, 5' 말단에 존재) 및 AP1을 프라이머 (즉, 프라이머 MA Elong 3'1 및 AP1 대신에)로서 사용한 것 외에는 제1회의 PCR을 상기한 동일한 성분 및 조건으로 수행하였다. 제2회의 PCR은 MA Elong 5'2 (서열 425, 5' 말단에 존재) 및 AP2를 프라이머로서 사용하였다. 2223 bp DNA 단편 (서열 90)을 수득하였다. 이를 정제하고, pCR2.1-TOPO 내로 클로닝하고, 서열결정하였다. 서열 분석은 단편이 ELO3 유전자의 5'-영역을 함유하였음을 보여주었다.

따라서, 모르티에렐라 알피나 ELO3 (서열 91)의 전체 cDNA 서열은 원래 부분적 cDNA 서열 (서열 88)을 게놈 워킹 (서열 90 및 89, 각각 도 20에 도시함)에 의해 수득한 중복 5' 및 3' 서열과 합하여 수득하였다. 이로써 ELO3의 추정적 'ATG' 번역 개시 코돈의 상류 2091 bp; ELO3 ORF 828 bp (즉, 서열 86, 서열 91의 염기 2092 내지 2919에 상응함); 및 ELO3 정지 코돈의 하류 638 bp (서열 91의 염기 2920 내지 3557에 상응함)를 포함하는, 본원에서 서열 91로서 동정된 3557 bp 서열을 수득하였다.

모르티에렐라 알피나 ELO3의 상응하는 게놈 서열은 서열 91로서 제공된 cDNA 단편보다 길다. 구체적으로, 542 bp 인트론 (서열 92)을 ORF의 318 bp에서 ELO3 유전자를 함유한 게놈 DNA에서 발견하였으며; 따라서, 본원에서 서열 93으로서 제공된 게놈 DNA 단편은 4,099 bp이다 (도 20).

번역된 ELO3 단백질 서열 (서열 87)은 BLAST 프로그램 분석에 기초하여 하기 상동성을 가졌다: 칸디다 알비칸스 SC5314로부터의 잠재적인 지방산 엘론가제 (진뱅크 관리 번호 EAL04510.1)와의 37% 동일성 및 51% 유사성, 기대값 4e-43. 추가로, 번역된 ELO3은 뉴로스포라 크라싸로부터의 XP_331368 ("가상 단백질"로서 본원에 주석을 달음)의 단백질 서열과 33% 동일성 및 44% 유사성을 공유하며, 기대값이 3e-44이었다.

FBAIN::ELO3::PEX16-3' 키메라 유전자를 포함하는 플라스미드 pZUF6S-E3WT의 구축

모르티에렐라 알피나 지방산 ELO3 ORF를 하기와 같이 클로닝하였다. 프라이머 MA Elong 5' NcoI 3 및 MA Elong 3' NotI (서열 426 및 427)을 사용하여 ELO3 ORF를 모르티에렐라 알피나의 cDNA (실시예 16)로부터 PCR에 의해 증폭시켰다. 반응 혼합물은 cDNA 1 ㎕, 각 프라이머 1 ㎕, 물 22 ㎕ 및 ExTaq 프리믹스 2X Taq PCR 용액 (다까라) 25 ㎕를 함유하였다. 증폭을 하기와 같이 수행하였다: 94℃에서 30초 동안 초기 변성, 이어서 94℃에서 30초 동안 변성, 55℃에서 30초 어닐링, 72℃에서 120초 동안 신장의 32 주기. 72℃에서 7분 동안의 최종 신장 주기를 수행하고, 이어서 4℃에서 반응을 종결하였다. ~830 bp DNA 단편을 PCR 반응으로부터 수득하였다. 이를 퀴아젠 (미국 캘리포니아주 발렌시아 소재) PCR 정제 키트로 제조자의 프로토콜에 따라 정제하였다. 정제된 PCR 생성물을 2개 분취액으로 나누었으며, 여기서 하나는 NcoI 및 NspI로 소화시키는 반면에, 다른 하나는 NspI 및 NotI로 소화시켰다. ~270 bp NcoI-NspI 및 ~560 bp NspI-NotI 단편을 NcoI-NotI 절단 pZF5T-PPC 벡터 (도 13B, 서열 170) 내로 3-조각 라이게이션에 의해 클로닝하여 유전자가 야로위아 리폴리티카에서의 발현을 위한 자가-복제 벡터 중에서 야로위아 리폴리티카 FBAIN 프로모터 및 PEX16-3' 종결자 영역 (진뱅크 관리 번호 U75433)의 조절 하에 놓이게 하였다. 정확한 형질전환체를 미니프렙 분석에 의해 확인하고, 생성된 플라스미드를 "pZF5T-PPC-E3" (서열 196)이라고 칭하였다.

플라스미드 pZF5T-PPC-E3을 ClaI 및 PacI로 소화시키고, ~2.2 kB 밴드 (즉, FBAIN::ELO3::PEX16-3' 단편)를 아가로스 겔로부터 퀴아젠 겔 추출 키트로 정제하였다. 상기 단편을 Pex20-3' 종결자와 FBAIN 프로모터의 조절 하의 모르티에렐라 알피나 Δ6 데새투라제 ORF ("D6S"; 진뱅크 관리 번호 AF465281) (즉, FBAIN::D6S::Pex20 키메라 유전자) 및 Ura3 유전자를 함유한 자가-복제 플라스미드인 ClaI-PacI 절단 pZUF6S (도 21A, 서열 197) 내로 클로닝하였다. 정확한 형질전환체를 미니프렙 분석에 의해 확인하고, 생성된 플라스미드를 "pZUF6S-E3WT" (도 21B, 서열 198)라고 칭하였다.

모르티에렐라 알피나 ELO3을 과발현하는 형질전환체 야로위아 리폴리티카 중의 지질 조성의 분석

야로위아 리폴리티카 균주 Y2031 (실시예 7)을 플라스미드 pZUF6S (대조군, FBAIN::D6S::Pex20 키메라 유전자 포함) 및 플라스미드 pZUF6S-E3WT (FBAIN::D6S::Pex20 키메라 유전자 및 FBAIN::ELO3::PEX16 키메라 유전자 포함)로 통상의 방법에 따라 형질전환시켰다. 형질전환체를 2일 동안 아미노산 보충 합성 MM에서, 이어서 4일 동안 HGM에서 성장시켰다. pZUF6S를 함유한 6개의 클론 (클론 #1-6, 단일 형질전환으로부터) 및 pZUF6S-E3WT를 함유한 22개의 클론 (4개의 상이한 형질전환 [즉, #3, 5, 6, 및 7]으로부터)의 지방산 프로파일을 (상기 '통상의 방법' 항목에 기재한 바와 같이) GC 분석에 기초하여 하기 표 49에 나타낸다. 지방산을 16:0 (팔미테이트), 16:1 (팔미트올레산), 18:0, 18:1 (올레산), 18:2 (LA) 및 GLA로서 확인하고; 각각의 조성을 총 지방산(%)으로서 나타낸다.

[표 49]

모르티에렐라 알피나 EL03을
과발현하도록 유전자 조작된
야로위아 균주 Y2031 중의 지질 조성

야로위아 리폴리티카 균주 Y2031 형질전환체 및/또는 클론 번호	지방산 조성 (총 지방산 (%))					
	16:0	16:1	18:0	18:1	18:2	GLA
pZUF6S #1 (대조군)	9.0	23.2	1.2	38.2	19.8	6.9
pZUF6S #2 (대조군)	10.1	23.4	1.4	39.0	17.5	7.1
pZUF6S #3 (대조군)	9.7	22.7	1.4	39.0	20.2	7.0
pZUF6S #4 (대조군)	8.5	24.1	0.0	40.8	19.8	6.9
pZUF6S #5 (대조군)	9.8	22.4	1.7	39.1	20.2	6.8
pZUF6S #6 (대조군)	9.1	22.7	1.9	39.9	19.7	6.6
pZUF6S-E3WT #3-1	8.9	17.3	4.1	36.5	21.6	11.6
pZUF6S-E3WT #3-2	8.8	17.8	3.7	36.9	21.3	11.5
pZUF6S-E3WT #3-3	***8.9***	***18.3***	***3.5***	***33.8***	***35.4***	***0.0***
pZUF6S-E3WT #3-6	8.5	19.9	4.4	37.8	17.1	12.3
pZUF6S-E3WT #5-1	8.6	17.6	4.0	37.6	21.1	11.1
pZUF6S-E3WT #5-2	8.8	17.1	3.9	37.6	21.3	11.2
pZUF6S-E3WT #5-3	9.1	17.1	3.5	37.6	21.5	11.1
pZUF6S-E3WT #5-4	8.8	17.9	4.3	38.0	19.3	11.7
pZUF6S-E3WT #5-5	9.2	16.1	4.4	37.0	21.6	11.7
pZUF6S-E3WT #5-6	***8.7***	***21.5***	***4.2***	***30.3***	***35.3***	***0.0***
pZUF6S-E3WT #6-1	9.4	16.9	4.6	36.6	21.5	11.0
pZUF6S-E3WT #6-2	9.8	16.2	4.1	36.5	21.9	11.6
pZUF6S-E3WT #6-3	9.4	17.0	4.4	36.2	21.8	11.3
pZUF6S-E3WT #6-4	8.3	16.6	4.2	36.9	21.9	12.2
pZUF6S-E3WT #6-5	8.8	18.5	5.5	36.0	17.8	13.4
pZUF6S-E3WT #6-6	8.7	19.5	5.2	35.4	18.1	13.2
pZUF6S-E3WT #7-1	***0.0***	***30.6***	***0.0***	***35.5***	***18.2***	***15.8***
pZUF6S-E3WT #7-2	8.0	17.7	4.0	37.7	20.9	11.7
pZUF6S-E3WT #7-3	***0.0***	***26.7***	***4.2***	***36.0***	***21.4***	***11.7***
pZUF6S-E3WT #7-4	***0.0***	***28.1***	***4.3***	***37.0***	***16.9***	***13.6***
pZUF6S-E3WT #7-5	8.3	17.0	4.7	36.7	21.2	12.1
pZUF6S-E3WT #7-6	8.0	18.0	4.8	36.3	20.8	12.1

일부 샘플 (볼드체 및 이탤릭체로 표기함)은 예측 판독에서 벗어났다. 구체적으로, Y2031+ pZUF6S-E3WT #3-3 또는 Y2031+ pZUF6S-E3WT #5-6은 GLA를 생산하지 않았다. 유사하게, Y2031+ pZUF6S-E3WT #7-1, #7-3 및 #7-4는 GC 에러를 가졌으며, 여기서 16:0 및 16:1 피크는 GC에 의해 단일 피크로서 판독되었다. 이들 상이한 결과로서, 표 50은 대조군 및 ELO3을 발현하는 형질전환체 균주에서의 평균 지질을 보고한다. 구체적으로, 표 50은 표 49에 부정확 것인 볼드체 및 이탤릭체로 나타낸 줄을 이들 평균 계산 시 포함하지 않은채 표 49에서의 지방산 프로파일로부터의 평균을 나타낸다. "총 C16"은 16:0 및 16:1의 평균 면적의 합을 나타내고, "총 C18"은 18:0, 18:1, 18:2 및 GLA의 평균 면적의 합을 반영한다.

[표 50]

모르티에렐라 알피나 EL03을
과발현하도록 유전자 조작된
야로위아 균주 Y2031 중의 평균 지질 조성

야로위아 리폴리티카 균주 Y2031 형질전환체	평균 지방산 조성 (총 지방산(%))						총 C16	총 C18
	16:0	16:1	18:0	18:1	18:2	GLA		
pZUF6S (대조군)	9.4	23.1	1.3	39.3	19.5	6.9	32.4	67.1
pZUF6S-E3WT #3	8.7	18.3	4.1	37.1	20.0	11.8	27.0	73.0
pZUF6S-E3WT #5	8.9	17.2	4.0	37.6	21.0	11.4	26.1	73.9
pZUF6S-E3WT #6	9.1	17.5	4.6	36.3	20.5	12.1	26.5	73.5
pZUF6S-E3WT #7	8.1	17.6	4.5	36.9	21.0	12.0	25.6	74.4

상기 보고한 데이타를 기초하여, 모르티에렐라 알피나 ELO3의 과발현은 모르티에렐라 알피나 Δ6 데새투라제만을 과발현하는 Y2031의 대조군 균주에 상대적으로 야로위아 리폴리티카 균주 Y2031에서 모르티에렐라 알피나 Δ6 데새투라제와 동시 발현하는 경우에 C18(%)을 증가시키고, C16(%)을 저하시켰다. 이는 모르티에렐라 알피나 ELO3이 실제로 $C_{16/18}$ 지방산 엘론가제이었음을 나타냈다.

실시예 22

야로위아 $C_{16/18}$ 지방산 엘론가제 "YE2"는 PUFA(%)를 증가시킨다

본 실시예는 야로위아 리폴리티카 $C_{16/18}$ 지방산 엘론가제 ("YE2", 서열 94)를 동시 발현하도록 형질전환된 야로위아 리폴리티카 균주 Y2031 (실시예 7)에서 증가된 GLA 생합성 및 축적을 기재한다. YE2 엘론가제가 원하는 PUFA, 즉 DHA의 생성을 증가시키기 위한 수단으로서 유전자 조작된 Δ6 데새투라제/Δ6 엘론가제 경로 또는 Δ9 엘론가제/Δ8 데새투라제 경로 내로 탄소 플럭스를 가할 수 있음을 고려한다. 예를 들어, 이 $C_{16/18}$ 지방산 엘론가제를 포함하는 키메라 유전자는 예를 들어, 균주 Y3000 내로 쉽게 도입될 수 있었다.

야로위아 리폴리티카 $C_{16/18}$ 지방산 엘론가제의 서열 확인

야로위아 리폴리티카로부터의 신규 지방산 엘론가제 후보는 래트 Elo2 $C_{16}/_{18}$ 지방산 엘론가제 단백질 서열 (진뱅크 관리 번호 AB071986, 서열 84)을 조회 서열로서 사용하는 서열 비교에 의해 동정하였다. 구체적으로, 이 rElo2 조회 서열을 사용하여 진뱅크 및 "효모 프로젝트 게놀레부레스"의 공개 야로위아 리폴리티카 단백질 데이타베이스 (프랑스 탈렌스 세덱스 LaBRI 바이오인포르마틱스 센터) (또한, 문헌 [Dujon, B. et al., Nature 430 (6995):35-44 (2004)] 참조)를 조사하였다. 이로써 "무명의 단백질 생성물"로서 주석을 달은 상동성 서열인 진뱅크 관리 번호 CAG77901 (서열 94 및 95)을 동정하였다. 이 유전자를 YE2라고 칭하였다.

BLAST 알고리즘 (문헌 [Altschul, S. F., et al., Nucleic Acids Res. 25:3389-3402 (1997)] 참조)을 사용해 야로위아 YE2 아미노산 서열을 공개 데이타베이스와 비교하여 가장 유사한 공지 아미노산 서열이 칸디다 알비칸스 SC5314 (서열 96, 진뱅크 관리 번호 EAL04510)의 것 (가능한 지방산 엘론가제로서 주석을 달음)임을 밝혀냈다. 상기 단백질들은 약 40% 동일성을 공유하고, 스코어가 236이며, E 값은 7e-61이었다.

야로위아 YE2 유전자의 단리

YE2 유전자의 코딩 영역은 야로위아 게놈 DNA를 주형으로서 사용하고, 올리고뉴클레오티드 YL597 및 YL598 (서열 428 및 429)을 프라이머로서 사용하여 PCR에 의해 증폭시켰다. PCR 반응을 상기 '통상의 방법' 항목에 기재한 바와 같이 총 부피 50 ㎕로 수행하였다. 열순환기 조건은 95℃에서 1분 동안, 56℃에서 30초 동안, 72℃에서 1분 동안의 35주기, 이어서 72℃에서 10분 동안의 최종 신장으로 설정하였다. YE2 코딩 영역의 PCR 생성물을 정제하고, NcoI/NotI로 소화시키고, 이어서 NcoI/NotI 소화 pZKUGPYE1-N (하기 참조, 실시예 23; 또한 도 21C, 서열 199 참조)와 라이게이션하여 pZKUGPYE2 (도 21D, 서열 200)를 생성하였다. 'ATG' 번역 개시 코돈 주변의 NcoI 부위의 첨가는 YE2의 제2 아미노산을 L에서 V로 변화시켰다.

pZKUGPYE2의 ClaI/NotI 단편 (GPAT 프로모터 및 YE2 코딩 영역 함유) 및 Aco 종결자를 함유하는 NotI/PacI 단편 (ACO 3' 종결자를 프라이머 YL325 및 YL326 [서열 430 및 431]으로 PCR 증폭시키고, 이어서 NotI/PacI로 소화시켜 제조함)을 ClaI/PacI 소화 벡터 pZUF6S와 정해진 방향으로 라이게이션하여 pZUF6YE2를 생산하였다. pZKUT16의 ClaI/NcoI 단편 (TEF 프로모터 함유) 및 pZUF6YE2의 NcoI/PacI 단편 (YE2의 코딩 영역 및 Aco 종결자 함유)을 이후 ClaI/PacI 소화 벡터 pZUF6S와 정해진 방향으로 라이게이션하여 pZUF6TYE2 (서열 201)를 생산하였다.

YE2를 과발현하는 형질전환체 야로위아 리폴리티카 중의 지질 조성의 분석

플라스미드 pZUF6S (도 21A, 서열 197) 및 pZUF6TYE2 (서열 201)를 사용하여 야로위아 균주 Y2031을 따로 형질전환시켰다. 이들 2개 플라스미드의 성분을 표 51 및 52에 기재한다.

[표 51]

플라스미드 pZUF6S (서열 197)의 기재	
서열 197 중의 RE 부위 및 뉴클레오티드	단편 및 키메라 유전자 성분의 기재
EcoRI/ClaI (3114-4510)	야로위아 자율 복제 서열 18 (ARS18; 진뱅크 관리 번호 M91600)
SalI/PacI (6022-4530)	야로위아 Ura3 유전자 (진뱅크 관리 번호 AJ306421)
EcoRI/BsiWI (6063-318)	· FBAIN: FBAIN 프로모터 (서열 214) · Δ6S: 코돈-최적화된 Δ6 데새투라제 유전자 (서열 3) (모르티에렐라 알피나 (진뱅크 관리 번호 AF465281)에서 유래됨) · Pex20: 야로위아 Pex20 유전자 (진뱅크 관리 번호 AF054613)로부터의 Pex20 종결자 서열 을 포함하는 FBAIN::Δ6S::Pex20

[표 52]

플라스미드 pZUF6TYE2 (서열 201)의 기재	
서열 201 중의 RE 부위 및 뉴클레오티드	단편 및 키메라 유전자 성분의 기재
EcoRI/ClaI (7461-8857)	야로위아 자율 복제 서열 18 (ARS18; 진뱅크 관리 번호M91600)
SalI/PacI (1907-415)	야로위아 Ura3 유전자 (진뱅크 관리 번호 AJ306421)
EcoRI/BsiWI (1948-4665)	FBAIN::Δ6S::Pex20: pZUF6에 기재된 바와 같음 (상기 참조)
ClaI/PacI (8857 -415)	· TEF: TEF 프로모터 (진뱅크 관리 번호 AF054508) · YE2: 야로위아 YE2 유전자 (서열 94; 진뱅크 관리 번호 CAG77901)의 코딩 영역 · Aco: 야로위아 Aco3 유전자 (진뱅크 관리 번호 AJ001301)의 Aco3 종결자 서열 을 포함하는 TEF::YE2::Aco

야로위아 리폴리티카 균주 Y2031 (실시예 7)을 플라스미드 pZUF6S (대조군) 및 플라스미드 pZUF6TYE2로 통상의 방법에 따라 형질전환시켰다. 형질전환체를 2일 동안 액체 MM에서 성장시켰다. pZUF6S 또는 pZUF6YE2를 각각 함유한 8개 콜로니의 지방산 프로파일을 (상기 '통상의 방법' 항목에 기재한 바와 같이) GC 분석에 기초하여 하기 표 53에 나타낸다. 지방산을 16:0 (팔미테이트), 16:1 (팔미트올레산), 18:0, 18:1 (올레산), 18:2 (LA) 및 GLA로서 확인하고; 각각의 조성을 총 지방산(%)으로서 나타낸다.

[표 53]

pZUF6S 및 pZUF6TYE2로 형질전환시킨 야로위아 균주 Y2031 중의 지방산 조성의 비교

야로위아 리폴리티카 균주 Y2031 형질전환체	지방산 조성 (총 지방산(%))					
	16:0	16:1	18:0	18:1	18:2	GLA
pZUF6S #1 (대조군)	15.4	13.8	2.5	34.1	16.8	8.3
pZUF6S #2 (대조군)	15.2	12.8	3.0	36.5	16.4	8.3
pZUF6S #3 (대조군)	15.1	12.2	3.2	36.5	17.1	8.5
pZUF6S #4 (대조군)	15.2	12.8	3.1	36.3	16.6	8.4
pZUF6S #5 (대조군)	14.9	10.9	3.6	37.4	18.0	8.7
pZUF6S #6 (대조군)	14.8	10.1	4.2	37.6	18.7	8.6
pZUF6S #7 (대조군)	14.7	11.9	3.0	36.0	17.8	9.1
pZUF6S #8 (대조군)	14.9	12.6	2.9	35.9	17.3	8.8
평균	**15.0**	**12.1**	**3.2**	**36.3**	**17.3**	**8.6**
pZUF6TYE2 #1	13.1	8.4	4.4	42.4	16.8	9.7
pZUF6TYE2 #2	13.1	7.6	5.3	40.8	18.6	9.8
pZUF6TYE2 #3	13.5	8.1	4.6	39.2	19.0	10.6
pZUF6TYE2 #4	13.4	7.4	5.7	39.9	18.7	9.8
pZUF6TYE2 #5	13.4	8.4	5.5	45.2	14.3	7.6
pZUF6TYE2 #6	13.4	7.4	5.5	39.3	19.2	10.5
pZUF6TYE2 #7	13.4	8.6	4.4	40.6	17.9	9.9
pZUF6TYE2 #8	13.2	7.5	5.4	41.2	18.0	9.7
평균	**13.3**	**8.0**	**5.0**	**41.1**	**17.8**	**9.7**

GC 분석은 pZUF6S를 갖는 Y2031 형질전환체에서 생산된 것이 총 지질의 약 27.1% C16 (C16:0 및 C16:1) 및 62.2% C18 (C18:0, C18:1, C18:2 및 GLA)임을 보여주었으며; pZUF6TYE2를 갖는 Y2031 형질전환체에서는 약 21.3% C16 및 73.6% C18이 생산되었다. 따라서, pZUF6TYE2 형질전환체에서 (pZUF6S를 갖는 형질전환체에 비해) C16의 총량은 약 21.4% 저하되었고, C18의 총량은 약 18% 증가되었다. 이들 데이타는 YE2가 $C_{16/18}$ 지방산 엘론가제로서 기능하여 야로위아에서 C18 지방산을 생산한다는 것을 입증하였다. 추가로, pZUF6S 형질전환체에서 생산되는 GLA에 상대적으로 pZUF6TYE2 형질전환체에서 약 12.8% GLA를 더 생산하였다. 이들 데이타는 YE2 엘론가제가 탄소 플럭스를 유전자 조작된 PUFA 경로 내로 가하여 최종 생성물 (즉, GLA)을 더 생산할 수 있었음을 제안하였다.

실시예 23

야로위아 $C_{14/16}$ 지방산 엘론가제 "YE1"이 PUFA(%)를 증가시킨다.

본 실시예는 야로위아 리폴리티카 $C_{14/16}$ 지방산 엘론가제 ("YE1", 서열 97)를 동시 발현하도록 형질전환된 야로위아 리폴리티카 균주 Y2031 (실시예 7)에서의 증가된 GLA 생합성 및 축적을 기재한다. YE1 엘론가제가 원하는 PUFA, 즉 DHA의 생성을 증가시키기 위한 수단으로서 유전자 조작된 Δ6 데새투라제/Δ6 엘론가제 경로 또는 Δ9 엘론가제/Δ8 데새투라제 경로 내로 탄소 플럭스를 가할 수 있음을 고려한다. 구체적으로, 이 $C_{14/16}$ 지방산 엘론가제를 포함하는 키메라 유전자는 예를 들어, 균주 Y3000 내로 쉽게 도입될 수 있었다.

야로위아 리폴리티카 $C_{14/16}$ 지방산 엘론가제의 서열 확인

야로위아 리폴리티카로부터의 신규 지방산 엘론가제 후보는 실시예 22에서 사용한 것과 유사한 방식으로 래트 Elo2 $C_{16/18}$ 지방산 엘론가제 단백질 서열 (진뱅크 관리 번호 AB071986, 서열 84)을 조회 서열로서 사용하는 서열 비교에 의해 동정하였다. 이로써 "무명의 단백질 생성물"로서 주석을 달은 상동성 서열인 진뱅크 관리 번호 CAG83378 (서열 97 및 98)을 동정하였다. 이 유전자를 "YE1"이라고 칭하였다.

BLAST 알고리즘 (문헌 [Altschul, S. F., et al., Nucleic Acids Res. 25:3389-3402 (1997)] 참조)을 사용해 야로위아 YE1 아미노산 서열을 공개 데이타베이스와 비교하여 가장 유사한 공지 서열이 뉴로스포라 크라싸로부터의 FEN1 (진뱅크 관리 번호 CAD70918, 서열 99)(YE1과 약 60% 동일성을 갖는 가능한 지방산 엘론가제)임을 밝혀냈다.

야로위아 YE1 유전자의 단리

YE1 유전자 (서열 97)의 DNA 서열은 내부 NcoI 부위를 보유한다. 야로위아 번역 모티프를 YE1 유전자의 'ATG' 번역 개시 코돈 주변에 혼입하기 위해 2-단계 전략을 사용하여 야로위아로부터의 전체 YE1 유전자를 PCR하였다. 구체적으로, 야로위아 게놈 DNA를 주형으로서 사용하여, YE1의 제1 절반은 올리고뉴클레오티드 YL567 및 YL568 (서열 432 및 433)을 프라이머로서 사용하는 PCR에 의해 증폭시키는 반면에, YE1 유전자의 제2 절반은 올리고뉴클레오티드 YL569 및 YL570 (서열 434 및 435)을 프라이머로서 유사하게 사용하여 증폭시켰다. PCR 반응을 상기 '통상의 방법' 항목에 기재한 바와 같이 총 부피 50 ㎕로 수행하였다. 열순환기 조건은 95℃에서 1분 동안, 56℃에서 30초 동안, 72℃에서 1분 동안의 35주기, 이어서 72℃에서 10분 동안의 최종 신장으로 설정하였다. YE1의 5' 부분에 상응하는 PCR 생성물을 정제하고, 이어서 NcoI 및 SacI로 소화시켜 YE1-1 단편을 수득한 반면에, YE1의 3' 부분의 PCR 생성물을 정제하고, SacI 및 NotI로 소화시켜 YE1-2 단편을 수득하였다. YE1-1 및 YE1-2 단편을 NcoI/NotI 소화 pZKUGPE1S (상기 참조, 실시예 14)와 직접 라이게이션하여 pZKUGPYE1 (도 22A, 서열 202)를 생성하였다. YE1의 내부 NcoI 부위는 이어서 pZKUGPYE1을 주형으로서 사용하고 올리고뉴클레오티드 YL571 및 YL572 (서열 436 및 437)를 프라이머로서 사용하는 부위-지정 돌연변이유발에 의해 돌연변이시킴으로써 pZKUGPYE1-N (서열 199)을 생성하였다. 서열 분석은 돌연변이가 YE1의 아미노산 서열을 변화시키지 않았음을 보여주었다. ATG 번역 개시 코돈 주변의 NcoI 부위의 첨가는 YE1의 제2 아미노산을 S에서 A로 변화시켰다.

pZF5T-PPC의 ClaI/NcoI 단편 (FBAIN 프로모터 함유) 및 pZKUGPYE1-N의 NcoI/PacI 단편 (YE1의 코딩 영역 및 Aco 종결자 함유)을 ClaI/PacI 소화 벡터 pZUF6S와 정해진 방향으로 라이게이션하여 pZUF6FYE1 (서열 203)을 생산하였다.

YE1을 과발현하는 형질전환체 야로위아 리폴리티카 중의 지질 조성의 분석

플라스미드 pZUF6S 및 pZUF6FYE1 (서열 203)을 사용하여 야로위아 균주 Y2031 (실시예 7로부터)을 통상의 방법에 따라 따로 형질전환시켰다. 대조군 플라스미드 pZUF6S (도 21A, 서열 197; FBAIN::D6S::Pex20 키메라 유전자 포함)의 성분을 실시예 22에 기재한다. pZUF6FYE1 (도 22B, 서열 203, FBAIN::D6S::Pex20 키메라 유전자 및 FBAIN::YE1::Aco 키메라 유전자 포함)의 성분을 하기 표 54에 기재한다.

[표 54]

플라스미드 pZUF6FYE1 (서열 203)의 기재	
서열 203 중의 RE 부위 및 뉴클레오티드	단편 및 키메라 유전자 성분의 기재
EcoRI/ClaI (7047-8445)	야로위아 자율 복제 서열 18 (ARS18; 진뱅크 관리 번호 M91600)
SalI/PacI (1493-1)	야로위아 Ura3 유전자 (진뱅크 관리 번호 AJ306421)
EcoRI/BsiWI (1534-4251)	FBAIN::Δ6S::Pex20: pZUF6에 기재된 바와 같음 (상기 참조, 실시예 22)
ClaI/PacI (8443-1)	· FBAIN: FBAIN 프로모터 (서열 214) · YE1: 야로위아 YE1 유전자 (서열 97; 진뱅크 관리 번호 CAG83378)의 코딩 영역 · Aco: 야로위아 Aco3 유전자 (진뱅크 관리 번호 AJ001301)로부터의 Aco3 종결자 서열을 포함하는 FBAIN::YE1::Aco

형질전환 후, 형질전환체를 2일 동안 아미노산 보충 합성 MM에서, 이어서 4일 동안 HGM에서 성장시켰다. pZUF6S를 함유한 6개 클론 및 pZUF6FYE1을 함유한 5개 클론의 지방산 프로파일을 (상기 '통상의 방법' 항목에 기재한 바와 같이) GC 분석에 기초하여 하기 표 55에 나타낸다. 지방산을 16:0 (팔미테이트), 16:1 (팔미트올레산), 18:0, 18:1 (올레산), 18:2 (LA) 및 GLA로서 확인하고; 각각의 조성을 총 지방산(%)으로서 나타낸다.

[표 55]

pZUF6S 및 pZUF6FYE1로 형질전환시킨 야로위아 균주 Y2031 중의 지방산 조성의 비교

형질전환체	지방산 조성 (총 지방산(%))				
	16:0	16:1	18:1	18:2	GLA
pZUF6S #1 (대조군)	12.9	18.2	29.6	23.5	10.7
pZUF6S #2 (대조군)	12.6	18.6	29.6	23.8	10.3
pZUF6S #3 (대조군)	13.0	17.8	29.8	23.9	10.6
pZUF6S #4 (대조군)	13.1	18.9	30.1	22.3	10.3
pZUF6S #5 (대조군)	13.0	17.8	29.6	23.4	10.9
pZUF6S #6 (대조군)	12.0	18.7	30.4	23.2	10.4
평균	**12.8**	**18.3**	**29.9**	**23.4**	**10.5**
pZUF6FYE1 #1	17.4	21.9	20.4	19.2	16.9
pZUF6FYE1 #2	16.7	22.8	21.1	19.1	16.1
pZUF6FYE1 #3	19.8	20.7	22.8	17.0	15.8
pZUF6FYE1 #4	16.8	22.4	23.7	16.1	16.8
pZUF6FYE1 #5	17.7	21.6	21.2	18.0	17.2
평균	**17.7**	**21.9**	**21.9**	**17.9**	**16.5**

GC 분석은 pZUF6S를 갖는 Y2031 형질전환체에서 생산된 것이 총 지질의 약 31.1% C16 (C16:0 + C16:1)임을 측정한 반면에, pZUF6FYE1을 갖는 Y2031 형질전환체에서는 약 39.6% C16이 생산되었다. pZUF6S를 갖는 형질전환체에 비해 pZUF6FYE1 형질전환체에서 C16의 총량은 약 26.7% 증가되었다. 따라서, 이들 데이타는 YE1이 $C_{14/16}$ 지방산 엘론가제로서 기능하여 야로위아에서 C16 지방산을 생산한다는 것을 입증하였다. 추가로, pZUF6S 형질전환체에서보다 pZUF6FYE1 형질전환체에서 57%의 GLA를 더 생산하였으며, 이는 YE1 엘론가제가 탄소 플럭스를 유전자 조작된 경로 내로 가하여 최종 생성물 (즉, GLA)을 더 생산할 수 있었음을 제안하였다.

실시예 24

야로위아 리폴리티카 CPT1 과발현은 PUFA(%)를 증가시킨다.

본 실시예는 야로위아 리폴리티카 CPT1 cDNA (서열 150)를 과발현하도록 형질전환된 야로위아 리폴리티카 균주 Y2067U (실시예 13)에서 증가된 EPA 생합성 및 축적을 기재한다. EPA의 합성을 유도하는 PUFA가 또한 증가하였다. Δ6 데새투라제/Δ6 엘론가제 경로 또는 Δ9 엘론가제/Δ8 데새투라제 경로를 통해 DHA를 생산하도록 유전자 조작된 야로위아 리폴리티카 숙주 균주가, 야로위아 리폴리티카 CPT1이 유사하게 동시 발현되는 경우에 (예컨대, 균주 Y3000에서) 증가된 DHA 생합성 및 축적을 입증할 수 있었음을 고려한다.

야로위아 리폴리티카 균주 ATCC #20326 cDNA를 하기 절차로 제조하였다. 세포를 YPD 배지 (2% 박토-효모 추출물, 3% 박토르-펩톤, 2% 글루코스) 200 mL에 1일 동안 30℃에서 성장시키고, 이어서 3750 rpm에서의 벡크만 GH3.8 로터에서 10분 동안 원심분리하여 펠렛화하고, HGM으로 2회 세척하였다. 세척된 세포를 HGM 200 mL에 재현탁하고, 추가 4시간 동안 30℃에서 성장시켰다. 이어서, 세포를 3750 rpm에서 10분 동안 4 x 50 mL 튜브 중에서 원심분리하여 수확하였다.

총 RNA를 퀴아젠 RNeasy 총 RNA 미디 키트로 단리하였다. 세포를 파괴하기 위해, 수확된 세포를 키트 완충제 RLT (제조자에 의해 지정된 바와 같이 β-머캅토에탄올을 보충함) 4 X 600 ㎕에 재현탁하고, 4개의 2 mL 스크루캡 튜브 중에서 동일한 부피의 0.5 mm 유리 비드와 혼합하였다. 바이오스펙 미니 비드 비터를 사용하여 2분 동안 균질화 셋팅에서 세포를 파괴하였다. 추가의 완충제 RLT 4 x 600 ㎕를 첨가하였다. 유리 비드 및 세포 파편을 원심분리로 제거하고, 상등액을 사용하여 총 RNA를 제조자의 프로토콜에 따라 단리하였다.

폴리A(+)RNA를 상기 총 RNA 샘플로부터 퀴아젠 올리고텍스 mRNA 정제 키트로 제조자의 프로토콜에 따라 단리하였다. 단리된 폴리A(+)RNA는 동일한 키트를 사용하여 추가로 1회 정제하여 mRNA 샘플의 순도를 보장하였다. 최종 정제된 폴리(A)+ RNA의 농도는 30.4 ng/㎕이었다.

제1 가닥 cDNA 합성을 위해 사용되는 PCR 열순환기 조건을 20초 동안 95℃, 이어서 5초 동안 95℃ 및 6분 동안 68℃의 20주기로 설정하는 것 외에 실시예 16에 기재한 바와 같이 BD-클론테크 및 폴리A(+) RNA 샘플 0.1 ㎍로 정해진 LD-PCR 방법으로 cDNA를 생성하였다. PCR 생성물을 아가로스 겔 전기영동 및 에티듐 브로마이드 염색으로 정량하였다.

야로위아 리폴리티카 CPT1 cDNA를 하기와 같이 클로닝하였다. 프라이머 CPT1-5'-NcoI 및 CPT1-3'-NotI (서열 438 및 439)를 사용하여 야로위아 리폴리티카의 cDNA로부터 야로위아 리폴리티카 ORF를 PCR에 의해 증폭시켰다. 반응 혼합물은 cDNA 0.5 ㎕, 각 프라이머 0.5 ㎕, 물 11 ㎕ 및 ExTaq 프리믹스 2X Taq PCR 용액 (다까라 바이오 인크., 일본 520-2193 시가현 오쯔 소재) 12.5 ㎕를 함유하였다. 증폭을 하기와 같이 수행하였다: 94℃에서 300초 동안 초기 변성, 이어서 94℃에서 30초 동안 변성, 55℃에서 30초 동안 어닐링 및 72℃에서 60초 동안 신장의 30주기. 72℃에서 10분 동안의 최종 신장 주기를 수행한 후에, 반응을 4℃에서 종결하였다. ~1190 bp DNA 단편을 PCR 반응으로부터 수득하였다. 이를 퀴아젠 PCR 정제 키트로 제조자의 프로토콜에 따라 정제하였다. 정제된 PCR 생성물을 NcoI 및 NotI로 소화시키고, NcoI-NotI 절단 pZUF17 벡터 (서열 162; 도 9B) 내로 클로닝하여 유전자가 야로위아 리폴리티카 FBAIN 프로모터 및 PEX20-3' 종결자 영역의 조절 하에 놓이게 하였다. 정확한 형질전환체를 미니프렙 분석에 의해 확인하고, 생성된 플라스미드를 "pYCPT1-17" (서열 204)이라고 칭하였다.

키메라 FBAIN::CPT1::PEX20 유전자를 야로위아 리폴리티카의 게놈 내로 통합하기 위해, pYCPT1-17을 NcoI 및 NotI로 소화시키고, CPT1 ORF를 함유한 ~1190 bp 단편을 단리함으로써 플라스미드 pYCPT1-ZP217을 생성하였다. 이 단편을 이어서 NcoI 및 NotI로 소화시킨 pZP217 + Ura (서열 205) 내로 클로닝하였다. 도 22C에 나타난 바와 같이, 플라스미드 pZP217 + Ura는 키메라 TEF::합성 Δ17 데새투라제 (야로위아 리폴리티카에 코돈-최적화됨)::Pex20-3' 유전자 및 Ura3 유전자 (선별가능한 마커로서 사용하기 위함)를 포함하는 야로위아 리폴리티카 통합 플라스미드이다. 정확한 형질전환체를 미니프렙 분석에 의해 확인하고, 생성된 플라스미드를 "pYCPT1-ZP217" (서열 206)이라고 칭하였다.

야로위아 리폴리티카 균주 Y2067U (실시예 13으로부터)를 BssHII/BbuI 소화 pYCPT1-ZP217 및 pZUF-MOD-1 (상기 참조, 실시예 7)로 통상의 방법에 따라 각각 형질전환시켰다. 형질전환체를 2일 동안 아미노산 보충 합성 MM에서, 이어서 4일 동안 HGM에서 성장시켰다. pZUF-MOD-1을 함유한 2개의 형질전환체 및 pYCPT1-ZP217을 함유한 4개의 형질전환체 (게놈 내에 통합됨)의 지방산 프로파일을 (상기 '통상의 방법' 항목에 기재한 바와 같이) GC 분석에 기초하여 하기 표에 나타낸다. 지방산을 18:0, 18:1 (올레산), 18:2 (LA), GLA, DGLA, ARA, ETA 및 EPA로서 확인하고; 각각의 조성을 총 지방산(%)으로서 나타낸다.

[표 56]

야로위아 리폴리티카 CPT1을
과발현하도록 유전자 조작된
야로위아 균주 Y2067U 중의 지질 조성

균주	총 지방산							
	18:0	18:1	18:2	GLA	DGLA	ARA	ETA	EPA
Y2067U + pZUF-MOD-1 #1	1.3	6.9	12.0	23.1	5.7	1.1	3.8	13.2
Y2067U + pZUF-MOD-1 #2	1.4	6.8	12.1	22.0	5.8	1.1	3.8	13.5
Y2067U + pYCPT1-ZP2I7 #1	0.6	8.0	8.2	27.4	7.1	1.6	4.1	15.7
Y2067U + pYCPT1-ZP2I7 #2	0.6	8.1	8.2	27.2	7.0	1.6	4.0	15.7
Y2067U + pYCPT1-ZP2I7 #3	1.0	7.9	8.0	24.7	6.1	1.6	3.2	15.5
Y2067U + pYCPT1-ZP2I7 #4	0.6	7.1	8.6	25.5	6.9	1.8	4.0	16.0

상기 나타낸 바와 같이, 게놈 통합에 의한 강력한 FBAIN 프로모터 조절 하의 야로위아 리폴리티카 CPT1의 발현은 EPA (%)를 대조군" 균주에서의 13.4%에서 15.7 내지 16%로 증가시켰다. 추가로, GLA, DGLA 및 ARA 수준을 또한 증가시켰다.

실시예 25

사카로마이세스 세레비지애 ISC1은 PUFA(%)를 증가시킨다.

본 실시예는 사카로마이세스 세레비지애 ISC1 유전자 (서열 152)를 동시 발현하도록 형질전환된 야로위아 리폴리티카 균주 M4 (실시예 6)에서의 증가된 EPA 생합성 및 축적을 기재한다. Δ6 데새투라제/Δ6 엘론가제 경로 또는 Δ9 엘론가제/Δ8 데새투라제 경로를 통해 DHA를 생산하도록 유전자 조작된 야로위아 리폴리티카 숙주 균주는 사카로마이세스 세레비지애 ISC1이 유사하게 동시 발현되는 경우에 (예컨대, 균주 Y3000에서) 증가된 DHA 생합성 및 축적을 입증할 수 있었음을 고려한다.

사카로마이세스 세레비지애 ISC1 ORF를 하기와 같이 플라스미드 pZP217 + Ura 내로 클로닝하였다. 첫째, ORF는 사카로마이세스 세레비지애 균주 S288C (프로메가, 미국 위스콘신주 매디슨 소재)로부터의 게놈 DNA 및 프라이머 쌍 Isc1F 및 Isc1R (서열 440 및 441)을 사용하여 PCR-증폭하였다. 프라이머 Isc1F는 증폭된 ORF에서 ISC1의 야생형 5' 서열을 'ATGTACAA'에서 'ATGGACAA'로 변형시켰는데, 이는 NcoI 부위를 혼입하여 프레임에 ISC1을 유지하는데 필요하기 때문이다. 증폭을 하기와 같이 수행하였다: 94℃에서 120초 동안 초기 변성, 이어서 94℃에서 30초 동안 변성, 50℃에서 30초 동안 어닐링 및 68℃에서 120초 동안 신장의 35주기. 68℃에서 10분 동안의 최종 신장 주기를 수행하고, 이어서 4℃에서 반응을 종결하였다. 1455 bp DNA 단편을 ISC1에 대한 PCR 반응으로부터 수득하고, PCR 생성물 크기는 1% 아가로스 겔 (30분 동안 120 V) 및 인비트로젠 (미국 캘리포니아주 칼스배드 소재)로부터의 1 kB DNA 표준 래더(ladder)를 사용하는 전기영동에 의해 확인하였다.

DNA는 제조자 지시서에 따라 자이모 리써치 코포레이션 (미국 캘리포니아주 오렌지 소재)로부터의 DNA 클린 & 컨센트레이터-5 (DNA Clean & Concentrator-5) 키트로 정제하고, 이어서 NcoI/NotI로 소화시켰다. 이어서, ISC1 단편을 NcoI 및 NotI로 소화시킨 pZP217 + Ura (서열 205; 도 22C) 내로 개별적으로 클로닝하였다. 정확한 형질전환체를 겔 전기영동으로 확인하고, 생성된 플라스미드를 "pTEF::ISC1" (서열 207)이라고 칭하였다. 따라서, 이 플라스미드는 하기를 포함하는 DNA 카세트를 함유하였다: 3'-POX2, URA3, TEF::ISC1::Pex20 및 POX2 프로모터 영역.

"대조군" 벡터를 하기와 같이 제조하였다. 첫째, 사카로마이세스 세레비지애 pcl1 ORF (유사분열 세포 주기 내로의 진입 및 형태발생의 조절과 관련된 단백질을 코딩함)를 사카로마이세스 세레비지애 균주 S288C로부터의 게놈 DNA 및 프라이머 쌍 Pcl1F 및 Pcl1R (서열 442 및 443)로 PCR 증폭시켰다. 증폭을 상기와 같이 수행하였다. 861 bp DNA 단편을 pcl1에 대한 PCR 반응으로부터 수득하였다 (전기영동에 의해 확인함, 상기 참조). DNA를 DNA 클린 & 컨센트레이터-5 키트로 정제하고, 이어서 NcoI/NotI로 소화시켰다. 이어서, 단편을 유사하게 소화시킨 pZP217 + Ura 내로 클로닝하였다. 정확한 형질전환체를 겔 전기영동에 의해 확인하고, 생성된 플라스미드를 "pTEF::pcl1"이라고 칭하였다. 이어서, 플라스미드 pTEF::pcl1을 HincII로 소화시켜 pcl1 ORF를 제거하였다. 잔존 플라스미드를 재라이게이션하여 3'-POX2, URA3, TEF::Pex20 및 POX2 프로모터 영역을 포함하는 선형 DNA 카세트를 AscI/SphI의 소화 시에 생성하였다.

감응성 야로위아 리폴리티카 균주 M4 세포 (실시예 6으로부터)를 AscI/SphI-소화 pTEF::ISC1 및 "대조군"으로 각각 형질전환시켰다 (여기서, 각 플라스미드 5 ㎍을 소화시킴). 형질전환을 동결 EZ 효모 형질전환 II 키트 (자이모 리써치)로 달성하고, 형질전환체를 아미노산 없는 YNB (6.7 g/L; 벡톤, 딕킨손 앤드 코포레이션(Becton, Dickinson and Co.), 미국 메릴랜드주 스파크스 소재 [카탈로그 #291940]), 글루코스 (20 g/L) 및 아가 (20 g/L)를 함유하는 플레이트 상에서 선별하였다. 수백개 형질전환체 콜로니를 수득하였다. 각 DNA 카세트를 야로위아 리폴리티카 P0X2 유전자좌 내로 통합한 것을 ISC1에 대한 5개 독립성 형질전환체로부터의 게놈 DNA를 사용하는 PCR에 의해 확인하였다.

형질전환체를 2% 글루코스를 함유한 아미노산 없는 YNB에서 2일 동안 성장시켰다. 세포를 원심분리로 수확하고, 100 g/L 덱스트로즈, 2 g/L $MgSO_4$ 및 50 mM 포스페이트 완충제 (pH 6.5)를 함유한 배지에 추가 5일 성장 동안 재현탁하였다. 각 배양물 0.75 mL로부터의 세포를 원심분리로 수확하고, 그의 지방산 조성을 분석하였다. "대조군" 벡터를 포함한 3개 형질전환체의 지방산 프로파일 및 pTEF::ISC1을 포함한 5개 형질전환체를 (상기 '통상의 방법' 항목에 기재한 바와 같이) GC 분석에 기초하여 하기에 나타낸다. 지방산을 16:0, 16:1, 18:0, 18:1 (올레산), 18:2 (LA), GLA, DGLA, ARA, ETA 및 EPA로서 확인하고; 각각의 조성을 총 지방산(%)으로서 나타낸다.

[표 57]

사카로마이세스 세레비지애 ISC1을
과발현하도록 유전자 조작된
야로위아 균주 M4 중의 지질 조성

균주	총 지방산									
	16:0	16:1	18:0	18:1	18:2	GLA	DGLA	ARA	ETA	EPA
M4 + "대조군"	14.7	7.2	2.1	13.5	8.7	21.8	8.9	0.9	4.1	9.3
M4 + pTEF::ISC1	13.5	8.5	1.7	15.6	8.1	21.3	7.5	0.7	3.9	10.7

사카로마이세스 세레비지애 ISC1 유전자의 발현은 EPA(%)를 "대조군" 균주에서의 9.3%에서 10.7% ("M4 + pTEF::ISC1")로 개선시켰으며, 14.5% 증가를 나타냈다.

실시예 26

야로위아 리폴리티카 아실트랜스퍼라제 넉아웃의 생성

본 실시예는 PDAT, DGAT2, DGAT1, PDAT 및 DGAT2, PDAT 및 DGAT1, DGAT1 및 DGAT2, 또는 PDAT, DGAT1 및 DGAT2 유전자를 파괴한 야로위아 리폴리티카의 단일, 이중 및 삼중 넉아웃 균주의 생성을 기재한다. 각 넉아웃 균주 중의 유전자(들)의 파괴를 확인하고, 지방산 함량 및 조성에 대한 각 파괴의 분석을 실시예 27에서의 총 지질의 GC 분석에 의해 측정하였다.

야로위아 리폴리티카 DGAT2 유전자의 표적화된 파괴

야로위아 리폴리티카 ATCC #90812 중의 DGAT2 유전자의 표적화된 파괴는 플라스미드 pY21DGAT2라고 칭한 표적화 카세트를 사용하는 내인성 DGAT2 유전자의 상동성 재조합-매개 대체에 의해 달성되었다. pY21DGAT2는 플라스미드 pY20 (도 22D, 서열 208)으로부터 유래되었다. 구체적으로, pY21DGAT2는 570 bp HindIII/EcoRI 단편을 유사하게 선형화된 pY20 내로 삽입하여 생성하였다. 570 bp DNA 단편은 (5'에서 3'로의 배향에서) (서열 130 중의 코딩 서열 (ORF)의) 위치 + 1090 내지 + 1464로부터의 3' 상동성 서열, BglII 제한 부위 및 (서열 130 중의 코딩 서열 (ORF)의) 위치 + 906 내지 + 1089로부터의 5' 상동성 서열을 함유하였다. 상기 단편은 2쌍의 PCR 프라이머 P95 및 P96 (서열 444 및 445), 및 P97 및 P98 (서열 446 및 447)을 각각 사용하는 PCR 증폭에 의해 제조하였다.

pY21DGAT2를 BglII 제한 소화에 의해 선형화하고, 통상의 방법에 따라 중간-로그 단계 야로위아 리폴리티카 ATCC # 90812 세포로 형질전환시켰다. 세포를 YPD 하이그로마이신 선별 플레이트 상에 플레이팅하고, 30℃에서 2 내지 3일 동안 유지시켰다.

14개 야로위아 리폴리티카 ATCC #90812 하이그로마이신-내성 콜로니를 단리하고, 표적화된 파괴에 대해 PCR로 스크리닝하였다. 1세트의 PCR 프라이머 (P115 및 P116 [서열 448 및 449])를 디자인하여 특이적 연결 단편을 증폭시킨 후 상동성 재조합하였다. 또다른 쌍의 PCR 프라이머 (P115 및 P112 [서열 450])를 디자인하여 천연 유전자를 검출하였다.

ATCC #90812 균주의 14개 하이그로마이신-내성 콜로니 중 2개는 연결 단편에 대해 양성이었고, 천연 단편에 대해 음성이었다. 따라서, 표적화된 통합을 이들 2개 균주에서 확인하였으며, 이들 중 하나를 "S-D2"라고 칭하였다.

야로위아 리폴리티카 PDAT 유전자의 표적화된 파괴

야로위아 리폴리티카 ATCC #90812 중의 PDAT 유전자의 표적화된 파괴는 pLV13 (도 22E, 서열 209)이라고 칭한 표적화 카세트를 사용하는 내인성 PDAT 유전자의 상동성 재조합-매개 대체에 의해 달성되었다. pLV13은 플라스미드 pY20 (도 22D, 서열 208)으로부터 유래되었다. 구체적으로, pY20의 하이그로마이신 내성 유전자를 야로위아 Ura3 유전자로 대체하여 플라스미드 pLV5를 생성하였다. 이어서, pLV13은 992 bp BamHI/EcoRI 단편을 유사하게 선형화된 pLV5 내에 삽입함으로써 생성하였다. 992 bp DNA 단편은 (5'에서 3'로의 배향에서) (서열 117 중의 코딩 서열 (ORF)의) 위치 + 877 내지 + 1371로부터의 3' 상동성 서열, BglII 제한 부위 및 (서열 117 중의 코딩 서열 (ORF)의) 위치 + 390 내지 + 876으로부터의 5' 상동성 서열을 함유하였다. 상기 단편은 PCR 프라이머 P39 및 P41 (서열 451 및 452), 및 P40 및 P42 (서열 453 및 454)를 각각 사용하는 PCR 증폭에 의해 제조하였다.

pLV13을 BglII 제한 소화에 의해 선형화하고, 통상의 방법에 따라 중간-로그 단계 야로위아 리폴리티카 ATCC #90812 세포로 형질전환시켰다. 세포를 Bio101 DOB/CSM-Ura 선별 플레이트 상에 플레이팅하고, 30℃에서 2 내지 3일 동안 유지시켰다.

10개 야로위아 리폴리티카 ATCC #90812 콜로니를 단리하고, 표적화된 파괴에 대해 PCR로 스크리닝하였다. 1세트의 PCR 프라이머 (P51 및 P52 [서열 455 및 456])를 디자인하여 표적화 카세트를 증폭시켰다. 또다른 세트의 PCR 프라이머 (P37 및 P38 [서열 457 및 458])를 디자인하여 천연 유전자를 검출하였다. 10개 균주 중 10개는 연결 단편에 대해 양성이었고, 10개 균주 중 3개는 천연 단편에 대해 음성이어서 이들 3개 균주에서 성공적인 표적화 통합을 확인하였다. 이들 중 하나를 "S-D2"라고 칭하였다.

<u>야로위아 리폴리티카 DGAT1 유전자의 표적화된 파괴</u>

전장 Y1 DGAT1 ORF는 추정 PCR 프라이머 P201 및 P203 (각각 서열 459 및 460) 및 주형으로서의 야로위아 리폴리티카 ATCC #76982 게놈 DNA를 사용하는 PCR에 의해 클로닝하였다. Y1 DGAT1을 코딩하는 뉴클레오티드 서열이 공지되어 있지 않기 때문에 추정 프라이머가 요구되었다.

PCR은 증폭이 하기와 같이 수행되는 RoboCycler Gradient 40 PCR 머신에서 수행하였다: 95℃에서 1분 동안 초기 변성, 이어서 95℃에서 30초 동안 변성, 55℃에서 1분 동안 어닐링 및 72℃에서 1분 동안 신장의 30주기. 72℃에서 10분 동안의 최종 신장 주기를 수행하고, 이어서 4℃에서 반응을 종결하였다. 예측된 PCR 생성물 (약 1.6 kB)을 아가로스 겔 전기영동에 의해 검출하고, 단리하고, 정제하고, TOPO® 클로닝 벡터 (인비트로젠) 내로 클로닝하고, 부분적으로 서열결정하여 그의 정체를 확인하였다.

야로위아 리폴리티카 ATCC #90812 중의 추정적 DGAT1 유전자의 표적화된 파괴는 표적화 카세트를 사용하는 내인성 DGAT1 유전자의 상동성 재조합-매개 대체에 의해 수행하였다 (DGAT2에 대한 상기 방법을 이용함). 구체적으로, 1.6 kB 단리된 Y1 DGAT1 ORF (서열 122)를 PCR 주형 분자로서 사용하여 5' 상동성 Y1 DGAT1 서열 (프라이머 P214 및 P215 (서열 461 및 462)로 증폭시킴), 야로위아 루이신2 (Leu2; 진뱅크 관리 번호 AAA35244) 유전자, 및 3' 상동성 Y1 DGAT1 서열 (프라이머 P216 및 P217 (서열 463 및 464)로 증폭시킴)로 이루어진 Y1 DGAT1 표적화 카세트를 구축하였다. 표적화 카세트의 각 개개의 부분을 Pfu 울트라 폴리머라제 (스트라타진, 카탈로그 #600630) 및 상기한 열순환기 조건으로 증폭시킨 후에, 각 단편을 정제하였다. 3개의 정확한 크기의 정제된 단편은 PCR 프라이머 P214 및 P219 (서열 465)를 사용하는 제2 PCR 반응을 위해 주형 분자로서 함께 혼합하여 Y1 DGAT1 파괴 카세트를 수득하였다.

표적화 카세트를 겔 정제하고, 이를 사용하여 중간-로그 단계 야생형 야로위아 리폴리티카 (ATCC #90812)를 형질전환시켰다. 형질전환을 상기 '통상의 방법' 항목에 기재한 바와 같이 수행하였다. 형질전환체를 Bio101 DOB/CSM-Leu 선별 플레이트 상에 플레이팅하고, 30℃에서 2 내지 3일 동안 유지시켰다. 몇몇 루이신 원시영양체를 PCR에 의해 스크리닝하여 표적화된 DGAT1 파괴를 확인하였다. 구체적으로, 1세트의 PCR 프라이머 (P226 및 P227 [서열 466 및 467])를 디자인하여 파괴 카세트와 천연 표적 유전자 간의 연결부를 증폭시켰다. 또다른 세트의 PCR 프라이머 (P214 및 P217 [서열 461 및 464])를 디지인하여 천연 유전자를 검출하였다.

모든 루이신 원시영양체 콜로니는 연결 단편에 대해 양성이고, 천연 단편에 대해 음성이었다. 따라서, 표적화된 통합을 이들 균주에서 확인하였으며, 이들 중 하나를 "S-D1"이라고 칭하였다.

<u>PDAT 및/또는 DGAT2 및/또는 DGAT1 유전자 중에 파괴를 함유하는 야로위아 리폴리티카 이중 및 삼중 넉아웃 균주의 생성</u>

야로위아 리폴리티카 ATCC #90812 하이그로마이신-내성 "S-D2" 돌연변이체 (DGAT2 파괴 함유)를 플라스미드 pLV13 (PDAT 파괴 함유)로 형질전환시키고, 형질전환체를 단일 PDAT 파괴에 대해 기재한 바와 같이 PCR에 의해 스크리닝하였다. 12개 형질전환체 중 2개는 DGAT2 및 PDAT 유전자 모두에서 파괴되었음을 확인하였다. 이들 균주 중 하나를 "S-D2-P"라고 칭하였다.

유사하게, DGAT1 및 PDAT ("S-D1-P"), DGAT2 및 DGAT1 ("S-D2-D1") 중의 이중 넉아웃, 및 DGAT2, DGAT1 및 PDAT ("S-D2-D1-P") 중의 삼중 넉아웃을 갖는 균주를 생성시켰다.

실시예 27

야로위아 리폴리티카 아실트랜스퍼라제 넉아웃은 지질 함량을 저하시키고, PUFA(%)를 증가시킨다.

본 실시예는 지방산 함량 및 조성의 변화에 의해 측정하는 바와 같이, 야생형 야로위아 리폴리티카 및 EPA를 생산하도록 미리 유전자 조작된 야로위아 리폴리티카 균주에서의 단일 및/또는 이중 및/또는 삼중 아실트랜스퍼라제 넉아웃의 영향을 분석한다. Δ6 데새투라제/Δ6 엘론가제 경로 또는 Δ9 엘론가제/Δ8 데새투라제 경로를 통해 DHA를 생산하도록 유전자 조작된 야로위아 리폴리티카 숙주 균주가 숙주의 천연 아실트랜스퍼라제에 대한 유사한 조작물이 생성되는 경우에 (예컨대, 균주 Y3000 내에서) 증가된 DHA 생합성 및 축적을 입증할 수 있었음을 고려한다.

TAG 함량은 아실트랜스퍼라제 파괴를 갖는 야로위아 리폴리티카 ATCC #90812에서 저하된다.

첫째, TAG 함량은 (1) PDAT, DGAT2 및 DGAT1 중의 단일 파괴; (2) PDAT 및 DGAT2, DGAT1 및 PDAT, 및 DGAT1 및 DGAT2 중의 이중 파괴; 및 (3) PDAT, DGAT2 및 DGAT1 중의 삼중 파괴를 함유한 야생형 및 돌연변이체 야로위아 리폴리티카 ATCC #90812에서 비교하였다.

구체적으로, 야생형 및 돌연변이체 야로위아 리폴리티카 ATCC #90812 (즉, 균주 S-D1, S-D2, S-P, S-D1-D2, S-D1-P, S-D2-P, 및 S-D1-D2-P)를 함유한 플레이트로부터 1백금이의 세포를 YPD 배지 3 mL에 각각 개별적으로 접종시키고, 밤새 진탕기 (300 rpm) 상에서 30℃에서 성장시켰다. 세포를 수확하고, 0.9% NaCl로 1회 세척하고, HGM 50 mL에 재현탁하였다. 이어서, 세포를 48시간 동안 진탕기 상에 성장시켰다. 세포를 물로 세척하고, 세포 펠렛을 동결건조시켰다. 세포 건중량 20 mg을 GC에 의한 총 지방산 분석에 사용하고, 오일 분획을 TLC (하기 참조) 및 GC 분석하였다.

TLC에 사용된 방법을 다음 5단계로 하기에 기재한다: (1) 15:0 지방산 (10 ㎕의 10 mg/mL) 내부 표준을 건조 세포 덩어리 2 내지 3 mg에 첨가한 후에, 총 지질을 메탄올/클로로포름 방법으로 추출하였다. (2) 5 x 20 cm 실리카 겔 60 플레이트의 하부로부터 대략 1 인치 그은 연한 연필 줄을 따라 추출된 지질 (50 ㎕)을 25 내지 50 ㎕ 마이크로피펫으로 블럿팅하였다. (3) 이어서 TLC 플레이트를 N_2 하에 건조시키고, 약 ~100 mL의 헥산:에틸 에테르:아세트산 (80:20:1) 용매를 함유한 탱크 내에 삽입하였다. (4) 밴드 분리 후에, 요오드 증기를 플레이트의 한 면 상에 취입하여 밴드를 확인하였다. 이는 추가 분석을 위해 면도날을 사용하여 플레이트의 다른 면 상의 샘플을 긁어 낼 수 있게 한다. (5) 긁어 낸 샘플의 염기성 에스테르교환반응 및 GC 분석을 상기 '통상의 방법' 항목에 기재한 바와 같이 수행하였다.

GC 결과를 하기 표 58에 나타낸다. 배양물을 "S" 균주 (야생형), "S-P" (PDAT 넉아웃), "S-D1" (DGAT1 넉아웃), "S-D2" (DGAT2 넉아웃), "S-D1-D2" (DGAT1 및 DGAT2 넉아웃), "S-P-D1" (PDAT 및 DGAT1 넉아웃), "S-P-D2" (PDAT 및 DGAT2 넉아웃) 및 "S-P-D1-D2" (PDAT, DGAT1 및 DGAT2 넉아웃)로서 기재한다. 약어는 하기와 같이 사용한다: "WT" = 야생형; "FA" = 지방산; "dcw" = 세포 건중량; 및 "FA% dcw, % WT" = 야생형 (%)에 상대적인 FA% dcw (여기서 "S" 균주는 야생형임).

[표 58]

PDAT, DGAT2 및 DGAT1에서의 단일, 이중, 또는 삼중 파괴를 갖는 야로위아 ATCC #90812 균주 중의 지질 함량

			총 지방산			TAG 분획		
균주	잔류 DAG AT	dcw, mg	FA, μg	FA % dcw	FA % dcw, % WT	FA, μg	FA % dcw	FA % dcw, % WT
S	D1, D2, P	32.0	797	15.9	100	697	13.9	100
S-D1	D2, P	78.8	723	13.6	86	617	11.6	83
S-D2	D1, P	37.5	329	6.4	40	227	4.4	32
S-P	D1, D2	28.8	318	6.0	38	212	4.0	29
S-D1-D2	P	64.6	219	4.1	26	113	2.1	15
S-D1-P	D2	76.2	778	13.4	84	662	11.4	82
S-D2-P	D1	31.2	228	4.3	27	122	2.3	17
S-D1-D2-P	없음	52.2	139	2.4	15	25	0.4	3

표 58의 결과는 오일 생합성에 대한 3개의 DAG AT의 상대적 기여도를 나타낸다. DGAT2가 최고로 기여하지만, PDAT 및 DGAT1은 DGAT2와 동일하거나 그 미만으로 기여한다. 삼중 넉아웃 균주 중의 잔류 오일 함량 약 3%는 야로위아 리폴리티카 아실-CoA:스테롤-아실트랜스퍼라제 효소 (ARE2에 의해 코딩됨) (서열 119 및 120)의 기여도일 수 있다.

파괴된 DGAT2 유전자를 갖는 야로위아 리폴리티카 균주 EU에서 TAG 함량은 저하되고, EPA(%)는 증가된다.

야생형 야로위아 리폴리티카 ATCC #90812 (상기 문헌)에서의 다양한 아실트랜스퍼라제 넉아웃의 영향을 시험한 후에, 이어서 TAG 함량 및 지방산 조성을 EU 균주의 DGAT2 넉아웃 균주 (즉, 10% EPA를 생산하도록 유전자 조작됨, 실시예 13 참조)에서 연구하였다.

구체적으로, 균주 EU 중의 DGAT2 유전자를 실시예 26에서 S 균주 (ATCC #90812)에 대해 기재한 바와 같이 파괴하였다. DGAT2-파괴 균주를 EU-D2라고 칭하였다. EU 및 EU-D2 균주를 2개의 상이한 조건에 따른 성장 후에 수확하고, 분석하였다. 하기 표에 "3 mL"로서 나타낸 조건에서는, 세포를 1일 동안 MM 배지 3 mL에서 성장시키고, 세척하고, 이어서 3일 동안 HGM 3 mL에서 성장시켰다. 별법으로, 하기 표에 "51 mL"로서 나타낸 조건에서는, 세포를 1일 동안 MM 배지 51 mL에서 성장시키고, 세척하고, 이어서 3일 동안 HGM 51 mL에서 성장시켰다. 포스파티딜콜린 (PC), 포스파티딜에탄올아민 (PE), 및 트리아실글리세롤 (TAG 또는 오일)의 지방산 조성을 배양물 51 mL의 추출물에서 TLC 분리 ("분획") 후에 측정하였다.

GC 결과를 하기 표 59에 나타낸다. 배양물을 "EU" 균주 (야생형) 및 "EU-D2" 균주 (DGAT2 넉아웃)라고 기재한다. 지방산을 16:0, 16:1, 18:0, 18:1 (올레산), 18:2 (LA), GLA, DGLA, ARA, ETA 및 EPA로서 확인하고; 각각의 조성을 총 지방산(%)으로서 나타낸다.

[표 59]

DGAT2에서의 파괴를 갖는
야로위아 균주 EU 중의 지질 함량 및 조성

균주 및 성장	분획	TFAs % dcw	% 16:0	% 16:1	% 18:0	% 18:1	% 18:2	% GLA	% DGLA	% ARA	% ETA	% EPA
EU, 3 mL	총		19	10	2	16	12	19	6	0	3	10
EU-D2, 3 mL	총		17	10	1	6	7	24	5	0	6	19
EU, 51 mL	총	37	18	11	3	19	31	5	1		1	4
	PC	2	12	9	1	8	43	7	3		5	4
	PE	1	24	14	0	14	37	5	0		0	1
	TAG	34	18	12	3	21	29	5	1		1	4
EU-D2, 51 mL	총	18	18	8	1	5	7	25	5		5	20
	PC	1	18	6	1	2	4	26	5		11	22
	PE	1	25	7	0	2	5	14	2		3	8
	TAG	15	16	9	1	6	5	26	6		5	21

결과는 DGAT2 넉아웃이 (총 지방산의) EPA(%)를 배가시키고, 지질 함량 (dcw(%))을 반감시켰음을 나타낸다. 추가로, 지질 함량에서 관찰된 거의 모든 변화는 TAG 분획에서의 변화 때문이다. 균주 EU의 배양물 51 mL에서의 예측보다 낮은 EPA(%)는 불안정성 때문인 것으로 여겨진다.

파괴된 아실트랜스퍼라제 유전자를 갖는 야로위아 리폴리티카 균주 MU에서 TAG 함량은 저하되고, EPA(%)는 증가된다.

최종적으로, 균주 EU-D2에서 단일 DGAT2 넉아웃으로부터 얻어진 증가된 EPA(%) 및 저하된 지질 함량을 기초로 하여, TAG 함량 및 지방산 조성을 이어서 균주 MU의 다양한 아실트랜스퍼라제 넉아웃 균주 (즉, 14% EPA를 생산하도록 유전자 조작됨, 실시예 15 참조)에서 연구하였다. 구체적으로, PDAT, DGAT2 및 DGAT1 중의 단일 파괴 및 PDAT 및 DGAT2 중의 이중 파괴를 균주 MU에서 생성하였다. 지질 함량 및 조성은 4개의 상이한 성장 조건에서의 성장 후에 각각의 이들 균주에서 비교하였다.

보다 구체적으로, PDAT, DGAT2, DGAT1 중의 단일 파괴를 실시예 26에 기재한 방법 (URA3 유전자에 기반한 DGAT1 파괴에 대한 선별은 제외함)으로 균주 MU에서 생성하였다. 이는 "MU-D1" (DGAT1에서 파괴됨), "MU-D2" (DGAT2에서 파괴됨), 및 "MU-P" (PDAT에서 파괴됨)라고 동정된 단일 넉아웃 균주를 생성하였다. 개개의 넉아웃 균주를 PCR로 확인하였다. 추가로, MU-D2 균주를 동일한 방법에 의해 PDAT 유전자에 대해 파괴하고, 이 파괴를 PCR에 의해 확인하였다. 생성된 이중 넉아웃 균주를 "MU-D2-P"라고 칭하였다.

MU-D1, MU-D2, MU-P, 및 M-D2-P 넉아웃 균주를 분석하여 하기에 기재하는 바와 같이 지질 함량 및 조성에 대한 각각의 넉아웃 효과를 측정하였다. 추가로, 유질성을 촉진하는 성장 조건을 총 지질 함량에 대한 그의 효과를 측정하기 위해 또한 조사하였다. 따라서, "실험 A", "실험 B", "실험 C" 및 "실험 E"로 나타내는 총 4개의 상이한 실험을 수행하였다. 구체적으로, 상기 각 균주를 함유한 플레이트로부터 3백금이의 세포를 MMU 배지 [실험 B 및 C의 경우 3 mL; 및 실험 A 및 E의 경우 50 mL]에 접종하고, 진탕기 중 30℃에서 24시간 (실험 A, B 및 C의 경우) 또는 48시간 (실험 E의 경우) 동안 성장시켰다. 세포를 수확하고, HGM으로 1회 세척하고, HGM 배지 (실험 A 및 E의 경우 50 mL; 및 실험 B의 경우 3 mL) 또는 HGM 배지 (우라실 ("HGMU") 함유) (실험 C의 경우 3 mL) 중에 재현탁하고, 4일 동안 상기와 같이 배양하였다. 하나의 분취액 (1 mL)을 상기 '통상의 방법' 항목에 기재한 바와 같이 GC에 의한 지질 분석에 사용한 반면에, 제2 분취액을 600 nm에서의 배양물 OD를 측정하는데 사용하였다. 실험 A 및 E의 잔존 배양물을 수확하고, 물로 1회 세척하고, 세포 건중량 (dew) 측정을 위해 동결건조시켰다. 반대로, 실험 B 및 C의 dcw를 그의 관계를 나타내는 방정식으로 그의 OD_{600}으로부터 측정하였다. 또한, 실험 A, B, C 및 E 중의 각각의 상이한 균주의 지방산 조성을 측정하였다.

결과를 하기 표 60에 나타낸다. 배양물을 "MU" 균주 (모 EPA 생산 균주), "MU-P" (PDAT 넉아웃), "MU-D1" (DGAT1 넉아웃), "MU-D2" (DGAT2 넉아웃) 및 "MU-D2-P" (DGAT2 및 PDAT 넉아웃)으로 기재한다. 약어는 하기와 같이 사용한다: "WT" = 야생형 (즉, MU); "OD" = 광학 밀도; "dcw" = 세포 건중량; "TFA" = 총 지방산; 및 "TFA% dcw, % WT" = 야생형 ("MU") 균주에 상대적인 TFA% dcw. 지방산을 16:0, 16:1, 18:0, 18:1 (올레산), 18:2 (LA), GLA, DGLA, ARA, ETA 및 EPA로서 확인하고; 각각의 조성을 총 지방산(%)으로서 나타낸다.

[표 60]

다양한 아실트랜스퍼라제 파괴를 갖는 야로위아 MU 균주 중의 지질 함량 및 조성

실험	균주	잔류 DAG AT	제1상 성장 조건	제2상 성장 조건	OD	dcw (mg)	TFA (μg)	TFA % dcw	TFA % dcw, % WT	% 16:0	% 16:1	% 18:0	% 18:1	% 18:2	% GLA	% DGLA	% ARA	% ETA	% EPA
A	MU	D1, D2, P	1 일, 50 mL MMU	4 일, 50 mL HGM	4.0	91	374	20.1	100	17	10	2	18	10	22	7	1	3	9.7
A	MU-D2	D1, P			3.1	75	160	10.4	52	16	12	0	8	9	23	7	0	8	17.4
A	MU-D1	D2, P			4.3	104	217	10.2	51	15	10	2	11	10	22	7	0	7	17.4
A	MU-P	D1, D2			4.4	100	238	11.7	58	16	9	2	11	7	24	7	1	6	17.5
B	MU	D1, D2, P	1 일, 3 mL MMU	4 일, 3 mL HGM	5.9	118	581	24.1	100	17	9	3	18	10	22	8	1	3	9.1
B	MU-D2	D1, P			4.6	102	248	11.9	50	16	10	0	7	10	24	7	1	7	17.8
B	MU-D1	D2, P			6.1	120	369	15.0	62	18	9	3	14	11	20	7	1	5	12.0
B	MU-P	D1, D2			6.4	124	443	17.5	72	15	8	3	16	10	25	6	1	4	11.9
C	MU	D1, D2, P	1 일, 3 mL MMU	4 일, 3 mL HGMU	6.8	129	522	19.9	100	16	10	2	13	11	21	10	1	4	12.6
C	MU-D2	D1, P			5.6	115	239	10.2	51	17	9	1	6	11	21	8	1	7	18.9
C	MU-D1	D2, P			6.9	129	395	15.0	75	15	9	2	12	12	20	10	1	5	13.5
C	MU-P	D1, D2			7.1	131	448	16.8	84	17	8	3	14	11	20	10	1	4	11.3
E	MU	D1, D2, P	2 일, 50 mL MM	4 일, 50 mL HGM	4.6	89	314	17.3	100	16	12	2	18	9	22	7	1	4	11.2
E	MU-D2	D1, P			2.8	62	109	8.5	49	14	12	1	6	8	25	6	0	7	20.0
E	MU-P	D2, P			5.0	99	232	11.5	66	16	10	2	14	7	24	7	1	5	15.8
E	MU-D2-P	D1			4.2	98	98	4.9	28	18	10	0	7	12	20	5	0	6	22.5

상기 데이타는 형질전환된 세포 내의 지질 함량이 성장 조건에 따라 변하였음을 보여주었다. 추가로, 지질 함량에 대한 각각의 아실트랜스퍼라제의 기여도도 변하였다. 구체적으로, 실험 B, C 및 E에서 DGAT2는 PDAT 또는 DGAT1보다 오일 생합성에 더 기여하였다. 반대로, 실험 A에서 입증된 바와 같이, DGAT2, DGAT1 및 PDAT에서의 단일 넉아웃은 지질 함량에서 대략 동등한 손실을 발생하였다 (즉, 각각 48%, 49% 및 42% 손실 ["TFA% dcw, % WT" 참조]).

지방산 조성과 관련하여, 상기 데이타는 각각의 개개 DAG AT 유전자의 넉아웃이 오일 함량을 낮추고, EPA(%)를 증가시켰음을 보여준다. 예를 들어, DGAT2 넉아웃은 지질 함량을 대략 반감시키고, 총 지방산에서 EPA(%)를 대략 배가시켰다 (균주 EU-D2에서 관찰된 결과와 유사함, 상기 참조). DAGAT2 및 PDAT 모두의 넉아웃은 최소 오일 및 최고 EPA(%)를 생산하였다.

본원에 보고한 결과를 기초로 하여, 천연 DGAT2 및/또는 DGAT1 및/또는 PDAT의 파괴가 DHA를 비롯한 고농도의 PUFA를 생산하도록 유전자 조작된 야로위아 리폴리티카의 균주 (예컨대, 균주 Y3000 내에서)에서 실질적으로 PUFA (%)를 증가시키는 유용한 수단임을 고려한다. 사실, 야로위아 리폴리티카 균주 Y2214 중의 천연 DGAT2 유전자의 파괴

(Δ9 엘론가제/Δ8 데새투라제 경로를 통해 14% ARA를 생산함; 야생형 야로위아 리폴리티카 ATCC #20362에 대한 이 균주의 최종 유전자형은 다음과 같음: Aco2-, Lys5-, 2X GPAT::IgD9e::PEX20, 2X TEF::IgD9e::LIP1, FBAINm::IgD9e::0CT, 2X FBAIN::D8SF::PEX16, GPD::D8SF::PEX16, GPAT::MAΔ5::PEX20, FBAIN::MAΔ5::PEX20, YAT1::I.D5S::LIP1, GPM/FBAIN::I.D5S::OCT, FBAIN::F.D12S::PEX20 및 GPM/FBAIN::rELO2S::OCT)는 ARA(%)를 1.7배 증가시켰다 (데이타는 나타내지 않음).

도면의 간단한 설명

<도면의 간단한 설명 및 서열의 기재>

도 1은 ω-3/ω-6 지방산 생합성 경로를 도시한다.

도 2는 유질 효모에서의 지질 축적에 대한 생화학적 메카니즘의 개략도이다.

도 3A는 메갈린(Megalign) DNASTAR 소프트웨어를 이용하여 생성한, 상이한 섬사상 진균으로부터의 Δ12 데새투라제 및 Δ15 데새투라제 단백질의 계통수를 도시한다. 도 3B는 pY57.YI.AHAS.w497l에 대한 플라스미드 지도를 제공한다.

도 4는 유질 효모에서의 지질 축적에서 다양한 아실트랜스퍼라제의 역할을 기재하는 개략도이다.

도 5는 전체 지질 분획 중 다양한 지방산 (DHA 포함)을 생산하는 본 발명의 일부 야로위아 리폴리티카 균주의 발달에 대한 약도이다.

도 6A는 pY5-30에 대한 플라스미드 지도를 제공한다. 도 6B는 야로위아 리폴리티카 ATCC #76982 균주에서 조직화학적 염색에 의해 측정된 TEF, GPD, GPM, FBA 및 FBAIN의 상대적인 프로모터 활성을 나타낸다. 도 6C는 다양한 배지에서 성장한 야로위아 리폴리티카에서 조직화학적 염색에 의해 측정된 YAT1, TEF, GPAT 및 FBAIN의 상대적인 프로모터 활성을 나타낸다.

도 7A는 야로위아 리폴리티카 ATCC #76982 균주에서 형광측정에 의해 측정된 GPD, GPM, FBA 및 FBAIN의 프로모터 활성을 비교하는 그래프이다. 도 7B는 실시간 PCR 상대적인 정량화의 결과를 그래프로 요약한 것이며, 여기서 야로위아 리폴리티카 ATCC #76982 균주의 GUS mRNA (즉, GPD::GUS, GPDIN::GUS, FBA::GUS 또는 FBAIN::GUS 키메라 유전자를 발현)는 pY5-30을 발현하는 야로위아 리폴리티카 균주의 mRNA (즉, 키메라 TEF::GUS 유전자) 수준에 대해 정량화되었다.

도 8는 (A) pKUNF12T6E, (B) pDMW271, (C) pZP3L37, (D) pZKUT16 및 (E) pKO2UM25E에 대한 플라스미드 지도를 제공한다.

도 9는 (A) pDMW303, (B) pZUF17, (C) pZUF4, (D) pFOE2S 및 (E) pZP2F0EN4에 대한 플라스미드 지도를 제공한다.

도 10은 (A) pKUNT2; (B) pDMW237; (C) pDMW240; (D) 효모 발현 벡터 pY89-5; 및 (E) pKUNFmKF2에 대한 플라스미드 지도를 제공한다.

도 11는 유글레나 그라실리스(*Euglena gracilis*) 세포 추출물의 지질 프로파일의 크로마토그램을 도시한다.

도 12은 다양한 유글레나 그라실리스 Δ8 데새투라제 폴리펩티드 서열의 정렬을 도시한다. 이용된 정렬 방법은 "클러스탈(Clustal) V 정렬 방법"에 상응한다.

도 13은 (A) pDMW277, (B) pZF5T-PPC, (C) pDMW287F 및 (D) pDMW297에 대한 플라스미드 지도를 제공한다.

도 14는 (A) pZP2C16M899, (B) pDMW314, (C) pDM325 및 (D) pZKL5598에 대한 플라스미드 지도를 제공한다.

도 15는 (A) pY72 [또는 "pY72.2loxp.Hyg.Fba.F15"], (B) pY80 [또는 "pY80.loxp.2F15"], (C) pY79 [또는 "pY79.Cre.AHASw497L"] 및 (D) pY86 [또는 "pY86.loxp.Ura3.Hyg.F12"]에 대한 플라스미드 지도를 제공한다.

도 16은 (A) pY94 [또는"pY94.loxp.D9ED8.Ura3"], (B) pY91M [또는"pY91.Dr.D6M (천연)"], (C) pDMW232 및 (D) pY37/F15에 대한 플라스미드 지도를 제공한다.

도 17은 (A) pKO2UF2PE, (B) pZKUGPI5S, (C) pDMW302T16 및 (D) pKO2UM26E에 대한 플라스미드 지도를 제공한다.

도 18은 (A) pZUF-Mod-1, (B) pMDAGAT1-17 및 (C) pMGPAT-17에 대한 플라스미드 지도를 제공한다.

도 19는 각각 모르티에렐라 알피나의 글리세롤-3-포스페이트 o-아실트랜스퍼라제 (GPAT)와 관련된 서열 138, 139, 140, 141, 142, 143, 144, 145, 146, 147 및 148 사이의 관계를 그래프로 나타낸다.

도 20은 각각 모르티에렐라 알피나의 $C_{16/18}$ 지방산 엘론가제 효소 (ELO3)와 관련된 서열 86, 87, 88, 89, 90, 91, 92 및 93 사이의 관계를 그래프로 나타낸다.

도 21은 (A) pZUF6S, (B) pZUF6S-E3WT, (C) pZKUGPYE1-N 및 (D) pZKUGPYE2에 대한 플라스미드 지도를 제공한다.

도 22는 (A) pZKUGPYE1, (B) pZUF6FYE1, (C) pZP217 + Ura, (D) pY20 및 (E) pLV13에 대한 플라스미드 지도를 제공한다.

본 발명은 본 출원의 일부분인 하기 발명의 상세한 설명 및 첨부된 서열 기재로부터 더욱 잘 이해될 것이다.

하기 서열은 37 C.F.R.§1.821-1.825 ("뉴클레오티드 서열 및/또는 아미노산 서열을 포함하는 특허 출원에 대한 기재 요건 - 서열 규칙")에 부합하며, WIPO 표준 ST.25 (1998) 및 EPO 및 PCT의 서열 목록 요건 (규칙 5.2 및 49.5(a-bis), 및 행정 지침서 섹션 208 및 첨부 C)에 일치한다. 뉴클레오티드 및 아미노산 서열 데이타에 사용된 기호 및 포맷은 37 C.F.R.§1.822의 규칙에 부합한다.

서열 1-153 및 210-221는 표 1에 나타낸 바와 같이 프로모터, 유전자 또는 단백질 (또는 그의 단편)을 코딩하는 ORF이다.

[표 1]

유전자 및 단백질 서열에 대한 요약		
기재	**핵산 서열**	**단백질 서열**
모르티에렐라 알피나 Δ6 데새투라제	1 (1374 bp)	2 (457 AA)
야로위아 리폴리티카에서의 발현을 위해 코돈-최적화된, 모르티에렐라 알피나로부터 유래된 합성 Δ6 데새투라제	3 (1374 bp)	2 (457 AA)
모르티에렐라 알피나 Δ6 데새투라제 "B"	4 (1521 bp)	5 (458 AA)
모르티에렐라 알피나 Δ5 데새투라제	6 (1341 bp)	7 (446 AA)
이소크리시스 갈바나 Δ5 데새투라제	8 (1329 bp)	9 (442 AA)
야로위아 리폴리티카에서의 발현을 위해 코돈-최적화된, 이소크리시스 갈바나로부터 유래된 합성 Δ5 데새투라제	10 (1329 bp)	9 (442 AA)
호모 사피엔스(*Homo sapiens*) Δ5 데새투라제	11 (1335 bp)	12 (444 AA)
야로위아 리폴리티카에서의 발현을 위해 코돈-최적화된, 호모 사피엔스로부터 유래된 합성 Δ5 데새투라제	13 (1335 bp)	12 (444 AA)
다니오 레리오(*Danio rerio*) Δ5/Δ6 데새투라제 Drd6/d5(V) (진뱅크 관리 번호 AF309556)	14 (1590 bp)	15 (444 AA)
다니오 레리오 Δ5/Δ6 데새투라제 (진뱅크 관리 번호 BC068224)	16 (1946 bp)	---
다니오 레리오 Δ5/Δ6 데새투라제 돌연변이체 Drd6/d5(M)	17 (1335 bp)	18 (444 AA)
사프롤레그니아 디클리나(*Saprolegnia diclina*) Δ17 데새투라제	19 (1077 bp)	20 (358 AA)
야로위아 리폴리티카에서의 발현을 위해 코돈-최적화된, 사프롤레그니아 디클리나로부터 유래된 합성 Δ17 데새투라제 유전자	21 (1077 bp)	20 (358 AA)

모르티에렐라 알피나 $C_{18/20}$ 엘론가제	22 (957 bp)	23 (318 AA)
야로위아 리폴리티카에서의 발현을 위해 코돈-최적화된, 모르티에렐라 알피나로부터 유래된 합성 $C_{18/20}$ 엘론가제 유전자	24 (957 bp)	23 (318 AA)
트라우스토키트륨 아우레움 $C_{18/20}$ 엘론가제	25 (819 bp)	26 (272 AA)
야로위아 리폴리티카에서의 발현을 위해 코돈-최적화된, 트라우스토키트륨 아우레움으로부터 유래된 합성 $C_{18/20}$ 엘론가제 유전자	27 (819 bp)	26 (272 AA)

[표 1a]

야로위아 리폴리티카 Δ12 데새투라제	28 (1936 bp)	29 (419 AA)
모르티에렐라 이자벨리나(*Mortierella isabellina*) Δ12 데새투라제	30 (1203 bp)	31 (400 AA)
푸사륨 모닐리포르메(*Fusarium moniliforme*) Δ12 데새투라제	32 (1434 bp)	33 (477 AA)
아스페르길루스 니둘란스(*Aspergillus nidulans*) Δ12 데새투라제	34 (1416 bp)	35 (471 AA)
아스페르길루스 플라부스(*Aspergillus flavus*) Δ12 데새투라제	--	36 (466 AA)
아스페르길루스 푸미가투스(*Aspergillus fumigatus*) Δ12 데새투라제	--	37 (424 AA)
마그나포르테 그리세아(*Magnaporthe grisea*) Δ12 데새투라제	38 (1656 bp)	39 (551 AA)
뉴로스포라 크라싸(*Neurospora crassa*) Δ12 데새투라제	40 (1446 bp)	41 (481 AA)
푸사륨 그라미네아룸(*Fusarium graminearium*) Δ12 데새투라제	42 (1371 bp)	43 (456 AA)
모르티에렐라 알피나 Δ12 데새투라제	44 (1403 bp)	45 (400 AA)
사카로마이세스 클루이베리(*Saccharomyces kluyveri*) Δ12 데새투라제	--	46 (416 AA)
클루이베로마이세스 락티스(*Kluyveromyces lactis*) Δ12 데새투라제	47 (1948 bp)	48 (415 AA)
칸디다 알비칸스(*Candida albicans*) Δ12 데새투라제	--	49 (436 AA)
데바리오마이세스 한세니이(*Debaryomyces hansenii*) CBS767 Δ12 데새투라제	--	50 (416 AA)
푸사륨 모닐리포르메 Δ15 데새투라제	51 (1209 bp)	52 (402 AA)
아스페르길루스 니둘란스 Δ15 데새투라제	53 (1206 bp)	54 (401 AA)
마그나포르테 그리세아 Δ15 데새투라제	55 (1185 bp)	56 (394 AA)
뉴로스포라 크라싸 Δ15 데새투라제	57 (1290 bp)	58 (429 AA)
푸사륨 그라미네아룸 Δ15 데새투라제	59 (1212 bp)	60 (403 AA)
모르티에렐라 알피나 Δ15 데새투라제	61 (1353 bp)	62 (403 AA)

[표 1b]

클루이베로마이세스 락티스 Δ15 데새투라제	63 (1248 bp)	64 (415 AA)
칸디다 알비칸스 Δ15 데새투라제	--	65 (433 AA)
사카로마이세스 클루이베리 Δ15 데새투라제	--	66 (419 AA)
데바리오마이세스 한세니이 CBS767 Δ15 데새투라제	--	67 (435 AA)
아스페르길루스 푸미가투스 Δ15 데새투라제	--	68 (396 AA)
이소크리시스 갈바나 Δ9 엘론가제	69 (792 bp)	70 (263 AA)
야로위아 리폴리티카에서의 발현을 위해 코돈-최적화된 합성 Δ9 엘론가제 유전자	71 (792 bp)	70 (263 AA)
유글레나 그라실리스 Δ8 데새투라제 유전자 (비-관능성; 진뱅크 관리 번호 AAD45877)	72 (1275 bp)	73 (419 AA)
유글레나 그라실리스 Δ8 데새투라제 유전자 (비-관능성; 문헌 [Wallis et al., Archives of Biochem. Biophys., 365:307-306 (1999)]; 제WO 00/34439호)	--	74 (422 A)
야로위아 리폴리티카에서의 발현을 위해 코돈-최적화된 합성 Δ8 데새투라제 유전자 (D8S-1)	75 (1270 bp)	--

야로위아 리폴리티카에서의 발현을 위해 코돈-최적화된 합성 Δ8 데새투라제 유전자 (D8S-3)	76 (1269 bp)	--
유글레나 그라실리스 Δ8 데새투라제 유전자 (Eg5)	77 (1271 bp)	78 (421 AA)
유글레나 그라실리스 Δ8 데새투라제 유전자 (Eg12)	79 (1271 bp)	80 (421 AA)
야로위아 리폴리티카에서의 발현을 위해 코돈-최적화된 합성 Δ8 데새투라제 유전자 (D8SF)	81 (1272 bp)	82 (422 AA)
라투스 노르베기쿠스(*Rattus norvegicus*) $C_{16/18}$ 엘론가제	83 (2628 bp)	84 (267 AA)
야로위아 리폴리티카에서의 발현을 위해 코돈-최적화된, 라투스 노르베기쿠스로부터 유래된 합성 $C_{16/18}$ 엘론가제 유전자	85 (804 bp)	84 (276 AA)
모르티에렐라 알피나 $C_{16/18}$ 엘론가제 (ELO3)	86 (828 bp)	87 (275 AA)
모르티에렐라 알피나 ELO3 - 부분 cDNA 서열	88 (607 bp)	--
모르티에렐라 알피나 ELO3 - 게놈 워킹(genome walking)에 의해 수득한 3' 서열	89 (1,042 bp)	--
모르티에렐라 알피나 ELO3 - 게놈 워킹에 의해 수득한 5' 서열	90 (2,223 bp)	--

[표 1c]

모르티에렐라 알피나 ELO3 - cDNA 콘티그	91 (3,557 bp)	--
모르티에렐라 알피나 ELO3 - 인트론	92 (542 bp)	--
모르티에렐라 알피나 ELO3 - 게놈 콘티그	93 (4,099 bp)	--
야로위아 리폴리티카 $C_{16/18}$ 엘론가제 유전자	94 (915 bp)	95 (304 AA)
칸디다 알비칸스 가능한 지방산 엘론가제 (진뱅크 관리 번호 EAL04510)	--	96 (353 AA)
야로위아 리폴리티카 $C_{14/16}$ 엘론가제 유전자	97 (978 bp)	98 (325 AA)
뉴로스포라 크라싸 FEN1 유전자 (진뱅크 관리 번호 CAD70918)	--	99 (337 AA)
오스트레오코쿠스 타우리(*Ostreococcus tauri*) $C_{20/22}$ 엘론가제	100 (903 bp)	101 (300 AA)
야로위아 리폴리티카에서의 발현을 위해 코돈-최적화된, 오스트레오코쿠스 타우리로부터 유래된 합성 $C_{20/22}$ 엘론가제 유전자	102 (903 bp)	103 (300 AA)
트라우스토키트륨 아우레움(*Traustochytrium aureum*) Δ4 데새투라세	104 (1548 bp)	105 (515 AA)
야로위아 리폴리티카에서의 발현을 위해 코돈-최적화된, 트라우스토키트륨 아우레움으로부터 유래된 합성 Δ4 데새투라세	106 (1545 bp)	107 (514 AA)
모르티에렐라 알피나 리소포스파티드산 아실트랜스퍼라제 (LPAAT1)	108 (945 bp)	109 (314 AA)
모르티에렐라 알피나 리소포스파티드산 아실트랜스퍼라제 (LPAAT2)	110 (927 bp)	111 (308 AA)
야로위아 리폴리티카 리소포스파티드산 아실트랜스퍼라제 (LPAAT1)	112 (1549 bp)	113 (282 AA)
야로위아 리폴리티카 리소포스파티드산 아실트랜스퍼라제 (LPAAT2) - 유전자를 포함하는 게놈 단편	114 (1495 bp)	--
야로위아 리폴리티카 리소포스파티드산 아실트랜스퍼라제 (LPAAT2)	115 (672 bp)	116 (223 AA)
야로위아 리폴리티카 포스포리피드:디아실글리세롤 아실트랜스퍼라제 (PDAT)	117 (2326 bp)	118 (648 AA)
야로위아 리폴리티카 아실-CoA:스테롤-아실트랜스퍼라제 (ARE2)	119 (1632 bp)	120 (543 AA)
캐노르합디티스 엘레강스(*Caenorhabditis elegans*) 아실-CoA:1-아실 리소포스파티딜콜린 아실트랜스퍼라제 (LPCAT)	--	121 (282 AA)
야로위아 리폴리티카 디아실글리세롤 아실트랜스퍼라제 (DGAT1)	122 (1578 bp)	123 (526 AA)
모르티에렐라 알피나 디아실글리세롤 아실트랜스퍼라제 (DGAT1)	124 (1578 bp)	125 (525 AA)
뉴로스포라 크라싸 디아실글리세롤 아실트랜스퍼라제 (DGAT1)	--	126 (533 AA)
지베렐라 제애(*Gibberella zeae*) PH-1 디아실글리세롤 아실트랜스퍼라제 (DGAT1)	--	127 (499 AA)
마그나포르테 그리세아 디아실글리세롤 아실트랜스퍼라제 (DGAT1)	--	128 (503 AA)

[표 1d]

아스페르길루스 니둘란스 디아실글리세롤 아실트랜스퍼라제 (DGAT1)	--	129 (458 AA)
야로위아 리폴리티카 디아실글리세롤 아실트랜스퍼라제 (DGAT2)	130 (2119 bp) 132 (1380 bp) 134 (1068 bp)	131 (514 AA) 133 (459 AA) 135 (355 AA)
모르티에렐라 알피나 디아실글리세롤 아실트랜스퍼라제 (DGAT2)	136 (996 bp)	137 (331 AA)
모르티에렐라 알피나 글리세롤-3-포스페이트 아실트랜스퍼라제 (GPAT)	138 (2151 bp)	139 (716 AA)
모르티에렐라 알피나 GPAT - 부분 cDNA 서열	140 (1212 bp)	--
모르티에렐라 알피나 GPAT - -1050 bp 내지 +2886 bp 영역을 포함하는 게놈 단편	141 (3935 bp)	--
모르티에렐라 알피나 GPAT - 게놈 워킹에 의해 수득한 3' cDNA 서열	142 (965 bp)	--
모르티에렐라 알피나 GPAT - 게놈 워킹에 의해 수득한 5' cDNA 서열	143 (1908 bp)	--
모르티에렐라 알피나 GPAT - 게놈 워킹에 의해 수득한 내부 서열	144 (966 bp)	--
모르티에렐라 알피나 GPAT - 인트론 #1	145 (275 bp)	--
모르티에렐라 알피나 GPAT - 인트론 #2	146 (255 bp)	--
모르티에렐라 알피나 GPAT - 인트론 #3	147 (83 bp)	--
모르티에렐라 알피나 GPAT - 인트론 #4	148 (99 bp)	--
야로위아 리폴리티카 디아실글리세롤 콜린포스포트랜스퍼라제 (CPT1) - 유전자를 포함하는 게놈 단편	149 (2133 bp)	--
야로위아 리폴리티카 디아실글리세롤 콜린포스포트랜스퍼라제 (CPT1)	150 (1185 bp)	151 (394 AA)
사카로마이세스 세레비지애 이노시톨 스핑고인지질-특이적 포스포리파제 C (ISC1)	152 (1434 bp)	153 (477 AA)
야로위아 리폴리티카 글리세르알데히드-3-포스페이트 데히드로게나제 프로모터 (GPD)	210 (971 bp)	--
야로위아 리폴리티카 글리세르알데히드-3-포스페이트 데히드로게나제 + 인트론 프로모터 (GPDIN)	211 (1174 bp)	--
야로위아 리폴리티카 포스포글리세레이트 뮤타제 프로모터 (GPM)	212 (878 bp)	--
야로위아 리폴리티카 프럭토스-비스포스페이트 알돌라제 프로모터 (FBA)	1213 (1001 bp)	--

[표 1e]

야로위아 리폴리티카 프럭토스-비스포스페이트 알돌라제 + 인트론 프로모터 (FBAIN)	214 (973 bp)	--
야로위아 리폴리티카 프럭토스-비스포스페이트 알돌라제 + 개질된 인트론 프로모터 (FBAINm)	215 (924 bp)	--
야로위아 리폴리티카 글리세롤-3-포스페이트 아실트랜스퍼라제 (GPAT)	216 (1130 bp)	--
야로위아 리폴리티카 암모늄 수송자 프로모터 (YAT1)	217 (778 bp)	--
야로위아 리폴리티카 번역 연장 인자 ET1-α 프로모터 (TEF)	218 (436 bp)	--
야로위아 리폴리티카 키메라 GPM::FBA 인트론 프로모터 (GPM::FBAIN)	219 (1020 bp)	--
야로위아 리폴리티카 키메라 GPM::GPD 인트론 프로모터 (GPM::GPDIN)	220 (1052 bp)	--
야로위아 리폴리티카 방출 단백질 프로모터 (EXP1)	221 (1000 bp)	--

서열 154 및 156-209는 표 2에 나타낸 플라스미드이다.

[표 2]

플라스미드 서열의 요약

플라스미드	상응하는 도면	서열
pY5-30	6A	154 (8,953 bp)
pKUNF12T6E	8A	156 (12,649 bp)
pDMW271	8B	157 (13,034 bp)
pZP3L37	8C	158 (12,690 bp)
pZKUT16	8D	159 (5,833 bp)
pKO2UM25E	8E	160 (12,663 bp)
pDMW303	9A	161 (15,996 bp)
pZUF17	9B	162 (8,165 bp)
pZUF4	--	163 (8,633 bp)
pZUF4S	9C	164 (8,633 bp)
pZP2FOEN4	9E	165 (10,660 bp)
pKUNT2	10A	166 (6,457 bp)
pDMW237	10B	167 (7,879 bp)
pY54PC	--	168 (8,502 bp)
pKUNFmkF2	10E	169 (7,145 bp)
pZF5T-PPC	13B	170 (5,553 bp)
pDMW297	13D	171 (10,448 bp)
pZP2C16M899	14A	172 (15,543 bp)
pDMW314	14B	173 (13,295 bp)
pDMW325	14C	174 (15,559 bp)
pZKSL5598	14D	175 (16,325 bp)
pY72	15A	176 (10,189 bp)
pY80	15B	177 (12,558 bp)
pY79	15C	178 (8,982 bp)
pY86	15D	179 (10,424 bp)
pY94	16A	180 (10,485 bp)
pY91M	16B	181 (8,423 bp)
pDMW232	16C	182 (10,945 bp)
pY37/F15	16D	183 (8,194 bp)
pKO2UF2PE	17A	184 (10,838 bp)
pZKUGPI5S	17B	185 (6,912 bp)
pDMW302T16	17C	186 (14,864 bp)
pZKUGPE1S	--	187 (6,540 bp)
pKO2UM26E	17D	188 (13,321 bp)
pZKUM	--	189 (4,313 bp)
pMLPAT-17	--	190 (8,015 bp)
pMLPAT-Int	--	191 (8,411 bp)
pZUF-MOD-1	18A	192 (7,323 bp)
pMDGAT1-17	18B	193 (8,666 bp)
pMDGAT2-17	--	194 (8,084 bp)
pMGPAT-17	18C	195 (9,239 bp)
pZF5T-PPC-E3	--	196 (5,031 bp)
pZUF6S	21A	197 (8,462 bp)
pZUF6S-E3WT	21B	198 (11,046 bp)
pZKUGPYE1-N	21C	199 (6,561 bp)
pZKUGPYE2	21D	200 (6,498 bp)
pZUF6TYE2	--	201 (10,195 bp)
pZKUGPYE1	22A	202 (6,561 bp)
pZUF6FYE1	22B	203 (10,809 bp)
pYCPT1-17	--	204 (8,273 bp)
pZP2I7 + Ura	22C	205 (7,822 bp)
pYCPT1-ZP2I7	--	206 (7,930 bp)
pTEF::ISC1	--	207 (8,179 bp)
pY20	22D	208 (8,196 bp)
pLV13	22E	209 (5,105 bp)

서열 155은 야로위아 종에서 최적으로 발현되는 유전자를 위한 코돈-최적화된 번역 개시 부위에 상응한다.

서열 222는 진균 Δ15 및 Δ12 데새투라제에서 발견되는 His 박스 1 모티프에 상응한다.

서열 223은 Δ15 데새투라제 활성을 갖는 진균 단백질임을 표시하는 모티프에 상응하고, 서열 224는 Δ12 데새투라제 활성을 갖는 진균 단백질임을 표시하는 모티프에 상응한다.

서열 225-238은 각각 야로위아 리폴리티카 프로모터 영역을 증폭시키는데 사용되는 프라이머 YL211, YL212, YL376, YL377, YL203, YL204, GPAT-5-1, GPAT-5-2, ODMW314, YL341, ODMW320, ODMW341, 27203-F 및 27203-R에 상응한다.

서열 239-242는 각각 실시간 분석에 사용되는 올리고뉴클레오티드 YL-URA-16F, YL-URA-78R, GUS-767F 및 GUS-891R이다.

서열 243은 W497L 돌연변이를 포함하는 돌연변이체 AHAS 유전자이다.

서열 244-249는 각각 돌연변이를 포함하는 돌연변이체 야로위아 리폴리티카 AHAS 유전자의 합성에 사용되는, 프라이머 410, 411, 412, 413, 414 및 415에 상응한다.

서열 250-281은 전체 코돈-최적화된 코딩 영역을 함께 포함하는 16쌍의 올리고뉴클레오티드 (즉, D4-1A, D4-1B, D4-2A, D4-2B, D4-3A, D4-3B, D4-4A, D4-4B, D4-5A, D4-5B, D4-6A, D4-6B, D4-7A, D4-7B, D4-8A, D4-8B, D4-9A, D4-9B, D4-10A, D4-10B, D4-11A, D4-11B, D4-12A, D4-12B, D4-13A, D4-13B, D4-14A, D4-14B, D4-15A, D4-15B, D4-16A 및 D4-16B)에 상응한다.

서열 282-289는 각각 코돈-최적화된 Δ4 데새투라제 유전자의 합성 동안에 PCR 증폭을 위해 사용되는 프라이머 D4-1F, D4-4R, D4-5F, D4-8R, D4-9F, D4-12R, D4-13 및 D4-16R에 상응한다.

서열 290 및 291은 각각 코돈-최적화된 Δ4 데새투라제 유전자의 합성 동안 사용되는 프라이머 YL251 및 YL252에 상응한다.

서열 292-307은 이소크리시스 갈바나 Δ9 엘론가제의 전체 코돈-최적화된 코딩 영역을 함께 포함하는 8쌍의 올리고뉴클레오티드 (즉, 각각 IL3-1A, IL3-1B, IL3-2A, IL3-2B, IL3-3A, IL3-3B, IL3-4A, IL3-4B, IL3-5A, IL3-5B, IL3-6A, IL3-6B, IL3-7A, IL3-7B, IL3-8A 및 IL3-8B)에 상응한다.

서열 308-311은 각각 코돈-최적화된 Δ9 엘론가제 유전자의 합성 동안에 PCR 증폭을 위해 사용되는 프라이머 IL3-1F, IL3-4R, IL3-5F 및 IL3-8R에 상응한다.

서열 312는 pT9(1-4)에 기재된 417 bp *NcoI/PstI* 단편이고, 서열 313은 pT9(5-8)에 기재된 377 bp *PstI/NotI* 단편이다.

서열 314-339는 유글레나 그라실리스 Δ8 데새투라제의 전체 코돈-최적화된 코딩 영역을 함께 포함하는 13쌍의 올리고뉴클레오티드 (즉, 각각 D8-1A, D8-1B, D8-2A, D8-2B, D8-3A, D8-3B, D8-4A, D8-4B, D8-5A, D8-5B, D8-6A, D8-6B, D8-7A, D8-7B, D8-8A, D8-8B, D8-9A, D8-9B, D8-10A, D8-10B, D8-11A, D8-11B, D8-12A, D8-12B, D8-13A 및 D8-13B)에 상응한다.

서열 340-347은 각각 코돈-최적화된 Δ8 데새투라제 유전자의 합성 동안에 PCR 증폭을 위해 사용되는 프라이머 D8-1F, D8-3R, D8-4F, D8-6R, D8-7F, D8-9R, D8-10F 및 D8-13R에 상응한다.

서열 348은 pT8(1-3)에 기재된 309 bp *Nco/BglII* 단편이고, 서열 349는 pT8(4-6)에 기재된 321 bp *BglII/XhoI* 단편이고, 서열 350은 pT8(7-9)에 기재된 264 bp *XhoI/SacI* 단편이고, 서열 351은 pT8(10-13)에 기재된 369 bp *SacI/NotI* 단편이다.

서열 352 및 353은 각각 pDMW255에서 D8S-2의 합성 동안에 사용되는 프라이머 ODMW390 및 ODMW391에 상응한다.

서열 354 및 355는 실시예 9에 기재된 키메라 D8S-1::XPR 및 D8S-2::XPR 유전자이다.

서열 356 및 357은 D8S-3의 합성 동안에 사용되는 프라이머 ODMW392 및 ODMW393에 상응한다.

서열 358 및 359는 각각 유글레나 그라실리스로부터의 Δ8 데새투라제의 증폭을 위해 사용되는 프라이머 Eg5-1 및 Eg3-3에 상응한다.

서열 360-363은 각각 Δ8 데새투라제 클론의 서열분석을 위해 사용되는 프라이머 T7, M13-28Rev, Eg3-2 및 Eg5-2에 상응한다.

서열 364는 D8S-3의 증폭을 위해 사용되는 프라이머 ODMW404이다.

서열 365는 D8S-3을 포함하는 1272 bp 키메라 유전자이다.

서열 366 및 367은 각각 클로닝된 D8S-3 유전자에서 새로운 제한 효소 부위를 생성하기 위해 사용되는 프라이머 YL521 및 YL522에 상응한다.

서열 368-381은 각각 D8SF를 제공하는 부위 지정 돌연변이유발 반응에 사용되는 프라이머 YL525, YL526, YL527, YL528, YL529, YL530, YL531, YL532, YL533, YL534, YL535, YL536, YL537 및 YL538에 상응한다.

서열 382는 Cre 레콤비나제 효소에 의해 인식되는 LoxP 재조합 부위에 상응한다.

서열 383 및 384는 각각 플라스미드 pY80의 합성 동안에 GPD::Fm1::XPR2를 증폭시키기 위해 사용되는 프라이머 436 및 437에 상응한다.

서열 385-388은 각각 이작용성 Δ5/Δ6 데새투라제의 클로닝을 위해 사용되는 프라이머 475, 477, 478 및 476에 상응한다.

서열 389 및 390은 각각 부위-특이적 돌연변이유발에 의해 플라스미드 pY91M으로부터 플라스미드 pY91V를 생성하기 위해 사용되는 프라이머 505 및 506에 상응한다.

서열 391-393는 각각 BD-클론테크 크리에이터 스마트(BD-Clontech Creator Smart®) cDNA 라이브러리 키트 프라이머 SMART IV 올리고뉴클레오티드, CDSIII/3' PCR 프라이머 및 5'-PCR 프라이머에 상응한다.

서열 394는 모르티에렐라 알피나 cDNA 라이브러리 서열분석에 사용되는 M13 전진 프라이머에 상응한다.

서열 395-398 및 400-401은 각각 모르티에렐라 알피나 LPAAT2 ORF의 클로닝에 사용되는 프라이머 MLPAT-F, MLPAT-R, LPAT-Re-5-1, LPAT-Re-5-2, LPAT-Re-3-1 및 LPAT-Re-3-2에 상응한다.

서열 399 및 402는 각각 야로위아 리폴리티카 LPAAT1 ORF의 5' (1129 bp) 및 3' (938 bp) 영역에 상응한다.

서열 403 및 404는 각각 "대조군" 플라스미드 pZUF-MOD-1의 생성을 위해 사용되는 프라이머 pzuf-mod1 및 pzuf-mod2에 상응한다.

서열 405 및 406은 각각 모르티에렐라 알피나 DGAT1 ORF의 클로닝을 위해 사용되는 프라이머 MACAT-F1 및 MACAT-R에 상응한다.

서열 407 및 408은 각각 모르티에렐라 알피나 DGAT2 ORF의 클로닝을 위해 사용되는 프라이머 MDGAT-F 및 MDGAT-R1에 상응한다.

서열 409 및 410은 각각 모르티에렐라 알피나 GPAT의 증폭을 위한 추정(degenerate) PCR에 사용되는 프라이머 MGPAT-N1 및 MGPAT-NR5에 상응한다.

서열 411-413은 각각 모르티에렐라 알피나 GPAT의 3'-말단의 증폭을 위해 사용되는 프라이머 MGPAT-5N1, MGPAT-5N2 및 MGPAT-5N3에 상응한다.

서열 414 및 415는 게놈-워킹에 사용되는 클론테크 유니버셜 게놈워커(Universal GenomeWalker^{TM}) 키트로부터의 게놈 워커 어댑터에 상응한다.

서열 416-419는 각각 게놈-워킹에 사용되는 PCR 프라이머 MGPAT-5-1A, 어댑터-1 (AP1), MGPAT-3N1 및 네스티드(Nested) 어댑터 프라이머 2 (AP2)에 상응한다.

서열 420 및 421은 각각 모르티에렐라 알피나 GPAT의 증폭을 위해 사용되는 프라이머 mgpat-cdna-5 및 mgpat-cdna-R에 상응한다.

서열 422 및 423은 각각 모르티에렐라 알피나 ELO3의 3'-말단 영역의 단리를 위한 게놈-워킹에 사용되는 프라이머 MA Elong 3'1 및 MA Elong 3'2에 상응한다.

서열 424 및 425는 각각 모르티에렐라 알피나 ELO3의 5'-말단 영역의 단리를 위한 게놈-워킹에 사용되는 프라이머 MA Elong 5'1 및 MA Elong 5'2에 상응한다.

서열 426 및 427은 각각 모르티에렐라 알피나 cDNA로부터의 완전한 ELO3을 증폭시키기 위해 사용되는 프라이머 MA ELONG 5' NcoI 3 및 MA ELONG 3' NotI 1에 상응한다.

서열 428 및 429는 야로위아 리폴리티카 YE2의 코딩 영역을 증폭시키기 위해 사용되는 프라이머 YL597 및 YL598에 상응한다.

서열 430 및 431은 각각 Aco 3' 종결자를 함유하는 *NotI/PacI* 단편의 증폭을 위해 사용되는 프라이머 YL325 및 YL326에 상응한다.

서열 432-435는 야로위아 리폴리티카 YE1의 코딩 영역을 증폭시키기 위해 사용되는 프라이머 YL567, YL568, YL569 및 YL570에 상응한다.

서열 436 및 437은 각각 야로위아 리폴리티카 YE1의 클로닝 동안에 부위-지정 돌연변이유발을 위해 사용되는 프라이머 YL571 및 YL572에 상응한다.

서열 438 및 439는 각각 야로위아 리폴리티카 CPT1 ORF의 클로닝을 위해 사용되는 프라이머 CPT1-5'-*NcoI* 및 CPT1-3'-*NotI*에 상응한다.

서열 440 및 441은 각각 사카로마이세스 세레비지애 ISC1 ORF의 클로닝을 위해 사용되는 프라이머 Isc1F 및 Isc1R에 상응한다.

서열 442 및 443은 각각 사카로마이세스 세레비지애 PCL1 ORF의 클로닝을 위해 사용되는 프라이머 Pcl1F 및 Pcl1R에 상응한다.

서열 444-447은 각각 야로위아 리폴리티카 DGAT2 유전자의 표적화된 파괴를 위해 사용되는 프라이머 P95, P96, P97 및 P98에 상응한다.

서열 448-450은 각각 야로위아 리폴리티카 DGAT2 유전자의 표적화된 통합을 스크리닝하기 위해 사용되는 프라이머 P115, P116 및 P112에 상응한다.

서열 451-454는 각각 야로위아 리폴리티카 PDAT 유전자의 표적화된 파괴를 위해 사용되는 프라이머 P39, P41, P40 및 P42에 상응한다.

서열 455-458은 각각 파괴된 야로위아 리폴리티카 PDAT 유전자의 표적화된 통합을 스크리닝하기 위해 사용되는 프라이머 P51, P52, P37 및 P38에 상응한다.

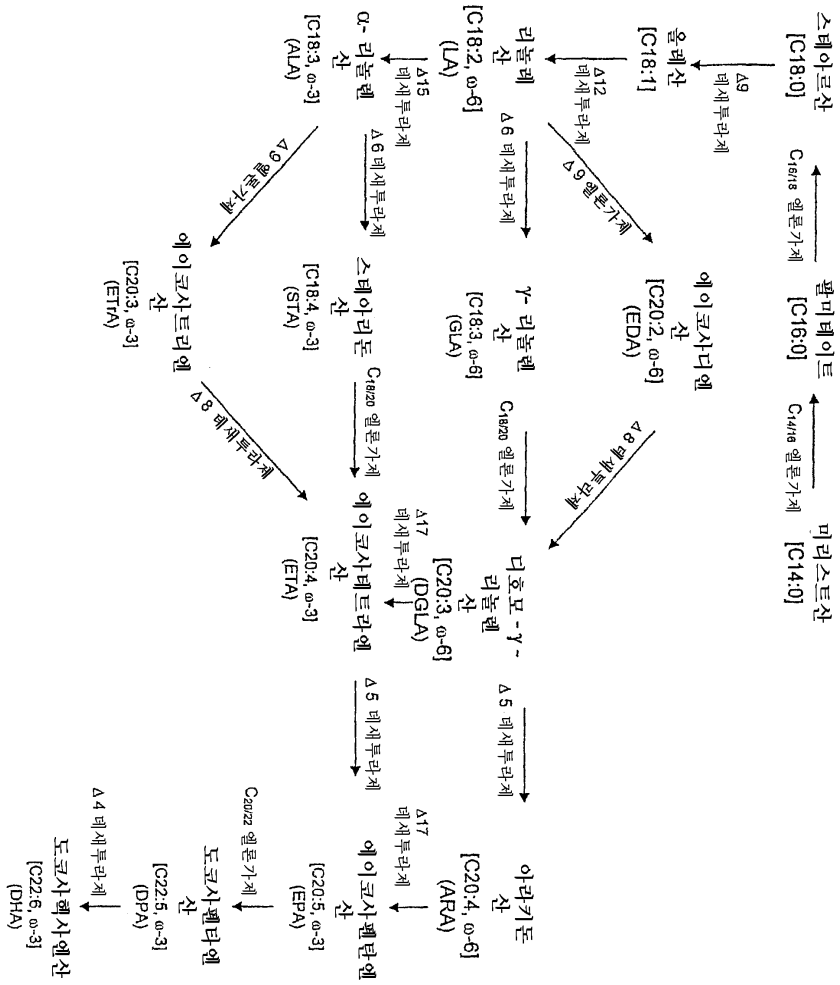
서열 459 및 460은 각각 야로위아 리폴리티카 DGAT1의 단리를 위해 사용되는 추정 프라이머 P201 및 P203에 상응한다.

서열 461-465는 각각 야로위아 리폴리티카에서 추정적 DGAT1 유전자의 표적화된 파괴를 위한 표적화 카세트의 생성에 사용되는 프라이머 P214, P215, P216, P217 및 P219에 상응한다.

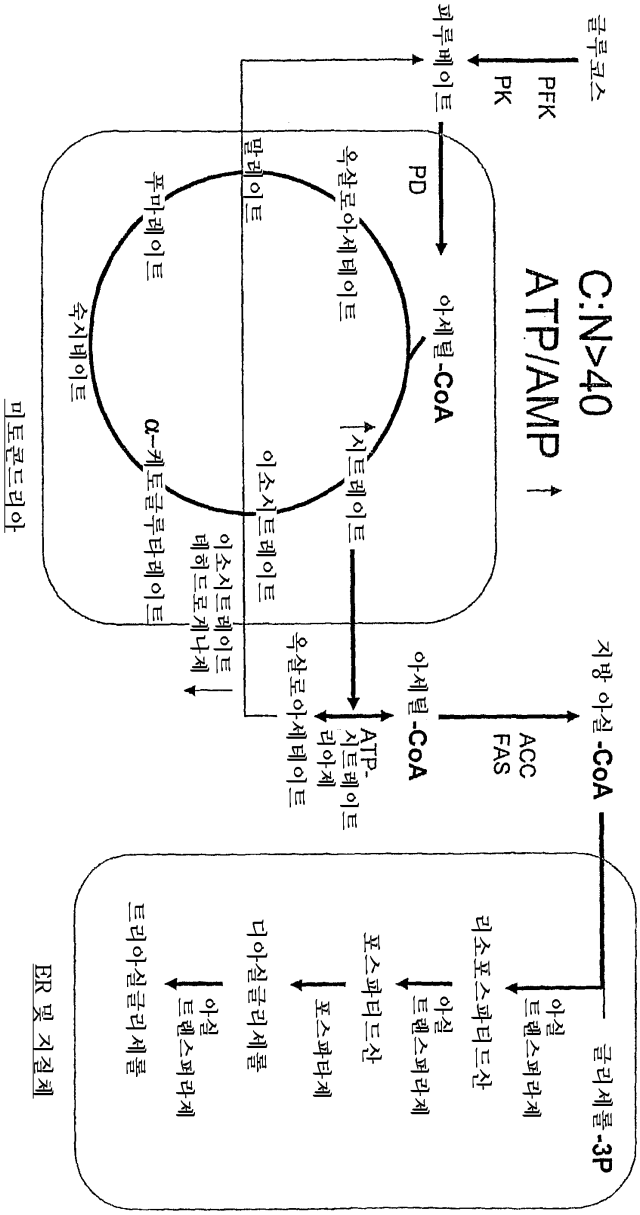
서열 466 및 467은 각각 파괴된 야로위아 리폴리티카 DGAT1 유전자의 표적화된 통합을 스크리닝하기 위해 사용되는 프라이머 P226 및 P227에 상응한다.

도면

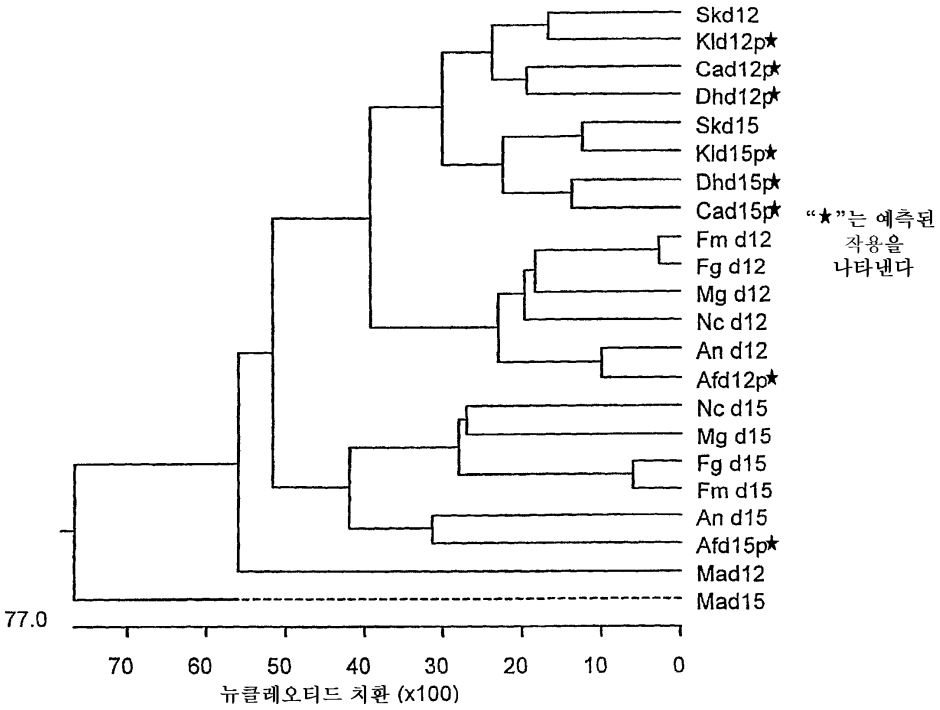
도면1



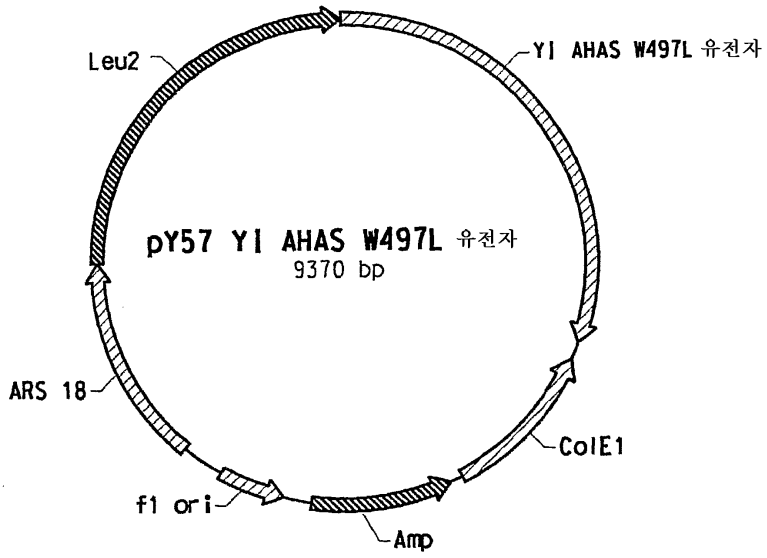
도면2


글루코스
PFK
PK
피루베이트
PD
아세틸-CoA
C:N>40
ATP/AMP ↑
말레이트
옥살로아세테이트
푸마레이트
숙시네이트
α-케토글루타레이트
시트레이트
이소시트레이트
이소시트레이트 데히드로게나제
미토콘드리아
지방 아실-CoA
ACC
FAS
아세틸-CoA
ATP-시트레이트 리아제
옥살로아세테이트
글리세롤-3P
아실 트랜스퍼라제
리소포스파티드산
아실 트랜스퍼라제
포스파티드산
포스파타제
디아실글리세롤
아실 트랜스퍼라제
트리아실글리세롤
ER 및 지질체

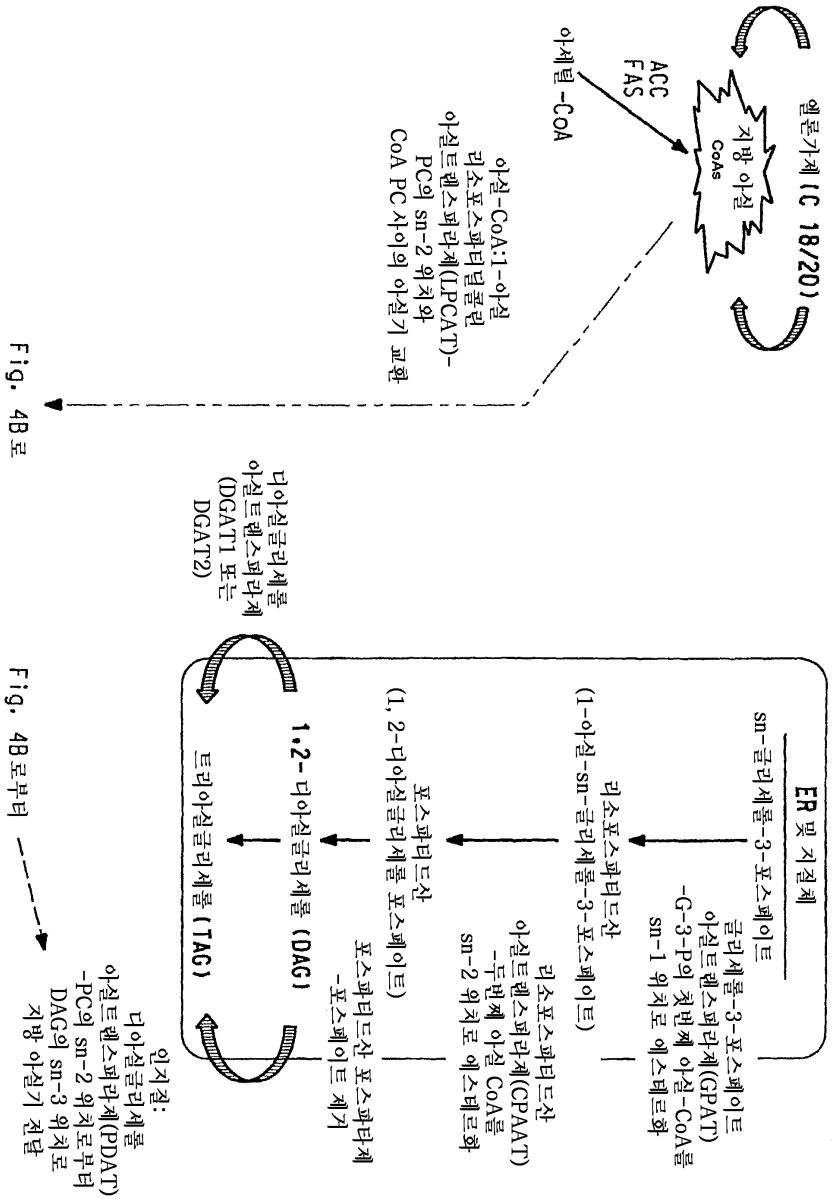
도면3A


Skd12
Kld12p★
Cad12p★
Dhd12p★
Skd15
Kld15p★
Dhd15p★
Cad15p★
Fm d12
Fg d12
Mg d12
Nc d12
An d12
Afd12p★
Nc d15
Mg d15
Fg d15
Fm d15
An d15
Afd15p★
Mad12
Mad15
"★"는 예측된 작용을 나타낸다
77.0
70
60
50
40
30
20
10
0
뉴클레오티드 치환 (x100)

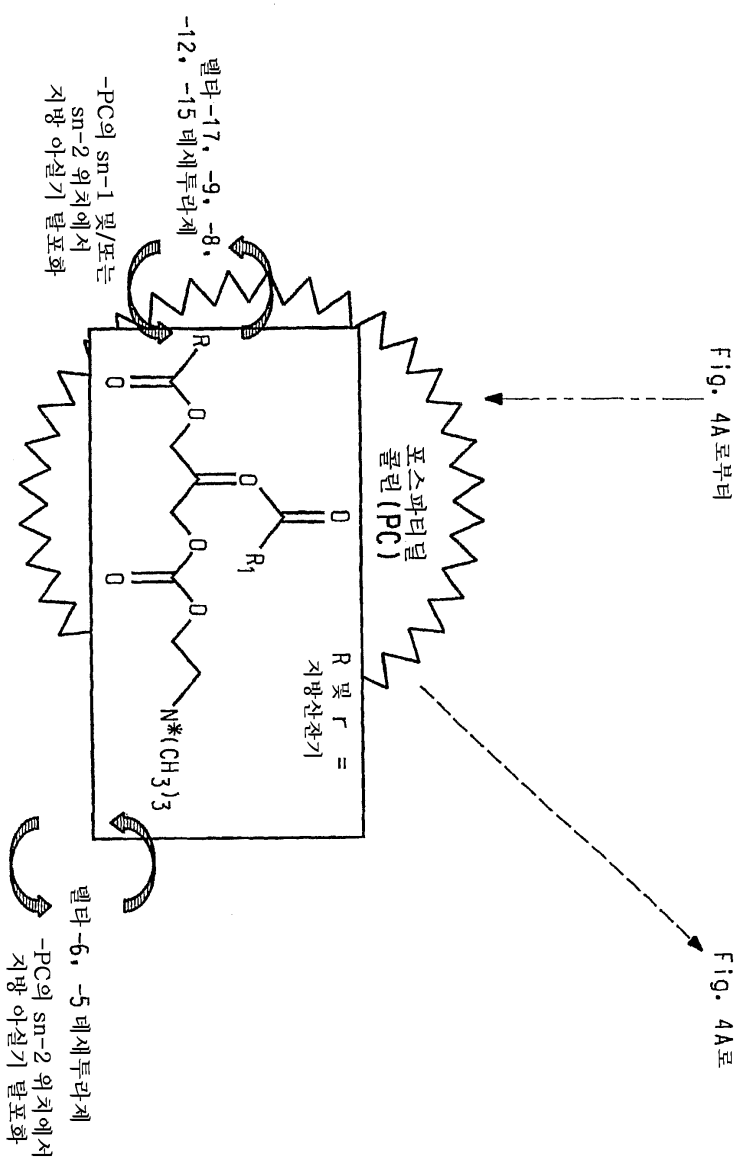
도면3B


Leu2
Y1 AHAS W497L 유전자
pY57 Y1 AHAS W497L 유전자
9370 bp
ARS 18
ColE1
f1 ori
Amp

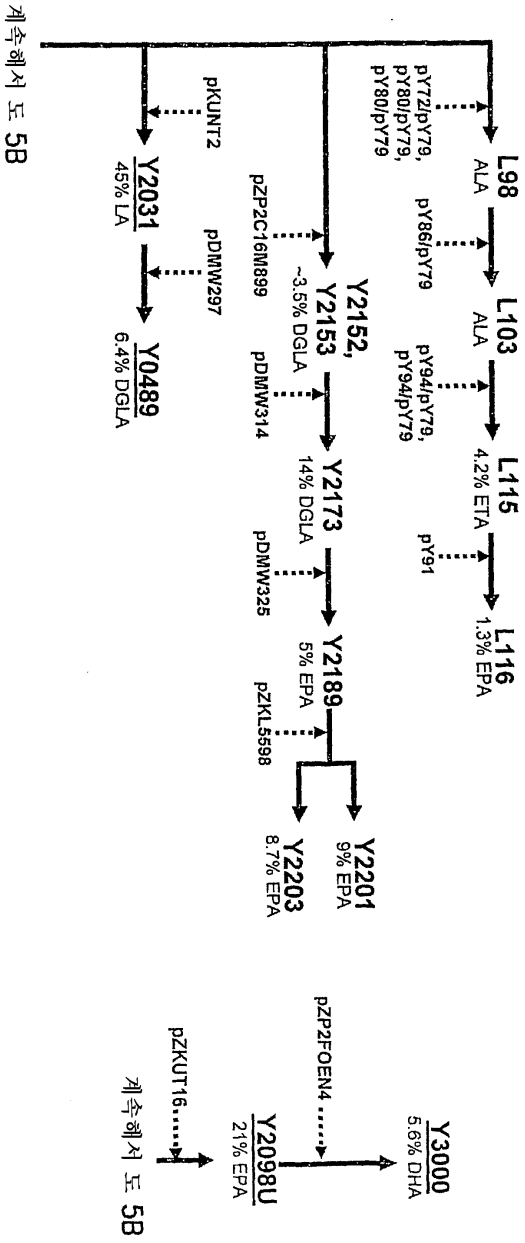
도면4A


엘론가제 (C 18/20)
지방 아실
CoAs
ACC
FAS
아세틸-CoA
아실-CoA:1-아실
리소포스파티딜콜린
아실트랜스퍼라제(LPCAT)-
PC의 sn-2 위치와
CoA PC 사이의 아실기 교환
Fig. 4B로
디아실글리세롤
아실트랜스퍼라제
(DGAT1 또는
DGAT2)
ER 및 지질체
sn-글리세롤-3-포스페이트
글리세롤-3-포스페이트
아실트랜스퍼라제(GPAT)
-G-3-P의 첫번째 아실-CoA를
sn-1 위치로 에스테르화
리소포스파티드산
(1-아실-sn-글리세롤-3-포스페이트)
리소포스파티드산
아실트랜스퍼라제(CPAAT)
-두번째 아실 CoA를
sn-2 위치로 에스테르화
포스파티드산
(1, 2-디아실글리세롤 포스페이트)
포스파티드산 포스파타제
-포스페이트 제거
1,2-디아실글리세롤 (DAG)
트리아실글리세롤 (TAG)
Fig. 4B로부터
인지질:
디아실글리세롤
아실트랜스퍼라제(PDAT)
-PC의 sn-2 위치로부터
DAG의 sn-3 위치로
지방 아실기 전달

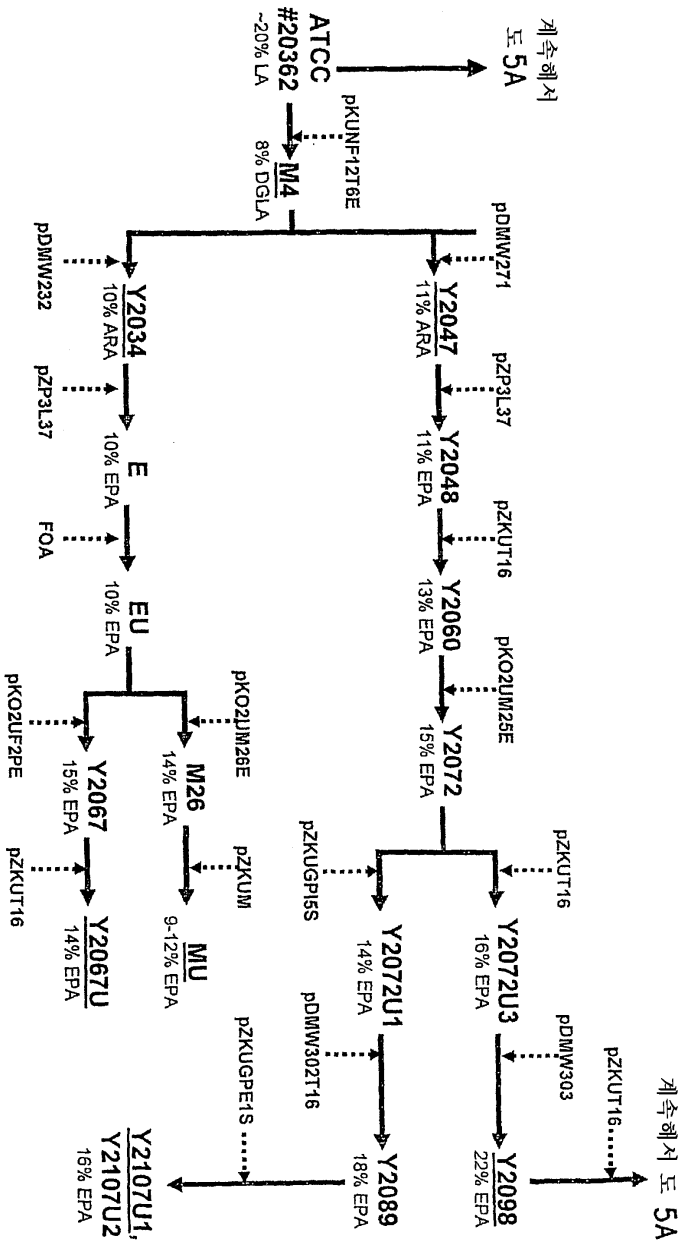
도면4B


Fig. 4A로부터
포스파티딜
콜린(PC)
R 및 r = 지방산잔기
델타-17, -9, -8, -12, -15 데새투라제
-PC의 sn-1 및/또는 sn-2 위치에서 지방 아실기 탈포화
델타-6, -5 데새투라제
-PC의 sn-2 위치에서 지방 아실기 탈포화
Fig. 4A로

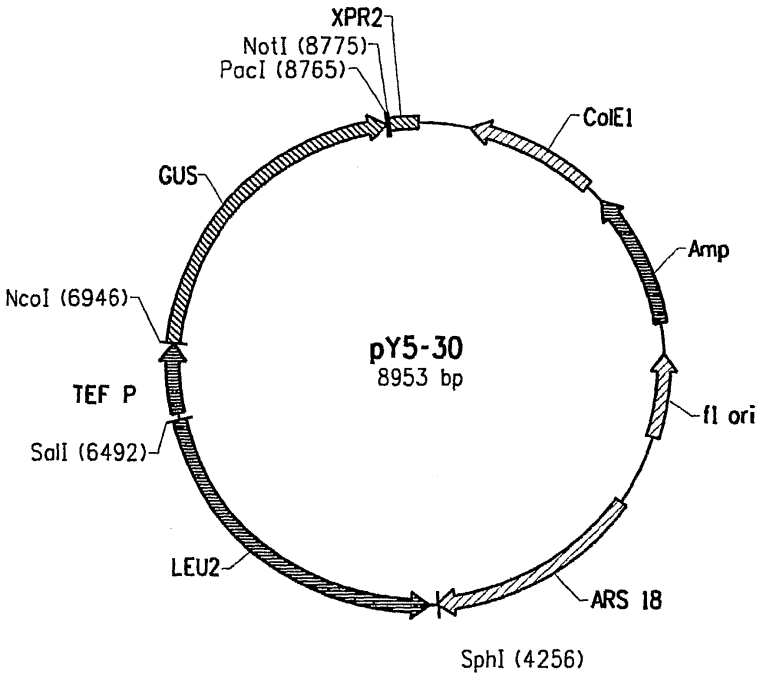
도면5A


계속해서 도 5B
pY72/pY79,
pY80/pY79,
pY80/pY79
L98
ALA
pY86/pY79
L103
ALA
pY94/pY79,
pY94/pY79
L115
4.2% ETA
pY91
L116
1.3% EPA
pKUNT2
Y2031
45% LA
pDMW297
Y0489
6.4% DGLA
pZP2C16M899
Y2152,
Y2153
~3.5% DGLA
pDMW314
Y2173
14% DGLA
pDMW325
Y2189
5% EPA
pZKL5598
Y2203
8.7% EPA
Y2201
9% EPA
pZP2FOEN4
Y2098U
21% EPA
Y3000
5.6% DHA
pZKUT16
계속해서 도 5B

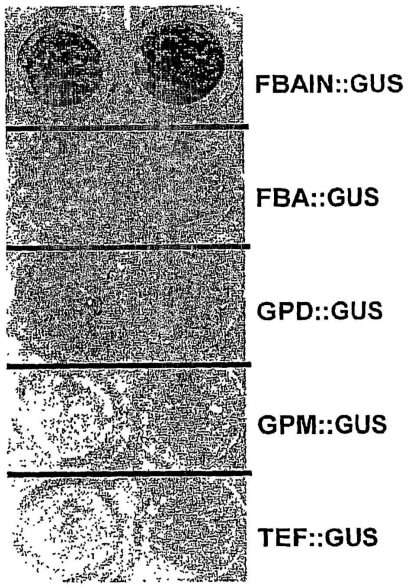
도면5B


계속해서 도 5A
ATCC #20362
~20% LA
pKUNF12T6E
M4
8% DGLA
pDMW271
Y2047
11% ARA
pZP3L37
Y2048
11% EPA
pZKUT16
Y2060
13% EPA
pKO2UM25E
Y2072
15% EPA
pZKUT16
Y2072U3
16% EPA
pDMW303
Y2098
22% EPA
pZKUT16
계속해서 도 5A
pZKUGPI5S
Y2072U1
14% EPA
pDMW302T16
Y2089
18% EPA
pZKUGPE1S
Y2107U1,
Y2107U2
16% EPA
pDMW232
Y2034
10% ARA
pZP3L37
E
10% EPA
FOA
EU
10% EPA
pKO2UF2PE
Y2067
15% EPA
pZKUT16
Y2067U
14% EPA
pKO2UM26E
M26
14% EPA
pZKUM
MU
9-12% EPA

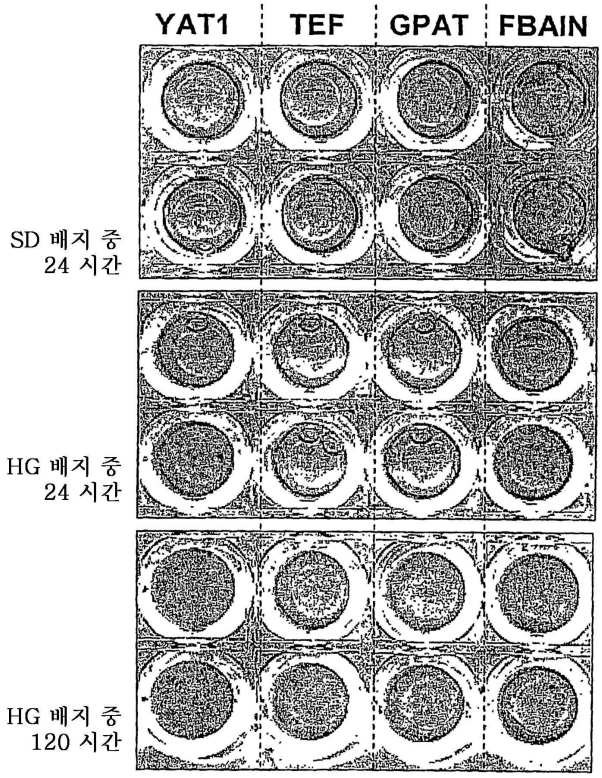
도면6A


XPR2
NotI (8775)
PacI (8765)
ColE1
GUS
Amp
NcoI (6946)
pY5-30
8953 bp
TEF P
f1 ori
SalI (6492)
LEU2
ARS 18
SphI (4256)

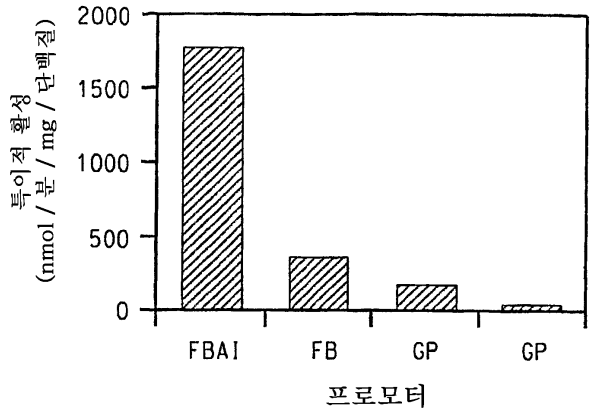
도면6B


FBAIN::GUS
FBA::GUS
GPD::GUS
GPM::GUS
TEF::GUS

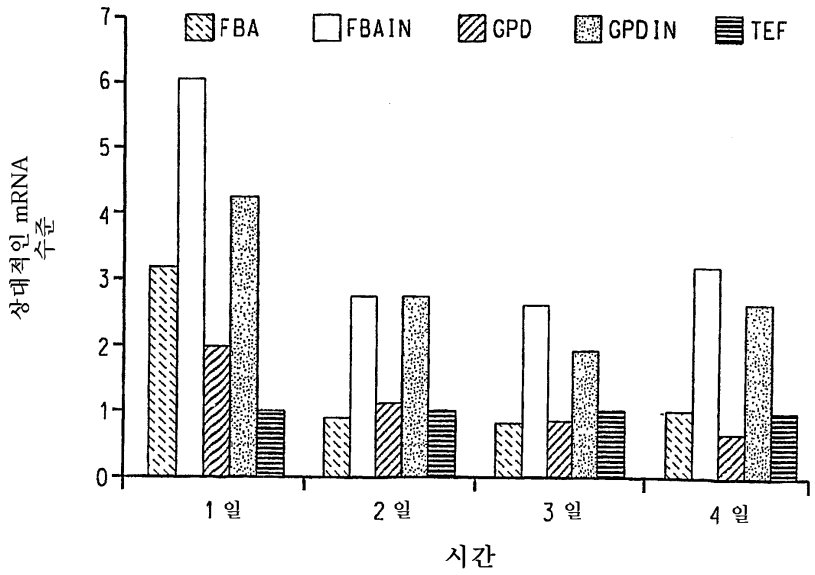
도면6C


YAT1
TEF
GPAT
FBAIN
SD 배지 중
24 시간
HG 배지 중
24 시간
HG 배지 중
120 시간

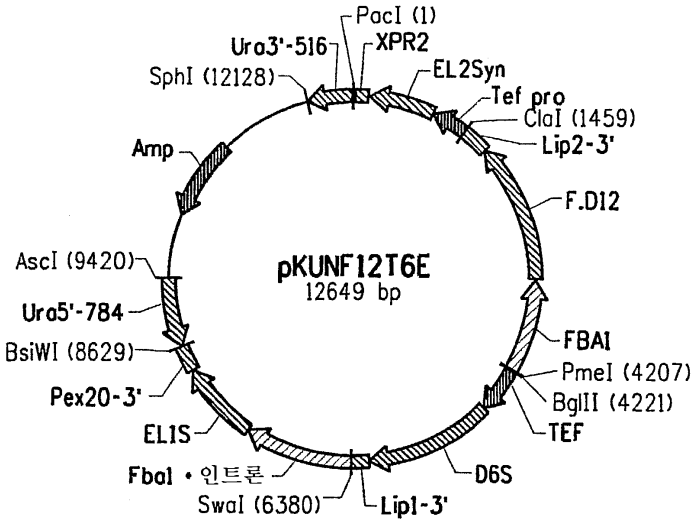
도면7A


특이적 활성
(nmol / 분 / mg / 단백질)
2000
1500
1000
500
0
FBAI
FB
GP
GP
프로모터

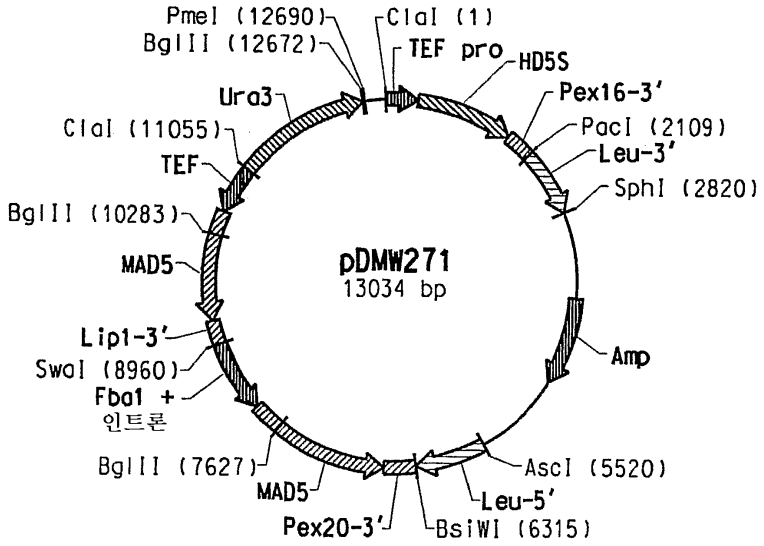
도면7B


FBA
FBAIN
GPD
GPDIN
TEF
상대적인 mRNA 수준
0
1
2
3
4
5
6
7
1 일
2 일
3 일
4 일
시간

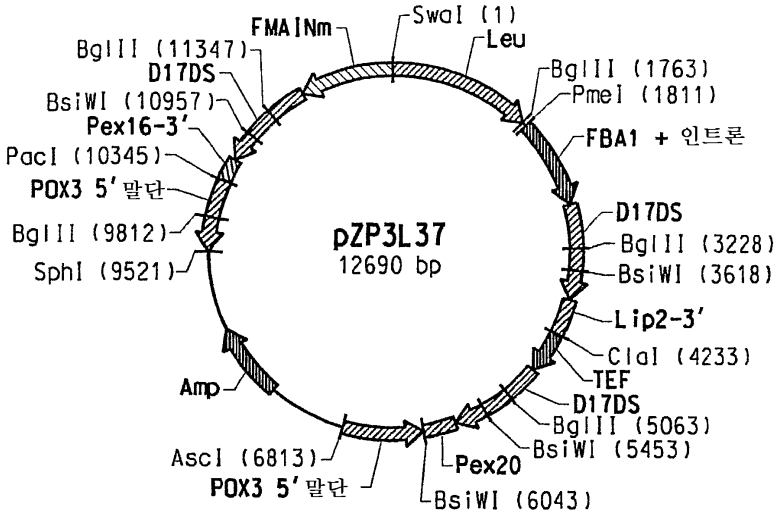
도면8A


PacI (1)
Ura3'-516
XPR2
SphI (12128)
EL2Syn
Tef pro
ClaI (1459)
Lip2-3'
Amp
F.D12
AscI (9420)
pKUNF12T6E
12649 bp
Ura5'-784
FBAI
BsiWI (8629)
PmeI (4207)
BglII (4221)
Pex20-3'
TEF
EL1S
D6S
FbaI • 인트론
SwaI (6380)
Lip1-3'

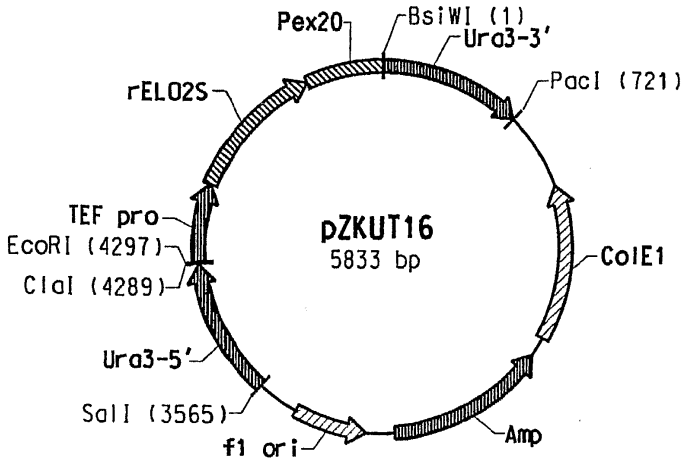
도면8B



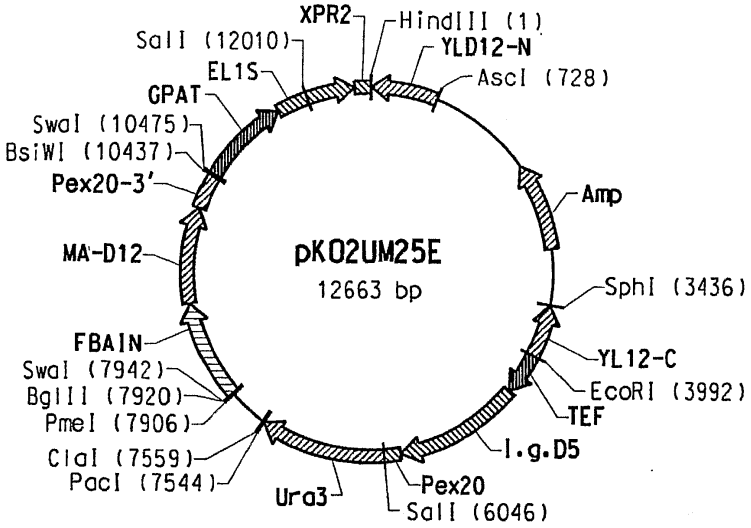
도면8C



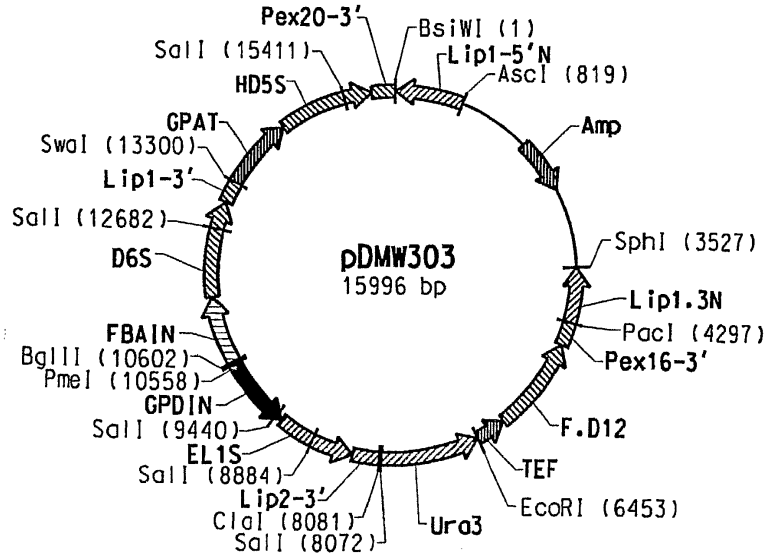
도면8D



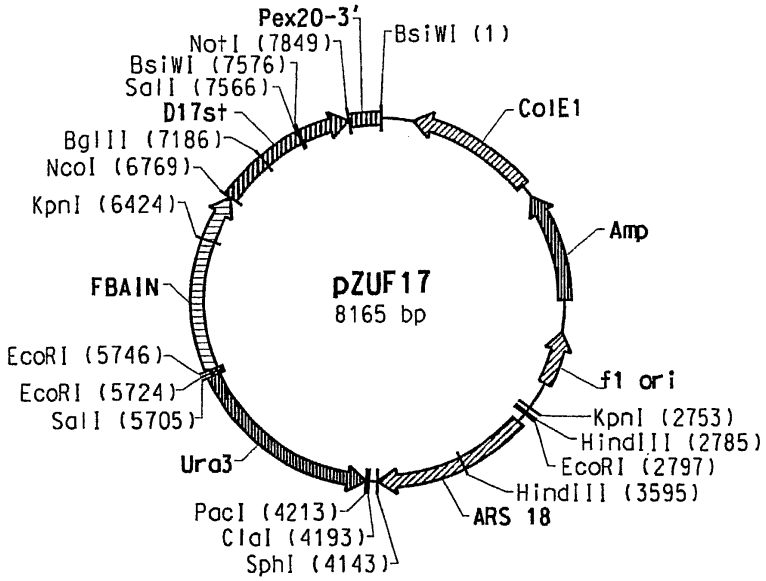
도면8E


XPR2
HindIII (1)
YLD12-N
AscI (728)
Amp
SphI (3436)
YL12-C
EcoRI (3992)
TEF
l.g.D5
Pex20
SalI (6046)
Ura3
PacI (7544)
ClaI (7559)
PmeI (7906)
BglII (7920)
SwaI (7942)
FBAIN
MA-D12
Pex20-3'
BsiWI (10437)
SwaI (10475)
GPAT
EL1S
SalI (12010)
pKO2UM25E
12663 bp

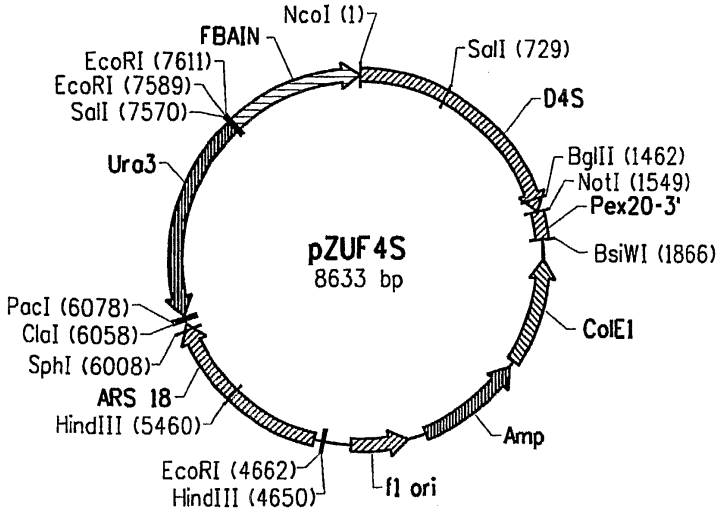
도면9A


Pex20-3'
BsiWI (1)
Lip1-5'N
AscI (819)
Amp
SphI (3527)
Lip1.3N
PacI (4297)
Pex16-3'
F.D12
TEF
EcoRI (6453)
Ura3
Sall (8072)
ClaI (8081)
Lip2-3'
EL1S
Sall (8884)
GPDIN
Sall (9440)
PmeI (10558)
BglII (10602)
FBAIN
D6S
Sall (12682)
Lip1-3'
SwaI (13300)
GPAT
HD5S
Sall (15411)
pDMW303
15996 bp

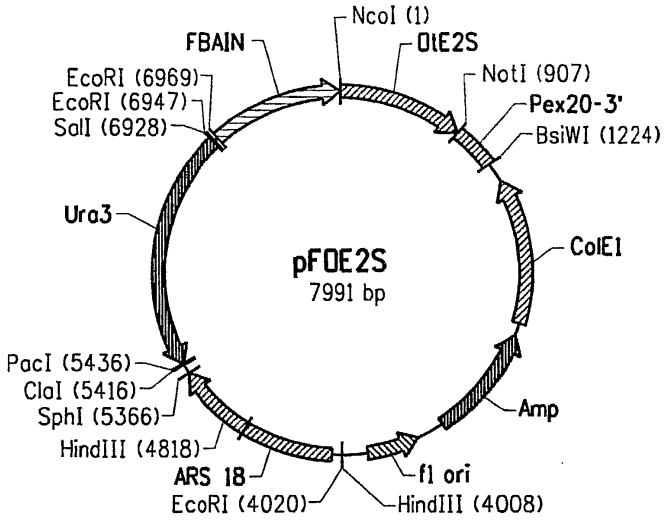
도면9B


Pex20-3'
NotI (7849)
BsiWI (7576)
SalI (7566)
D17st
BglII (7186)
NcoI (6769)
KpnI (6424)
FBAIN
EcoRI (5746)
EcoRI (5724)
SalI (5705)
Ura3
PacI (4213)
ClaI (4193)
SphI (4143)
BsiWI (1)
ColE1
Amp
f1 ori
KpnI (2753)
HindIII (2785)
EcoRI (2797)
HindIII (3595)
ARS 18
pZUF17
8165 bp

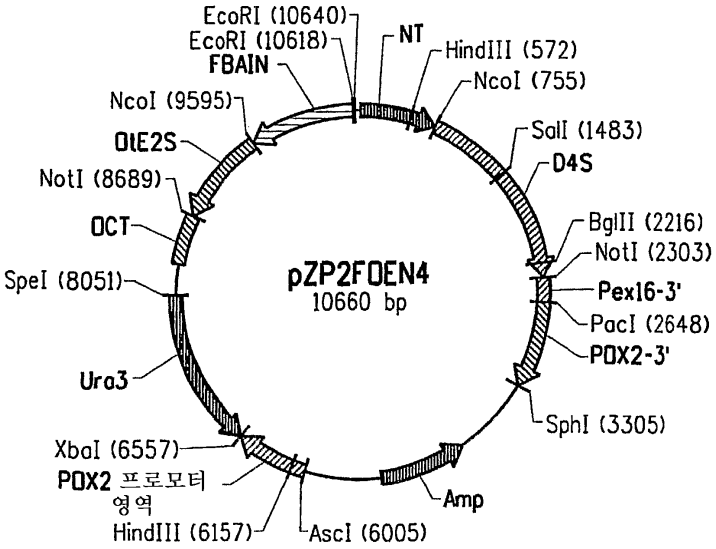
도면9C


NcoI (1)
FBAIN
EcoRI (7611)
EcoRI (7589)
SalI (7570)
Ura3
PacI (6078)
ClaI (6058)
SphI (6008)
ARS 18
HindIII (5460)
EcoRI (4662)
HindIII (4650)
SalI (729)
D4S
BglII (1462)
NotI (1549)
Pex20-3'
BsiWI (1866)
ColE1
Amp
f1 ori
pZUF4S
8633 bp

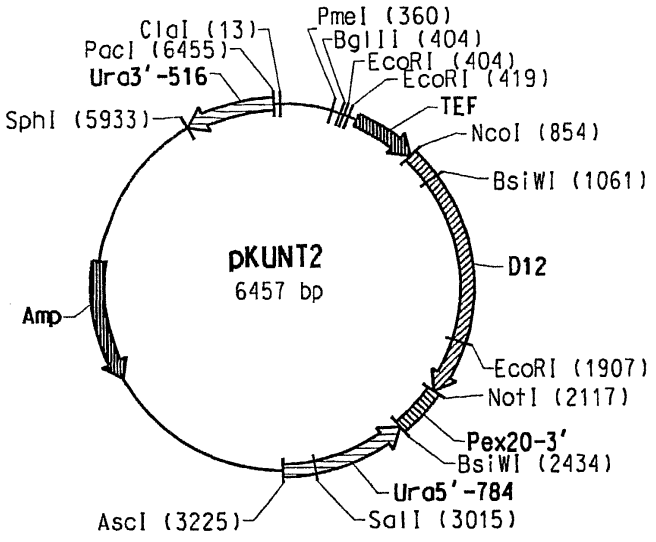
도면9D



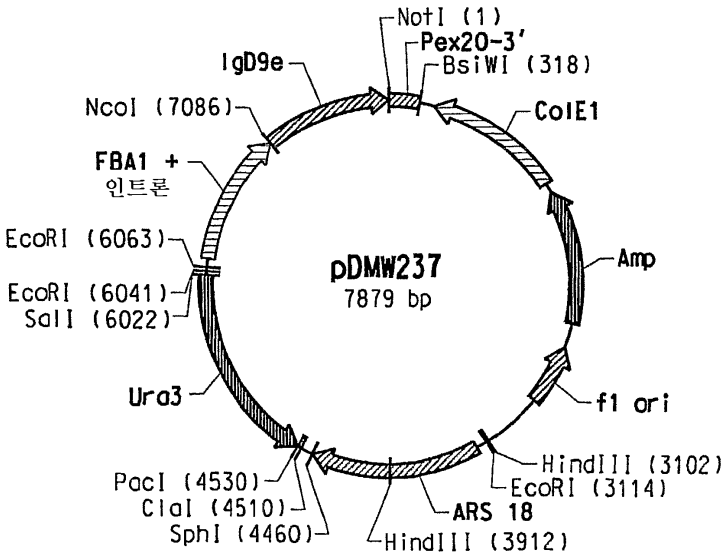
도면9E



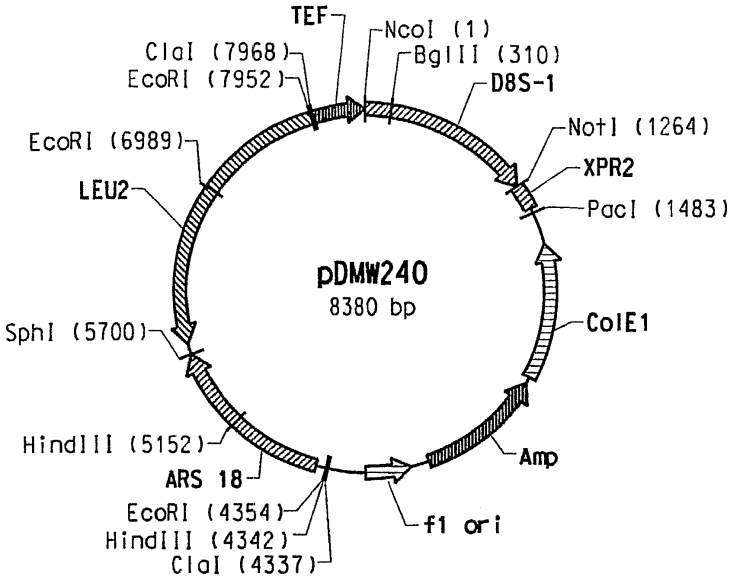
도면10A



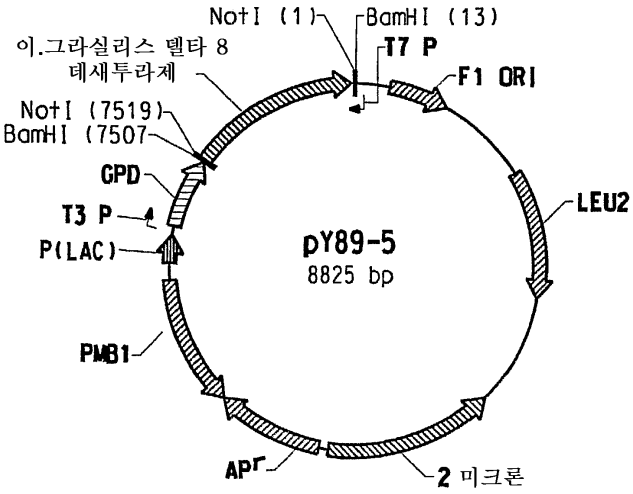
도면10B


NotI (1)
Pex20-3'
BsiWI (318)
ColE1
Amp
f1 ori
HindIII (3102)
EcoRI (3114)
ARS 18
HindIII (3912)
PacI (4530)
ClaI (4510)
SphI (4460)
Ura3
EcoRI (6063)
EcoRI (6041)
SalI (6022)
FBA1 +
인트론
NcoI (7086)
IgD9e
pDMW237
7879 bp

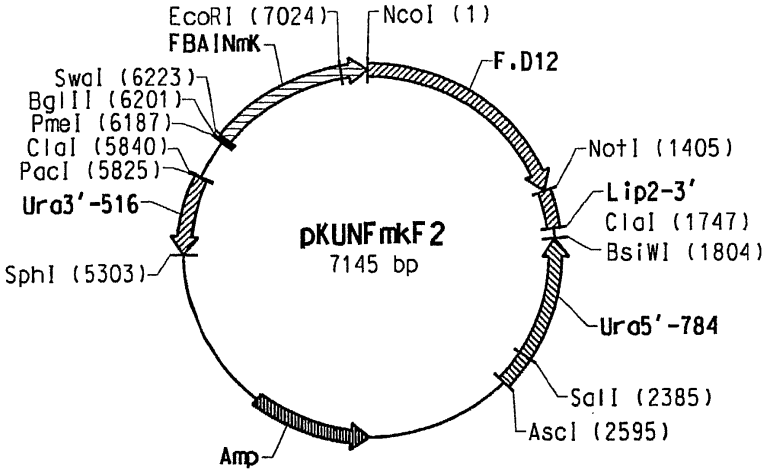
도면10C


TEF
NcoI (1)
BglII (310)
D8S-1
NotI (1264)
XPR2
PacI (1483)
ColE1
Amp
f1 ori
EcoRI (4354)
HindIII (4342)
ClaI (4337)
ARS 18
HindIII (5152)
SphI (5700)
LEU2
EcoRI (6989)
EcoRI (7952)
ClaI (7968)
pDMW240
8380 bp

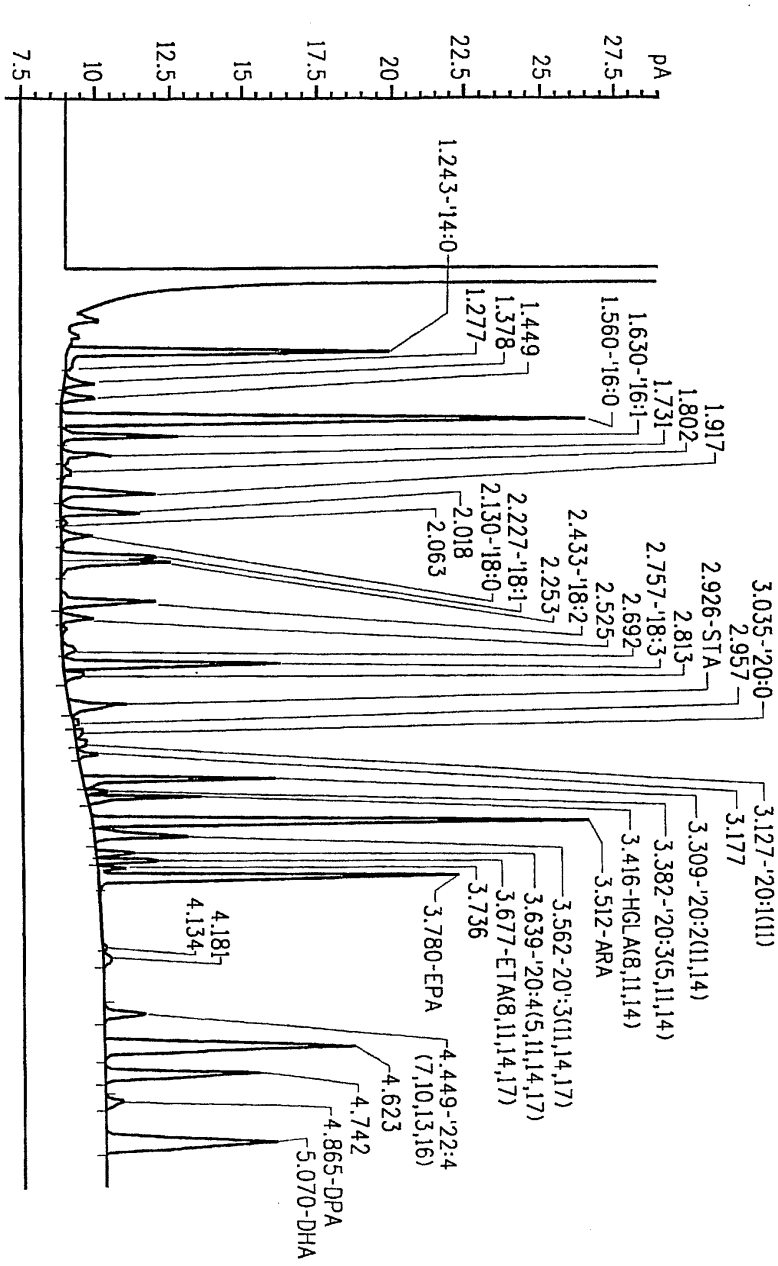
도면10D


NotI (1)
BamHI (13)
이.그라실리스 텔타 8
데새투라제
T7 P
F1 ORI
NotI (7519)
BamHI (7507)
GPD
T3 P
P(LAC)
pY89-5
8825 bp
LEU2
PMB1
APr
2 미크론

도면10E


EcoRI (7024)
NcoI (1)
FBAINmK
F.D12
SwaI (6223)
BglII (6201)
PmeI (6187)
ClaI (5840)
PacI (5825)
Ura3'-516
NotI (1405)
Lip2-3'
ClaI (1747)
BsiWI (1804)
pKUNFmkF2
7145 bp
SphI (5303)
Ura5'-784
SalI (2385)
AscI (2595)
Amp

도면11


pA
27.5
25
22.5
20
17.5
15
12.5
10
7.5
1.243-'14:0
1.449
1.378
1.277
1.630-'16:1
1.560-'16:0
1.917
1.802
1.731
2.227-'18:1
2.130-'18:0
2.018
2.063
2.433-'18:2
2.253
2.525
2.692
2.757-'18:3
2.813
2.926-STA
2.957
3.035-'20:0
3.127-'20:1(11)
3.177
3.309-'20:2(11,14)
3.382-'20:3(5,11,14)
3.416-HGLA(8,11,14)
3.512-ARA
3.562-20':3(11,14,17)
3.639-'20:4(5,11,14,17)
3.677-ETA(8,11,14,17)
3.736
3.780-EPA
4.181
4.134
4.449-'22:4
(7,10,13,16)
4.623
4.742
4.865-DPA
5.070-DHA

도면12

```
        ******** **     ****   *************************** *********
서열 78 MKSKRQAL-PLTIDGTTYDVSAWVNFHPGGAEIIENYQGRDATDAFMVMHSQEAFDKLKR
서열 80 MKSKRQAL-PLTIDGTTYDVSAWVNFHPGGAEIIENYQGRDATDAFMVMHSQEAFDKLKR
서열 73 MKSKRQALSPLQLMEQTYDV---VNFHPGGAEIIENYQGRDATDAFMVMHFQEAFDKLKR
서열 74 MKSKRQALSPLQLMEQTYDVSAWVNFHPGGAEIIENYQGRDATDAFMVMHFQEAFDKLKR

        ******* ************************************************
서열 78 MPKINPSSELPPQAAVNEAQEDFRKLREELIATGMFDASPLWYSYKISTTLGLGVLGYFL
서열 80 MPKINPSSELPPQAAVNEAQEDFRKLREELIATGMFDASPLWYSYKISTTLGLGVLGYFL
서열 73 MPKINPSFELPPQAAVNEAQEDFRKLREELIATGMFDASPLWYSYKISTTLGLGVLGYFL
서열 74 MPKINPSFELPPQAAVNEAQEDFRKLREELIATGMFDASPLWYSYKISTTLGLGVLGYFL

        ********************************************************* **
서열 78 MVQYQMYFIGAVLLGMHYQQMGWLSHDICHHQTFKNRNWNNLVGLVFGNGLQGFSVTWWK
서열 80 MVQYQMYFIGAVLLGMHYQQMGWLSHDICHHQTFKNRNWNNLVGLVFGNGLQGFSVTWWK
서열 73 MVQYQMYFIGAVLLGMHYQQMGWLSHDICHHQTFKNRNWNNLVGLVFGNGLQGFSVTCWK
서열 74 MVQYQMYFIGAVLLGMHYQQMGWLSHDICHHQTFKNRNWNNLVGLVFGNGLQGFSVTCWK

        ********************** ***********************************
서열 78 DRHNAHHSATNVQGHDPDIDNLPLLAWSEDDVTRASPISRKLIQFQQYYFLVICILLRFI
서열 80 DRHNAHHSATNVQGHDPDIDNLPLLAWSEDDVTRASPISRKLIQFQQYYFLVICILLRFI
서열 73 DRHNAHHSATNVQGHDPDIDNLPPLAWSEDDVTRASPISRKLIQFQQYYFLVICILLRFI
서열 74 DRHNAHHSATNVQGHDPDIDNLPPLAWSEDDVTRASPISRKLIQFQQYYFLVICILLRFI

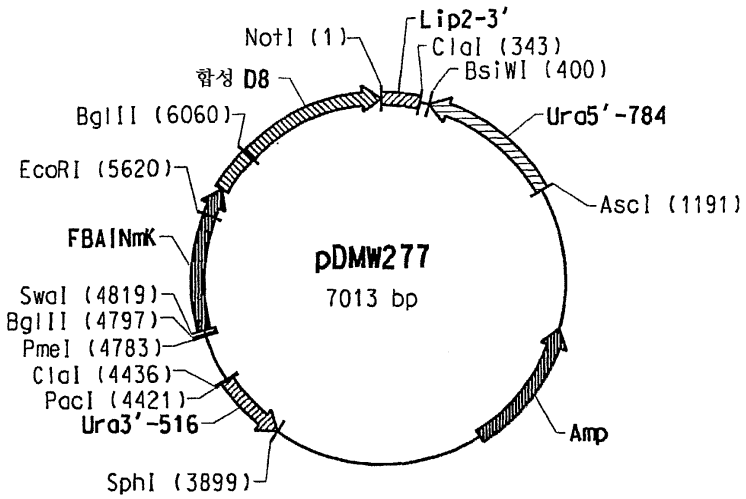
        ************************************ *********************
서열 78 WCFQSVLTVRSLKDRDNQFYRSQYKKEAIGLALHWTLKTLFHLFFMPSILTSLLVFFVSE
서열 80 WCFQSVLTVRSLKDRDNQFYRSQYKKEAIGLALHWTLKALFHLFFMPSILTSLLVFFVSE
서열 73 WCFQCVLTVRSLKDRDNQFYRSQYKKEAIGLALHWTLKALFHLFFMPSILTSLLVFFVSE
서열 74 WCFQCVLTVRSLKDRDNQFYRSQYKKEAIGLALHWTLKALFHLFFMPSILTSLLVFFVSE

        ********************** ***********************************
서열 78 LVGGFGIAIVVFMNHYPLEKIGDSVWDGHGFSVGQIHETMNIRRGIITDWFFGGLNYQIE
서열 80 LVGGFGIAIVVFMNHYPLEKIGDSVWDGHGFSVGQIHETMNIRRGIITDWFFGGLNYQIE
서열 73 LVGGFGIAIVVFMNHYPLEKIGDPVWDGHGFSVGQIHETMNIRRGIITDWFFGGLNYQIE
서열 74 LVGGFGIAIVVFMNHYPLEKIGDPVWDGHGFSVGQIHETMNIRRGIITDWFFGGLNYQIE

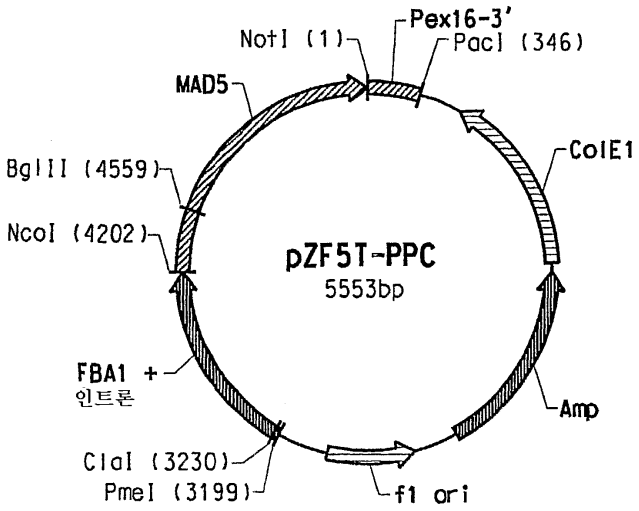
        ************************************************************
서열 78 HHLWPTLPRHNLTAVSYQVEQLCQKHNLPYRNPLPHEGLVILLRYLAVFARMAEKQPAGK
서열 80 HHLWPTLPRHNLTAVSYQVEQLCQKHNLPYRNPLPHEGLVILLRYLAVFARMAEKQPAGK
서열 73 HHLWPTLPRHNLTAVSYQVEQLCQKHNLPYRNPLPHEGLVILLRYLAVFARMAEKQPAGK
서열 74 HHLWPTLPRHNLTAVSYQVEQLCQKHNLPYRNPLPHEGLVILLRYLAVFARMAEKQPAGK

        *
서열 78 AL.
서열 80 AL.
서열 73 A-L
서열 74 A-L
```

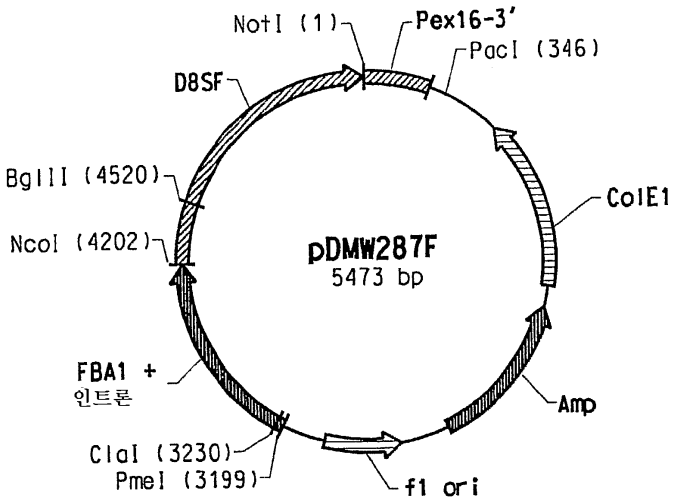
도면13A



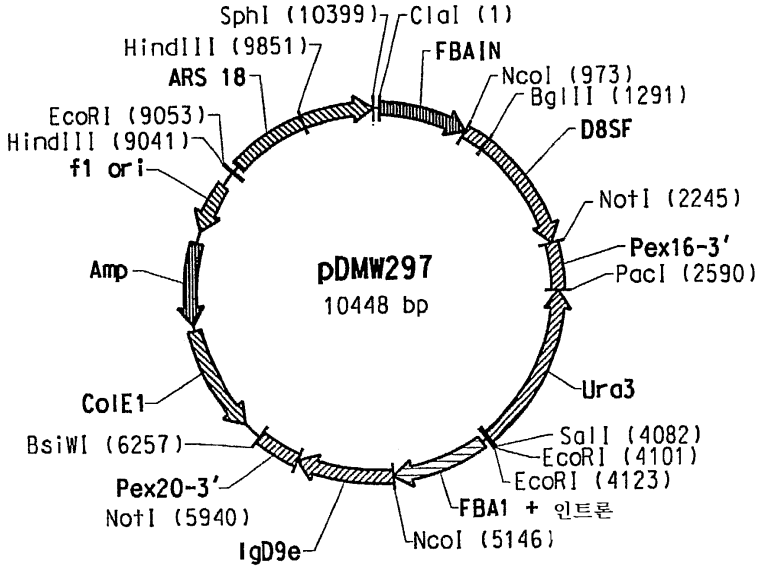
도면13B



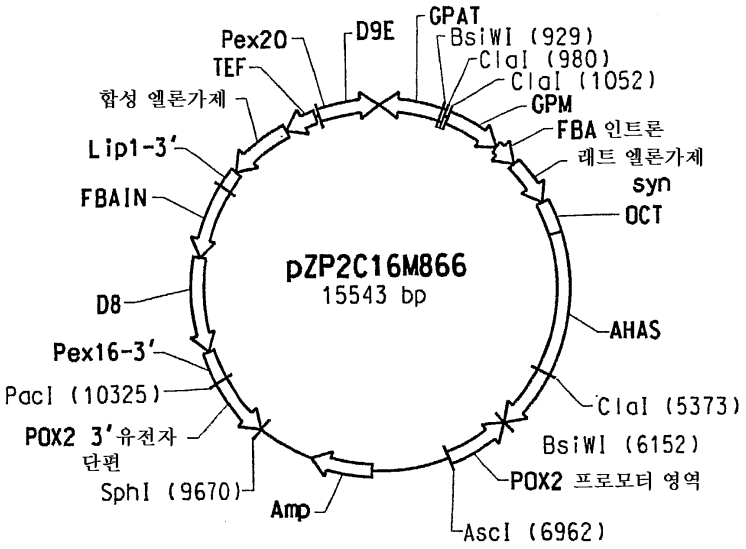
도면13C



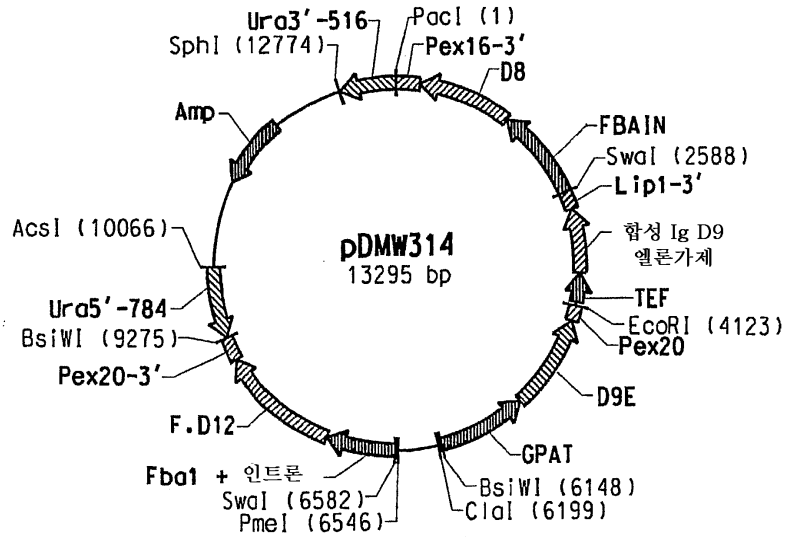
도면13D



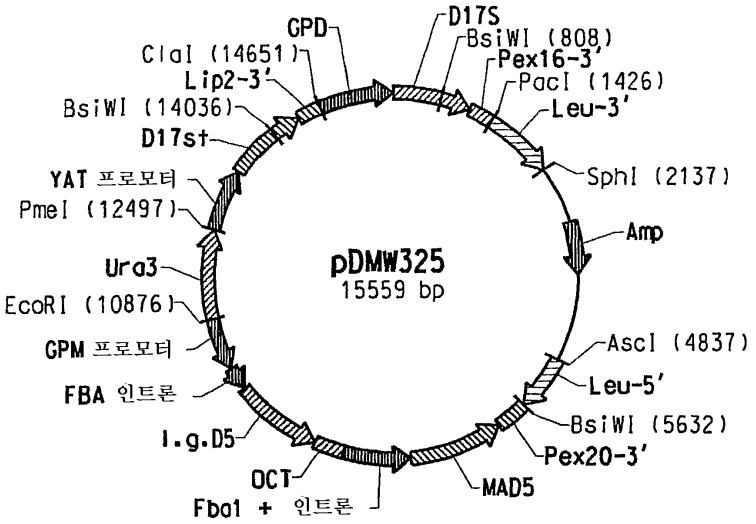
도면14A


pZP2C16M866
15543 bp
Pex20
D9E
GPAT
BsiWI (929)
ClaI (980)
ClaI (1052)
GPM
FBA 인트론
레트 엘론가제
syn
OCT
AHAS
ClaI (5373)
BsiWI (6152)
POX2 프로모터 영역
AscI (6962)
Amp
SphI (9670)
POX2 3'유전자 단편
PacI (10325)
Pex16-3'
D8
FBAIN
Lip1-3'
TEF
합성 엘론가제

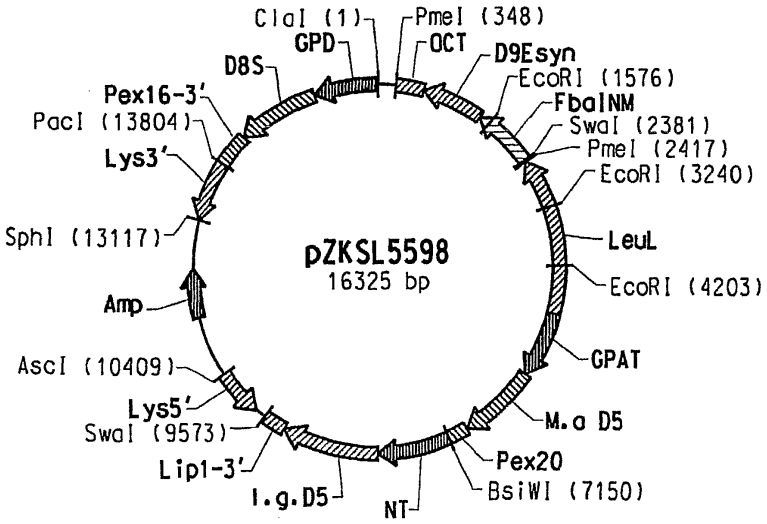
도면14B


pDMW314
13295 bp
Ura3'-516
PacI (1)
Pex16-3'
D8
FBAIN
SwaI (2588)
Lip1-3'
합성 Ig D9 엘론가제
TEF
EcoRI (4123)
Pex20
D9E
GPAT
BsiWI (6148)
ClaI (6199)
Amp
SphI (12774)
AcsI (10066)
Ura5'-784
BsiWI (9275)
Pex20-3'
F.D12
Fba1 + 인트론
SwaI (6582)
PmeI (6546)

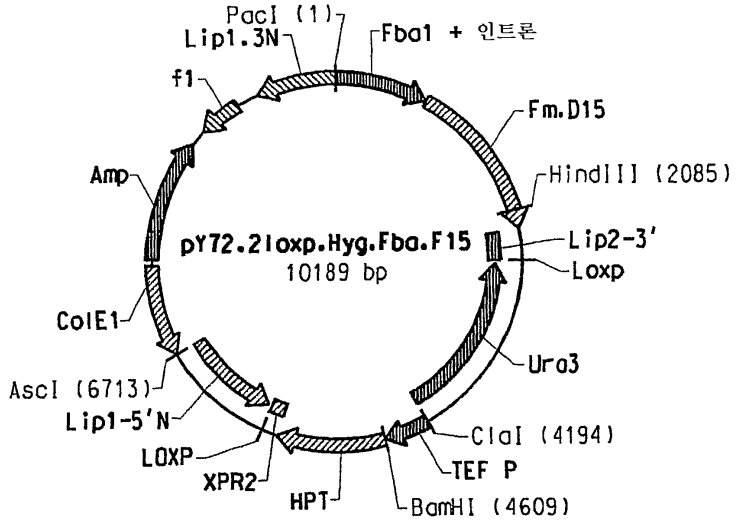
도면14C



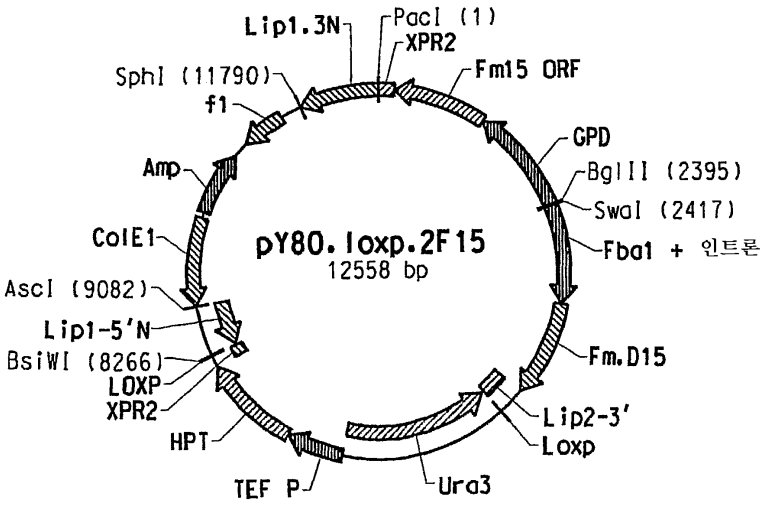
도면14D



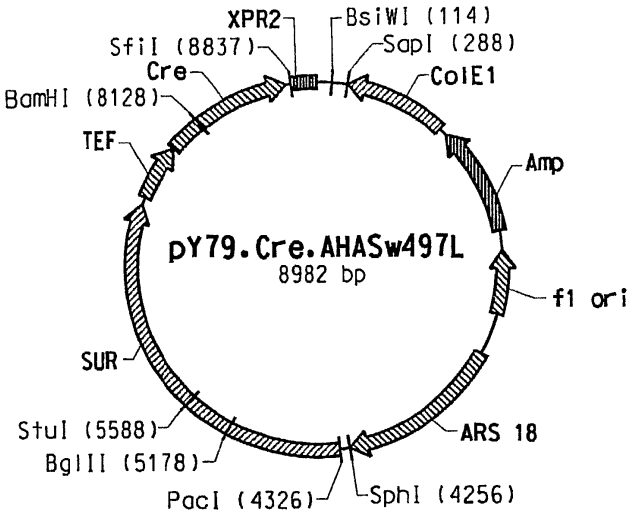
도면15A



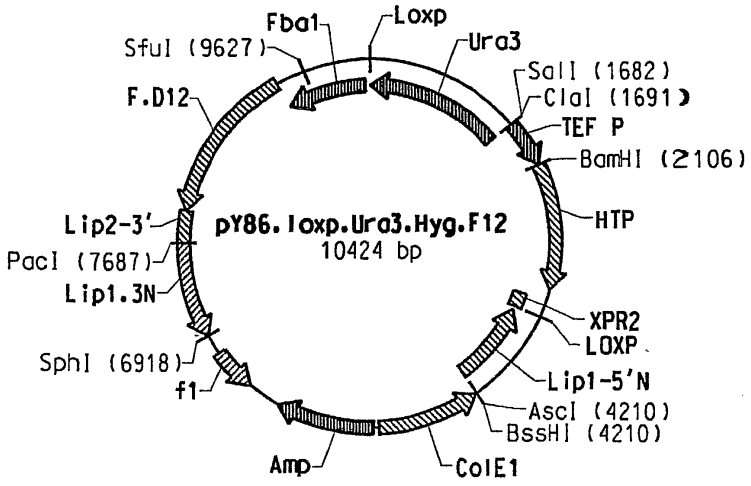
도면15B



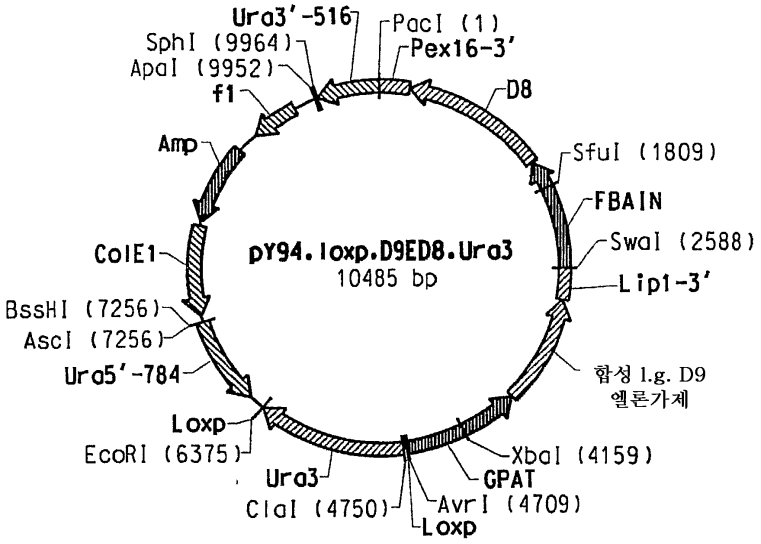
도면15C



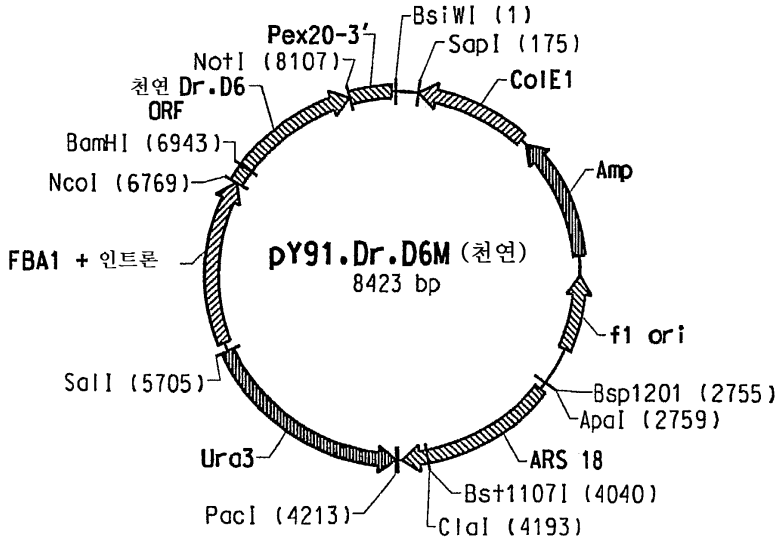
도면15D



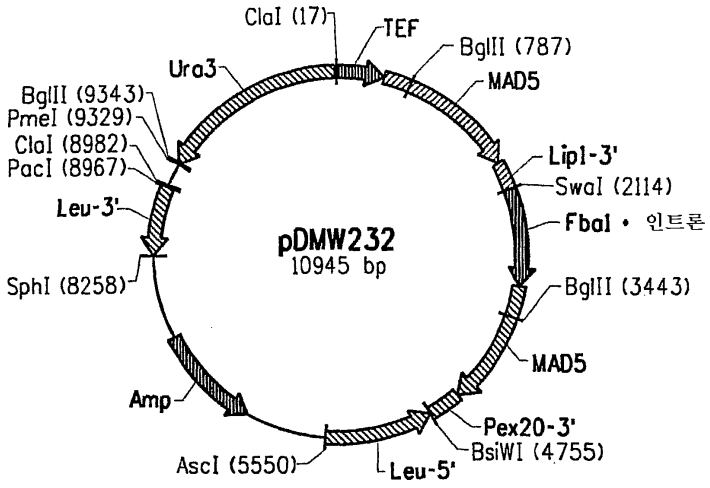
도면16A



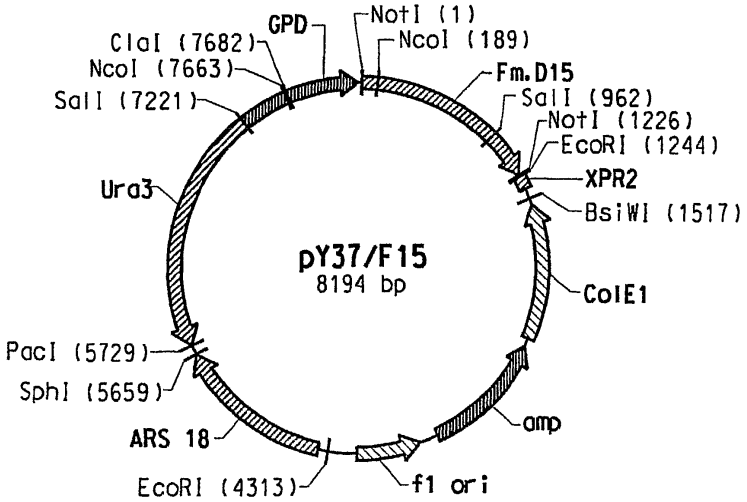
도면16B



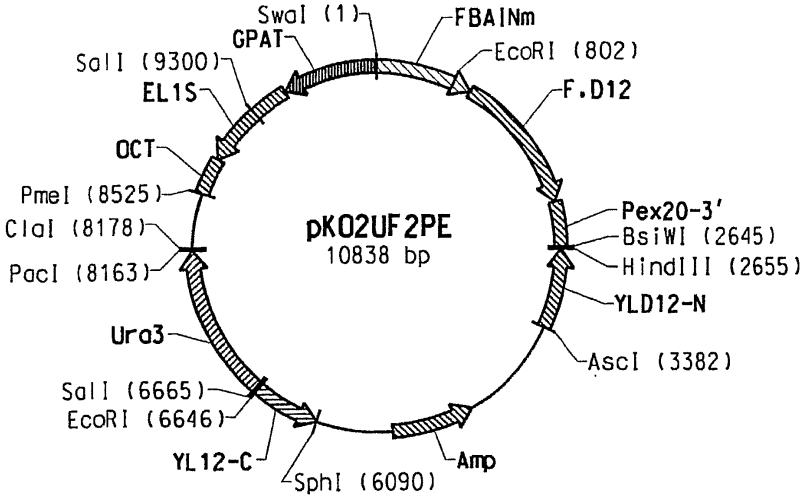
도면16C



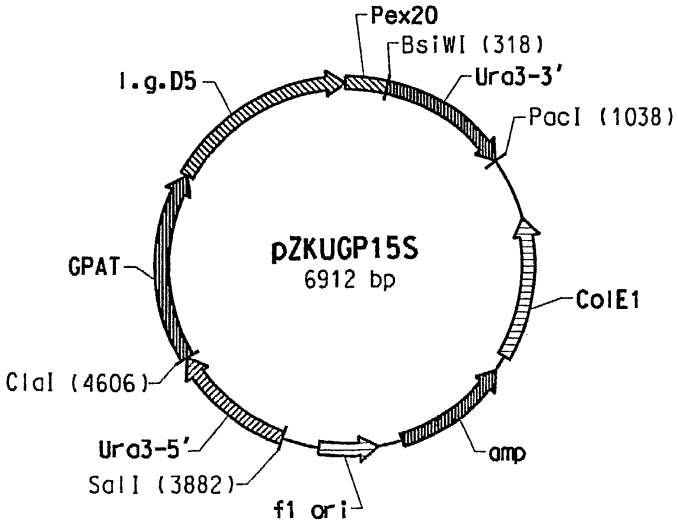
도면16D



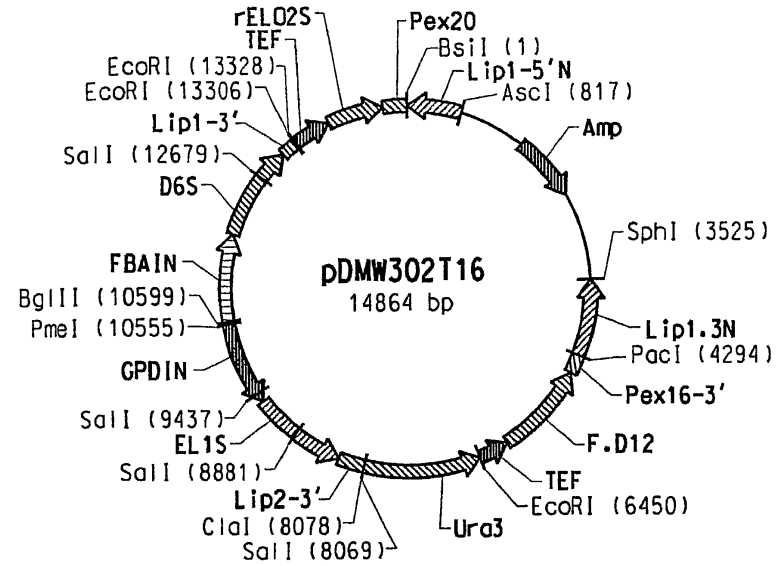
도면17A



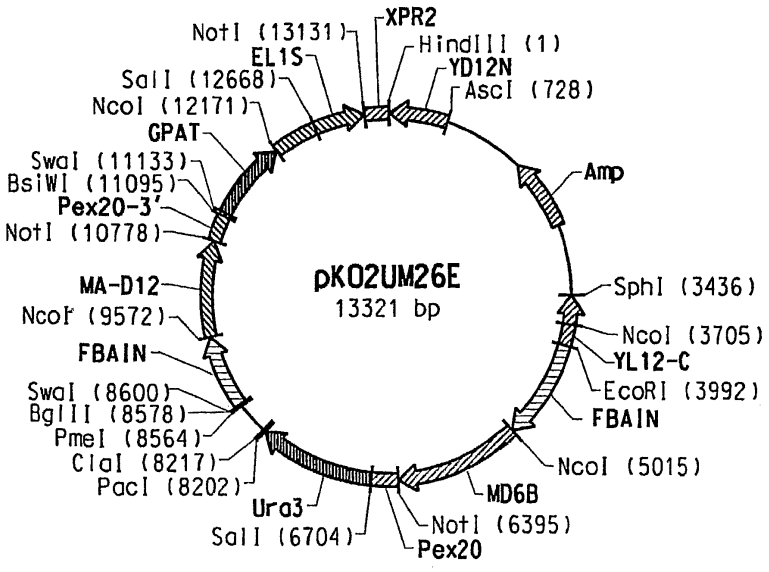
도면17B



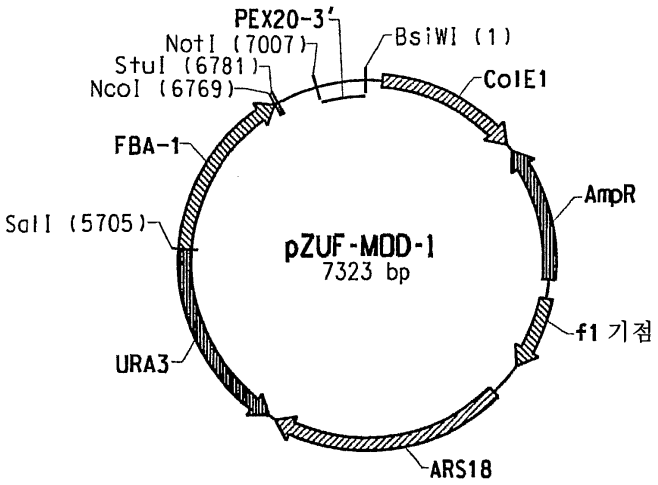
도면17C


rELO2S
TEF
EcoRI (13328)
EcoRI (13306)
Lip1-3'
SalI (12679)
D6S
FBAIN
BglII (10599)
PmeI (10555)
GPDIN
SalI (9437)
EL1S
SalI (8881)
Lip2-3'
ClaI (8078)
SalI (8069)
Pex20
BsiI (1)
Lip1-5'N
AscI (817)
Amp
SphI (3525)
pDMW302T16
14864 bp
Lip1.3N
PacI (4294)
Pex16-3'
F.D12
TEF
EcoRI (6450)
Ura3

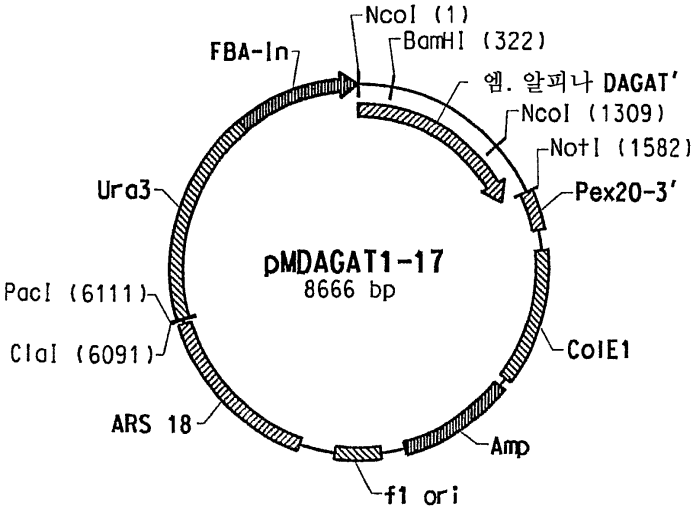
도면17D


NotI (13131)
EL1S
SalI (12668)
NcoI (12171)
GPAT
SwaI (11133)
BsiWI (11095)
Pex20-3'
NotI (10778)
MA-D12
NcoI (9572)
FBAIN
SwaI (8600)
BglII (8578)
PmeI (8564)
ClaI (8217)
PacI (8202)
Ura3
SalI (6704)
XPR2
HindIII (1)
YD12N
AscI (728)
Amp
pKO2UM26E
13321 bp
SphI (3436)
NcoI (3705)
YL12-C
EcoRI (3992)
FBAIN
NcoI (5015)
MD6B
NotI (6395)
Pex20

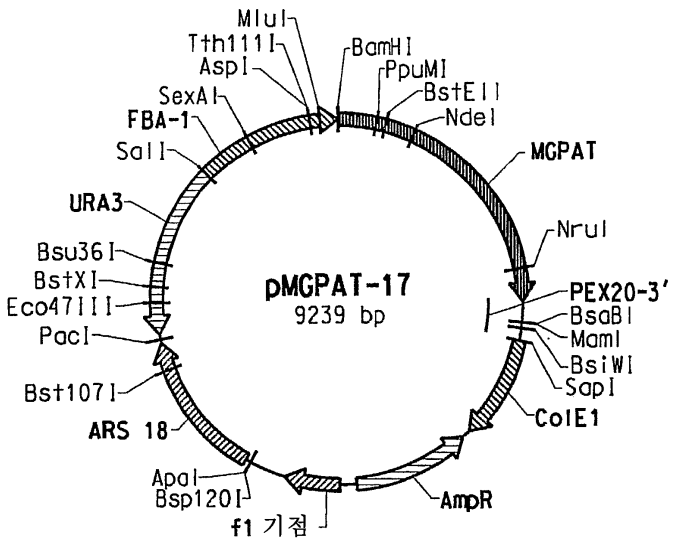
도면18A



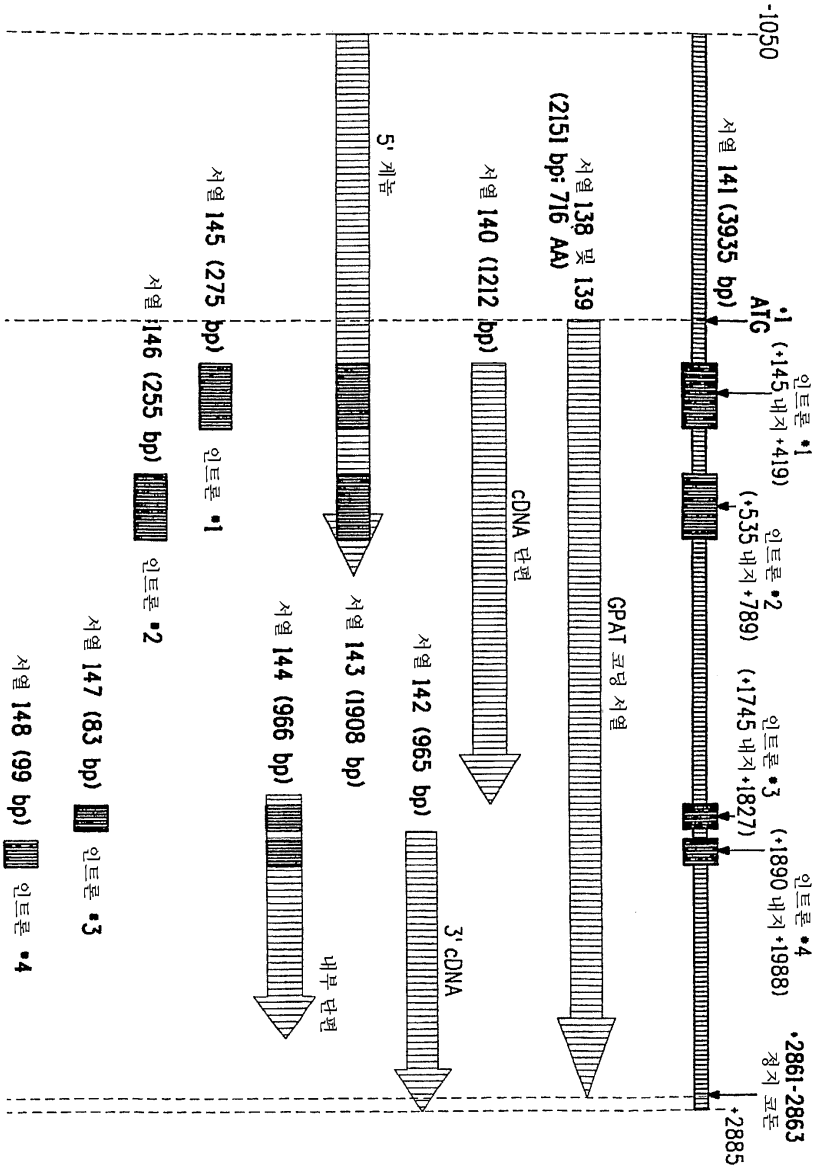
도면18B



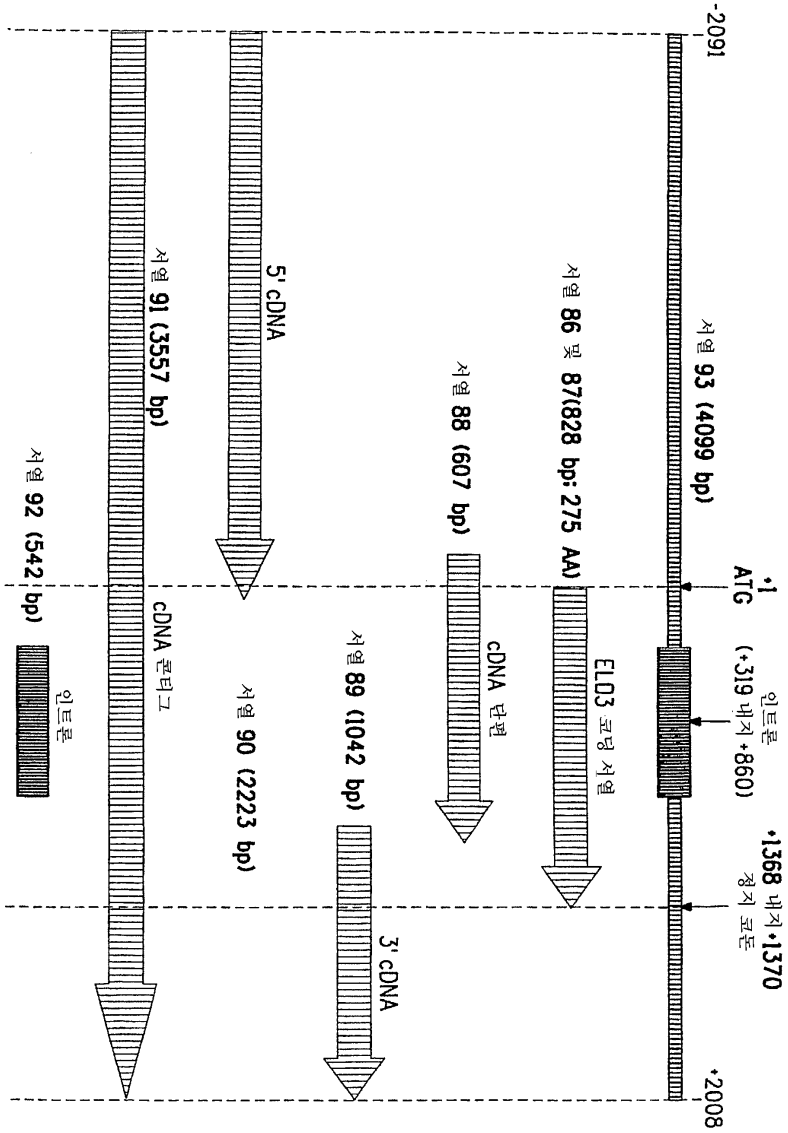
도면18C



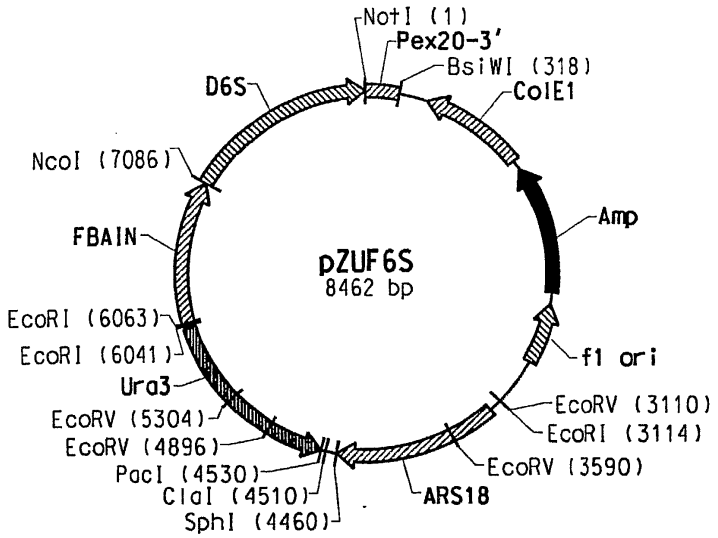
도면19


-1050
서열 141 (3935 bp)
•1
ATG
(+145 내지 +419)
인트론 •1
(+535 내지 +789)
인트론 •2
(+1745 내지 +1827)
인트론 •3
(+1890 내지 +1988)
인트론 •4
•2861-2863
정지 코돈
+2885
서열 138 및 139
(2151 bp; 716 AA)
GPAT 코딩 서열
서열 140 (1212 bp)
cDNA 단편
서열 142 (965 bp)
3' cDNA
5' 게놈
서열 143 (1908 bp)
서열 144 (966 bp)
내부 단편
서열 145 (275 bp)
인트론 •1
서열 146 (255 bp)
인트론 •2
서열 147 (83 bp)
인트론 •3
서열 148 (99 bp)
인트론 •4

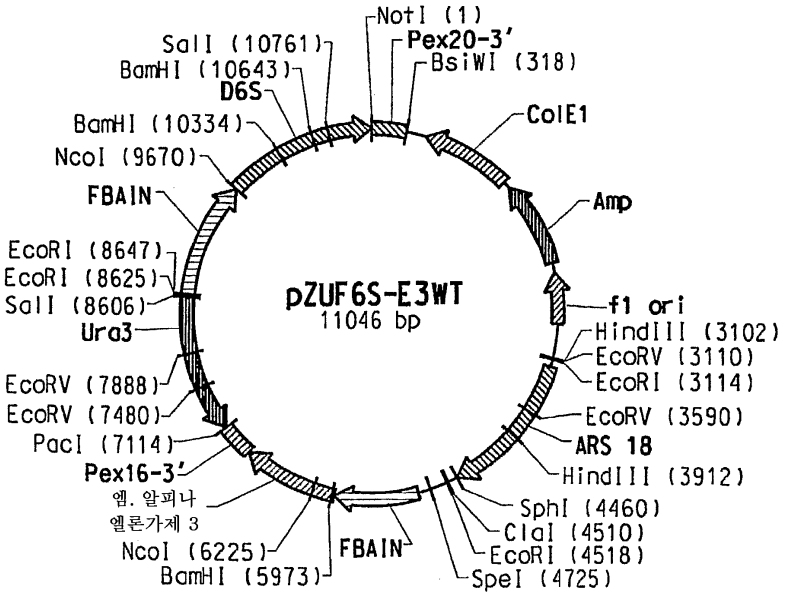
도면20


-2091
서열 91 (3557 bp)
5' cDNA
서열 86 및 87(828 bp; 275 AA)
서열 88 (607 bp)
서열 93 (4099 bp)
서열 92 (542 bp)
+1
ATG
cDNA 콘터그
서열 90 (2223 bp)
서열 89 (1042 bp)
cDNA 단편
ELO3 코딩 서열
인트론
(+319 내지 +860)
인트론
+1368 내지 +1370
경계 코돈
3' cDNA
+2008

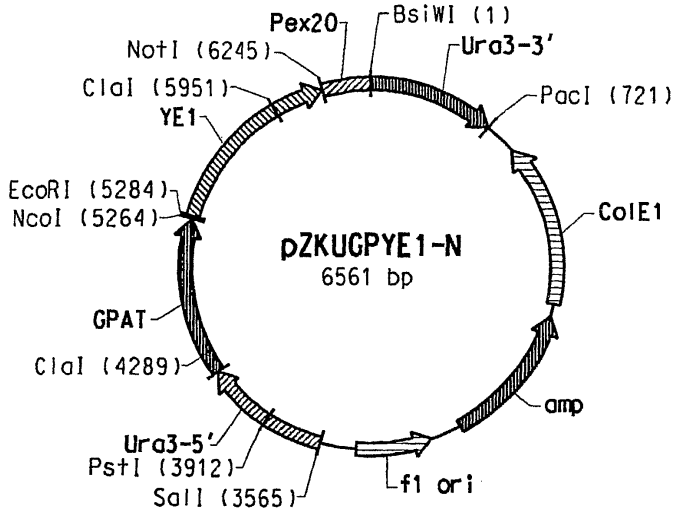
도면21A


NotI (1)
Pex20-3'
BsiWI (318)
ColE1
D6S
NcoI (7086)
Amp
FBAIN
pZUF6S
8462 bp
EcoRI (6063)
EcoRI (6041)
f1 ori
Ura3
EcoRV (3110)
EcoRV (5304)
EcoRI (3114)
EcoRV (4896)
EcoRV (3590)
PacI (4530)
ClaI (4510)
ARS18
SphI (4460)

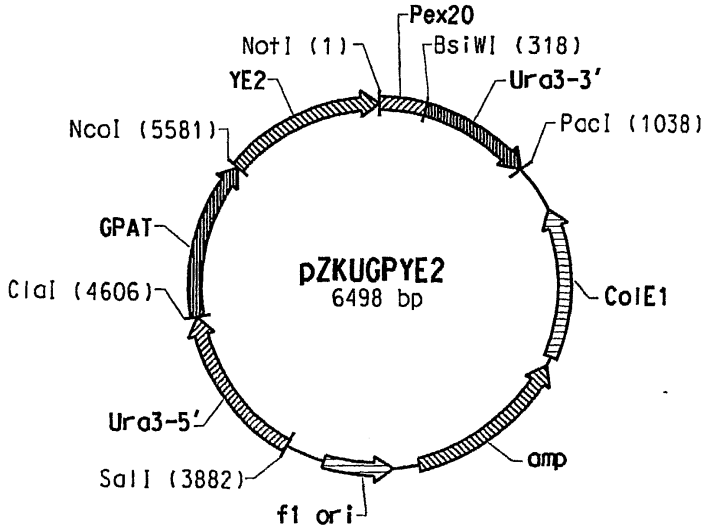
도면21B


NotI (1)
Pex20-3'
BsiWI (318)
ColE1
Amp
f1 ori
HindIII (3102)
EcoRV (3110)
EcoRI (3114)
EcoRV (3590)
ARS 18
HindIII (3912)
SphI (4460)
ClaI (4510)
EcoRI (4518)
SpeI (4725)
FBAIN
NcoI (6225)
BamHI (5973)
Pex16-3'
엠. 알피나
엘론가제 3
Ura3
EcoRV (7888)
EcoRV (7480)
PacI (7114)
EcoRI (8647)
EcoRI (8625)
SalI (8606)
FBAIN
NcoI (9670)
BamHI (10334)
D6S
BamHI (10761)
SalI (10761)
pZUF6S-E3WT
11046 bp

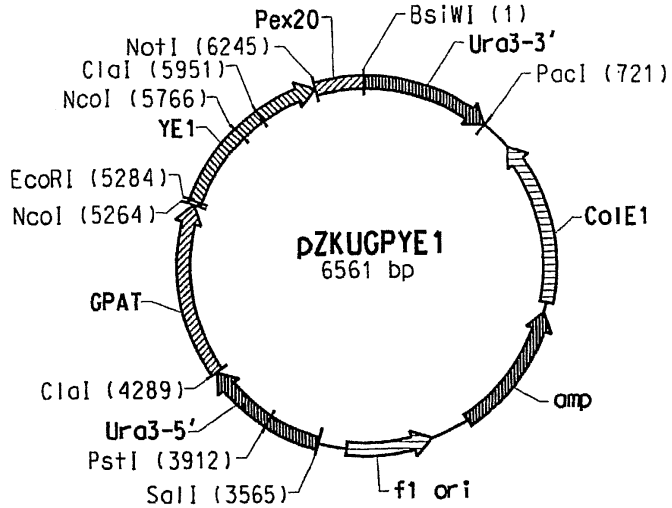
도면21C


Pex20
BsiWI (1)
Ura3-3'
PacI (721)
ColE1
amp
f1 ori
SalI (3565)
PstI (3912)
Ura3-5'
ClaI (4289)
GPAT
EcoRI (5284)
NcoI (5264)
YE1
ClaI (5951)
NotI (6245)
pZKUGPYE1-N
6561 bp

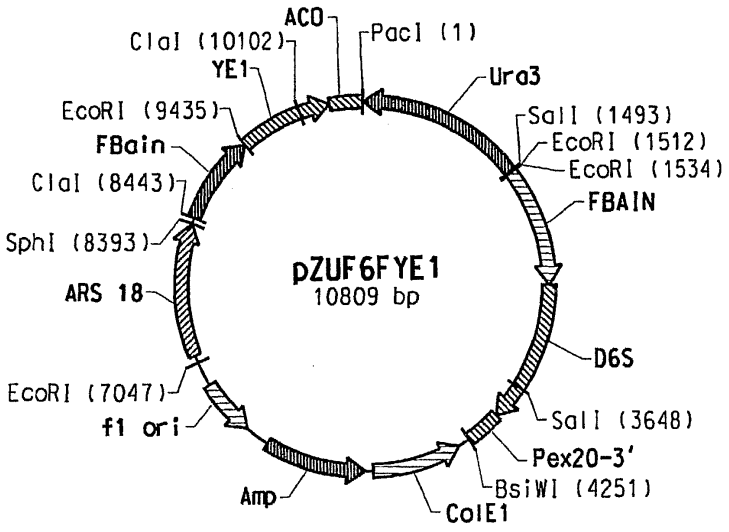
도면21D



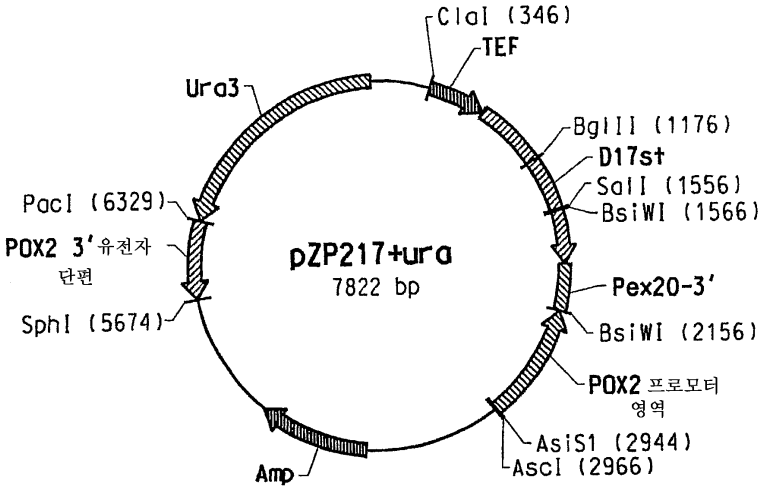
도면22A



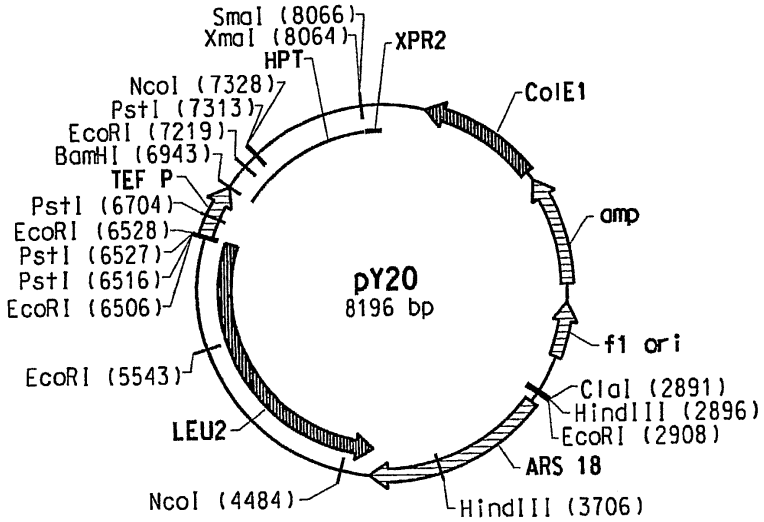
도면22B



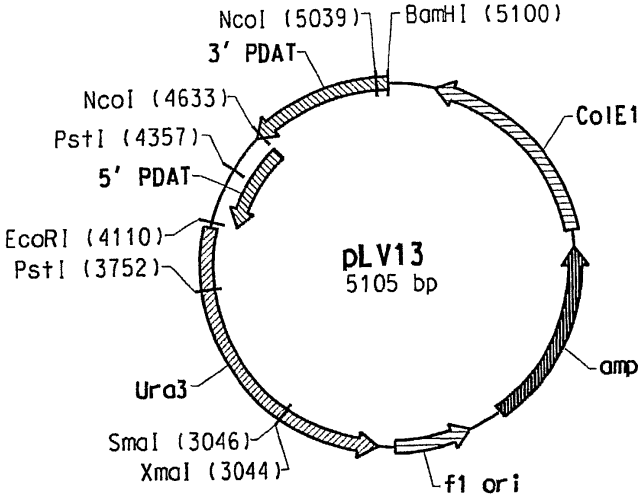
도면22C



도면22D



도면22E



서열목록

SEQUENCE LISTING

<110> E.I. du Pont de Nemours and Company

<120> DOCOSAHEXAENOIC ACID PRODUCING STRAINS OF YARROWIA LIPOLYTICA

<130> CL3160 PCT

<140> US 60/624,812
<141> 2004-11-04

<160> 467

<170> PatentIn version 3.3

<210> 1
<211> 1374
<212> DNA
<213> Mortierella alpina AF465281

<220>
<221> misc_feature
<223> delta-6 desaturase

<400> 1
atggctgctg ctcccagtgt gaggacgttt actcgggccg aggttttgaa tgccgaggct 60

ctgaatgagg gcaagaagga tgccgaggca cccttcttga tgatcatcga caacaaggtg 120

tacgatgtcc gcgagttcgt ccctgatcat cccggtggaa gtgtgattct cacgcacgtt 180

ggcaaggacg gcactgacgt ctttgacact tttcaccccg aggctgcttg ggagactctt 240

gccaactttt acgttggtga tattgacgag agcgaccgcg atatcaagaa tgatgacttt 300

gcggccgagg tccgcaagct gcgtaccttg ttccagtctc ttggttacta cgattcttcc 360

aaggcatact acgccttcaa ggtctcgttc aacctctgca tctggggttt gtcgacggtc 420

attgtggcca agtggggcca gacctcgacc ctcgccaacg tgctctcggc tgcgcttttg 480

ggtctgttct ggcagcagtg cggatggttg gctcacgact ttttgcatca ccaggtcttc 540

caggaccgtt tctggggtga tcttttcggc gccttcttgg gaggtgtctg ccagggcttc 600

tcgtcctcgt ggtggaagga caagcacaac actcaccacg ccgcccccaa cgtccacggc 660

gaggatcccg acattgacac ccaccctctg ttgacctgga gtgagcatgc gttggagatg 720

ttctcggatg tcccagatga ggagctgacc cgcatgtggt cgcgtttcat ggtcctgaac 780

cagacctggt tttacttccc cattctctcg tttgcccgtc tctcctggtg cctccagtcc 840

attctctttg tgctgcctaa cggtcaggcc cacaagccct cgggcgcgcg tgtgcccatc 900

tcgttggtcg agcagctgtc gcttgcgatg cactggacct ggtacctcgc caccatgttc 960

ctgttcatca aggatcccgt caacatgctg gtgtactttt tggtgtcgca ggcggtgtgc 1020

ggaaacttgt tggcgatcgt gttctcgctc aaccacaacg gtatgcctgt gatctcgaag 1080

gaggaggcgg tcgatatgga tttcttcacg aagcagatca tcacgggtcg tgatgtccac 1140

ccgggtctat ttgccaactg gttcacgggt ggattgaact atcagatcga gcaccacttg 1200

ttcccttcga tgcctcgcca caacttttca aagatccagc ctgctgtcga gaccctgtgc 1260

aaaaagtaca atgtccgata ccacaccacc ggtatgatcg agggaactgc agaggtcttt 1320

agccgtctga acgaggtctc caaggctacc tccaagatgg gtaaggcgca gtaa 1374

<210> 2
<211> 457
<212> PRT
<213> Mortierella alpina AF465281

<400> 2

Met Ala Ala Ala Pro Ser Val Arg Thr Phe Thr Arg Ala Glu Val Leu
1 5 10 15

Asn Ala Glu Ala Leu Asn Glu Gly Lys Lys Asp Ala Glu Ala Pro Phe
20 25 30

Leu Met Ile Ile Asp Asn Lys Val Tyr Asp Val Arg Glu Phe Val Pro
35 40 45

Asp His Pro Gly Gly Ser Val Ile Leu Thr His Val Gly Lys Asp Gly
50 55 60

Thr Asp Val Phe Asp Thr Phe His Pro Glu Ala Ala Trp Glu Thr Leu
65 70 75 80

Ala Asn Phe Tyr Val Gly Asp Ile Asp Glu Ser Asp Arg Asp Ile Lys
85 90 95

```
Asn Asp Asp Phe Ala Ala Glu Val Arg Lys Leu Arg Thr Leu Phe Gln
            100                 105                 110

Ser Leu Gly Tyr Tyr Asp Ser Ser Lys Ala Tyr Tyr Ala Phe Lys Val
        115                 120                 125

Ser Phe Asn Leu Cys Ile Trp Gly Leu Ser Thr Val Ile Val Ala Lys
    130                 135                 140

Trp Gly Gln Thr Ser Thr Leu Ala Asn Val Leu Ser Ala Ala Leu Leu
145                 150                 155                 160

Gly Leu Phe Trp Gln Gln Cys Gly Trp Leu Ala His Asp Phe Leu His
                165                 170                 175

His Gln Val Phe Gln Asp Arg Phe Trp Gly Asp Leu Phe Gly Ala Phe
            180                 185                 190

Leu Gly Gly Val Cys Gln Gly Phe Ser Ser Ser Trp Trp Lys Asp Lys
        195                 200                 205

His Asn Thr His His Ala Ala Pro Asn Val His Gly Glu Asp Pro Asp
    210                 215                 220

Ile Asp Thr His Pro Leu Leu Thr Trp Ser Glu His Ala Leu Glu Met
225                 230                 235                 240

Phe Ser Asp Val Pro Asp Glu Glu Leu Thr Arg Met Trp Ser Arg Phe
                245                 250                 255

Met Val Leu Asn Gln Thr Trp Phe Tyr Phe Pro Ile Leu Ser Phe Ala
            260                 265                 270

Arg Leu Ser Trp Cys Leu Gln Ser Ile Leu Phe Val Leu Pro Asn Gly
        275                 280                 285

Gln Ala His Lys Pro Ser Gly Ala Arg Val Pro Ile Ser Leu Val Glu
    290                 295                 300
```

```
Gln Leu Ser Leu Ala Met His Trp Thr Trp Tyr Leu Ala Thr Met Phe
305                 310                 315                 320

Leu Phe Ile Lys Asp Pro Val Asn Met Leu Val Tyr Phe Leu Val Ser
                325                 330                 335

Gln Ala Val Cys Gly Asn Leu Leu Ala Ile Val Phe Ser Leu Asn His
            340                 345                 350

Asn Gly Met Pro Val Ile Ser Lys Glu Glu Ala Val Asp Met Asp Phe
        355                 360                 365

Phe Thr Lys Gln Ile Ile Thr Gly Arg Asp Val His Pro Gly Leu Phe
    370                 375                 380

Ala Asn Trp Phe Thr Gly Gly Leu Asn Tyr Gln Ile Glu His His Leu
385                 390                 395                 400

Phe Pro Ser Met Pro Arg His Asn Phe Ser Lys Ile Gln Pro Ala Val
                405                 410                 415

Glu Thr Leu Cys Lys Lys Tyr Asn Val Arg Tyr His Thr Thr Gly Met
            420                 425                 430

Ile Glu Gly Thr Ala Glu Val Phe Ser Arg Leu Asn Glu Val Ser Lys
        435                 440                 445

Ala Thr Ser Lys Met Gly Lys Ala Gln
    450                 455
```

<210> 3
<211> 1374
<212> DNA
<213> Mortierella alpina

<220>
<221> misc_feature
<223> synthetic delta-6 desaturase (codon-optimized)

<400> 3

atggctgccg ctccctctgt gcgaaccttt acccgagccg aggttctgaa cgctgaggct 60

ctgaacgagg gcaagaagga cgctgaggct cccttcctga tgatcatcga caacaaggtg 120

tacgacgtcc gagagttcgt ccctgaccat cctggaggct ccgtgattct cacccacgtt 180

ggcaaggacg gcaccgacgt ctttgacacc tttcatcccg aggctgcttg ggagactctc 240

gccaacttct acgttggaga cattgacgag tccgaccgag acatcaagaa cgatgacttt 300

gccgctgagg tccgaaagct gcgaaccctg ttccagtctc tcggctacta cgactcctct 360

aaggcctact acgccttcaa ggtctccttc aacctctgca tctggggact gtccaccgtc 420

attgtggcca agtggggtca gacctccacc ctcgccaacg tgctctctgc tgccctgctc 480

ggcctgttct ggcagcagtg cggatggctg gctcacgact ttctgcacca ccaggtcttc 540

caggaccgat tctggggtga tctcttcgga gccttcctgg gaggtgtctg ccagggcttc 600

tcctcttcct ggtggaagga caagcacaac actcaccatg ccgctcccaa cgtgcatggc 660

gaggatcctg acattgacac ccaccctctc ctgacctggt ccgagcacgc tctggagatg 720

ttctccgacg tccccgatga ggagctgacc cgaatgtggt ctcgattcat ggtcctgaac 780

cagacctggt tctacttccc cattctctcc ttcgctcgac tgtcttggtg cctccagtcc 840

attctctttg tgctgcccaa cggtcaggct cacaagccct ccggagctcg agtgcccatc 900

tccctggtcg agcagctgtc cctcgccatg cactggacct ggtacctcgc taccatgttc 960

ctgttcatca aggatcctgt caacatgctc gtgtacttcc tggtgtctca ggctgtgtgc 1020

ggaaacctgc tcgccatcgt gttctccctc aaccacaacg gtatgcctgt gatctccaag 1080

gaggaggctg tcgacatgga tttctttacc aagcagatca tcactggtcg agatgtccat 1140

cctggactgt tcgccaactg gttcaccggt ggcctgaact accagatcga gcatcacctg 1200

ttcccttcca tgcctcgaca caacttctcc aagatccagc ctgccgtcga gaccctgtgc 1260

aagaagtaca acgtccgata ccacaccact ggtatgatcg agggaactgc cgaggtcttc 1320

tcccgactga acgaggtctc caaggccacc tccaagatgg gcaaggctca gtaa 1374

<210> 4
<211> 1521
<212> DNA
<213> Mortierella alpina (GenBank Accession No. AB070555)

<220>
<221> misc_feature
<223> delta-6 desaturase

<400> 4
acaacgtgct cttgtcctca gaagactgtt gtgctcttcc ttccaccacc ccaagcactc 60

tctcacccca agcactgcct gccatgacca ccagcgaccc atctgtcaga gcgttcacac 120

gctcagaagt gttgcacgcc gatgccttga acgagggcaa aaagaacgcc gaggcaccgt 180

ttctcatgat catcgacaac aaggtctacg atgtgcgcga gtttatcccc gaccatcctg 240

gtgggagcgt cattttgacc cacgtaggca aggacggcac cgacgttttc gagaccttcc 300

atcctgaggc tgcttgggag acgctcgcca atttttatgt cggtgacatt gtagaatccg 360

atcgcgccat cgagaacgac gagtttgcag ctgaggttcg taagctgcgg acattgtttt 420

attctttggg ctactacgac tcatccaagg tttactacgc cttcaaggtc tcgttcaacc 480

tctgcatctg gggcctgtct gcattcattg ttgccaaatg gggccagacc tcgaccctcg 540

caaacgtgat atcagcctca ctcctgggtg tcttttggca acagtgcggt tggctcgccc 600

atgatttctt gcaccatcag gtctttcacg atcgattctg gggcgatctg ttcggtgcat 660

ttctcggcgg agtctgtcaa ggtttctcct cgtcctggtg gaaggacaaa cacaacaccc 720

accacgcggc gcccaatgtc catggagagg atcccgatat cgacacacat ccgcttttga 780

cgtggagtga gcatgcgctc gagatgtttt cggatgtgcc cgatgaggag cttacccaaa 840

tgtggtcccg gtttatggtt ctgaaccagg cctggtttta ctttcccatt ctgtcatttg 900

cccgcctgtc ctggtgcatc cagtcgattc tttttgtgct accgaacgga caggcacaca 960

aacctgcggg ggctcgggtt cccatctcgc tggtggagca attgtcgttg gcgatgcact 1020

ggacctggta cctggcaacc atgttcctgt tcatcaagga tcccgtcaac atgatggtgt 1080

atttcttggt ctcgcaagct gtctgcggca acctgttagc gattgtgttc tcgctgaacc 1140

ataacggtat gcctgtgatc tcgcaggagg aagcggtcga gatggatttc ttcacaaagc 1200

agatcattac gggtcgtgat gtctacccgg gttggtttgc agactggttc acgggtggat 1260

tgaactatca gattgaacac catctgttcc cgtcgatgcc tcgacaccat ttctcaaaga 1320

tccagcccgc ggttgaatcg ctgtgcaaga agtacggggt ccgataccat acgacgggga 1380

tgattgctgg caccgcagag gtcttttcgc gactgaacga ggtgtcccag gctgcaagca 1440

agctcggcaa gtctgcttga gtctttcatg tcctcaagtt gattctagat acttattttc 1500

gcagacttct atcgataaat t 1521

<210> 5
<211> 458
<212> PRT
<213> Mortierella alpina (GenBank Accession No. AB070555)

<400> 5

Met Thr Thr Ser Asp Pro Ser Val Arg Ala Phe Thr Arg Ser Glu Val
1 5 10 15

Leu His Ala Asp Ala Leu Asn Glu Gly Lys Lys Asn Ala Glu Ala Pro
20 25 30

Phe Leu Met Ile Ile Asp Asn Lys Val Tyr Asp Val Arg Glu Phe Ile
35 40 45

Pro Asp His Pro Gly Gly Ser Val Ile Leu Thr His Val Gly Lys Asp
50 55 60

Gly Thr Asp Val Phe Glu Thr Phe His Pro Glu Ala Ala Trp Glu Thr
65 70 75 80

Leu Ala Asn Phe Tyr Val Gly Asp Ile Val Glu Ser Asp Arg Ala Ile
85 90 95

Glu Asn Asp Glu Phe Ala Ala Glu Val Arg Lys Leu Arg Thr Leu Phe
100 105 110

Tyr Ser Leu Gly Tyr Tyr Asp Ser Ser Lys Val Tyr Tyr Ala Phe Lys
115 120 125

Val Ser Phe Asn Leu Cys Ile Trp Gly Leu Ser Ala Phe Ile Val Ala
130 135 140

Lys Trp Gly Gln Thr Ser Thr Leu Ala Asn Val Ile Ser Ala Ser Leu
145 150 155 160

```
Leu Gly Val Phe Trp Gln Gln Cys Gly Trp Leu Ala His Asp Phe Leu
               165                 170                 175

His His Gln Val Phe His Asp Arg Phe Trp Gly Asp Leu Phe Gly Ala
            180                 185                 190

Phe Leu Gly Gly Val Cys Gln Gly Phe Ser Ser Ser Trp Trp Lys Asp
        195                 200                 205

Lys His Asn Thr His His Ala Ala Pro Asn Val His Gly Glu Asp Pro
    210                 215                 220

Asp Ile Asp Thr His Pro Leu Leu Thr Trp Ser Glu His Ala Leu Glu
225                 230                 235                 240

Met Phe Ser Asp Val Pro Asp Glu Glu Leu Thr Gln Met Trp Ser Arg
               245                 250                 255

Phe Met Val Leu Asn Gln Ala Trp Phe Tyr Phe Pro Ile Leu Ser Phe
            260                 265                 270

Ala Arg Leu Ser Trp Cys Ile Gln Ser Ile Leu Phe Val Leu Pro Asn
        275                 280                 285

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

Glu Gln Leu Ser Leu Ala Met His Trp Thr Trp Tyr Leu Ala Thr Met
305                 310                 315                 320

Phe Leu Phe Ile Lys Asp Pro Val Asn Met Met Val Tyr Phe Leu Val
               325                 330                 335

Ser Gln Ala Val Cys Gly Asn Leu Leu Ala Ile Val Phe Ser Leu Asn
            340                 345                 350

His Asn Gly Met Pro Val Ile Ser Gln Glu Glu Ala Val Glu Met Asp
        355                 360                 365
```

Phe Phe Thr Lys Gln Ile Ile Thr Gly Arg Asp Val Tyr Pro Gly Trp
370 375 380

Phe Ala Asp Trp Phe Thr Gly Gly Leu Asn Tyr Gln Ile Glu His His
385 390 395 400

Leu Phe Pro Ser Met Pro Arg His His Phe Ser Lys Ile Gln Pro Ala
405 410 415

Val Glu Ser Leu Cys Lys Lys Tyr Gly Val Arg Tyr His Thr Thr Gly
420 425 430

Met Ile Ala Gly Thr Ala Glu Val Phe Ser Arg Leu Asn Glu Val Ser
435 440 445

Gln Ala Ala Ser Lys Leu Gly Lys Ser Ala
450 455

<210> 6
<211> 1341
<212> DNA
<213> Mortierella alpina AF067654

<220>
<221> misc_feature
<223> delta-5 desaturase

<400> 6
atgggaacgg accaaggaaa aaccttcacc tgggaagagc tggcggccca taacaccaag 60

gacgacctac tcttggccat ccgcggcagg gtgtacgatg tcacaaagtt cttgagccgc 120

catcctggtg gagtggacac tctcctgctc ggagctggcc gagatgttac tccggtcttt 180

gagatgtatc acgcgtttgg ggctgcagat gccattatga agaagtacta tgtcggtaca 240

ctggtctcga atgagctgcc catcttcccg gagccaacgg tgttccacaa aaccatcaag 300

acgagagtcg agggctactt tacggatcgg aacattgatc ccaagaatag accagagatc 360

tggggacgat acgctcttat ctttggatcc ttgatcgctt cctactacgc gcagctcttt 420

gtgcctttcg ttgtcgaacg cacatggctt caggtggtgt ttgcaatcat catgggattt 480

gcgtgcgcac aagtcggact caaccctctt catgatgcgt ctcacttttc agtgacccac 540

aaccccactg tctggaagat tctgggagcc acgcacgact ttttcaacgg agcatcgtac 600

ctggtgtgga tgtaccaaca tatgctcggc catcacccct acaccaacat tgctggagca 660

gatcccgacg tgtcgacgtc tgagcccgat gttcgtcgta tcaagcccaa ccaaaagtgg 720

tttgtcaacc acatcaacca gcacatgttt gttcctttcc tgtacggact gctggcgttc 780

aaggtgcgca ttcaggacat caacattttg tactttgtca agaccaatga cgctattcgt 840

gtcaatccca tctcgacatg gcacactgtg atgttctggg gcggcaaggc tttctttgtc 900

tggtatcgcc tgattgttcc cctgcagtat ctgcccctgg gcaaggtgct gctcttgttc 960

acggtcgcgg acatggtgtc gtcttactgg ctggcgctga ccttccaggc gaaccacgtt 1020

gttgaggaag ttcagtggcc gttgcctgac gagaacggga tcatccaaaa ggactgggca 1080

gctatgcagg tcgagactac gcaggattac gcacacgatt cgcacctctg gaccagcatc 1140

actggcagct tgaactacca ggctgtgcac catctgttcc ccaacgtgtc gcagcaccat 1200

tatcccgata ttctggccat catcaagaac acctgcagcg agtacaaggt tccatacctt 1260

gtcaaggata cgttttggca agcatttgct tcacatttgg agcacttgcg tgttcttgga 1320

ctccgtccca aggaagagta g 1341

<210> 7
<211> 446
<212> PRT
<213> Mortierella alpina AF067654

<400> 7

Met Gly Thr Asp Gln Gly Lys Thr Phe Thr Trp Glu Glu Leu Ala Ala
1 5 10 15

His Asn Thr Lys Asp Asp Leu Leu Leu Ala Ile Arg Gly Arg Val Tyr
20 25 30

Asp Val Thr Lys Phe Leu Ser Arg His Pro Gly Gly Val Asp Thr Leu
35 40 45

Leu Leu Gly Ala Gly Arg Asp Val Thr Pro Val Phe Glu Met Tyr His
50 55 60

```
Ala Phe Gly Ala Ala Asp Ala Ile Met Lys Lys Tyr Tyr Val Gly Thr
65                  70                  75                  80

Leu Val Ser Asn Glu Leu Pro Ile Phe Pro Glu Pro Thr Val Phe His
                85                  90                  95

Lys Thr Ile Lys Thr Arg Val Glu Gly Tyr Phe Thr Asp Arg Asn Ile
            100                 105                 110

Asp Pro Lys Asn Arg Pro Glu Ile Trp Gly Arg Tyr Ala Leu Ile Phe
        115                 120                 125

Gly Ser Leu Ile Ala Ser Tyr Tyr Ala Gln Leu Phe Val Pro Phe Val
    130                 135                 140

Val Glu Arg Thr Trp Leu Gln Val Val Phe Ala Ile Ile Met Gly Phe
145                 150                 155                 160

Ala Cys Ala Gln Val Gly Leu Asn Pro Leu His Asp Ala Ser His Phe
                165                 170                 175

Ser Val Thr His Asn Pro Thr Val Trp Lys Ile Leu Gly Ala Thr His
            180                 185                 190

Asp Phe Phe Asn Gly Ala Ser Tyr Leu Val Trp Met Tyr Gln His Met
        195                 200                 205

Leu Gly His His Pro Tyr Thr Asn Ile Ala Gly Ala Asp Pro Asp Val
    210                 215                 220

Ser Thr Ser Glu Pro Asp Val Arg Arg Ile Lys Pro Asn Gln Lys Trp
225                 230                 235                 240

Phe Val Asn His Ile Asn Gln His Met Phe Val Pro Phe Leu Tyr Gly
                245                 250                 255

Leu Leu Ala Phe Lys Val Arg Ile Gln Asp Ile Asn Ile Leu Tyr Phe
            260                 265                 270
```

```
Val Lys Thr Asn Asp Ala Ile Arg Val Asn Pro Ile Ser Thr Trp His
        275                 280                 285

Thr Val Met Phe Trp Gly Gly Lys Ala Phe Phe Val Trp Tyr Arg Leu
    290                 295                 300

Ile Val Pro Leu Gln Tyr Leu Pro Leu Gly Lys Val Leu Leu Leu Phe
305                 310                 315                 320

Thr Val Ala Asp Met Val Ser Ser Tyr Trp Leu Ala Leu Thr Phe Gln
                325                 330                 335

Ala Asn His Val Val Glu Glu Val Gln Trp Pro Leu Pro Asp Glu Asn
            340                 345                 350

Gly Ile Ile Gln Lys Asp Trp Ala Ala Met Gln Val Glu Thr Thr Gln
        355                 360                 365

Asp Tyr Ala His Asp Ser His Leu Trp Thr Ser Ile Thr Gly Ser Leu
    370                 375                 380

Asn Tyr Gln Ala Val His His Leu Phe Pro Asn Val Ser Gln His His
385                 390                 395                 400

Tyr Pro Asp Ile Leu Ala Ile Ile Lys Asn Thr Cys Ser Glu Tyr Lys
                405                 410                 415

Val Pro Tyr Leu Val Lys Asp Thr Phe Trp Gln Ala Phe Ala Ser His
            420                 425                 430

Leu Glu His Leu Arg Val Leu Gly Leu Arg Pro Lys Glu Glu
        435                 440                 445
```

<210> 8
<211> 1329
<212> DNA
<213> Isochrysis galbana CCMP1323

<220>
<221> misc_feature
<223> delta-5 desaturase

<400> 8
atggtggcag gcaaatcagg cgctgcggcg cacgtgactc acagctcgac attgccccgt 60

gagtaccatg gcgcgaccaa cgactcgcgc tctgaggcgg ccgacgtcac cgtctctagc 120

atcgatgctg aaaaggagat gatcatcaac ggccgcgtgt atgacgtgtc gtcatttgtg 180

aagcggcacc caggtggctc ggtgatcaag ttccagctgg gcgccgacgc gagcgacgcg 240

tacaacaact ttcacgtccg ctccaagaag gcggacaaga tgctgtattc gctcccgtcc 300

cggccggccg aggccggcta cgcccaggac gacatctccc gcgactttga gaagctgcgc 360

ctcgagctga aggaggaggg ctacttcgag cccaacctgg tgcacgtgag ctacaggtgt 420

gtggaggttc ttgccatgta ctgggctggc gtccagctca tctggtccgg gtactggttc 480

ctcggcgcga tcgtggccgg cattgcgcag ggccgctgcg gctggctcca gcatgagggt 540

gggcactact cgctcaccgg caacatcaag atcgaccggc atctgcagat ggccatctat 600

gggcttggct gcggcatgtc gggctgctac tggcgcaacc agcacaacaa gcaccacgcc 660

acgccgcaga agctcgggac cgaccccgac ctgcagacga tgccgctggt ggccttccac 720

aagatcgtcg gcgccaaggc gcgaggcaag ggcaaggcgt ggctggcgtg gcaggcgccg 780

ctcttctttg gcgggatcat ctgctcgctc gtctctttcg gctggcagtt cgtgctccac 840

cccaaccacg cgctgcgcgt gcacaatcac ctggagctcg cgtacatggg cctgcggtac 900

gtgctgtggc acctggcctt tggccacctc gggctgctga gctcgctccg cctgtacgcc 960

ttttacgtgg ccgtgggcgg cacctacatc ttcaccaact tcgccgtctc gcacacccac 1020

aaggacgtcg tcccgcccac caagcacatc tcgtgggcac tctactcggc caaccacacg 1080

accaactgct ccgactcgcc ctttgtcaac tggtggatgg cctacctcaa cttccagatc 1140

gagcaccacc tcttcccgtc gatgccgcag tacaaccacc ccaagatcgc cccgcgggtg 1200

cgcgcgctct tcgagaagca cggggtcgag tatgacgtcc ggccatacct ggagtgtttt 1260

cgggtcacgt acgtcaacct gctcgccgta ggcaacccgg agcactccta ccacgagcac 1320

acgcactag 1329

<210> 9
<211> 442
<212> PRT
<213> Isochrysis galbana CCMP1323

<400> 9

Met Val Ala Gly Lys Ser Gly Ala Ala Ala His Val Thr His Ser Ser
1 5 10 15

Thr Leu Pro Arg Glu Tyr His Gly Ala Thr Asn Asp Ser Arg Ser Glu
20 25 30

Ala Ala Asp Val Thr Val Ser Ser Ile Asp Ala Glu Lys Glu Met Ile
35 40 45

Ile Asn Gly Arg Val Tyr Asp Val Ser Ser Phe Val Lys Arg His Pro
50 55 60

Gly Gly Ser Val Ile Lys Phe Gln Leu Gly Ala Asp Ala Ser Asp Ala
65 70 75 80

Tyr Asn Asn Phe His Val Arg Ser Lys Lys Ala Asp Lys Met Leu Tyr
85 90 95

Ser Leu Pro Ser Arg Pro Ala Glu Ala Gly Tyr Ala Gln Asp Asp Ile
100 105 110

Ser Arg Asp Phe Glu Lys Leu Arg Leu Glu Leu Lys Glu Glu Gly Tyr
115 120 125

Phe Glu Pro Asn Leu Val His Val Ser Tyr Arg Cys Val Glu Val Leu
130 135 140

Ala Met Tyr Trp Ala Gly Val Gln Leu Ile Trp Ser Gly Tyr Trp Phe
145 150 155 160

Leu Gly Ala Ile Val Ala Gly Ile Ala Gln Gly Arg Cys Gly Trp Leu
165 170 175

Gln His Glu Gly Gly His Tyr Ser Leu Thr Gly Asn Ile Lys Ile Asp
180 185 190

```
Arg His Leu Gln Met Ala Ile Tyr Gly Leu Gly Cys Gly Met Ser Gly
        195                 200                 205

Cys Tyr Trp Arg Asn Gln His Asn Lys His His Ala Thr Pro Gln Lys
    210                 215                 220

Leu Gly Thr Asp Pro Asp Leu Gln Thr Met Pro Leu Val Ala Phe His
225                 230                 235                 240

Lys Ile Val Gly Ala Lys Ala Arg Gly Lys Gly Lys Ala Trp Leu Ala
                245                 250                 255

Trp Gln Ala Pro Leu Phe Phe Gly Gly Ile Ile Cys Ser Leu Val Ser
            260                 265                 270

Phe Gly Trp Gln Phe Val Leu His Pro Asn His Ala Leu Arg Val His
        275                 280                 285

Asn His Leu Glu Leu Ala Tyr Met Gly Leu Arg Tyr Val Leu Trp His
    290                 295                 300

Leu Ala Phe Gly His Leu Gly Leu Leu Ser Ser Leu Arg Leu Tyr Ala
305                 310                 315                 320

Phe Tyr Val Ala Val Gly Gly Thr Tyr Ile Phe Thr Asn Phe Ala Val
                325                 330                 335

Ser His Thr His Lys Asp Val Val Pro Pro Thr Lys His Ile Ser Trp
            340                 345                 350

Ala Leu Tyr Ser Ala Asn His Thr Thr Asn Cys Ser Asp Ser Pro Phe
        355                 360                 365

Val Asn Trp Trp Met Ala Tyr Leu Asn Phe Gln Ile Glu His His Leu
    370                 375                 380

Phe Pro Ser Met Pro Gln Tyr Asn His Pro Lys Ile Ala Pro Arg Val
385                 390                 395                 400
```

Arg Ala Leu Phe Glu Lys His Gly Val Glu Tyr Asp Val Arg Pro Tyr
405 410 415

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

Pro Glu His Ser Tyr His Glu His Thr His
435 440

<210> 10
<211> 1329
<212> DNA
<213> Isochrysis galbana

<220>
<221> misc_feature
<223> synthetic delta-5 desaturase (codon-optimized)

<400> 10
atggtcgctg gcaagtccgg agctgcagcc cacgtgaccc actcttccac tctccctcga 60

gagtaccacg gtgctaccaa cgactcccga tctgaggctg ccgatgtcac cgtctcgtct 120

atcgacgccg agaaggagat gattatcaac ggacgagtgt acgacgtctc ctcgttcgtg 180

aagcgacacc ctggtggctc cgtcatcaag ttccagctcg gagcagatgc ttctgacgct 240

tacaacaact tccacgtccg atcgaagaag gctgacaaga tgctgtactc tcttccctcc 300

cgacctgccg aggctggcta tgcacaggac gacatctcta gagactttga gaagctgcga 360

ctggaactga aagaggaagg ttacttcgag cccaatctgg tgcacgtgtc ctaccgatgt 420

gtcgaggtgc ttgccatgta ctgggctggc gtccagctga tctggtccgg atactggttc 480

ctcggtgcca tcgttgctgg aattgctcaa ggtcgatgcg gatggctcca gcatgaaggc 540

ggacactact cgctcactgg caacatcaag attgaccgac atctccagat ggccatctat 600

ggactgggct gtggtatgtc tggctgctac tggagaaacc agcacaacaa acatcacgcc 660

actcctcaga agctcggaac cgatcccgac ctgcagacca tgcctctcgt tgccttccac 720

aagattgtcg gagccaaggc acgaggcaag ggtaaagcct ggcttgcttg gcaagctccc 780

ctcttctttg gaggcatcat ttgctccctg gtctctttcg gctggcagtt cgttctccac 840

cccaatcatg cactgcgagt gcacaaccat ctcgaactgg cctacatggg tctccgatac 900

gttctctggc accttgcctt tggccatctg ggactcctgt cctctcttcg actgtatgcc 960

ttctacgtgg ctgtcggtgg cacctacatc ttcaccaact tcgccgtctc ccatactcac 1020

aaggatgtcg ttcctcccac caagcacatt tcgtgggctc tgtactctgc caaccacact 1080

accaactgtt ccgactctcc ctttgtcaac tggtggatgg cctacctcaa cttccagatc 1140

gagcaccatc tgttcccctc catgcctcag tacaaccacc ccaagattgc tcctcgagtg 1200

cgagcactct tcgagaagca cggagtcgag tacgacgtcc gaccctatct ggaatgcttt 1260

cgagtgacct acgtcaacct ccttgctgtt ggcaaccctg agcactccta ccacgagcat 1320

actcactaa 1329

<210> 11
<211> 1335
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> delta-5 desaturase

<400> 11
atggcccccg acccggtggc cgccgagacc gcggctcagg gacctacccc gcgctacttc 60

acctgggacg aggtggccca gcgctcaggg tgcgaggagc ggtggctagt gatcgaccgt 120

aaggtgtaca acatcagcga gttcacccgc cggcatccag ggggctcccg ggtcatcagc 180

cactacgccg ggcaggatgc cacggatccc tttgtggcct tccacatcaa caagggcctt 240

gtgaagaagt atatgaactc tctcctgatt ggagaactgt ctccagagca gcccagcttt 300

gagcccacca agaataaaga gctgacagat gagttccggg agctgcgggc cacagtggag 360

cggatggggc tcatgaaggc caaccatgtc ttcttcctgc tgtacctgct gcacatcttg 420

ctgctggatg gtgcagcctg gctcaccctt tgggtctttg ggacgtcctt tttgcccttc 480

ctcctctgtg cggtgctgct cagtgcagtt caggcccagg ctggctggct gcagcatgac 540

tttgggcacc tgtcggtctt cagcacctca aagtggaacc atctgctaca tcattttgtg 600

attggccacc tgaagggggc ccccgccagt tggtggaacc acatgcactt ccagcaccat 660

gccaagccca actgcttccg caaagaccca gacatcaaca tgcatccctt cttctttgcc 720

ttggggaaga tcctctctgt ggagcttggg aaacagaaga aaaaatatat gccgtacaac 780

caccagcaca aatacttctt cctaattggg cccccagcct tgctgcctct ctacttccag 840

tggtatattt tctattttgt tatccagcga aagaagtggg tggacttggc ctggatgatt 900

accttctacg tccgcttctt cctcacttat gtgccactat tggggctgaa agccttcctg 960

ggccttttct tcatagtcag gttcctggaa agcaactggt ttgtgtgggt gacacagatg 1020

aaccatattc ccatgcacat tgatcatgac cggaacatgg actgggtttc cacccagctc 1080

caggccacat gcaatgtcca caagtctgcc ttcaatgact ggttcagtgg acacctcaac 1140

ttccagattg agcaccatct ttttcccacg atgcctcgac acaattacca caaagtggct 1200

cccctggtgc agtccttgtg tgccaagcat ggcatagagt accagtccaa gcccctgctg 1260

tcagccttcg ccgacatcat ccactcacta aaggagtcag ggcagctctg gctagatgcc 1320

tatcttcacc aataa 1335

<210> 12
<211> 444
<212> PRT
<213> Homo sapiens

<400> 12

Met Ala Pro Asp Pro Val Ala Ala Glu Thr Ala Ala Gln Gly Pro Thr
1 5 10 15

Pro Arg Tyr Phe Thr Trp Asp Glu Val Ala Gln Arg Ser Gly Cys Glu
20 25 30

Glu Arg Trp Leu Val Ile Asp Arg Lys Val Tyr Asn Ile Ser Glu Phe
35 40 45

Thr Arg Arg His Pro Gly Gly Ser Arg Val Ile Ser His Tyr Ala Gly
50 55 60

Gln Asp Ala Thr Asp Pro Phe Val Ala Phe His Ile Asn Lys Gly Leu
65 70 75 80

Val Lys Lys Tyr Met Asn Ser Leu Leu Ile Gly Glu Leu Ser Pro Glu
85 90 95

```
Gln Pro Ser Phe Glu Pro Thr Lys Asn Lys Glu Leu Thr Asp Glu Phe
            100                 105                 110

Arg Glu Leu Arg Ala Thr Val Glu Arg Met Gly Leu Met Lys Ala Asn
        115                 120                 125

His Val Phe Phe Leu Leu Tyr Leu Leu His Ile Leu Leu Leu Asp Gly
    130                 135                 140

Ala Ala Trp Leu Thr Leu Trp Val Phe Gly Thr Ser Phe Leu Pro Phe
145                 150                 155                 160

Leu Leu Cys Ala Val Leu Leu Ser Ala Val Gln Ala Gln Ala Gly Trp
                165                 170                 175

Leu Gln His Asp Phe Gly His Leu Ser Val Phe Ser Thr Ser Lys Trp
            180                 185                 190

Asn His Leu Leu His His Phe Val Ile Gly His Leu Lys Gly Ala Pro
        195                 200                 205

Ala Ser Trp Trp Asn His Met His Phe Gln His His Ala Lys Pro Asn
    210                 215                 220

Cys Phe Arg Lys Asp Pro Asp Ile Asn Met His Pro Phe Phe Phe Ala
225                 230                 235                 240

Leu Gly Lys Ile Leu Ser Val Glu Leu Gly Lys Gln Lys Lys Lys Tyr
                245                 250                 255

Met Pro Tyr Asn His Gln His Lys Tyr Phe Phe Leu Ile Gly Pro Pro
            260                 265                 270

Ala Leu Leu Pro Leu Tyr Phe Gln Trp Tyr Ile Phe Tyr Phe Val Ile
        275                 280                 285

Gln Arg Lys Lys Trp Val Asp Leu Ala Trp Met Ile Thr Phe Tyr Val
    290                 295                 300
```

```
Arg Phe Phe Leu Thr Tyr Val Pro Leu Leu Gly Leu Lys Ala Phe Leu

305                 310                 315                 320

Gly Leu Phe Phe Ile Val Arg Phe Leu Glu Ser Asn Trp Phe Val Trp
                325                 330                 335

Val Thr Gln Met Asn His Ile Pro Met His Ile Asp His Asp Arg Asn
            340                 345                 350

Met Asp Trp Val Ser Thr Gln Leu Gln Ala Thr Cys Asn Val His Lys
        355                 360                 365

Ser Ala Phe Asn Asp Trp Phe Ser Gly His Leu Asn Phe Gln Ile Glu
    370                 375                 380

His His Leu Phe Pro Thr Met Pro Arg His Asn Tyr His Lys Val Ala
385                 390                 395                 400

Pro Leu Val Gln Ser Leu Cys Ala Lys His Gly Ile Glu Tyr Gln Ser
                405                 410                 415

Lys Pro Leu Leu Ser Ala Phe Ala Asp Ile Ile His Ser Leu Lys Glu
            420                 425                 430

Ser Gly Gln Leu Trp Leu Asp Ala Tyr Leu His Gln
        435                 440
```

<210> 13
<211> 1335
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> synthetic delta-5 desaturase (codon-optimized)

<400> 13
atggctcccg accctgtcgc tgccgagacc gctgcccagg gtcccactcc ccgatacttc 60

acctgggacg aggtcgccca gcgatccggt tgcgaggaac gatggctggt catcgaccga 120

aaggtgtaca acatctctga gttcacccga cgacatcccg gtggctcccg agtgatctcg 180

cactacgctg gacaggacgc cactgacccc ttcgttgcct ttcacattaa caagggcctg 240

gttaagaagt acatgaactc cctgctcatt ggagagctgt ctcccgaaca gccttcgttt 300

gagcctacca agaacaagga gctgaccgac gagtttcgag agctccgagc caccgttgag 360

cgaatgggac tgatgaaggc caaccatgtc ttctttctgc tctacctgct ccacattctt 420

ctccttgacg gagctgcctg gcttaccctg tgggtcttcg gcacttcctt tctgcccttt 480

cttctctgcg ccgtcctgct ctctgccgtg caggctcagg ctggttggct tcagcatgac 540

tttggtcacc tttccgtgtt ctctacctcc aagtggaacc acctgctcca tcacttcgtg 600

atcggccacc tcaagggtgc tcctgcctcg tggtggaacc acatgcattt ccagcaccat 660

gccaagccca actgttttcg aaaggatccc gacatcaaca tgcacccctt ctttttcgct 720

cttggcaaga tcctgtccgt cgagctcgga aagcagaaga agaagtacat gccctacaac 780

caccagcaca agtacttctt cctgattgga cctcccgctc tcctgcctct ttactttcag 840

tggtacatct tttactttgt tattcagcga aagaagtggg ttgatcttgc ctggatgatc 900

accttctacg tccgattctt cctgacctac gtccctctcc ttggactgaa ggcctttctc 960

ggtctgttct ttatcgtccg attcctggag tccaactggt tcgtgtgggt gacccagatg 1020

aaccacattc ccatgcacat tgaccatgat cgaaacatgg actgggtgtc gactcagctg 1080

caggccacct gcaacgttca caagtctgct ttcaacgact ggttttccgg tcacctcaac 1140

tttcagattg agcaccatct gtttcccacc atgcctcgac acaactacca caaggttgct 1200

cccctggtcc agtcgctctg tgccaagcat ggcatcgagt accagtccaa gcccctgctc 1260

tctgccttcg ctgacatcat tcactcgctg aaggaatctg gccagctctg gctcgatgcc 1320

tacctgcacc agtaa 1335

<210> 14
<211> 1590
<212> DNA
<213> Danio rerio (GenBank Accession No. AF309556)

<220>
<221> CDS

<222> (1)..(1335)

<220>
<221> misc_feature
<222> (1)..(1335)
<223> Drd6/d5(V)

<400> 14

```
atg ggt ggc gga gga cag cag aca gac cga atc acc gac acc aac ggc       48
Met Gly Gly Gly Gly Gln Gln Thr Asp Arg Ile Thr Asp Thr Asn Gly
1               5                  10                  15

aga ttc agc agc tac acc tgg gag gag gtg cag aaa cac acc aaa cat       96
Arg Phe Ser Ser Tyr Thr Trp Glu Glu Val Gln Lys His Thr Lys His
            20                  25                  30

gga gat cag tgg gtg gtg gtg gag agg aag gtt tat aac gtc agc cag      144
Gly Asp Gln Trp Val Val Val Glu Arg Lys Val Tyr Asn Val Ser Gln
        35                  40                  45

tgg gtg aag aga cac ccc gga gga ctg agg atc ctc gga cac tat gct      192
Trp Val Lys Arg His Pro Gly Gly Leu Arg Ile Leu Gly His Tyr Ala
    50                  55                  60

gga gaa gac gcc acg gag gcg ttc act gcg ttt cat cca aac ctt cag      240
Gly Glu Asp Ala Thr Glu Ala Phe Thr Ala Phe His Pro Asn Leu Gln
65                  70                  75                  80

ctg gtg agg aaa tac ctg aag ccg ctg cta atc gga gag ctg gag gcg      288
Leu Val Arg Lys Tyr Leu Lys Pro Leu Leu Ile Gly Glu Leu Glu Ala
                85                  90                  95

tct gaa ccc agt cag gac cgg cag aaa aac gct gct ctc gtg gag gat      336
Ser Glu Pro Ser Gln Asp Arg Gln Lys Asn Ala Ala Leu Val Glu Asp
            100                 105                 110

ttc cga gcc ctg cgt gag cgt ctg gag gct gaa ggc tgt ttt aaa acg      384
Phe Arg Ala Leu Arg Glu Arg Leu Glu Ala Glu Gly Cys Phe Lys Thr
        115                 120                 125

cag ccg ctg ttt ttc gct ctg cat ttg ggc cac att ctg ctc ctg gag      432
Gln Pro Leu Phe Phe Ala Leu His Leu Gly His Ile Leu Leu Leu Glu
    130                 135                 140

gcc atc gct ttc atg atg gtg tgg tat ttc ggc acc ggt tgg atc aac      480
Ala Ile Ala Phe Met Met Val Trp Tyr Phe Gly Thr Gly Trp Ile Asn
145                 150                 155                 160

acg ctc atc gtc gct gtt att ctg gct act gca cag tca caa gct gga      528
Thr Leu Ile Val Ala Val Ile Leu Ala Thr Ala Gln Ser Gln Ala Gly
                165                 170                 175
```

```
tgg ttg cag cat gac ttc ggt cat ctg tcc gtg ttt aaa acc tct gga      576
Trp Leu Gln His Asp Phe Gly His Leu Ser Val Phe Lys Thr Ser Gly
            180                 185                 190

atg aat cat ttg gtg cac aaa ttt gtc atc gga cac ctg aag gga gcg      624
Met Asn His Leu Val His Lys Phe Val Ile Gly His Leu Lys Gly Ala
        195                 200                 205

tct gcg ggc tgg tgg aac cat cgg cac ttc cag cat cac gct aaa ccc      672
Ser Ala Gly Trp Trp Asn His Arg His Phe Gln His His Ala Lys Pro
    210                 215                 220

aac atc ttc aag aag gac ccg gac gtc aac atg ctg aac gcc ttt gtg      720
Asn Ile Phe Lys Lys Asp Pro Asp Val Asn Met Leu Asn Ala Phe Val
225                 230                 235                 240

gtg gga aac gtg cag ccc gtg gag tat ggc gtt aag aag atc aag cat      768
Val Gly Asn Val Gln Pro Val Glu Tyr Gly Val Lys Lys Ile Lys His
                245                 250                 255

ctg ccc tac aac cat cag cac aag tac ttc ttc ttc att ggt cct ccc      816
Leu Pro Tyr Asn His Gln His Lys Tyr Phe Phe Phe Ile Gly Pro Pro
            260                 265                 270

ctg ctc atc cca gtg tat ttc cag ttc caa atc ttt cac aat atg atc      864
Leu Leu Ile Pro Val Tyr Phe Gln Phe Gln Ile Phe His Asn Met Ile
        275                 280                 285

agt cat ggc atg tgg gtg gac ctg ctg tgg tgt atc agc tac tac gtc      912
Ser His Gly Met Trp Val Asp Leu Leu Trp Cys Ile Ser Tyr Tyr Val
    290                 295                 300

cga tac ttc ctt tgt tac acg cag ttc tac ggc gtc ttt tgg gct att      960
Arg Tyr Phe Leu Cys Tyr Thr Gln Phe Tyr Gly Val Phe Trp Ala Ile
305                 310                 315                 320

atc ctc ttt aat ttc gtc agg ttt atg gag agc cac tgg ttt gtt tgg     1008
Ile Leu Phe Asn Phe Val Arg Phe Met Glu Ser His Trp Phe Val Trp
                325                 330                 335

gtc aca cag atg agc cac atc ccc atg aac att gac tat gag aaa aat     1056
Val Thr Gln Met Ser His Ile Pro Met Asn Ile Asp Tyr Glu Lys Asn
            340                 345                 350

cag gac tgg ctc agc atg cag ctg gtc gcg acc tgt aac atc gag cag     1104
Gln Asp Trp Leu Ser Met Gln Leu Val Ala Thr Cys Asn Ile Glu Gln
        355                 360                 365

tct gcc ttc aac gac tgg ttc agc gga cac ctc aac ttc cag atc gag     1152
Ser Ala Phe Asn Asp Trp Phe Ser Gly His Leu Asn Phe Gln Ile Glu
    370                 375                 380
```

```
cat cat ctc ttt ccc aca gtg cct cgg cac aac tac tgg cgc gcc gct     1200
His His Leu Phe Pro Thr Val Pro Arg His Asn Tyr Trp Arg Ala Ala
385                 390                 395                 400

cca cgg gtg cga gcg ttg tgt gag aaa tac gga gtc aaa tac caa gag     1248
Pro Arg Val Arg Ala Leu Cys Glu Lys Tyr Gly Val Lys Tyr Gln Glu
                405                 410                 415

aag acc ttg tac gga gca ttt gcg gat atc att agg tct ttg gag aaa     1296
Lys Thr Leu Tyr Gly Ala Phe Ala Asp Ile Ile Arg Ser Leu Glu Lys
            420                 425                 430

tct ggc gag ctc tgg ctg gat gcg tat ctc aac aaa taa agcatgactt      1345
Ser Gly Glu Leu Trp Leu Asp Ala Tyr Leu Asn Lys
        435                 440

ctccctaaat agaaaaaaag cggagccgaa tcctccaatc agagcgagcc ttcacagact   1405

agacatcagg aagaatcgcg actaaacccc cagtcaatca gttcagtaaa atgacgcgta   1465

gcactttgtt ttttgatatc tgcaagataa ttgtggcact ttgaatgcgt tttaaagaga   1525

cggctcaaag atgtgtgtct cgctcctaat aaaggaagaa tctgaaagga aaaaaaaaaa   1585

aaaaa                                                               1590
```

<210> 15
<211> 444
<212> PRT
<213> Danio rerio (GenBank Accession No. AF309556)

<400> 15

```
Met Gly Gly Gly Gly Gln Gln Thr Asp Arg Ile Thr Asp Thr Asn Gly
1               5                   10                  15

Arg Phe Ser Ser Tyr Thr Trp Glu Glu Val Gln Lys His Thr Lys His
            20                  25                  30

Gly Asp Gln Trp Val Val Val Glu Arg Lys Val Tyr Asn Val Ser Gln
        35                  40                  45

Trp Val Lys Arg His Pro Gly Gly Leu Arg Ile Leu Gly His Tyr Ala
    50                  55                  60

Gly Glu Asp Ala Thr Glu Ala Phe Thr Ala Phe His Pro Asn Leu Gln
65                  70                  75                  80
```

```
Leu Val Arg Lys Tyr Leu Lys Pro Leu Leu Ile Gly Glu Leu Glu Ala
                85                  90                  95

Ser Glu Pro Ser Gln Asp Arg Gln Lys Asn Ala Ala Leu Val Glu Asp
            100                 105                 110

Phe Arg Ala Leu Arg Glu Arg Leu Glu Ala Glu Gly Cys Phe Lys Thr
        115                 120                 125

Gln Pro Leu Phe Phe Ala Leu His Leu Gly His Ile Leu Leu Leu Glu
    130                 135                 140

Ala Ile Ala Phe Met Met Val Trp Tyr Phe Gly Thr Gly Trp Ile Asn
145                 150                 155                 160

Thr Leu Ile Val Ala Val Ile Leu Ala Thr Ala Gln Ser Gln Ala Gly
                165                 170                 175

Trp Leu Gln His Asp Phe Gly His Leu Ser Val Phe Lys Thr Ser Gly
            180                 185                 190

Met Asn His Leu Val His Lys Phe Val Ile Gly His Leu Lys Gly Ala
        195                 200                 205

Ser Ala Gly Trp Trp Asn His Arg His Phe Gln His His Ala Lys Pro
    210                 215                 220

Asn Ile Phe Lys Lys Asp Pro Asp Val Asn Met Leu Asn Ala Phe Val
225                 230                 235                 240

Val Gly Asn Val Gln Pro Val Glu Tyr Gly Val Lys Lys Ile Lys His
                245                 250                 255

Leu Pro Tyr Asn His Gln His Lys Tyr Phe Phe Phe Ile Gly Pro Pro
            260                 265                 270

Leu Leu Ile Pro Val Tyr Phe Gln Phe Gln Ile Phe His Asn Met Ile
        275                 280                 285
```

```
Ser His Gly Met Trp Val Asp Leu Leu Trp Cys Ile Ser Tyr Tyr Val
    290                 295                 300

Arg Tyr Phe Leu Cys Tyr Thr Gln Phe Tyr Gly Val Phe Trp Ala Ile
305                 310                 315                 320

Ile Leu Phe Asn Phe Val Arg Phe Met Glu Ser His Trp Phe Val Trp
                325                 330                 335

Val Thr Gln Met Ser His Ile Pro Met Asn Ile Asp Tyr Glu Lys Asn
            340                 345                 350

Gln Asp Trp Leu Ser Met Gln Leu Val Ala Thr Cys Asn Ile Glu Gln
        355                 360                 365

Ser Ala Phe Asn Asp Trp Phe Ser Gly His Leu Asn Phe Gln Ile Glu
    370                 375                 380

His His Leu Phe Pro Thr Val Pro Arg His Asn Tyr Trp Arg Ala Ala
385                 390                 395                 400

Pro Arg Val Arg Ala Leu Cys Glu Lys Tyr Gly Val Lys Tyr Gln Glu
                405                 410                 415

Lys Thr Leu Tyr Gly Ala Phe Ala Asp Ile Ile Arg Ser Leu Glu Lys
            420                 425                 430

Ser Gly Glu Leu Trp Leu Asp Ala Tyr Leu Asn Lys
        435                 440
```

<210> 16
<211> 1946
<212> DNA
<213> Danio rerio (GenBank Accession No. BC068224)

<400> 16

```
cacaacagta ctgtatgagt gtgtgatcaa agtgcctgct gagacgtgat acgcgagcag     60

cactgttcag agatcagcga tgggtggcgg aggacagcag acagaccgaa tcaccgacac    120
```

caacggcaga ttcagcagct acacctggga ggaggtgcag aaacacacca aacatggaga 180

tcagtgggtg gtggtggaga ggaaggttta taacgtcagc cagtgggtga agagacaccc 240

cggaggactg aggatcctcg gacactatgc tggagaagac gccacggagg cgttcactgc 300

gtttcatcca aaccttcagc tggtgaggaa atacctgaag ccgctgctaa tcggagagct 360

ggaggcgtct gaacccagtc aggaccggca gaaaaacgct gctctcgtgg aggatttccg 420

agccctgcgt gagcgtctgg aggctgaagg ctgttttaaa acgcagccgc tgtttttcgc 480

tctgcatttg ggccacattc tgctcctgga ggccatcgct ttcatgatgg tgtggtattt 540

cggcaccggt tggatcaaca cgctcatcgt cgctgttatt ctggctactg cacagtcaca 600

agctggatgg ttgcagcatg acttcggtca tctgtccgtg tttaaaacct ctggaatgaa 660

tcatttggtg cacaaatttg tcatcggaca cctgaaggga gcgtctgcgg gctggtggaa 720

ccatcggcac ttccagcatc acgctaaacc caacatcttc aagaaggacc cggacgtcaa 780

catgctgaac gcctttgtgg tgggaaacgt gcagcccgtg gagtatggcg ttaagaagat 840

caagcatctg ccctacaacc atcagcacaa gtacttcttc ttcattggtc ctcccctgct 900

catcccagtg tatttccagt tccaaatctt tcacaatatg atcagtcatg gcatgtgggt 960

ggacctgctg tggtgtatca gctactacgt ccgatacttc ctttgttaca cgcagttcta 1020

cggcgtcttt tgggctatta tcctctttaa tttcgtcagg ttatggagag ccactggttt 1080

gtttgggtca cacagatgag ccacatcccc atgaacattg actatgagaa aaatcaggac 1140

tggctcagca tgcagctggt cgcgacctgt aacatcgagc agtctgcctt caacgactgg 1200

ttcagcggac acctcaactt ccagatcgag catcatctct ttcccacaat gcctcggcac 1260

aactactggc gcgccgctcc acgggtgcga gcgttgtgtg agaaatacgg agtcaaatac 1320

caagagaaga ccttgtacgg agcatttgcg gatatcatta ggtctttgga gaaatctggc 1380

gagctctggc tggatgcgta tctcaacaaa taaagcatga cttctcccta aatagaaaaa 1440

aagcggagcc gaatcctcca atcagagcga gccttcacag actagacatc aggaagaatc 1500

gcaactaaac ccccagtcaa tcagttcagt aaaatgacgc gtagcacttt gtttttgat 1560

atctgcaaga taattgtggc actttgaatg cgttttaaag agacggctca aagatgtgtg 1620

tctcgctcct aataaaggaa aaatctgaaa gtacatgtac gtgcatgaat gtgtgtgaaa 1680

gttgtgattg ctggtttaca atcgccgagt ttgctaggta atttcgaagt ttttatttag 1740

taaatttgca attcagatgt tgtttttata actaacattc ggacaaatga gatgacgata 1800

acactagtgg ttgcactttg tggtttatta tgaaatttaa ttggatttcc tactattggt 1860

catcgattaa acggattaaa cctcaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 1920

aaaaaaaaaa aaaaaaaaaa aaaaaa 1946

<210> 17
<211> 1335
<212> DNA
<213> Danio rerio

<220>
<221> misc_feature
<223> Drd6/d5(M)

<400> 17
atgggtggcg gaggacagca gacagaccga atcaccgaca ccaacggcag attcagcagc 60

tacacctggg aggaggtgca gaaacacacc aaacatggag atcagtgggt ggtggtggag 120

aggaaggttt ataacgtcag ccagtgggtg aagagacacc ccggaggact gaggatcctc 180

ggacactatg ctggagaaga cgccacggag gcgttcactg cgtttcatcc aaaccttcag 240

ctggtgagga aatacctgaa gccgctgcta atcggagagc tggaggcgtc tgaacccagt 300

caggaccggc agaaaaacgc tgctctcgtg gaggatttcc gagccctgcg tgagcgtctg 360

gaggctgaag gctgttttaa aacgcagccg ctgtttttcg ctctgcattt gggccacatt 420

ctgctcctgg aggccatcgc tttcatgatg gtgtggtatt tcggcaccgg ttggatcaac 480

acgctcatcg tcgctgttat tctggctact gcacagtcac aagctggatg gttgcagcat 540

gacttcggtc atctgtccgt gtttaaaacc tctggaatga atcatttggt gcacaaattt 600

gtcatcggac acctgaaggg agcgtctgcg ggctggtgga accatcggca cttccagcat 660

cacgctaaac ccaacatctt caagaaggac ccggacgtca acatgctgaa cgcctttgtg 720

gtgggaaacg tgcagcccgt ggagtatggc gttaagaaga tcaagcatct gccctacaac 780

catcagcaca agtacttctt cttcattggt cctcccctgc tcatcccagt gtatttccag 840

ttccaaatct ttcacaatat gatcagtcat ggcatgtggg tggacctgct gtggtgtatc 900

agctactacg tccgatactt cctttgttac acgcagttct acggcgtctt ttgggctatt 960

atcctcttta atttcgtcag gtttatggag agccactggt ttgtttgggt cacacagatg 1020

agccacatcc ccatgaacat tgactatgag aaaaatcagg actggctcag catgcagctg 1080

gtcgcgacct gtaacatcga gcagtctgcc ttcaacgact ggttcagcgg acacctcaac 1140

ttccagatcg agcatcatct ctttcccaca atgcctcggc acaactactg gcgcgccgct 1200

ccacgggtgc gagcgttgtg tgagaaatac ggagtcaaat accaagagaa gaccttgtac 1260

ggagcatttg cggatatcat taggtctttg gagaaatctg gcgagctctg gctggatgcg 1320

tatctcaaca aataa 1335

<210> 18
<211> 444
<212> PRT
<213> Danio rerio

<220>
<221> MISC_FEATURE
<223> Drd6/d5(M)

<400> 18

Met Gly Gly Gly Gly Gln Gln Thr Asp Arg Ile Thr Asp Thr Asn Gly
1 5 10 15

Arg Phe Ser Ser Tyr Thr Trp Glu Glu Val Gln Lys His Thr Lys His
20 25 30

Gly Asp Gln Trp Val Val Val Glu Arg Lys Val Tyr Asn Val Ser Gln
35 40 45

Trp Val Lys Arg His Pro Gly Gly Leu Arg Ile Leu Gly His Tyr Ala
50 55 60

Gly Glu Asp Ala Thr Glu Ala Phe Thr Ala Phe His Pro Asn Leu Gln
65 70 75 80

Leu Val Arg Lys Tyr Leu Lys Pro Leu Leu Ile Gly Glu Leu Glu Ala
85 90 95

```
Ser Glu Pro Ser Gln Asp Arg Gln Lys Asn Ala Ala Leu Val Glu Asp
            100                 105                 110

Phe Arg Ala Leu Arg Glu Arg Leu Glu Ala Glu Gly Cys Phe Lys Thr
        115                 120                 125

Gln Pro Leu Phe Phe Ala Leu His Leu Gly His Ile Leu Leu Leu Glu
    130                 135                 140

Ala Ile Ala Phe Met Met Val Trp Tyr Phe Gly Thr Gly Trp Ile Asn
145                 150                 155                 160

Thr Leu Ile Val Ala Val Ile Leu Ala Thr Ala Gln Ser Gln Ala Gly
                165                 170                 175

Trp Leu Gln His Asp Phe Gly His Leu Ser Val Phe Lys Thr Ser Gly
            180                 185                 190

Met Asn His Leu Val His Lys Phe Val Ile Gly His Leu Lys Gly Ala
        195                 200                 205

Ser Ala Gly Trp Trp Asn His Arg His Phe Gln His His Ala Lys Pro
    210                 215                 220

Asn Ile Phe Lys Lys Asp Pro Asp Val Asn Met Leu Asn Ala Phe Val
225                 230                 235                 240

Val Gly Asn Val Gln Pro Val Glu Tyr Gly Val Lys Lys Ile Lys His
                245                 250                 255

Leu Pro Tyr Asn His Gln His Lys Tyr Phe Phe Phe Ile Gly Pro Pro
            260                 265                 270

Leu Leu Ile Pro Val Tyr Phe Gln Phe Gln Ile Phe His Asn Met Ile
        275                 280                 285

Ser His Gly Met Trp Val Asp Leu Leu Trp Cys Ile Ser Tyr Tyr Val
    290                 295                 300
```

```
Arg Tyr Phe Leu Cys Tyr Thr Gln Phe Tyr Gly Val Phe Trp Ala Ile
305                 310                 315                 320

Ile Leu Phe Asn Phe Val Arg Phe Met Glu Ser His Trp Phe Val Trp
                325                 330                 335

Val Thr Gln Met Ser His Ile Pro Met Asn Ile Asp Tyr Glu Lys Asn
            340                 345                 350

Gln Asp Trp Leu Ser Met Gln Leu Val Ala Thr Cys Asn Ile Glu Gln
        355                 360                 365

Ser Ala Phe Asn Asp Trp Phe Ser Gly His Leu Asn Phe Gln Ile Glu
    370                 375                 380

His His Leu Phe Pro Thr Met Pro Arg His Asn Tyr Trp Arg Ala Ala
385                 390                 395                 400

Pro Arg Val Arg Ala Leu Cys Glu Lys Tyr Gly Val Lys Tyr Gln Glu
                405                 410                 415

Lys Thr Leu Tyr Gly Ala Phe Ala Asp Ile Ile Arg Ser Leu Glu Lys
            420                 425                 430

Ser Gly Glu Leu Trp Leu Asp Ala Tyr Leu Asn Lys
        435                 440
```

<210> 19
<211> 1077
<212> DNA
<213> Saprolegnia diclina (ATCC #56851)

<220>
<221> misc_feature
<223> delta-17 desaturase

<400> 19
atgactgagg ataagacgaa ggtcgagttc ccgacgctca cggagctcaa gcactcgatc 60

ccgaacgcgt gctttgagtc gaacctcggc ctctcgctct actacacggc ccgcgcgatc 120

ttcaacgcgt cggcctcggc ggcgctgctc tacgcggcgc gctcgacgcc gttcattgcc 180

gataacgttc tgctccacgc gctcgtttgc gccacctaca tctacgtgca gggcgtcatc 240

ttctggggct tcttcacggt cggccacgac tgcggccact cggccttctc gcgctaccac 300

agcgtcaact ttatcatcgg ctgcatcatg cactctgcga ttttgacgcc gttcgagagc 360

tggcgcgtga cgcaccgcca ccaccacaag aacacgggca acattgataa ggacgagatc 420

ttttacccgc accggtcggt caaggacctc caggacgtgc gccaatgggt ctacacgctc 480

ggcggtgcgt ggtttgtcta cttgaaggtc gggtatgccc cgcgcacgat gagccacttt 540

gacccgtggg acccgctcct ccttcgccgc gcgtcggccg tcatcgtgtc gctcggcgtc 600

tgggccgcct tcttcgccgc gtacgcgtac ctcacatact cgctcggctt tgccgtcatg 660

ggcctctact actatgcgcc gctctttgtc tttgcttcgt tcctcgtcat tacgaccttc 720

ttgcaccaca acgacgaagc gacgccgtgg tacggcgact cggagtggac gtacgtcaag 780

ggcaacctct cgagcgtcga ccgctcgtac ggcgcgttcg tggacaacct gagccaccac 840

attggcacgc accaggtcca ccacttgttc ccgatcattc cgcactacaa gctcaacgaa 900

gccaccaagc actttgcggc cgcgtacccg cacctcgtgc gcaggaacga cgagcccatc 960

atcacggcct tcttcaagac cgcgcacctc tttgtcaact acggcgctgt gcccgagacg 1020

gcgcagatct tcacgctcaa agagtcggcc gcggccgcca aggccaagtc ggactaa 1077

<210> 20
<211> 358
<212> PRT
<213> Saprolegnia diclina (ATCC #56851)

<400> 20

Met Ala Glu Asp Lys Thr Lys Val Glu Phe Pro Thr Leu Thr Glu Leu
1 5 10 15

Lys His Ser Ile Pro Asn Ala Cys Phe Glu Ser Asn Leu Gly Leu Ser
20 25 30

Leu Tyr Tyr Thr Ala Arg Ala Ile Phe Asn Ala Ser Ala Ser Ala Ala
35 40 45

Leu Leu Tyr Ala Ala Arg Ser Thr Pro Phe Ile Ala Asp Asn Val Leu
50 55 60

```
Leu His Ala Leu Val Cys Ala Thr Tyr Ile Tyr Val Gln Gly Val Ile
65                  70                  75                  80

Phe Trp Gly Phe Phe Thr Val Gly His Asp Cys Gly His Ser Ala Phe
                85                  90                  95

Ser Arg Tyr His Ser Val Asn Phe Ile Ile Gly Cys Ile Met His Ser
            100                 105                 110

Ala Ile Leu Thr Pro Phe Glu Ser Trp Arg Val Thr His Arg His His
        115                 120                 125

His Lys Asn Thr Gly Asn Ile Asp Lys Asp Glu Ile Phe Tyr Pro His
    130                 135                 140

Arg Ser Val Lys Asp Leu Gln Asp Val Arg Gln Trp Val Tyr Thr Leu
145                 150                 155                 160

Gly Gly Ala Trp Phe Val Tyr Leu Lys Val Gly Tyr Ala Pro Arg Thr
                165                 170                 175

Met Ser His Phe Asp Pro Trp Asp Pro Leu Leu Leu Arg Arg Ala Ser
            180                 185                 190

Ala Val Ile Val Ser Leu Gly Val Trp Ala Ala Phe Phe Ala Ala Tyr
        195                 200                 205

Ala Tyr Leu Thr Tyr Ser Leu Gly Phe Ala Val Met Gly Leu Tyr Tyr
    210                 215                 220

Tyr Ala Pro Leu Phe Val Phe Ala Ser Phe Leu Val Ile Thr Thr Phe
225                 230                 235                 240

Leu His His Asn Asp Glu Ala Thr Pro Trp Tyr Gly Asp Ser Glu Trp
                245                 250                 255

Thr Tyr Val Lys Gly Asn Leu Ser Ser Val Asp Arg Ser Tyr Gly Ala
            260                 265                 270
```

Phe Val Asp Asn Leu Ser His His Ile Gly Thr His Gln Val His His
275 280 285

Leu Phe Pro Ile Ile Pro His Tyr Lys Leu Asn Glu Ala Thr Lys His
290 295 300

Phe Ala Ala Ala Tyr Pro His Leu Val Arg Arg Asn Asp Glu Pro Ile
305 310 315 320

Ile Thr Ala Phe Phe Lys Thr Ala His Leu Phe Val Asn Tyr Gly Ala
325 330 335

Val Pro Glu Thr Ala Gln Ile Phe Thr Leu Lys Glu Ser Ala Ala Ala
340 345 350

Ala Lys Ala Lys Ser Asp
355

<210> 21
<211> 1077
<212> DNA
<213> Saprolegnia diclina

<220>
<221> misc_feature
<223> synthetic delta-17 desaturase (codon-optimized)

<400> 21
atggctgagg ataagaccaa ggtcgagttc cctaccctga ctgagctgaa gcactctatc 60

cctaacgctt gctttgagtc caacctcgga ctctcgctct actacactgc ccgagcgatc 120

ttcaacgcat ctgcctctgc tgctctgctc tacgctgccc gatctactcc cttcattgcc 180

gataacgttc tgctccacgc tctggtttgc gccacctaca tctacgtgca gggtgtcatc 240

ttctggggtt tctttaccgt cggtcacgac tgtggtcact ctgccttctc ccgataccac 300

tccgtcaact tcatcattgg ctgcatcatg cactctgcca ttctgactcc cttcgagtcc 360

tggcgagtga cccaccgaca ccatcacaag aacactggca acattgataa ggacgagatc 420

ttctaccctc atcggtccgt caaggacctc caggacgtgc gacaatgggt ctacaccctc 480

ggaggtgctt ggtttgtcta cctgaaggtc ggatatgctc ctcgaaccat gtcccacttt 540

gacccctggg accctctcct gcttcgacga gcctccgctg tcatcgtgtc cctcggagtc 600

tgggctgcct tcttcgctgc ctacgcctac ctcacatact cgctcggctt tgccgtcatg 660

ggcctctact actatgctcc tctctttgtc tttgcttcgt tcctcgtcat tactaccttc 720

ttgcatcaca acgacgaagc tactccctgg tacggtgact cggagtggac ctacgtcaag 780

ggcaacctga gctccgtcga ccgatcgtac ggagctttcg tggacaacct gtctcaccac 840

attggcaccc accaggtcca tcacttgttc cctatcattc cccactacaa gctcaacgaa 900

gccaccaagc actttgctgc cgcttaccct cacctcgtga gacgtaacga cgagcccatc 960

attactgcct tcttcaagac cgctcacctc tttgtcaact acggagctgt gcccgagact 1020

gctcagattt tcaccctcaa agagtctgcc gctgcagcca aggccaagag cgactaa 1077

<210> 22
<211> 957
<212> DNA
<213> Mortierella alpina AX464731

<220>
<221> misc_feature
<223> elongase

<400> 22
atggagtcga ttgcgccatt cctcccatca aagatgccgc aagatctgtt tatggacctt 60

gccaccgcta tcggtgtccg ggccgcgccc tatgtcgatc ctctcgaggc cgcgctggtg 120

gcccaggccg agaagtacat ccccacgatt gtccatcaca cgcgtgggtt cctggtcgcg 180

gtggagtcgc ctttggcccg tgagctgccg ttgatgaacc cgttccacgt gctgttgatc 240

gtgctcgctt atttggtcac ggtctttgtg ggcatgcaga tcatgaagaa ctttgagcgg 300

ttcgaggtca agacgttttc gctcctgcac aacttttgtc tggtctcgat cagcgcctac 360

atgtgcggtg ggatcctgta cgaggcttat caggccaact atggactgtt tgagaacgct 420

gctgatcata ccttcaaggg tcttcctatg gccaagatga tctggctctt ctacttctcc 480

aagatcatgg agtttgtcga caccatgatc atggtcctca agaagaacaa ccgccagatc 540

tccttcttgc acgtttacca ccacagctcc atcttcacca tctggtggtt ggtcaccttt 600

gttgcaccca acggtgaagc ctacttctct gctgcgttga actcgttcat ccatgtgatc 660

atgtacggct actacttctt gtcggccttg ggcttcaagc aggtgtcgtt catcaagttc 720

tacatcacgc gctcgcagat gacacagttc tgcatgatgt cggtccagtc ttcctgggac 780

atgtacgcca tgaaggtcct tggccgcccc ggataccct tcttcatcac ggctctgctt 840

tggttctaca tgtggaccat gctcggtctc ttctacaact tttacagaaa gaacgccaag 900

ttggccaagc aggccaaggc cgacgctgcc aaggagaagg caaggaagtt gcagtaa 957

<210> 23
<211> 318
<212> PRT
<213> Mortierella alpina AX464731

<400> 23

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> 24
<211> 957
<212> DNA
<213> Mortierella alpina

<220>
<221> misc_feature
<223> synthetic elongase 1 (codon-optimized)

<400> 24
atggagtcca ttgctccctt cctgccctcc aagatgcctc aggacctgtt catggacctc 60

gccagcgcta tcggtgtccg agctgctccc tacgtcgatc ccctggaggc tgccctggtt 120

gcccaggccg agaagtacat tcccaccatt gtccatcaca ctcgaggctt cctggttgcc 180

gtggagtctc ccctggctcg agagctgcct ctgatgaacc ccttccacgt gctcctgatc 240

gtgctcgcct acctggtcac cgtgtttgtg ggtatgcaga tcatgaagaa ctttgaacga 300

ttcgaggtca agaccttctc cctcctgcac aacttctgtc tggtctccat ctccgcctac 360

atgtgcggtg gcatcctgta cgaggcttat caggccaact atggactgtt tgagaacgct 420

gccgatcaca ccttcaaggg tctccctatg gctaagatga tctggctctt ctacttctcc 480

aagatcatgg agtttgtcga caccatgatc atggtcctca agaagaacaa ccgacagatt 540

tcctttctgc acgtgtacca ccactcttcc atcttcacca tctggtggct ggtcaccttc 600

gttgctccca acggtgaagc ctacttctct gctgccctga actccttcat ccacgtcatc 660

atgtacggct actactttct gtctgccctg ggcttcaagc aggtgtcgtt catcaagttc 720

tacatcactc gatcccagat gacccagttc tgcatgatgt ctgtccagtc ttcctgggac 780

atgtacgcca tgaaggtcct tggccgacct ggataccct tcttcatcac cgctctgctc 840

tggttctaca tgtggaccat gctcggtctc ttctacaact tttaccgaaa gaacgccaag 900

ctcgccaagc aggccaaggc tgacgctgcc aaggagaagg ccagaaagct ccagtaa 957

<210> 25
<211> 819
<212> DNA
<213> Thraustochytrium aureum

<220>
<221> misc_feature
<223> elongase 2

<400> 25
atggcaaaca gcagcgtgtg ggatgatgtg gtgggccgcg tggagaccgg cgtggaccag 60

tggatggatg gcgccaagcc gtacgcactc accgatgggc tcccgatgat ggacgtgtcc 120

accatgctgg cattcgaggt gggatacatg gccatgctgc tcttcggcat cccgatcatg 180

aagcagatgg agaagccttt tgagctcaag accatcaagc tcttgcacaa cttgtttctc 240

ttcggacttt ccttgtacat gtgcgtggag accatccgcc aggctatcct cggaggctac 300

aaagtgtttg gaaacgacat ggagaagggc aacgagtctc atgctcaggg catgtctcgc 360

atcgtgtacg tgttctacgt gtccaaggca tacgagttct tggataccgc catcatgatc 420

ctttgcaaga agttcaacca ggtttccttc ttgcatgtgt accaccatgc caccattttt 480

gccatctggt gggctatcgc caagtacgct ccaggaggtg atgcgtactt ttcagtgatc 540

ctcaactctt tcgtgcacac cgtcatgtac gcatactact tcttctcctc ccaagggttc 600

gggttcgtga agccaatcaa gccgtacatc accacccttc agatgaccca gttcatggca 660

atgcttgtgc agtccttgta cgactacctc ttcccatgcg actacccaca ggctcttgtg 720

cagctccttg gagtgtacat gatcaccttg cttgccctct tcggcaactt ttttgtgcag 780

agctatctta aaaagccaaa aaagagcaag accaactaa 819

<210> 26
<211> 272
<212> PRT
<213> Thraustochytrium aureum

<400> 26

Met Ala Asn Ser Ser Val Trp Asp Asp Val Val Gly Arg Val Glu Thr
1 5 10 15

Gly Val Asp Gln Trp Met Asp Gly Ala Lys Pro Tyr Ala Leu Thr Asp
20 25 30

Gly Leu Pro Met Met Asp Val Ser Thr Met Leu Ala Phe Glu Val Gly
35 40 45

Tyr Met Ala Met Leu Leu Phe Gly Ile Pro Ile Met Lys Gln Met Glu
50 55 60

Lys Pro Phe Glu Leu Lys Thr Ile Lys Leu Leu His Asn Leu Phe Leu
65 70 75 80

Phe Gly Leu Ser Leu Tyr Met Cys Val Glu Thr Ile Arg Gln Ala Ile
85 90 95

```
Leu Gly Gly Tyr Lys Val Phe Gly Asn Asp Met Glu Lys Gly Asn Glu
            100                 105                 110

Ser His Ala Gln Gly Met Ser Arg Ile Val Tyr Val Phe Tyr Val Ser
        115                 120                 125

Lys Ala Tyr Glu Phe Leu Asp Thr Ala Ile Met Ile Leu Cys Lys Lys
    130                 135                 140

Phe Asn Gln Val Ser Phe Leu His Val Tyr His His Ala Thr Ile Phe
145                 150                 155                 160

Ala Ile Trp Trp Ala Ile Ala Lys Tyr Ala Pro Gly Gly Asp Ala Tyr
                165                 170                 175

Phe Ser Val Ile Leu Asn Ser Phe Val His Thr Val Met Tyr Ala Tyr
            180                 185                 190

Tyr Phe Phe Ser Ser Gln Gly Phe Gly Phe Val Lys Pro Ile Lys Pro
        195                 200                 205

Tyr Ile Thr Thr Leu Gln Met Thr Gln Phe Met Ala Met Leu Val Gln
    210                 215                 220

Ser Leu Tyr Asp Tyr Leu Phe Pro Cys Asp Tyr Pro Gln Ala Leu Val
225                 230                 235                 240

Gln Leu Leu Gly Val Tyr Met Ile Thr Leu Leu Ala Leu Phe Gly Asn
                245                 250                 255

Phe Phe Val Gln Ser Tyr Leu Lys Lys Pro Lys Lys Ser Lys Thr Asn
            260                 265                 270
```

<210> 27
<211> 819
<212> DNA
<213> Thraustochytrium aureum

<220>
<221> misc_feature
<223> synthetic elongase (codon-optimized)

<400> 27
atggccaact cctctgtctg ggacgacgtg gtcggacgag tcgagaccgg tgtcgaccag 60

tggatggacg gagctaagcc ctacgctctg accgacggtc tgcccatgat ggacgtctcc 120

accatgctcg ccttcgaggt cggctacatg gccatgctgc tcttcggcat tcccatcatg 180

aagcagatgg agaagccctt cgagctgaag accatcaagc tgctccacaa cctgttcctc 240

ttcggactgt ccctctacat gtgcgtcgag accatccgac aggctatcct gggtggctac 300

aaggtcttcg gcaacgacat ggagaagggc aacgagtccc acgctcaggg catgtcccga 360

atcgtctacg tgttctacgt ctccaaggcc tacgagttcc tggacaccgc tatcatgatc 420

ctgtgcaaga agttcaacca ggtctccttc ctgcacgtgt accaccatgc caccatcttc 480

gccatctggt gggctattgc caagtacgct cctggtggcg acgcctactt ctccgtcatc 540

ctcaactcct tcgtccacac cgtcatgtac gcctactact tcttttcctc tcagggcttc 600

ggcttcgtca agcccatcaa gccctacatc accactctgc agatgaccca gttcatggct 660

atgctggtgc agtccctgta cgactacctc ttcccctgcg actaccctca ggctctggtc 720

cagctgctcg gcgtgtacat gatcaccctg ctcgctctgt tcggcaactt ctttgtccag 780

tcctacctga agaagcccaa gaagtccaag accaactaa 819

<210> 28
<211> 1936
<212> DNA
<213> Yarrowia lipolytica

<220>
<221> CDS
<222> (283)..(1539)
<223> delta-12 desaturase

<400> 28
cgtagttata tacaagaggt agatgcgtgc tggtgttaga ggggctctca ggattaggag 60

gaaaatttga cattggccct caacatataa cctcgggtgt gcctctgttt accctcagct 120

tttgcttgtc cccaagtcag tcacgccagg ccaaaaaggt tggtggattg acagggagaa 180

```
aaaaaaaagc ctagtgggtt taaactcgag gtaagacatt gaaatatata ccggtcggca    240

tcctgagtcc ctttctcgta ttccaacaga ccgaccatag aa atg gat tcg acc       294
                                                 Met Asp Ser Thr
                                                 1

acg cag acc aac acc ggc acc ggc aag gtg gcc gtg cag ccc ccc acg      342
Thr Gln Thr Asn Thr Gly Thr Gly Lys Val Ala Val Gln Pro Pro Thr
5                   10                  15                  20

gcc ttc att aag ccc att gag aag gtg tcc gag ccc gtc tac gac acc      390
Ala Phe Ile Lys Pro Ile Glu Lys Val Ser Glu Pro Val Tyr Asp Thr
                25                  30                  35

ttt ggc aac gag ttc act cct cca gac tac tct atc aag gat att ctg      438
Phe Gly Asn Glu Phe Thr Pro Pro Asp Tyr Ser Ile Lys Asp Ile Leu
            40                  45                  50

gat gcc att ccc cag gag tgc tac aag cgg tcc tac gtt aag tcc tac      486
Asp Ala Ile Pro Gln Glu Cys Tyr Lys Arg Ser Tyr Val Lys Ser Tyr
        55                  60                  65

tcg tac gtg gcc cga gac tgc ttc ttt atc gcc gtt ttt gcc tac atg      534
Ser Tyr Val Ala Arg Asp Cys Phe Phe Ile Ala Val Phe Ala Tyr Met
    70                  75                  80

gcc tac gcg tac ctg cct ctt att ccc tcg gct tcc ggc cga gct gtg      582
Ala Tyr Ala Tyr Leu Pro Leu Ile Pro Ser Ala Ser Gly Arg Ala Val
85                  90                  95                  100

gcc tgg gcc atg tac tcc att gtc cag ggt ctg ttt ggc acc ggt ctg      630
Ala Trp Ala Met Tyr Ser Ile Val Gln Gly Leu Phe Gly Thr Gly Leu
                105                 110                 115

tgg gtt ctt gcc cac gag tgt ggc cac tct gct ttc tcc gac tct aac      678
Trp Val Leu Ala His Glu Cys Gly His Ser Ala Phe Ser Asp Ser Asn
            120                 125                 130

acc gtc aac aac gtc acc gga tgg gtt ctg cac tcc tcc atg ctg gtc      726
Thr Val Asn Asn Val Thr Gly Trp Val Leu His Ser Ser Met Leu Val
        135                 140                 145

cct tac tac gcc tgg aag ctg acc cac tcc atg cac cac aag tcc act      774
Pro Tyr Tyr Ala Trp Lys Leu Thr His Ser Met His His Lys Ser Thr
    150                 155                 160

ggt cac ctc acc cgt gat atg gtg ttt gtg ccc aag gac cga aag gag      822
Gly His Leu Thr Arg Asp Met Val Phe Val Pro Lys Asp Arg Lys Glu
165                 170                 175                 180

ttt atg gag aac cga ggc gcc cat gac tgg tct gag ctt gct gag gac      870
Phe Met Glu Asn Arg Gly Ala His Asp Trp Ser Glu Leu Ala Glu Asp
```

```
                185                 190                 195

gct ccc ctc atg acc ctc tac ggc ctc atc acc cag cag gtg ttt gga      918
Ala Pro Leu Met Thr Leu Tyr Gly Leu Ile Thr Gln Gln Val Phe Gly
            200                 205                 210

tgg cct ctg tat ctg ctg tct tac gtt acc gga cag aag tac ccc aag      966
Trp Pro Leu Tyr Leu Leu Ser Tyr Val Thr Gly Gln Lys Tyr Pro Lys
        215                 220                 225

ctc aac aaa tgg gct gtc aac cac ttc aac ccc aac gcc ccg ctg ttt     1014
Leu Asn Lys Trp Ala Val Asn His Phe Asn Pro Asn Ala Pro Leu Phe
    230                 235                 240

gag aag aag gac tgg ttc aac atc tgg atc tct aac gtc ggt att ggt     1062
Glu Lys Lys Asp Trp Phe Asn Ile Trp Ile Ser Asn Val Gly Ile Gly
245                 250                 255                 260

atc acc atg tcc gtc atc gca tac tcc atc aac cga tgg ggc ctg gct     1110
Ile Thr Met Ser Val Ile Ala Tyr Ser Ile Asn Arg Trp Gly Leu Ala
                265                 270                 275

tcc gtc acc ctc tac tac ctg atc ccc tac ctg tgg gtc aac cac tgg     1158
Ser Val Thr Leu Tyr Tyr Leu Ile Pro Tyr Leu Trp Val Asn His Trp
            280                 285                 290

ctc gtg gcc atc acc tac ctg cag cac acc gac ccc act ctg ccc cac     1206
Leu Val Ala Ile Thr Tyr Leu Gln His Thr Asp Pro Thr Leu Pro His
        295                 300                 305

tac cac gcc gac cag tgg aac ttc acc cga gga gcc gcc gcc acc atc     1254
Tyr His Ala Asp Gln Trp Asn Phe Thr Arg Gly Ala Ala Ala Thr Ile
    310                 315                 320

gac cga gag ttt ggc ttc atc ggc tcc ttc tgc ttc cat gac atc atc     1302
Asp Arg Glu Phe Gly Phe Ile Gly Ser Phe Cys Phe His Asp Ile Ile
325                 330                 335                 340

gag acc cac gtt ctg cac cac tac gtg tct cga att ccc ttc tac aac     1350
Glu Thr His Val Leu His His Tyr Val Ser Arg Ile Pro Phe Tyr Asn
                345                 350                 355

gcc cga atc gcc act gag aag atc aag aag gtc atg ggc aag cac tac     1398
Ala Arg Ile Ala Thr Glu Lys Ile Lys Lys Val Met Gly Lys His Tyr
            360                 365                 370

cga cac gac gac acc aac ttc atc aag tct ctt tac act gtc gcc cga     1446
Arg His Asp Asp Thr Asn Phe Ile Lys Ser Leu Tyr Thr Val Ala Arg
        375                 380                 385

acc tgc cag ttt gtt gaa ggt aag gaa ggc att cag atg ttt aga aac     1494
Thr Cys Gln Phe Val Glu Gly Lys Glu Gly Ile Gln Met Phe Arg Asn
```

```
    390                 395                 400

gtc aat gga gtc gga gtt gct cct gac ggc ctg cct tct aaa aag         1539
Val Asn Gly Val Gly Val Ala Pro Asp Gly Leu Pro Ser Lys Lys
405                 410                 415

tagagctaga aatgttattt gattgtgttt taactgaaca gcaccgagcc cgaggctaag   1599

ccaagcgaag ccgaggggtt gtgtagtcca tggacgtaac gagtaggcga tatcaccgca   1659

ctcggcactg cgtgtctgcg ttcatgggcg aagtcacatt acgctgacaa ccgttgtagt   1719

ttccctttag tatcaatact gttacaagta ccggtctcgt actcgtactg atacgaatct   1779

gtgggaagaa gtcaccctta tcagaccttc atactgatgt ttcggatatc aatagaactg   1839

gcatagagcc gttaaagaag tttcacttaa tcactccaac cctcctactt gtagattcaa   1899

gcagatcgat aagatggatt tgatggtcag tgctagc                           1936


<210>  29
<211>  419
<212>  PRT
<213>  Yarrowia lipolytica

<400>  29

Met Asp Ser Thr Thr Gln Thr Asn Thr Gly Thr Gly Lys Val Ala Val
1               5                   10                  15


Gln Pro Pro Thr Ala Phe Ile Lys Pro Ile Glu Lys Val Ser Glu Pro
            20                  25                  30


Val Tyr Asp Thr Phe Gly Asn Glu Phe Thr Pro Pro Asp Tyr Ser Ile
        35                  40                  45


Lys Asp Ile Leu Asp Ala Ile Pro Gln Glu Cys Tyr Lys Arg Ser Tyr
    50                  55                  60


Val Lys Ser Tyr Ser Tyr Val Ala Arg Asp Cys Phe Phe Ile Ala Val
65                  70                  75                  80


Phe Ala Tyr Met Ala Tyr Ala Tyr Leu Pro Leu Ile Pro Ser Ala Ser
                85                  90                  95


Gly Arg Ala Val Ala Trp Ala Met Tyr Ser Ile Val Gln Gly Leu Phe
            100                 105                 110
```

```
Gly Thr Gly Leu Trp Val Leu Ala His Glu Cys Gly His Ser Ala Phe
        115                 120                 125

Ser Asp Ser Asn Thr Val Asn Asn Val Thr Gly Trp Val Leu His Ser
    130                 135                 140

Ser Met Leu Val Pro Tyr Tyr Ala Trp Lys Leu Thr His Ser Met His
145                 150                 155                 160

His Lys Ser Thr Gly His Leu Thr Arg Asp Met Val Phe Val Pro Lys
                165                 170                 175

Asp Arg Lys Glu Phe Met Glu Asn Arg Gly Ala His Asp Trp Ser Glu
            180                 185                 190

Leu Ala Glu Asp Ala Pro Leu Met Thr Leu Tyr Gly Leu Ile Thr Gln
        195                 200                 205

Gln Val Phe Gly Trp Pro Leu Tyr Leu Leu Ser Tyr Val Thr Gly Gln
    210                 215                 220

Lys Tyr Pro Lys Leu Asn Lys Trp Ala Val Asn His Phe Asn Pro Asn
225                 230                 235                 240

Ala Pro Leu Phe Glu Lys Lys Asp Trp Phe Asn Ile Trp Ile Ser Asn
                245                 250                 255

Val Gly Ile Gly Ile Thr Met Ser Val Ile Ala Tyr Ser Ile Asn Arg
            260                 265                 270

Trp Gly Leu Ala Ser Val Thr Leu Tyr Tyr Leu Ile Pro Tyr Leu Trp
        275                 280                 285

Val Asn His Trp Leu Val Ala Ile Thr Tyr Leu Gln His Thr Asp Pro
    290                 295                 300

Thr Leu Pro His Tyr His Ala Asp Gln Trp Asn Phe Thr Arg Gly Ala
305                 310                 315                 320
```

```
Ala Ala Thr Ile Asp Arg Glu Phe Gly Phe Ile Gly Ser Phe Cys Phe
                325                 330                 335

His Asp Ile Ile Glu Thr His Val Leu His His Tyr Val Ser Arg Ile
            340                 345                 350

Pro Phe Tyr Asn Ala Arg Ile Ala Thr Glu Lys Ile Lys Lys Val Met
        355                 360                 365

Gly Lys His Tyr Arg His Asp Asp Thr Asn Phe Ile Lys Ser Leu Tyr
    370                 375                 380

Thr Val Ala Arg Thr Cys Gln Phe Val Glu Gly Lys Glu Gly Ile Gln
385                 390                 395                 400

Met Phe Arg Asn Val Asn Gly Val Gly Val Ala Pro Asp Gly Leu Pro
                405                 410                 415

Ser Lys Lys
```

<210> 30
<211> 1203
<212> DNA
<213> Mortieralla isabellina (GenBank Accession No. AF417245)

<220>
<221> CDS
<222> (1)..(1203)
<223> delta-12 desaturase

<400> 30

```
atg gca cct ccc aac act atc gat gcc ggc ttg acc cag cgt cat atc       48
Met Ala Pro Pro Asn Thr Ile Asp Ala Gly Leu Thr Gln Arg His Ile
1               5                   10                  15

acc acc acg gcc gcc cca acc tcg gcc aag ccc gct ttc gag cgc aac       96
Thr Thr Thr Ala Ala Pro Thr Ser Ala Lys Pro Ala Phe Glu Arg Asn
            20                  25                  30

tac cag ctc ccc gag ttc act atc aag gag atc cga gag tgc atc cct      144
Tyr Gln Leu Pro Glu Phe Thr Ile Lys Glu Ile Arg Glu Cys Ile Pro
```

```
        35                  40                  45

gcc cac tgc ttt gag cgc tcc ggt ctt cgt ggt ctc tgc cac gtt gcc      192
Ala His Cys Phe Glu Arg Ser Gly Leu Arg Gly Leu Cys His Val Ala
    50                  55                  60

att gat ctg acc tgg gcc tcg ctc ttg ttc ctg gct gca acc cag atc      240
Ile Asp Leu Thr Trp Ala Ser Leu Leu Phe Leu Ala Ala Thr Gln Ile
65                  70                  75                  80

gac aag ttc gag aac ccc ttg atc cgc tat ctg gcc tgg cct gcg tac      288
Asp Lys Phe Glu Asn Pro Leu Ile Arg Tyr Leu Ala Trp Pro Ala Tyr
                85                  90                  95

tgg atc atg cag ggc att gtc tgc acc ggc ata tgg gtg ctg gcc cac      336
Trp Ile Met Gln Gly Ile Val Cys Thr Gly Ile Trp Val Leu Ala His
            100                 105                 110

gag tgc ggt cac cag tcc ttc tcg acc tcc aag act ctc aac aac acc      384
Glu Cys Gly His Gln Ser Phe Ser Thr Ser Lys Thr Leu Asn Asn Thr
        115                 120                 125

gtc ggc tgg atc ctg cac tcg atg ctc ttg gtc ccc tac cac tcc tgg      432
Val Gly Trp Ile Leu His Ser Met Leu Leu Val Pro Tyr His Ser Trp
    130                 135                 140

aga atc tcg cac tcg aag cac cac aag gcc act ggc cac atg acc aag      480
Arg Ile Ser His Ser Lys His His Lys Ala Thr Gly His Met Thr Lys
145                 150                 155                 160

gac cag gtc ttt gtt ccc aag acc cgc tcc cag gtt ggt ttg cct ccc      528
Asp Gln Val Phe Val Pro Lys Thr Arg Ser Gln Val Gly Leu Pro Pro
                165                 170                 175

aag gag agc gct gct gct gcc gtt caa gag gag gac atg tcc gtg cac      576
Lys Glu Ser Ala Ala Ala Ala Val Gln Glu Glu Asp Met Ser Val His
            180                 185                 190

ctg gat gag gag gcc cct att gtg act ttg ttc tgg atg gtg atc cag      624
Leu Asp Glu Glu Ala Pro Ile Val Thr Leu Phe Trp Met Val Ile Gln
        195                 200                 205

ttc ctg ttc gga tgg cct gca tac ctg atc atg aac gcc tct ggt cag      672
Phe Leu Phe Gly Trp Pro Ala Tyr Leu Ile Met Asn Ala Ser Gly Gln
    210                 215                 220

gac tat ggc cgc tgg acc tcg cac ttc cac act tac tcg ccc atc ttt      720
Asp Tyr Gly Arg Trp Thr Ser His Phe His Thr Tyr Ser Pro Ile Phe
225                 230                 235                 240

gag ccc cgc aac ttc ttc gac att atc atc tcg gat ctc ggt gtg ttg      768
Glu Pro Arg Asn Phe Phe Asp Ile Ile Ile Ser Asp Leu Gly Val Leu
```

```
                245                 250                 255

gct gcc ctc ggt gcc ctg atc tac gct tcc atg cag ctg tcg ctc ttg      816
Ala Ala Leu Gly Ala Leu Ile Tyr Ala Ser Met Gln Leu Ser Leu Leu
            260                 265                 270

acc gtg acc aag tac tac atc atc ccg tac ctg ttt gtc aac ttt tgg      864
Thr Val Thr Lys Tyr Tyr Ile Ile Pro Tyr Leu Phe Val Asn Phe Trp
        275                 280                 285

ttg gtc ctg att act ttc ttg cag cac acc gac ccc aag ctg ccc cat      912
Leu Val Leu Ile Thr Phe Leu Gln His Thr Asp Pro Lys Leu Pro His
    290                 295                 300

tac cgt gag ggt gcc tgg aac ttc cag cgt gga gcc ctc tgc acc gtt      960
Tyr Arg Glu Gly Ala Trp Asn Phe Gln Arg Gly Ala Leu Cys Thr Val
305                 310                 315                 320

gac cgc tcg ttt ggc aag ttc ttg gac cat atg ttc cac ggc atc gtc     1008
Asp Arg Ser Phe Gly Lys Phe Leu Asp His Met Phe His Gly Ile Val
                325                 330                 335

cat acc cat gtg gcc cat cac ttg ttc tcg cag atg ccg ttc tac cat     1056
His Thr His Val Ala His His Leu Phe Ser Gln Met Pro Phe Tyr His
            340                 345                 350

gct gaa gaa gct acc tac cat ctc aag aaa ctg ctg gga gag tac tac     1104
Ala Glu Glu Ala Thr Tyr His Leu Lys Lys Leu Leu Gly Glu Tyr Tyr
        355                 360                 365

gtt tac gac cca tcc ccg atc gtc gtt gcg gtc tgg agg tcg ttc cgc     1152
Val Tyr Asp Pro Ser Pro Ile Val Val Ala Val Trp Arg Ser Phe Arg
    370                 375                 380

gag tgc cga ttc gtg gag gat cat gga gac gtg gtc ttt ttc aag aag     1200
Glu Cys Arg Phe Val Glu Asp His Gly Asp Val Val Phe Phe Lys Lys
385                 390                 395                 400

taa                                                                1203


<210>  31
<211>  400
<212>  PRT
<213>  Mortieralla isabellina (GenBank Accession No. AF417245)

<400>  31

Met Ala Pro Pro Asn Thr Ile Asp Ala Gly Leu Thr Gln Arg His Ile
1               5                   10                  15


Thr Thr Thr Ala Ala Pro Thr Ser Ala Lys Pro Ala Phe Glu Arg Asn
            20                  25                  30
```

```
Tyr Gln Leu Pro Glu Phe Thr Ile Lys Glu Ile Arg Glu Cys Ile Pro
        35                  40                  45

Ala His Cys Phe Glu Arg Ser Gly Leu Arg Gly Leu Cys His Val Ala
    50                  55                  60

Ile Asp Leu Thr Trp Ala Ser Leu Leu Phe Leu Ala Ala Thr Gln Ile
65                  70                  75                  80

Asp Lys Phe Glu Asn Pro Leu Ile Arg Tyr Leu Ala Trp Pro Ala Tyr
                85                  90                  95

Trp Ile Met Gln Gly Ile Val Cys Thr Gly Ile Trp Val Leu Ala His
            100                 105                 110

Glu Cys Gly His Gln Ser Phe Ser Thr Ser Lys Thr Leu Asn Asn Thr
        115                 120                 125

Val Gly Trp Ile Leu His Ser Met Leu Leu Val Pro Tyr His Ser Trp
    130                 135                 140

Arg Ile Ser His Ser Lys His His Lys Ala Thr Gly His Met Thr Lys
145                 150                 155                 160

Asp Gln Val Phe Val Pro Lys Thr Arg Ser Gln Val Gly Leu Pro Pro
                165                 170                 175

Lys Glu Ser Ala Ala Ala Ala Val Gln Glu Glu Asp Met Ser Val His
            180                 185                 190

Leu Asp Glu Glu Ala Pro Ile Val Thr Leu Phe Trp Met Val Ile Gln
        195                 200                 205

Phe Leu Phe Gly Trp Pro Ala Tyr Leu Ile Met Asn Ala Ser Gly Gln
    210                 215                 220

Asp Tyr Gly Arg Trp Thr Ser His Phe His Thr Tyr Ser Pro Ile Phe
225                 230                 235                 240
```

```
Glu Pro Arg Asn Phe Phe Asp Ile Ile Ile Ser Asp Leu Gly Val Leu
                245                 250                 255

Ala Ala Leu Gly Ala Leu Ile Tyr Ala Ser Met Gln Leu Ser Leu Leu
            260                 265                 270

Thr Val Thr Lys Tyr Tyr Ile Ile Pro Tyr Leu Phe Val Asn Phe Trp
        275                 280                 285

Leu Val Leu Ile Thr Phe Leu Gln His Thr Asp Pro Lys Leu Pro His
    290                 295                 300

Tyr Arg Glu Gly Ala Trp Asn Phe Gln Arg Gly Ala Leu Cys Thr Val
305                 310                 315                 320

Asp Arg Ser Phe Gly Lys Phe Leu Asp His Met Phe His Gly Ile Val
                325                 330                 335

His Thr His Val Ala His His Leu Phe Ser Gln Met Pro Phe Tyr His
            340                 345                 350

Ala Glu Glu Ala Thr Tyr His Leu Lys Lys Leu Leu Gly Glu Tyr Tyr
        355                 360                 365

Val Tyr Asp Pro Ser Pro Ile Val Val Ala Val Trp Arg Ser Phe Arg
    370                 375                 380

Glu Cys Arg Phe Val Glu Asp His Gly Asp Val Val Phe Phe Lys Lys
385                 390                 395                 400
```

<210> 32
<211> 1434
<212> DNA
<213> Fusarium monoliforme

<220>
<221> misc_feature
<223> delta-12 desaturase

<400> 32

atggcgtcca cttcggctct gcccaagcag aaccctgcgc ttagacgcac cgtcacctca 60

actactgtga cggattctga gtctgccgcc gtctctcctt cagactctcc ccgccactcg 120

gcctcttcca catcgctctc gtccatgtcc gaggttgata tcgccaagcc caagtccgag 180

tatggtgtca tgctcgacac ctacggcaac cagttcgagg ttcccgactt taccatcaag 240

gacatctaca atgccatccc taagcactgc ttcaagcgct ccgctctcaa gggatacggt 300

tatatcctcc gcgacattgt cctcctgact accactttca gcatctggta caactttgtg 360

acccccgaat atatcccctc cacccccgcc cgcgctggtc tgtgggccgt gtacaccgtt 420

cttcagggtc ttttcggtac tggtctctgg gttattgccc atgagtgcgg tcacggtgct 480

ttctccgatt ctcgcatcat caacgacatt actggctggg ttcttcactc ttccctcctt 540

gtcccctact tcagctggca aatctcccac cgaaagcacc acaaggccac tggcaacatg 600

gagcgtgaca tggtcttcgt tccccgaacc cgcgagcagc aggctactcg tctcggaaag 660

atgacccacg agctcgctca tcttactgag gagacccccg ctttcactct tctcatgctc 720

gtccttcagc agctcgttgg ctggcccaac tacctcatca ccaatgttac cggccacaac 780

taccacgagc gccagcgtga gggtcgcggc aagggcaagc ataacggcct cggcggtggt 840

gttaaccact tcgatccccg cagccctctg tacgagaaca gtgacgctaa gctcatcgtc 900

ctcagcgata ttggtatcgg tctgatggcc actgctctgt acttcctcgt tcagaagttc 960

ggtttctaca acatggccat ctggtacttt gttccctacc tctgggttaa ccactggctc 1020

gttgccatca ccttcctcca gcacaccgac cctacccttc cccactacac caacgacgag 1080

tggaacttcg tccgtggtgc cgctgctacc attgaccgtg agatgggctt catcggccgc 1140

caccttctcc acggcatcat cgagactcat gtcctccacc actacgtcag cagcatcccc 1200

ttctacaacg cggacgaggc caccgaggcc attaagccca tcatgggcaa gcactaccgg 1260

gctgatgtcc aggatggtcc tcgtggcttc atccgcgcca tgtaccgcag tgcgcgtatg 1320

tgccagtggg ttgagcccag cgctggtgcc gagggtgctg gtaagggtgt tctgttcttc 1380

cgcaaccgca acaacgtggg caccccccgc gctgttatca agcccgttgc ttaa 1434

<210> 33
<211> 477
<212> PRT

<213> Fusarium monoliforme

<400> 33

```
Met Ala Ser Thr Ser Ala Leu Pro Lys Gln Asn Pro Ala Leu Arg Arg
1               5                  10                  15

Thr Val Thr Ser Thr Thr Val Thr Asp Ser Glu Ser Ala Ala Val Ser
            20                  25                  30

Pro Ser Asp Ser Pro Arg His Ser Ala Ser Ser Thr Ser Leu Ser Ser
        35                  40                  45

Met Ser Glu Val Asp Ile Ala Lys Pro Lys Ser Glu Tyr Gly Val Met
    50                  55                  60

Leu Asp Thr Tyr Gly Asn Gln Phe Glu Val Pro Asp Phe Thr Ile Lys
65                  70                  75                  80

Asp Ile Tyr Asn Ala Ile Pro Lys His Cys Phe Lys Arg Ser Ala Leu
                85                  90                  95

Lys Gly Tyr Gly Tyr Ile Leu Arg Asp Ile Val Leu Leu Thr Thr Thr
            100                 105                 110

Phe Ser Ile Trp Tyr Asn Phe Val Thr Pro Glu Tyr Ile Pro Ser Thr
        115                 120                 125

Pro Ala Arg Ala Gly Leu Trp Ala Val Tyr Thr Val Leu Gln Gly Leu
    130                 135                 140

Phe Gly Thr Gly Leu Trp Val Ile Ala His Glu Cys Gly His Gly Ala
145                 150                 155                 160

Phe Ser Asp Ser Arg Ile Ile Asn Asp Ile Thr Gly Trp Val Leu His
                165                 170                 175

Ser Ser Leu Leu Val Pro Tyr Phe Ser Trp Gln Ile Ser His Arg Lys
            180                 185                 190

His His Lys Ala Thr Gly Asn Met Glu Arg Asp Met Val Phe Val Pro
        195                 200                 205
```

```
Arg Thr Arg Glu Gln Gln Ala Thr Arg Leu Gly Lys Met Thr His Glu
    210                 215                 220

Leu Ala His Leu Thr Glu Glu Thr Pro Ala Phe Thr Leu Leu Met Leu
225                 230                 235                 240

Val Leu Gln Gln Leu Val Gly Trp Pro Asn Tyr Leu Ile Thr Asn Val
                245                 250                 255

Thr Gly His Asn Tyr His Glu Arg Gln Arg Glu Gly Arg Gly Lys Gly
            260                 265                 270

Lys His Asn Gly Leu Gly Gly Gly Val Asn His Phe Asp Pro Arg Ser
        275                 280                 285

Pro Leu Tyr Glu Asn Ser Asp Ala Lys Leu Ile Val Leu Ser Asp Ile
    290                 295                 300

Gly Ile Gly Leu Met Ala Thr Ala Leu Tyr Phe Leu Val Gln Lys Phe
305                 310                 315                 320

Gly Phe Tyr Asn Met Ala Ile Trp Tyr Phe Val Pro Tyr Leu Trp Val
                325                 330                 335

Asn His Trp Leu Val Ala Ile Thr Phe Leu Gln His Thr Asp Pro Thr
            340                 345                 350

Leu Pro His Tyr Thr Asn Asp Glu Trp Asn Phe Val Arg Gly Ala Ala
        355                 360                 365

Ala Thr Ile Asp Arg Glu Met Gly Phe Ile Gly Arg His Leu Leu His
    370                 375                 380

Gly Ile Ile Glu Thr His Val Leu His His Tyr Val Ser Ser Ile Pro
385                 390                 395                 400

Phe Tyr Asn Ala Asp Glu Ala Thr Glu Ala Ile Lys Pro Ile Met Gly
                405                 410                 415
```

Lys His Tyr Arg Ala Asp Val Gln Asp Gly Pro Arg Gly Phe Ile Arg
420 425 430

Ala Met Tyr Arg Ser Ala Arg Met Cys Gln Trp Val Glu Pro Ser Ala
435 440 445

Gly Ala Glu Gly Ala Gly Lys Gly Val Leu Phe Phe Arg Asn Arg Asn
450 455 460

Asn Val Gly Thr Pro Pro Ala Val Ile Lys Pro Val Ala
465 470 475

<210> 34
<211> 1416
<212> DNA
<213> Aspergillus nidulans

<220>
<221> misc_feature
<223> delta-12 desaturase

<400> 34
atgtcgtcta ctgccctccc gaagcgcgtt gcgctgcatc gcaacccgac taccgactct 60

tcggtcccca gctccgtctc ggtctccccg ctggactcgc cccgtcagtc tccgtcgtcg 120

acttcgctct cgtcaatggc ctcggatgcg ggcaagggag acttgggcaa gatgctcgac 180

acctatggca atgagttcaa gatccccgac tacaccatca aggatatccg tgatgccatt 240

ccgtcccact gctacaaccg gtctgctatc aggagtctgt cctatgtctt ccgtgatctc 300

gccgtcctcg cttccgtctt ctacgtcttc cacaaatacg tgaccccgga gaccgtccct 360

tcgtacccgg cgcgtgttgc gctgtggact ctctacactg tcgtccaggg tctgttcggt 420

accggtattt gggttcttgc tcacgagtgt ggacaccagg cgttctctac ttccaaggtg 480

ctcaacgaca ctgttggctg gatcctgcat tcggctctgc tggtccccta tttctcgtgg 540

aagatctctc acggcaagca ccacaaggcc accggtaacc tggctcgtga catggtcttc 600

gtccccaaga cccgcgaggt gtacgcctcc cgcatcaaga agaccatcta cgacctgaac 660

gaggtgatgg aggagacccc cttggccact gccacccact ccatcctgca gcagctgttc 720

ggctggccct tgtacctgct caccaacgtt accggtcacg acaaccacga gcgccagcct 780

gaaggccgcg gcaagggcaa gcgtaacggc tacttcaccg gcgtcaacca cttcaacccc 840

aacagccctc tgttcgaggc caaggacgcc aagctcatca ttctgagtga tatcggcctc 900

gccatcaccg ccagcatcct gtacctgatc ggctccaagt tcggctggat gaacttgctc 960

gtctggtacg gtatccccta cctctgggtg aaccactggc ttgttgccat cacctacctc 1020

cagcacaccg accccactct cccccactac cagcccgagt cctggacctt cgcccgcggt 1080

gccgctgcca ccattgaccg cgagttcggc ttcatcggcc gtcacattct ccacggcatc 1140

atcgagaccc acgtcctcca ccactacgtc agcaccatcc ccttctacca cgccgacgag 1200

gccagcgagg ctatcaagaa ggtcatgggc tcgcactacc gcagcgaggc acacaccggt 1260

cctctgggct tcctcaaggc tctctggacc agcgcccgtg tctgccactg ggtcgagccc 1320

accgaaggca ccaagggcga gaacgctggt gtcttgttct tccgcaacac caacggcatc 1380

ggtgttcctc ccattaagct gaccaagcct aactaa 1416

<210> 35
<211> 471
<212> PRT
<213> Aspergillus nidulans

<400> 35

Met Ser Ser Thr Ala Leu Pro Lys Arg Val Ala Leu His Arg Asn Pro
1 5 10 15

Thr Thr Asp Ser Ser Val Pro Ser Ser Val Ser Val Ser Pro Leu Asp
20 25 30

Ser Pro Arg Gln Ser Pro Ser Ser Thr Ser Leu Ser Ser Met Ala Ser
35 40 45

Asp Ala Gly Lys Gly Asp Leu Gly Lys Met Leu Asp Thr Tyr Gly Asn
50 55 60

Glu Phe Lys Ile Pro Asp Tyr Thr Ile Lys Asp Ile Arg Asp Ala Ile
65 70 75 80

Pro Ser His Cys Tyr Asn Arg Ser Ala Ile Arg Ser Leu Ser Tyr Val
85 90 95

```
Phe Arg Asp Leu Ala Val Leu Ala Ser Val Phe Tyr Val Phe His Lys
            100                 105                 110

Tyr Val Thr Pro Glu Thr Val Pro Ser Tyr Pro Ala Arg Val Ala Leu
        115                 120                 125

Trp Thr Leu Tyr Thr Val Val Gln Gly Leu Phe Gly Thr Gly Ile Trp
    130                 135                 140

Val Leu Ala His Glu Cys Gly His Gln Ala Phe Ser Thr Ser Lys Val
145                 150                 155                 160

Leu Asn Asp Thr Val Gly Trp Ile Leu His Ser Ala Leu Leu Val Pro
                165                 170                 175

Tyr Phe Ser Trp Lys Ile Ser His Gly Lys His His Lys Ala Thr Gly
            180                 185                 190

Asn Leu Ala Arg Asp Met Val Phe Val Pro Lys Thr Arg Glu Val Tyr
        195                 200                 205

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

Glu Thr Pro Leu Ala Thr Ala Thr His Ser Ile Leu Gln Gln Leu Phe
225                 230                 235                 240

Gly Trp Pro Leu Tyr Leu Leu Thr Asn Val Thr Gly His Asp Asn His
                245                 250                 255

Glu Arg Gln Pro Glu Gly Arg Gly Lys Gly Lys Arg Asn Gly Tyr Phe
            260                 265                 270

Thr Gly Val Asn His Phe Asn Pro Asn Ser Pro Leu Phe Glu Ala Lys
        275                 280                 285

Asp Ala Lys Leu Ile Ile Leu Ser Asp Ile Gly Leu Ala Ile Thr Ala
    290                 295                 300
```

```
Ser Ile Leu Tyr Leu Ile Gly Ser Lys Phe Gly Trp Met Asn Leu Leu
305                 310                 315                 320

Val Trp Tyr Gly Ile Pro Tyr Leu Trp Val Asn His Trp Leu Val Ala
                325                 330                 335

Ile Thr Tyr Leu Gln His Thr Asp Pro Thr Leu Pro His Tyr Gln Pro
            340                 345                 350

Glu Ser Trp Thr Phe Ala Arg Gly Ala Ala Ala Thr Ile Asp Arg Glu
        355                 360                 365

Phe Gly Phe Ile Gly Arg His Ile Leu His Gly Ile Ile Glu Thr His
    370                 375                 380

Val Leu His His Tyr Val Ser Thr Ile Pro Phe Tyr His Ala Asp Glu
385                 390                 395                 400

Ala Ser Glu Ala Ile Lys Lys Val Met Gly Ser His Tyr Arg Ser Glu
                405                 410                 415

Ala His Thr Gly Pro Leu Gly Phe Leu Lys Ala Leu Trp Thr Ser Ala
            420                 425                 430

Arg Val Cys His Trp Val Glu Pro Thr Glu Gly Thr Lys Gly Glu Asn
        435                 440                 445

Ala Gly Val Leu Phe Phe Arg Asn Thr Asn Gly Ile Gly Val Pro Pro
    450                 455                 460

Ile Lys Leu Thr Lys Pro Asn
465                 470
```

<210> 36
<211> 466
<212> PRT
<213> Aspergillus flavus

<220>
<221> misc_feature
<223> delta-12 desaturase

<400> 36

```
Met Ser Ser Thr Ala Ile Pro Lys Arg Met Ala Leu Asn Arg Asn Pro
1               5                  10                  15

Gly Thr Asp Ser Ser Val Pro Ser Val Ser Val Ser Pro Phe Asp Ser
            20                  25                  30

Pro Arg His Ser Pro Ser Ser Thr Ser Leu Ser Ser Leu Ala Ser Glu
        35                  40                  45

Ser Glu Asn Lys Gly Lys Met Leu Asp Thr Tyr Gly Asn Glu Phe Lys
    50                  55                  60

Ile Pro Asp Tyr Thr Ile Lys Gln Ile Arg Asp Ala Ile Pro Ala His
65                  70                  75                  80

Cys Tyr Glu Arg Lys Ala Leu Thr Ser Leu Tyr Tyr Val Phe Arg Asp
                85                  90                  95

Ile Ala Met Leu Gly Ser Ile Phe Tyr Val Phe His Asn Tyr Val Thr
            100                 105                 110

Pro Glu Thr Val Pro Ser Phe Pro Ala Arg Val Ala Leu Trp Ser Leu
        115                 120                 125

Tyr Thr Val Val Gln Gly Leu Ile Ala Thr Gly Val Trp Val Leu Ala
    130                 135                 140

His Glu Cys Gly His Gln Ala Phe Ser Pro Ser Lys Val Leu Asn Asp
145                 150                 155                 160

Thr Val Gly Trp Ile Cys His Ser Ala Leu Leu Val Pro Tyr Phe Ser
                165                 170                 175

Trp Lys Ile Ser His Gly Lys His His Lys Ala Thr Gly Asn Ile Ala
            180                 185                 190
```

```
Arg Asp Met Val Phe Val Pro Lys Thr Arg Glu Glu Tyr Ala Ser Arg
        195                 200                 205

Ile Gly Lys Thr Ile His Asp Leu Asn Glu Leu Met Glu Glu Thr Pro
    210                 215                 220

Ile Ala Thr Val Thr Asn Leu Ile Leu Gln Gln Leu Phe Gly Trp Pro
225                 230                 235                 240

Met Tyr Leu Leu Thr Asn Val Thr Gly His Asn Asn His Glu Arg Gln
                245                 250                 255

Pro Glu Gly Arg Gly Lys Gly Lys Arg Asn Gly Tyr Phe Gly Gly Val
            260                 265                 270

Asn His Phe Asn Pro Ser Ser Pro Leu Tyr Glu Ala Lys Asp Ala Lys
        275                 280                 285

Leu Ile Val Leu Ser Asp Leu Gly Leu Ala Ile Thr Gly Ser Val Leu
    290                 295                 300

Tyr Tyr Ile Gly Ser Thr Tyr Gly Trp Leu Asn Leu Leu Val Trp Tyr
305                 310                 315                 320

Gly Ile Pro Tyr Leu Trp Val Asn His Trp Leu Val Ala Ile Thr Tyr
                325                 330                 335

Leu Gln His Thr Asp Pro Thr Leu Pro His Tyr Gln Pro Glu Val Trp
            340                 345                 350

Asn Phe Ala Arg Gly Ala Ala Ala Thr Ile Asp Arg Asp Phe Gly Phe
        355                 360                 365

Val Gly Arg His Ile Leu His Gly Ile Ile Glu Thr His Val Leu His
    370                 375                 380

His Tyr Val Ser Thr Ile Pro Phe Tyr His Ala Asp Glu Ala Ser Glu
385                 390                 395                 400
```

Ala Ile Gln Lys Val Met Gly Ser His Tyr Arg Thr Glu Ala His Thr
405 410 415

Gly Trp Thr Gly Phe Phe Lys Ala Leu Phe Thr Ser Ala Arg Val Cys
420 425 430

His Trp Val Glu Pro Thr Glu Gly Ala Arg Gly Glu Ser Glu Gly Val
435 440 445

Leu Phe Tyr Arg Asn Thr Asn Gly Ile Gly Val Pro Pro Ala Lys Leu
450 455 460

Ser Lys
465

<210> 37
<211> 424
<212> PRT
<213> Aspergillus fumigatus

<220>
<221> misc_feature
<223> delta-12 desaturase

<400> 37

Met Ala Ser Asp Ala Glu Lys Thr Ser Ser Lys Met Ile Asp Thr Tyr
1 5 10 15

Gly Asn Glu Phe Lys Ile Pro Asp Tyr Thr Ile Lys Gln Ile Arg Asp
20 25 30

Ala Ile Pro Ala His Cys Tyr Gln Arg Ser Ala Ala Thr Ser Leu Tyr
35 40 45

Tyr Val Phe Arg Asp Met Ala Ile Leu Ala Ser Val Phe Tyr Val Phe
50 55 60

His Asn Tyr Val Thr Pro Glu Thr Val Pro Ser Met Pro Val Arg Val
65 70 75 80

```
Val Leu Trp Thr Ile Tyr Thr Val Val Gln Gly Leu Val Gly Thr Gly
                85                  90                  95

Val Trp Val Leu Ala His Glu Cys Gly His Gln Ala Phe Ser Thr Ser
            100                 105                 110

Lys Val Leu Asn Asp Thr Val Gly Trp Ile Cys His Ser Leu Leu Leu
        115                 120                 125

Val Pro Tyr Phe Ser Trp Lys Ile Ser His Gly Lys His His Lys Ala
    130                 135                 140

Thr Gly Asn Ile Ala Arg Asp Met Val Phe Val Pro Lys Thr Arg Glu
145                 150                 155                 160

Glu Tyr Ala Thr Arg Ile Gly Arg Ala Ala His Glu Leu Ser Glu Leu
                165                 170                 175

Met Glu Glu Thr Pro Ile Leu Thr Ala Thr Asn Leu Val Leu Gln Gln
            180                 185                 190

Leu Phe Gly Trp Pro Met Tyr Leu Leu Thr Asn Val Thr Gly His Asn
        195                 200                 205

Asn His Glu Arg Gln Pro Glu Gly Arg Gly Lys Gly Lys Arg Asn Gly
    210                 215                 220

Tyr Phe Gly Gly Val Asn His Phe Asn Pro Ser Ser Pro Leu Tyr Glu
225                 230                 235                 240

Ala Lys Asp Ala Lys Leu Ile Val Leu Ser Asp Leu Gly Leu Phe Leu
                245                 250                 255

Val Gly Ser Leu Leu Tyr Tyr Ile Gly Ser Thr Tyr Gly Trp Leu Asn
            260                 265                 270

Leu Leu Val Trp Tyr Gly Ile Pro Tyr Leu Trp Val Asn His Trp Leu
        275                 280                 285
```

```
Val Ala Ile Thr Phe Leu Gln His Thr Asp Pro Thr Leu Pro His Tyr
    290                 295                 300

Gln Pro Glu Ala Trp Asp Phe Thr Arg Gly Ala Ala Ala Thr Ile Asp
305                 310                 315                 320

Arg Asp Phe Gly Phe Val Gly Arg His Ile Phe His Gly Ile Ile Glu
                325                 330                 335

Thr His Val Leu His His Tyr Val Ser Thr Ile Pro Phe Tyr His Ala
            340                 345                 350

Asp Glu Ala Ser Glu Ala Ile Gln Lys Val Met Gly Pro His Tyr Arg
        355                 360                 365

Ser Glu Ala His Thr Gly Trp Thr Gly Phe Leu Lys Ala Leu Trp Thr
    370                 375                 380

Ser Ala Arg Thr Cys Gln Trp Val Glu Pro Thr Glu Gly Ala Lys Gly
385                 390                 395                 400

Glu Ser Gln Tyr Val Leu Phe Tyr Arg Asn Ile Asn Gly Ile Gly Val
                405                 410                 415

Pro Pro Ala Lys Ile Pro Ala Lys
            420
```

<210> 38
<211> 1656
<212> DNA
<213> Magnaporthe grisea

<220>
<221> misc_feature
<223> delta-12 desaturase

<400> 38

```
atgccttcca ccagatcgac cacatcgggc attgcccagg agaagactcc tatgaggcga     60

acgaccacct cggccactgt cgaatcggac gtctcagctc cgggaaccgc tgttcagtcg    120
```

cctatggact cgccccgcca ctctgcctcg tccacctcac tctcttcact ctcttccgtt 180

gatgcggcgg ccgagaagaa atccaacgag tctgtcggca aactcgtcga cacgtacggc 240

aacacctttg agatccctga cttcaccatc aaggacatcc acgatgccat tccaaagcac 300

tgctttgaac gctctgctat tcgtagcttg agctacgtcg cccgtgatat ggtcctcctg 360

gcgacgacct tctacgtgtt ccacaactac gtgacaccag agtacattcc ctcgaagccg 420

gctcgtgctg gtctgtgggc catttacacg gtgctccagg gcctcttcgg caccggaatc 480

tgggttcttg cccatgagtg tggccaccag gctttctcgc cttccaagac catcaacaac 540

acggttggct ggattctcca ctcgtctctg ctggttccgt acttcagctg gcagatgtca 600

cacagcaagc accacaaggc cactggccat attgagcgcg acatggtctt tgtgccccgc 660

acccgggagg agcacgccag caggatcggc cgcatggtcc acgagctgtc ggagttgacc 720

gaggagacgc ctattgccac ccttatccac ttggttgggc agcagctgat cggctggcct 780

ctgtacatca tcactaacaa gaccggtcac aactaccacg agcgccagcg tgagggccgt 840

ggcaagggca agaagaacgg tcttttcact ggcgtcaacc acttcaaccc cagcagccct 900

ctgtacgaga acaaggacgc cggaaaggtg cttctcagcg acctgggtgt cggccttgtt 960

atcgctggcc tcgtgtacct ttgccaaact ttcggcaccc agaacatgct ggtttggtac 1020

tttatcccct acctctgggt gaaccactgg ctcgttgcca ttacattcct tcagcacacc 1080

gacccctcgc ttccgcacta tactgccgag gaatggaact tcgtccgagg tgccgctgcc 1140

acgatcgatc gcgagtttgg cttcgtcggc cgccacctgc ttcacggtat cattgagacc 1200

cacgtcctgc accactatgt cagcacgatc ccctttttaca acgccgacga ggctactgat 1260

gccatcaaga aggtgatggg caagcactac cgcagcgaca ctgccggcgg ccctgctggc 1320

ttccttaagt cactctggac gagtagccgc atgtgccaat gggttgagcc cagcgccgag 1380

gctgagggta gtggcaaggg tgtcctgttc ttccgcaacc acaacaagat cggcactcct 1440

cctatcaaga tgtctgctca gaaaattaga ctatgcaatg accttcttgg catgcataag 1500

ggaaagaatc aaatgaatgg atcaagggag cgccgcggcg gacaaagtag tttaaagagg 1560

gtgagaaatc agcgatcgac aaatatgaac gaatcacaca tgacggtgtt ccgggcattc 1620

cgaacttgga gctcatgcac gcgcgcgtcc acatga 1656

<210> 39
<211> 551
<212> PRT
<213> Magnaporthe grisea

<400> 39

```
Met Pro Ser Thr Arg Ser Thr Thr Ser Gly Ile Ala Gln Glu Lys Thr
1               5                  10                  15

Pro Met Arg Arg Thr Thr Thr Ser Ala Thr Val Glu Ser Asp Val Ser
            20                  25                  30

Ala Pro Gly Thr Ala Val Gln Ser Pro Met Asp Ser Pro Arg His Ser
        35                  40                  45

Ala Ser Ser Thr Ser Leu Ser Ser Leu Ser Ser Val Asp Ala Ala Ala
    50                  55                  60

Glu Lys Lys Ser Asn Glu Ser Val Gly Lys Leu Val Asp Thr Tyr Gly
65                  70                  75                  80

Asn Thr Phe Glu Ile Pro Asp Phe Thr Ile Lys Asp Ile His Asp Ala
                85                  90                  95

Ile Pro Lys His Cys Phe Glu Arg Ser Ala Ile Arg Ser Leu Ser Tyr
            100                 105                 110

Val Ala Arg Asp Met Val Leu Leu Ala Thr Thr Phe Tyr Val Phe His
        115                 120                 125

Asn Tyr Val Thr Pro Glu Tyr Ile Pro Ser Lys Pro Ala Arg Ala Gly
    130                 135                 140

Leu Trp Ala Ile Tyr Thr Val Leu Gln Gly Leu Phe Gly Thr Gly Ile
145                 150                 155                 160

Trp Val Leu Ala His Glu Cys Gly His Gln Ala Phe Ser Pro Ser Lys
                165                 170                 175

Thr Ile Asn Asn Thr Val Gly Trp Ile Leu His Ser Ser Leu Leu Val
            180                 185                 190
```

```
Pro Tyr Phe Ser Trp Gln Met Ser His Ser Lys His His Lys Ala Thr
        195                 200                 205

Gly His Ile Glu Arg Asp Met Val Phe Val Pro Arg Thr Arg Glu Glu
    210                 215                 220

His Ala Ser Arg Ile Gly Arg Met Val His Glu Leu Ser Glu Leu Thr
225                 230                 235                 240

Glu Glu Thr Pro Ile Ala Thr Leu Ile His Leu Val Gly Gln Gln Leu
                245                 250                 255

Ile Gly Trp Pro Leu Tyr Ile Ile Thr Asn Lys Thr Gly His Asn Tyr
            260                 265                 270

His Glu Arg Gln Arg Glu Gly Arg Gly Lys Gly Lys Lys Asn Gly Leu
        275                 280                 285

Phe Thr Gly Val Asn His Phe Asn Pro Ser Ser Pro Leu Tyr Glu Asn
    290                 295                 300

Lys Asp Ala Gly Lys Val Leu Leu Ser Asp Leu Gly Val Gly Leu Val
305                 310                 315                 320

Ile Ala Gly Leu Val Tyr Leu Cys Gln Thr Phe Gly Thr Gln Asn Met
                325                 330                 335

Leu Val Trp Tyr Phe Ile Pro Tyr Leu Trp Val Asn His Trp Leu Val
            340                 345                 350

Ala Ile Thr Phe Leu Gln His Thr Asp Pro Ser Leu Pro His Tyr Thr
        355                 360                 365

Ala Glu Glu Trp Asn Phe Val Arg Gly Ala Ala Ala Thr Ile Asp Arg
    370                 375                 380

Glu Phe Gly Phe Val Gly Arg His Leu Leu His Gly Ile Ile Glu Thr
385                 390                 395                 400
```

```
His Val Leu His His Tyr Val Ser Thr Ile Pro Phe Tyr Asn Ala Asp
                405                 410                 415

Glu Ala Thr Asp Ala Ile Lys Lys Val Met Gly Lys His Tyr Arg Ser
            420                 425                 430

Asp Thr Ala Gly Gly Pro Ala Gly Phe Leu Lys Ser Leu Trp Thr Ser
        435                 440                 445

Ser Arg Met Cys Gln Trp Val Glu Pro Ser Ala Glu Ala Glu Gly Ser
    450                 455                 460

Gly Lys Gly Val Leu Phe Phe Arg Asn His Asn Lys Ile Gly Thr Pro
465                 470                 475                 480

Pro Ile Lys Met Ser Ala Gln Lys Ile Arg Leu Cys Asn Asp Leu Leu
                485                 490                 495

Gly Met His Lys Gly Lys Asn Gln Met Asn Gly Ser Arg Glu Arg Arg
            500                 505                 510

Gly Gly Gln Ser Ser Leu Lys Arg Val Arg Asn Gln Arg Ser Thr Asn
        515                 520                 525

Met Asn Glu Ser His Met Thr Val Phe Arg Ala Phe Arg Thr Trp Ser
    530                 535                 540

Ser Cys Thr Arg Ala Ser Thr
545                 550
```

<210> 40
<211> 1446
<212> DNA
<213> Neurospora crassa

<220>
<221> misc_feature
<223> delta-12 desaturase

<400> 40

atggcgtccg tctcctctgc ccttcccgag ggcaacaagc ctgccctgcg caggacccaa 60

accgaggcca cctccgactc ataccctggt accgctgatg cctctccctt cgactctccc 120

cttgagcgct cggcctccaa cacctcgctt tcttcccagg cctctgacaa cgtcaagacc 180

gacaaggccg agttcggcaa gctgctcgac acgtatggca acgagttcga ggtccccgac 240

ttcaccatca aggacatccg cgatgccatc cccgcccact gctttgagcg ttcggctctt 300

cacagcttgg cgcacgtcgt ccgcgacatc atttacctca ccgtcacttt ttacgtctgg 360

aacaagtatg tcactcccga gtacatcccc atgaaggctg cccgtgtcgt cctctggggt 420

ctgtacacct tcatgcaggg ccttttcggc accggtctct gggttcttgc ccatgagtgc 480

ggtcaccagg ctttctcccc gtccaggttg atcaacgaca ccgtcggctg ggtcctccac 540

tctgcccttc tcgtccccta cttctcgtgg aagttctccc acagcaagca ccacaaggcc 600

accggcaaca tcgagcgtga catggtcttc gttcctcgga cccgcgagca gtttgcgtct 660

cgcatcggcc gtttcgtcca tgagatttcc gagttgaccg aggagacccc catctacacc 720

ttgatccacc ttatcggtca gcagctcatc ggctggccca actacctcat gaccaacgtc 780

accggccaca acttccacga gaggcagcgc gagggtcgtg gcaagggcaa gaagaacggc 840

tggttcactg gtgtcaacca cttcaacccc agctctcccc tctatgagga gcgtgaggcc 900

ccctggatca tcgtctccga catcggtatc gctatcgccg ccaccgccct catctacctc 960

ggcaacacct tcggctggtc caacatgttc gtctggtact tccttcccta cctctgggtc 1020

aaccactggc ttgttgccat cacctacctc cagcacaccg acccctcgct cccccactac 1080

acccctgatc agtggaactt tgtccgtggt gccgccgcga ctattgaccg cgagttcggc 1140

ttcatcggcc gtcacctcct ccacggcatt atcgagaccc acgttctcca ccactacgtc 1200

agcaccattc ccttttacca cgccgacgag gcctccgagg ccatcaagaa ggtcatgggc 1260

cgtcactacc gcgctgacgt ccaagatggc cccatcggtt tcatcaaggc catgtggaag 1320

gctgctcgtt ggtgccagtg ggttgagcct accgagggcg ctgagggtaa gggcaagggc 1380

gtcttgttct accgcaacca gaacggtctc ggtgtcaagc ctgccaagct ccccaaaacc 1440

aactaa 1446

<210> 41

<211> 481
<212> PRT
<213> Neurospora crassa

<400> 41

```
Met Ala Ser Val Ser Ser Ala Leu Pro Glu Gly Asn Lys Pro Ala Leu
1               5                  10                  15

Arg Arg Thr Gln Thr Glu Ala Thr Ser Asp Ser Tyr Pro Gly Thr Ala
            20                  25                  30

Asp Ala Ser Pro Phe Asp Ser Pro Leu Glu Arg Ser Ala Ser Asn Thr
        35                  40                  45

Ser Leu Ser Ser Gln Ala Ser Asp Asn Val Lys Thr Asp Lys Ala Glu
    50                  55                  60

Phe Gly Lys Leu Leu Asp Thr Tyr Gly Asn Glu Phe Glu Val Pro Asp
65                  70                  75                  80

Phe Thr Ile Lys Asp Ile Arg Asp Ala Ile Pro Ala His Cys Phe Glu
                85                  90                  95

Arg Ser Ala Leu His Ser Leu Ala His Val Val Arg Asp Ile Ile Tyr
            100                 105                 110

Leu Thr Val Thr Phe Tyr Val Trp Asn Lys Tyr Val Thr Pro Glu Tyr
        115                 120                 125

Ile Pro Met Lys Ala Ala Arg Val Val Leu Trp Gly Leu Tyr Thr Phe
    130                 135                 140

Met Gln Gly Leu Phe Gly Thr Gly Leu Trp Val Leu Ala His Glu Cys
145                 150                 155                 160

Gly His Gln Ala Phe Ser Pro Ser Arg Leu Ile Asn Asp Thr Val Gly
                165                 170                 175

Trp Val Leu His Ser Ala Leu Leu Val Pro Tyr Phe Ser Trp Lys Phe
            180                 185                 190
```

```
Ser His Ser Lys His His Lys Ala Thr Gly Asn Ile Glu Arg Asp Met
        195                 200                 205

Val Phe Val Pro Arg Thr Arg Glu Gln Phe Ala Ser Arg Ile Gly Arg
    210                 215                 220

Phe Val His Glu Ile Ser Glu Leu Thr Glu Glu Thr Pro Ile Tyr Thr
225                 230                 235                 240

Leu Ile His Leu Ile Gly Gln Gln Leu Ile Gly Trp Pro Asn Tyr Leu
                245                 250                 255

Met Thr Asn Val Thr Gly His Asn Phe His Glu Arg Gln Arg Glu Gly
            260                 265                 270

Arg Gly Lys Gly Lys Lys Asn Gly Trp Phe Thr Gly Val Asn His Phe
        275                 280                 285

Asn Pro Ser Ser Pro Leu Tyr Glu Glu Arg Glu Ala Pro Trp Ile Ile
    290                 295                 300

Val Ser Asp Ile Gly Ile Ala Ile Ala Ala Thr Ala Leu Ile Tyr Leu
305                 310                 315                 320

Gly Asn Thr Phe Gly Trp Ser Asn Met Phe Val Trp Tyr Phe Leu Pro
                325                 330                 335

Tyr Leu Trp Val Asn His Trp Leu Val Ala Ile Thr Tyr Leu Gln His
            340                 345                 350

Thr Asp Pro Ser Leu Pro His Tyr Thr Pro Asp Gln Trp Asn Phe Val
        355                 360                 365

Arg Gly Ala Ala Ala Thr Ile Asp Arg Glu Phe Gly Phe Ile Gly Arg
    370                 375                 380

His Leu Leu His Gly Ile Ile Glu Thr His Val Leu His His Tyr Val
385                 390                 395                 400
```

Ser Thr Ile Pro Phe Tyr His Ala Asp Glu Ala Ser Glu Ala Ile Lys
405 410 415

Lys Val Met Gly Arg His Tyr Arg Ala Asp Val Gln Asp Gly Pro Ile
420 425 430

Gly Phe Ile Lys Ala Met Trp Lys Ala Ala Arg Trp Cys Gln Trp Val
435 440 445

Glu Pro Thr Glu Gly Ala Glu Gly Lys Gly Lys Gly Val Leu Phe Tyr
450 455 460

Arg Asn Gln Asn Gly Leu Gly Val Lys Pro Ala Lys Leu Pro Lys Thr
465 470 475 480

Asn

<210> 42
<211> 1371
<212> DNA
<213> Fusarium graminearium

<220>
<221> misc_feature
<223> delta-12 desaturase

<400> 42
tcaaccacgg cgacggatac tgagtctgct gccgtttctc cttcagactc tccccgccat 60

tcggcctctt ccacctcgct ctcgtctctt tccgagattg atatcgccaa gcccaaggcc 120

gaatatggtg ttatgcttga cacctatggc aacaagttcg aggttcccga cttcaccatc 180

aaggagatct acaatgccat ccccaagcac tgcttccagc gctccgctct caagggatac 240

ggatacatcc tccgcgacat tgtccttctt gctaccacct ttagcatctg gtacaactat 300

gtgacccccg agtacatccc tagcactccc gcccgcgctg gtctctgggc tgtctacact 360

gttctccagg gtcttttcgg taccggtctc tgggtcatcg ctcacgagtg tggccacggt 420

gctttctccg actctcgcct tatcaacgac atcaccggct gggtcctcca ctcttctctc 480

ctcgtcccct acttcagctg gcaaatctcc caccgaaagc accacaaggc taccggaaac 540

atggagcgtg acatggtctt tgttccccga actcgcgagc agcaggctac tcgtctcggc 600

aagatgaccc acgagcttgc tcacctcact gaggagaccc ccgtcttcac tctgatcatg 660

cttgttctcc agcagctcgt cggctggccc aactacctca tgaccaacgt tactggccac 720

aactaccacg agcgtcagaa ggagggccgt ggcaagggca agcacaacgg tctcggcggc 780

ggtgtcaacc actttgatcc ccgcagccct ctttacgagc acagcgatgc taagctcatt 840

gtcttgagtg atattggtat cggtctgatg ggtaccgctc tgtacttcct cgtccagaag 900

tttggctttt acaacatggc catctggtac tttgtccctt acctttgggt caaccactgg 960

ctcgtcgcca ttactttcct ccagcacacc gaccctaccc ttccccacta caccaacgac 1020

gagtggaact ttgtccgcgg tgctgctgct accatcgatc gtgagatggg tttcattggc 1080

cgacacctcc tccacggtat catcgagact cacgtcctcc accactacgt cagcagcatc 1140

cccttctaca acgccgacga ggctaccgag gctatcaagc ctgtcatggg caagcactac 1200

cgtgccgacg tccaggatgg tccccgtggt ttcattcgtg ccatgtaccg cagtgcccgt 1260

atgtgccagt gggttgagcc cagcgctgag gccgagggtg ctggcaaggg tgttctgttc 1320

ttccgcaacc gcaacaaggt tggcactgct cctgccgtcc tcaaggctta g 1371

<210> 43
<211> 456
<212> PRT
<213> Fusarium graminearium

<400> 43

Ser Thr Thr Ala Thr Asp Thr Glu Ser Ala Ala Val Ser Pro Ser Asp
1 5 10 15

Ser Pro Arg His Ser Ala Ser Ser Thr Ser Leu Ser Ser Leu Ser Glu
20 25 30

Ile Asp Ile Ala Lys Pro Lys Ala Glu Tyr Gly Val Met Leu Asp Thr
35 40 45

Tyr Gly Asn Lys Phe Glu Val Pro Asp Phe Thr Ile Lys Glu Ile Tyr
50 55 60

```
Asn Ala Ile Pro Lys His Cys Phe Gln Arg Ser Ala Leu Lys Gly Tyr
65                  70                  75                  80

Gly Tyr Ile Leu Arg Asp Ile Val Leu Leu Ala Thr Thr Phe Ser Ile
                85                  90                  95

Trp Tyr Asn Tyr Val Thr Pro Glu Tyr Ile Pro Ser Thr Pro Ala Arg
            100                 105                 110

Ala Gly Leu Trp Ala Val Tyr Thr Val Leu Gln Gly Leu Phe Gly Thr
        115                 120                 125

Gly Leu Trp Val Ile Ala His Glu Cys Gly His Gly Ala Phe Ser Asp
    130                 135                 140

Ser Arg Leu Ile Asn Asp Ile Thr Gly Trp Val Leu His Ser Ser Leu
145                 150                 155                 160

Leu Val Pro Tyr Phe Ser Trp Gln Ile Ser His Arg Lys His His Lys
                165                 170                 175

Ala Thr Gly Asn Met Glu Arg Asp Met Val Phe Val Pro Arg Thr Arg
            180                 185                 190

Glu Gln Gln Ala Thr Arg Leu Gly Lys Met Thr His Glu Leu Ala His
        195                 200                 205

Leu Thr Glu Glu Thr Pro Val Phe Thr Leu Ile Met Leu Val Leu Gln
    210                 215                 220

Gln Leu Val Gly Trp Pro Asn Tyr Leu Met Thr Asn Val Thr Gly His
225                 230                 235                 240

Asn Tyr His Glu Arg Gln Lys Glu Gly Arg Gly Lys Gly Lys His Asn
                245                 250                 255

Gly Leu Gly Gly Gly Val Asn His Phe Asp Pro Arg Ser Pro Leu Tyr
            260                 265                 270
```

```
Glu His Ser Asp Ala Lys Leu Ile Val Leu Ser Asp Ile Gly Ile Gly
        275                 280                 285

Leu Met Gly Thr Ala Leu Tyr Phe Leu Val Gln Lys Phe Gly Phe Tyr
    290                 295                 300

Asn Met Ala Ile Trp Tyr Phe Val Pro Tyr Leu Trp Val Asn His Trp
305                 310                 315                 320

Leu Val Ala Ile Thr Phe Leu Gln His Thr Asp Pro Thr Leu Pro His
                325                 330                 335

Tyr Thr Asn Asp Glu Trp Asn Phe Val Arg Gly Ala Ala Ala Thr Ile
            340                 345                 350

Asp Arg Glu Met Gly Phe Ile Gly Arg His Leu Leu His Gly Ile Ile
        355                 360                 365

Glu Thr His Val Leu His His Tyr Val Ser Ser Ile Pro Phe Tyr Asn
    370                 375                 380

Ala Asp Glu Ala Thr Glu Ala Ile Lys Pro Val Met Gly Lys His Tyr
385                 390                 395                 400

Arg Ala Asp Val Gln Asp Gly Pro Arg Gly Phe Ile Arg Ala Met Tyr
                405                 410                 415

Arg Ser Ala Arg Met Cys Gln Trp Val Glu Pro Ser Ala Glu Ala Glu
            420                 425                 430

Gly Ala Gly Lys Gly Val Leu Phe Phe Arg Asn Arg Asn Lys Val Gly
        435                 440                 445

Thr Ala Pro Ala Val Leu Lys Ala
    450                 455
```

<210> 44
<211> 1403
<212> DNA
<213> Mortierella alpina (GenBank Accession No. AB020033)

<220>
<221> misc_feature
<223> delta-12 desaturase

<400> 44
ctcctcccac caccctccga gacgactgca actgtaatca ggaaccgaaa aatatacgat 60

ttctttttac tcagcaccaa ctcaaaatct tcagcagcaa ccctttttca ggatggcacc 120

tcccaacact attgatgccg gtttgaccca gcgccatatc agcacctcgg ccgccccaac 180

ctctgccaag cccgccttcg agcgcaacta ccagctccct gagttcacca tcaaggagat 240

ccgtgagtgc atccctgcac actgctttga gcgctccggt ctccgtggtc tttgccacgt 300

tgctattgat ctgacctggg cctcgctctt gttcctggct gcgacccaga tcgacaagtt 360

cgagaaccct ttgatccgct acttggcctg gcctgcgtat tggatcatgc agggtattgt 420

ttgcaccggt atctgggtat tggcacacga atgtggtcat cagtccttct cgacctccaa 480

gacccttaac aacactgtcg gctggatctt gcactcgatg ctcttggtcc cttaccactc 540

ctggagaatc tcgcactcga agcaccacaa ggccactggc cacatgacca aggaccaggt 600

ctttgttccc aagacccgct ctcaggttgg cttgcccccc aaggagaatg ttgcagttgc 660

cgttcaggag gaggatatgt ccgtgcacct ggatgaggag gcccccattg tgactttgtt 720

ctggatggtg attcagttcc tgttcggatg gcctgcgtac cttattatga acgcctctgg 780

tcaagactat ggccgctgga cctcgcactt ccacacctac tctcctatct ttgagccccg 840

caacttttc gacattatca tttcggatct cggtgtgttg gctgctcttg gtaccttgat 900

ctacgcctcc atgcagctct cgctcttgac cgtgaccaag tactacattg tcccctactt 960

gtttgtcaac ttctggttgg tcctgatcac cttcttgcag cacaccgacc ctaagctgcc 1020

ccattaccgt gagggtgcct ggaacttcca gcgtggagcc ctctgcaccg ttgaccgctc 1080

gttcggcaag ttcttggacc atatgttcca cggcattgtc catacccatg tagcccatca 1140

cttgttctcg cagatgccgt tctaccatgc tgaggaagcc acccatcatc tcaagaaact 1200

gctgggagag tactacgtct atgacccatc gccgattgtt gttgcggtct ggaggtcgtt 1260

ccgtgaatgc cgattcgtgg aagaccatgg agacgtggtc tttttcaaga agtaaagaag 1320

catttaccgg acacaacttt gtctttacac gcctacgtac tgagtagcca tcccacttta 1380

taaaagattt cagttccagc agc 1403

<210> 45
<211> 400
<212> PRT
<213> Mortierella alpina (GenBank Accession No. AB020033)

<220>
<221> MISC_FEATURE
<223> delta-12 desaturase

<400> 45

Met Ala Pro Pro Asn Thr Ile Asp Ala Gly Leu Thr Gln Arg His Ile
1 5 10 15

Ser Thr Ser Ala Ala Pro Thr Ser Ala Lys Pro Ala Phe Glu Arg Asn
20 25 30

Tyr Gln Leu Pro Glu Phe Thr Ile Lys Glu Ile Arg Glu Cys Ile Pro
35 40 45

Ala His Cys Phe Glu Arg Ser Gly Leu Arg Gly Leu Cys His Val Ala
50 55 60

Ile Asp Leu Thr Trp Ala Ser Leu Leu Phe Leu Ala Ala Thr Gln Ile
65 70 75 80

Asp Lys Phe Glu Asn Pro Leu Ile Arg Tyr Leu Ala Trp Pro Ala Tyr
85 90 95

Trp Ile Met Gln Gly Ile Val Cys Thr Gly Ile Trp Val Leu Ala His
100 105 110

Glu Cys Gly His Gln Ser Phe Ser Thr Ser Lys Thr Leu Asn Asn Thr
115 120 125

Val Gly Trp Ile Leu His Ser Met Leu Leu Val Pro Tyr His Ser Trp
130 135 140

Arg Ile Ser His Ser Lys His His Lys Ala Thr Gly His Met Thr Lys
145 150 155 160

```
Asp Gln Val Phe Val Pro Lys Thr Arg Ser Gln Val Gly Leu Pro Pro
                165                 170                 175

Lys Glu Asn Val Ala Val Ala Val Gln Glu Glu Asp Met Ser Val His
            180                 185                 190

Leu Asp Glu Glu Ala Pro Ile Val Thr Leu Phe Trp Met Val Ile Gln
        195                 200                 205

Phe Leu Phe Gly Trp Pro Ala Tyr Leu Ile Met Asn Ala Ser Gly Gln
    210                 215                 220

Asp Tyr Gly Arg Trp Thr Ser His Phe His Thr Tyr Ser Pro Ile Phe
225                 230                 235                 240

Glu Pro Arg Asn Phe Phe Asp Ile Ile Ile Ser Asp Leu Gly Val Leu
                245                 250                 255

Ala Ala Leu Gly Thr Leu Ile Tyr Ala Ser Met Gln Leu Ser Leu Leu
            260                 265                 270

Thr Val Thr Lys Tyr Tyr Ile Val Pro Tyr Leu Phe Val Asn Phe Trp
        275                 280                 285

Leu Val Leu Ile Thr Phe Leu Gln His Thr Asp Pro Lys Leu Pro His
    290                 295                 300

Tyr Arg Glu Gly Ala Trp Asn Phe Gln Arg Gly Ala Leu Cys Thr Val
305                 310                 315                 320

Asp Arg Ser Phe Gly Lys Phe Leu Asp His Met Phe His Gly Ile Val
                325                 330                 335

His Thr His Val Ala His His Leu Phe Ser Gln Met Pro Phe Tyr His
            340                 345                 350

Ala Glu Glu Ala Thr His His Leu Lys Lys Leu Leu Gly Glu Tyr Tyr
        355                 360                 365
```

Val Tyr Asp Pro Ser Pro Ile Val Val Ala Val Trp Arg Ser Phe Arg
370 375 380

Glu Cys Arg Phe Val Glu Asp His Gly Asp Val Val Phe Phe Lys Lys
385 390 395 400

<210> 46
<211> 416
<212> PRT
<213> Saccharomyces kluyveri (GenBank Accession No. BAD08375)

<220>
<221> MISC_FEATURE
<223> delta-12 desaturase

<400> 46

Met Ser Ala Val Thr Val Thr Gly Ser Asp Pro Lys Asn Arg Gly Ser
1 5 10 15

Ser Ser Asn Thr Glu Gln Glu Val Pro Lys Val Ala Ile Asp Thr Asn
20 25 30

Gly Asn Val Phe Ser Val Pro Asp Phe Thr Ile Lys Asp Ile Leu Gly
35 40 45

Ala Ile Pro His Glu Cys Tyr Glu Arg Arg Leu Ala Thr Ser Leu Tyr
50 55 60

Tyr Val Phe Arg Asp Ile Phe Cys Met Leu Thr Thr Gly Tyr Leu Thr
65 70 75 80

His Lys Ile Leu Tyr Pro Leu Leu Ile Ser Tyr Thr Ser Asn Ser Ile
85 90 95

Ile Lys Phe Thr Phe Trp Ala Leu Tyr Thr Tyr Val Gln Gly Leu Phe
100 105 110

Gly Thr Gly Ile Trp Val Leu Ala His Glu Cys Gly His Gln Ala Phe
115 120 125

```
Ser Asp Tyr Gly Ile Val Asn Asp Phe Val Gly Trp Thr Leu His Ser
    130                 135                 140

Tyr Leu Met Val Pro Tyr Phe Ser Trp Lys Tyr Ser His Gly Lys His
145                 150                 155                 160

His Lys Ala Thr Gly His Met Thr Arg Asp Met Val Phe Val Pro Ala
                165                 170                 175

Thr Lys Glu Glu Phe Lys Lys Ser Arg Asn Phe Phe Gly Asn Leu Ala
            180                 185                 190

Glu Tyr Ser Glu Asp Ser Pro Leu Arg Thr Leu Tyr Glu Leu Leu Val
        195                 200                 205

Gln Gln Leu Gly Gly Trp Ile Ala Tyr Leu Phe Val Asn Val Thr Gly
    210                 215                 220

Gln Pro Tyr Pro Asp Val Pro Ser Trp Lys Trp Asn His Phe Trp Leu
225                 230                 235                 240

Thr Ser Pro Leu Phe Glu Gln Arg Asp Ala Leu Tyr Ile Phe Leu Ser
                245                 250                 255

Asp Leu Gly Ile Leu Thr Gln Gly Ile Val Leu Thr Leu Trp Tyr Lys
            260                 265                 270

Lys Phe Gly Gly Trp Ser Leu Phe Ile Asn Trp Phe Val Pro Tyr Ile
        275                 280                 285

Trp Val Asn His Trp Leu Val Phe Ile Thr Phe Leu Gln His Thr Asp
    290                 295                 300

Pro Thr Met Pro His Tyr Asn Ala Glu Glu Trp Thr Phe Ala Lys Gly
305                 310                 315                 320

Ala Ala Ala Thr Ile Asp Arg Lys Phe Gly Phe Ile Gly Pro His Ile
                325                 330                 335
```

Phe His Asp Ile Ile Glu Thr His Val Leu His His Tyr Cys Ser Arg
340 345 350

Ile Pro Phe Tyr Asn Ala Arg Pro Ala Ser Glu Ala Ile Lys Lys Val
355 360 365

Met Gly Lys His Tyr Arg Ser Ser Asp Glu Asn Met Trp Lys Ser Leu
370 375 380

Trp Lys Ser Phe Arg Ser Cys Gln Tyr Val Asp Gly Asp Asn Gly Val
385 390 395 400

Leu Met Phe Arg Asn Ile Asn Asn Cys Gly Val Gly Ala Ala Glu Lys
405 410 415

<210> 47
<211> 1948
<212> DNA
<213> Kluyveromyces lactis (gnl|GLV|KLLA0B00473g)

<220>
<221> misc_feature
<223> delta-12 desaturase

<400> 47
gaaaaaaaac aaacacagtc tcgaagaaga ttcagacaaa cagtagcaca agactatttt 60

ctacaccgac gaatgttcgt actaccgcgt tacttggcga acaccgccca tgcttaatgt 120

gaaaaagcac acacattgat gcgtgtagct agaccaattg caacccatta gatttactgc 180

agaaagagga aaagtgctac ttttttttcag tcaatggtat cacaattggc gatgagcctt 240

gacgaaagga agcagcgata aaaagaaaat ggattaaaaa aaatgtcctg cttagttatg 300

ggcaattgat agttttactt caaatatcgt ttgaagattt ccgttcagtc agtttccatt 360

cacatttggc ttcagcttat ataaatataa tatataaaga attttatgat catgatatat 420

gatcattact ttacctaagt ggaagattaa gtggtagttg tacatcatat acacccgcag 480

aacgtagcta gacagaccca atgagcaaaa gcaccggcgt cgagcatcat atcagtggtg 540

ttgctactac agaaacggca acagagacag ttactgttcc tcctgccaag acagctattg 600

acacacatgg taacatcttc aaggttccag actatactat caaagacatc ctaggtgcca 660

tcccaaagga gtgttacaag agagacactc tttggtcatt gcattatgtg gtaagagata 720

ttattgccat ctgtattatt gggtatgttg gtaccaacta cattccagtt tggtttccaa 780

atagtggtct cctaagattt gttgcctata tggtgcaatc atacttaatt gggttatttg 840

gttttgggtt gtggattttg gctcacgaat gcggtcatgg tgccttctct gattcaagat 900

taatcaacga taccgttggg tgggttttgc actcatggtg gatggtccca tacttttctt 960

ggaaattctc tcactctaag catcataaag ctaccggtca tttgactaga gatatggttt 1020

ttgttccata cacgaaaaaa gaatacttgg aaatgaaggg taaatcaaaa ctaagggaga 1080

tcaccgaaga agcaccaatt gtaactttat taactttgat tggccaacaa attggtggtt 1140

tgcagttgta tttagcaaca aatgccactg gccaatctta tcctggtgtt cctaaattct 1200

tcaaatccca ttactggcct acttctccgg tgttcgatac caaggacttc tggtacatca 1260

ttctgagtga cattggtatt atttccacac ttacaatcaa ttatttgtgg gcaaagacct 1320

atggctccca tgtcatgttg atcaactggt ttgttccatg gttatgggtt aaccactggt 1380

tagtatttgt cacatttttg caacacactg acccaaccat gcctcattac gaagccagtg 1440

aatggacctt cgctaaaggt gctgctgcca ccattgatag aaactttggc tttgttggtc 1500

aacatatttt ccacgatatt attgaaacgc atgttttgca tcattactgt tcaagaattc 1560

ctttctacaa cgctcgtgta gccactgaag ctattaagaa agttatgggt gaacattatc 1620

gttatgaagg cgaaaatatg tggcaatcat tatggaaggt ggcaagatcc tgtcagtttg 1680

tagatggtga caacggtgtt ttgatgttca gaaacacaaa cggagtcgga gcaccatgtc 1740

aagaatgagt gcacatagag tatcaatttt tcattcatca agttcattcc cttatcatta 1800

tcattcttat tattatttat cttaactgta cattaaaatt ttacccgatg caaaatatca 1860

accgaatgtg cttatgccac gtcggttaat tggaaaatat taatcttacc agtcaaattg 1920

ctttactctc aaaatttttg ataggagt 1948

<210> 48
<211> 415
<212> PRT
<213> Kluyveromyces lactis (gnl|GLV|KLLA0B00473g)

<220>
<221> MISC_FEATURE
<223> delta-12 desaturase

<400> 48

```
Met Ser Lys Ser Thr Gly Val Glu His His Ile Ser Gly Val Ala Thr
1               5                  10                  15

Thr Glu Thr Ala Thr Glu Thr Val Thr Val Pro Pro Ala Lys Thr Ala
            20                  25                  30

Ile Asp Thr His Gly Asn Ile Phe Lys Val Pro Asp Tyr Thr Ile Lys
        35                  40                  45

Asp Ile Leu Gly Ala Ile Pro Lys Glu Cys Tyr Lys Arg Asp Thr Leu
    50                  55                  60

Trp Ser Leu His Tyr Val Val Arg Asp Ile Ile Ala Ile Cys Ile Ile
65                  70                  75                  80

Gly Tyr Val Gly Thr Asn Tyr Ile Pro Val Trp Phe Pro Asn Ser Gly
                85                  90                  95

Leu Leu Arg Phe Val Ala Tyr Met Val Gln Ser Tyr Leu Ile Gly Leu
            100                 105                 110

Phe Gly Phe Gly Leu Trp Ile Leu Ala His Glu Cys Gly His Gly Ala
        115                 120                 125

Phe Ser Asp Ser Arg Leu Ile Asn Asp Thr Val Gly Trp Val Leu His
    130                 135                 140

Ser Trp Trp Met Val Pro Tyr Phe Ser Trp Lys Phe Ser His Ser Lys
145                 150                 155                 160

His His Lys Ala Thr Gly His Leu Thr Arg Asp Met Val Phe Val Pro
                165                 170                 175

Tyr Thr Lys Lys Glu Tyr Leu Glu Met Lys Gly Lys Ser Lys Leu Arg
            180                 185                 190
```

```
Glu Ile Thr Glu Glu Ala Pro Ile Val Thr Leu Leu Thr Leu Ile Gly
        195                 200                 205

Gln Gln Ile Gly Gly Leu Gln Leu Tyr Leu Ala Thr Asn Ala Thr Gly
    210                 215                 220

Gln Ser Tyr Pro Gly Val Pro Lys Phe Phe Lys Ser His Tyr Trp Pro
225                 230                 235                 240

Thr Ser Pro Val Phe Asp Thr Lys Asp Phe Trp Tyr Ile Ile Leu Ser
                245                 250                 255

Asp Ile Gly Ile Ile Ser Thr Leu Thr Ile Asn Tyr Leu Trp Ala Lys
            260                 265                 270

Thr Tyr Gly Ser His Val Met Leu Ile Asn Trp Phe Val Pro Trp Leu
        275                 280                 285

Trp Val Asn His Trp Leu Val Phe Val Thr Phe Leu Gln His Thr Asp
    290                 295                 300

Pro Thr Met Pro His Tyr Glu Ala Ser Glu Trp Thr Phe Ala Lys Gly
305                 310                 315                 320

Ala Ala Ala Thr Ile Asp Arg Asn Phe Gly Phe Val Gly Gln His Ile
                325                 330                 335

Phe His Asp Ile Ile Glu Thr His Val Leu His His Tyr Cys Ser Arg
            340                 345                 350

Ile Pro Phe Tyr Asn Ala Arg Val Ala Thr Glu Ala Ile Lys Lys Val
        355                 360                 365

Met Gly Glu His Tyr Arg Tyr Glu Gly Glu Asn Met Trp Gln Ser Leu
    370                 375                 380

Trp Lys Val Ala Arg Ser Cys Gln Phe Val Asp Gly Asp Asn Gly Val
385                 390                 395                 400
```

Leu Met Phe Arg Asn Thr Asn Gly Val Gly Ala Pro Cys Gln Glu
405 410 415

<210> 49
<211> 436
<212> PRT
<213> Candida albicans SC5314 (GenBank Accession No. EAK94955)

<220>
<221> MISC_FEATURE
<223> delta-12 desaturase

<400> 49

Met Ala Ala Ala Thr Thr Ser Phe Ser Ser Gly Phe Asn Asn Asn Asn
1 5 10 15

Asn Ala Asp Gln Ser Thr Asp Ser Ser Ala Thr Ile Ser Lys Ser Gly
20 25 30

Asn Val Ala Ser Phe Lys Thr Thr Ser Thr Thr Ser Thr Tyr Gln Thr
35 40 45

Asn Leu Thr Ala Ile Asp Thr Tyr Gly Asn Glu Phe Lys Val Pro Asp
50 55 60

Tyr Thr Ile Lys Asp Ile Leu Ser Ala Ile Pro Thr His Cys Tyr Glu
65 70 75 80

Arg Arg Leu Leu Gln Ser Leu Ser Tyr Val Phe Arg Asp Ile Phe Cys
85 90 95

Met Val Val Leu Gly Phe Ile Ala Asn Asn Tyr Ile His Leu Ile Pro
100 105 110

Asn Gln Phe Ile Arg Phe Ala Ala Trp Thr Gly Tyr Val Trp Cys Gln
115 120 125

Gly Leu Phe Gly Thr Gly Ile Trp Val Leu Ala His Glu Cys Gly His
130 135 140

```
Gln Ala Phe Ser Asp Tyr Gly Ser Val Asn Asp Phe Val Gly Trp Val
145                 150                 155                 160

Leu His Ser Tyr Leu Leu Val Pro Tyr Phe Ser Trp Lys Phe Ser His
                165                 170                 175

Gly Lys His His Lys Ala Thr Gly His Leu Thr Arg Asp Met Val Phe
            180                 185                 190

Val Pro Lys Thr Lys Glu Glu Phe Leu Gln Asn Arg Gly Val Lys Asp
        195                 200                 205

Leu Asp Asp Leu Leu Gly Asp Ser Pro Met Tyr Ser Leu Leu Thr Leu
    210                 215                 220

Ile Phe Gln Gln Thr Phe Gly Trp Ile Ser Tyr Leu Val Ala Asn Val
225                 230                 235                 240

Ser Gly Gln Lys Tyr Pro Gly Val Ser Phe Leu Lys Leu Asn His Phe
                245                 250                 255

Asn Pro Asn Ser Leu Ile Phe Asp Lys Lys Asp Tyr Trp Tyr Ile Leu
            260                 265                 270

Leu Ser Asp Leu Gly Ile Leu Leu Gln Phe Phe Asn Leu Tyr Val Trp
        275                 280                 285

Tyr Gln Ser Phe Gly Gly Phe Asn Leu Leu Val Asn Tyr Val Leu Pro
    290                 295                 300

Tyr Phe Leu Val Asn His Trp Leu Val Phe Ile Thr Tyr Leu Gln His
305                 310                 315                 320

Ser Asp Pro Gln Met Pro His Tyr Glu Ala Ser Gln Trp Thr Phe Ala
                325                 330                 335

Arg Gly Ala Ala Ala Thr Ile Asp Arg Glu Phe Gly Phe Val Gly Lys
            340                 345                 350
```

```
His Ile Phe His Asp Ile Ile Glu Thr His Val Leu His His Tyr Val
        355                 360                 365

Ser Arg Ile Pro Phe Tyr Asn Ala Arg Glu Ala Ser Glu Ala Ile Lys
    370                 375                 380

Lys Val Met Gly Ile His Tyr Gln His Ser Asp Glu Asn Met Trp Val
385                 390                 395                 400

Ser Leu Trp Lys Ser Ala Arg Trp Cys Gln Phe Val Asp Gly Asn Asn
                405                 410                 415

Gly Val Leu Met Tyr Arg Asn Thr Asn Gly Phe Gly Val Asp Pro Lys
            420                 425                 430

Lys Gln Thr His
        435
```

<210> 50
<211> 416
<212> PRT
<213> Debaryomyces hansenii CBS767 (GenBank Accession No. CAG90237)

<220>
<221> MISC_FEATURE
<223> delta-12 desaturase

<400> 50

```
Met Ala Ser Gln Ile Ala Ser Ser Thr Lys Ser Ser Gly Ile Gly Gly
1               5                   10                  15

Ser Ser Ser Ile Gln Lys Arg Gly Asn Val Ala Thr Leu Gln Thr Asn
            20                  25                  30

Glu Asn Leu Thr Ala Ile Asp Ala His Gly Asn Val Phe Lys Val Pro
        35                  40                  45

Asp Tyr Thr Ile Lys Asp Ile Leu Lys Ala Ile Pro Ala His Cys Tyr
    50                  55                  60
```

```
Glu Arg Arg Val Ala Glu Ser Met Tyr Tyr Val Phe Arg Asp Ile Phe
65                  70                  75                  80

Trp Leu Leu Thr Ile Gly Tyr Val Ala Asn Asn Tyr Ile Gln Leu Leu
                85                  90                  95

Pro Asn Ala Phe Thr Arg Phe Val Ala Trp Ser Gly Tyr Val Tyr Val
            100                 105                 110

Gln Ser Leu Phe Leu Thr Gly Leu Trp Val Leu Ala His Glu Cys Gly
        115                 120                 125

His Gln Ala Phe Ser Asp Tyr Gly Trp Val Asn Asp Thr Val Gly Trp
    130                 135                 140

Val Leu His Ser Tyr Leu Met Val Pro Tyr Phe Ser Trp Lys Phe Ser
145                 150                 155                 160

His Ser Lys His His Lys Ala Thr Gly His Leu Thr Arg Asp Met Val
                165                 170                 175

Phe Val Pro Tyr Thr Lys Glu Glu Phe Val Ala Ser Lys Asn Ala His
            180                 185                 190

His Ile Asp Asp Ile Val Gly Asp Ser Pro Ile Tyr Thr Leu Tyr Gln
        195                 200                 205

Leu Val Val Gln Gln Phe Gly Gly Trp Ile Ala Tyr Leu Phe Thr Asn
    210                 215                 220

Val Thr Gly Gln Gln Tyr Glu Asn Lys Ser Phe Trp Gly Val Ser His
225                 230                 235                 240

Phe Asn Pro Asn Ala Ala Ile Phe Glu Lys Lys Glu Gln Trp Tyr Val
                245                 250                 255

Leu Leu Ser Asp Ile Gly Ile Phe Ala Gln Gly Leu Val Leu His Thr
            260                 265                 270
```

```
Trp Tyr Lys Asn Phe Gly Gly Phe Asn Leu Leu Val Asn Trp Phe Leu
        275                 280                 285

Pro Tyr Ile Leu Val Asn His Trp Leu Val Phe Ile Thr Tyr Leu Gln
    290                 295                 300

His Thr Asp Ser Gln Met Pro His Tyr Glu Ser His Gln Trp Asn Phe
305                 310                 315                 320

Ala Arg Gly Ala Ala Ala Thr Ile Asp Arg Glu Phe Gly Phe Val Gly
                325                 330                 335

Lys Phe Met Phe His Asp Ile Ile Glu Thr His Val Leu His His Tyr
            340                 345                 350

Cys Ser Arg Ile Pro Phe Tyr Asn Gly Arg Glu Ala Ser Glu Ala Ile
        355                 360                 365

Lys Lys Val Met Gly Glu His Tyr Gln Tyr Ser Asp Glu Asn Met Trp
    370                 375                 380

Val Ser Leu Trp Lys Ser Ala Arg Thr Cys Gln Tyr Val Asp Gly Asp
385                 390                 395                 400

Asn Gly Val Leu Met Phe Arg Asn Ala Ala Phe Lys Gly Pro Lys Ala
                405                 410                 415


<210>  51
<211>  1209
<212>  DNA
<213>  Fusarium monoliforme


<220>
<221>  misc_feature
<223>  delta-15 desaturase

<400>  51
atggcgactc gacagcgaac tgccaccact gttgtggtcg aggaccttcc caaggtcact     60

cttgaggcca agtctgaacc tgtgttcccc gatatcaaga ccatcaagga tgccattccc    120
```

gcgcactgct tccagccctc gctcgtcacc tcattctact acgtcttccg cgattttgcc 180

atggtctctg ccctcgtctg ggctgctctc acctacatcc ccagcatccc cgaccagacc 240

ctccgcgtcg cagcttggat ggtctacggc ttcgtccagg gtctgttctg caccggtgtc 300

tggattctcg gccatgagtg cggccacggt gctttctctc tccacggaaa ggtcaacaat 360

gtgaccggct ggttcctcca ctcgttcctc ctcgtcccct acttcagctg gaagtactct 420

caccaccgcc accaccgctt caccggccac atggatctcg acatggcttt cgtccccaag 480

actgagccca agccctccaa gtcgctcatg attgctggca ttgacgtcgc cgagcttgtt 540

gaggacaccc ccgctgctca gatggtcaag ctcatcttcc accagctttt cggatggcag 600

gcgtacctct tcttcaacgc tagctctggc aagggcagca agcagtggga gcccaagact 660

ggcctctcca agtggttccg agtcagtcac ttcgagccta ccagcgctgt cttccgcccc 720

aacgaggcca tcttcatcct catctccgat atcggtcttg ctctaatggg aactgctctg 780

tactttgctt ccaagcaagt tggtgtttcg accattctct tcctctacct tgttccctac 840

ctgtgggttc accactggct cgttgccatt acctacctcc accaccacca caccgagctc 900

cctcactaca ccgctgaggg ctggacctac gtcaagggag ctctcgccac tgtcgaccgt 960

gagtttggct tcatcggaaa gcacctcttc cacggtatca ttgagaagca cgttgttcac 1020

catctcttcc ctaagatccc cttctacaag gctgacgagg ccaccgaggc catcaagccc 1080

gtcattggcg accactactg ccacgacgac cgaagcttcc tgggccagct gtggaccatc 1140

ttcggcacgc tcaagtacgt cgagcacgac cctgcccgac ccggtgccat gcgatggaac 1200

aaggactag 1209

<210> 52
<211> 402
<212> PRT
<213> Fusarium monoliforme

<400> 52

Met Ala Thr Arg Gln Arg Thr Ala Thr Thr Val Val Val Glu Asp Leu
1 5 10 15

Pro Lys Val Thr Leu Glu Ala Lys Ser Glu Pro Val Phe Pro Asp Ile
20 25 30

```
Lys Thr Ile Lys Asp Ala Ile Pro Ala His Cys Phe Gln Pro Ser Leu
        35                  40                  45

Val Thr Ser Phe Tyr Tyr Val Phe Arg Asp Phe Ala Met Val Ser Ala
    50                  55                  60

Leu Val Trp Ala Ala Leu Thr Tyr Ile Pro Ser Ile Pro Asp Gln Thr
65                  70                  75                  80

Leu Arg Val Ala Ala Trp Met Val Tyr Gly Phe Val Gln Gly Leu Phe
                85                  90                  95

Cys Thr Gly Val Trp Ile Leu Gly His Glu Cys Gly His Gly Ala Phe
            100                 105                 110

Ser Leu His Gly Lys Val Asn Asn Val Thr Gly Trp Phe Leu His Ser
        115                 120                 125

Phe Leu Leu Val Pro Tyr Phe Ser Trp Lys Tyr Ser His His Arg His
    130                 135                 140

His Arg Phe Thr Gly His Met Asp Leu Asp Met Ala Phe Val Pro Lys
145                 150                 155                 160

Thr Glu Pro Lys Pro Ser Lys Ser Leu Met Ile Ala Gly Ile Asp Val
                165                 170                 175

Ala Glu Leu Val Glu Asp Thr Pro Ala Ala Gln Met Val Lys Leu Ile
            180                 185                 190

Phe His Gln Leu Phe Gly Trp Gln Ala Tyr Leu Phe Phe Asn Ala Ser
        195                 200                 205

Ser Gly Lys Gly Ser Lys Gln Trp Glu Pro Lys Thr Gly Leu Ser Lys
    210                 215                 220

Trp Phe Arg Val Ser His Phe Glu Pro Thr Ser Ala Val Phe Arg Pro
225                 230                 235                 240
```

```
Asn Glu Ala Ile Phe Ile Leu Ile Ser Asp Ile Gly Leu Ala Leu Met
                245                 250                 255

Gly Thr Ala Leu Tyr Phe Ala Ser Lys Gln Val Gly Val Ser Thr Ile
            260                 265                 270

Leu Phe Leu Tyr Leu Val Pro Tyr Leu Trp Val His His Trp Leu Val
        275                 280                 285

Ala Ile Thr Tyr Leu His His His His Thr Glu Leu Pro His Tyr Thr
    290                 295                 300

Ala Glu Gly Trp Thr Tyr Val Lys Gly Ala Leu Ala Thr Val Asp Arg
305                 310                 315                 320

Glu Phe Gly Phe Ile Gly Lys His Leu Phe His Gly Ile Ile Glu Lys
                325                 330                 335

His Val Val His His Leu Phe Pro Lys Ile Pro Phe Tyr Lys Ala Asp
            340                 345                 350

Glu Ala Thr Glu Ala Ile Lys Pro Val Ile Gly Asp His Tyr Cys His
        355                 360                 365

Asp Asp Arg Ser Phe Leu Gly Gln Leu Trp Thr Ile Phe Gly Thr Leu
    370                 375                 380

Lys Tyr Val Glu His Asp Pro Ala Arg Pro Gly Ala Met Arg Trp Asn
385                 390                 395                 400

Lys Asp
```

<210> 53
<211> 1206
<212> DNA
<213> Aspergillus nidulans

<220>
<221> misc_feature
<223> delta-15 desaturase

<400> 53
atggctgcaa ctgcaacaac cctagcagag attgaaaaga aaaaagaaga aataactctg 60

cagacaatca aaaatgcgat tcccaaacac tgcttcaacc gctctctcct catttcctct 120

gcctacgtcg tccgcgatct cctctacgcc tccgtcctct tctactttgc cctgcacatt 180

gacaccctct tttcctcgca actcctccgc atcctcgcct ggaccgccta cggtttcatg 240

caaggctgcg tcggcaccgg aatctggatc ctcgcacacg aatgcggcca tggagctttc 300

tccccatacc aaacgtggaa cgatgtcgtc ggatggacat tgcactccct cctgatggtc 360

ccgtatttca gctggaagat cacgcacgct cgacaccacc ggtacacaaa caacacagag 420

cgagatacag catttgtccc ctggacagag aaggaatacg acactcgccc gcgctacttc 480

cctgcctggt ttgagatgtt tgaggacacg cccgtctaca accttattag cctactggcg 540

catcagatcg caggatggca gatgtatctc tgtttttacg ttagcgccgg cgcaaagagt 600

aagcctgtac cgcagggaaa acagagcggg tggtttggag gccagcagag cgccagccac 660

tttgatccgg gcagttcgct gtggacggaa aaccagcggc atctgattgc gatttcggac 720

ctggggttgc tgcttgttgc ggcggcaaat tggtaccttg cgcagcaagt gggcgtgctc 780

cgcatggtgc tgatctatgt tgtgccgtac ttctgggtgc accattggct tgtggcgatc 840

acgtacctcc accacacaca cccctcgatc ccgcactaca ctgatagcac ctggacgttc 900

accaaaggcg ctctgtccac cgtcgaccgc gacttcggtt tcatcgggcg gcatttcttc 960

caccatatca ttgaccacca tgtcgtgcat cacttgttta accggatccc gttctaccat 1020

gccgaggagg cgactaatgc cattattccc gtactcgggg acatgtatca tcgcgaagag 1080

accggcttct tgtggagttt aatggagacg tacaagaact gtcggtttgt aggcgttgaa 1140

aatgatgttg gaaaggaggg cgttttgcat tgggtttttg aggagaagaa gggtgccaaa 1200

gcggaa 1206

<210> 54
<211> 401
<212> PRT
<213> Aspergillus nidulans

<400> 54

```
Met Ala Ala Thr Ala Thr Thr Leu Ala Glu Ile Glu Lys Lys Lys Glu
1               5                  10                  15

Glu Ile Thr Leu Gln Thr Ile Lys Asn Ala Ile Pro Lys His Cys Phe
            20                  25                  30

Asn Arg Ser Leu Leu Ile Ser Ser Ala Tyr Val Val Arg Asp Leu Leu
        35                  40                  45

Tyr Ala Ser Val Leu Phe Tyr Phe Ala Leu His Ile Asp Thr Leu Phe
    50                  55                  60

Ser Ser Gln Leu Leu Arg Ile Leu Ala Trp Thr Ala Tyr Gly Phe Met
65                  70                  75                  80

Gln Gly Cys Val Gly Thr Gly Ile Trp Ile Leu Ala His Glu Cys Gly
                85                  90                  95

His Gly Ala Phe Ser Pro Tyr Gln Thr Trp Asn Asp Val Val Gly Trp
            100                 105                 110

Thr Leu His Ser Leu Leu Met Val Pro Tyr Phe Ser Trp Lys Ile Thr
        115                 120                 125

His Ala Arg His His Arg Tyr Thr Asn Asn Thr Glu Arg Asp Thr Ala
    130                 135                 140

Phe Val Pro Trp Thr Glu Lys Glu Tyr Asp Thr Arg Pro Arg Tyr Phe
145                 150                 155                 160

Pro Ala Trp Phe Glu Met Phe Glu Asp Thr Pro Val Tyr Asn Leu Ile
                165                 170                 175

Ser Leu Leu Ala His Gln Ile Ala Gly Trp Gln Met Tyr Leu Cys Phe
            180                 185                 190

Tyr Val Ser Ala Gly Ala Lys Ser Lys Pro Val Pro Gln Gly Lys Gln
        195                 200                 205
```

```
Ser Gly Trp Phe Gly Gly Gln Gln Ser Ala Ser His Phe Asp Pro Gly
    210                 215                 220

Ser Ser Leu Trp Thr Glu Asn Gln Arg His Leu Ile Ala Ile Ser Asp
225                 230                 235                 240

Leu Gly Leu Leu Leu Val Ala Ala Ala Asn Trp Tyr Leu Ala Gln Gln
                245                 250                 255

Val Gly Val Leu Arg Met Val Leu Ile Tyr Val Val Pro Tyr Phe Trp
            260                 265                 270

Val His His Trp Leu Val Ala Ile Thr Tyr Leu His His Thr His Pro
        275                 280                 285

Ser Ile Pro His Tyr Thr Asp Ser Thr Trp Thr Phe Thr Lys Gly Ala
    290                 295                 300

Leu Ser Thr Val Asp Arg Asp Phe Gly Phe Ile Gly Arg His Phe Phe
305                 310                 315                 320

His His Ile Ile Asp His His Val Val His His Leu Phe Asn Arg Ile
                325                 330                 335

Pro Phe Tyr His Ala Glu Glu Ala Thr Asn Ala Ile Ile Pro Val Leu
            340                 345                 350

Gly Asp Met Tyr His Arg Glu Glu Thr Gly Phe Leu Trp Ser Leu Met
        355                 360                 365

Glu Thr Tyr Lys Asn Cys Arg Phe Val Gly Val Glu Asn Asp Val Gly
    370                 375                 380

Lys Glu Gly Val Leu His Trp Val Phe Glu Glu Lys Lys Gly Ala Lys
385                 390                 395                 400

Ala
```

<210> 55
<211> 1185
<212> DNA
<213> Magnaporthe grisea

<220>
<221> misc_feature
<223> delta-15 desaturase

<400> 55
atgtccacca ccgtcactca gcggccgggc gccgctagcc gcgctgaagc caagcccaag 60

gagcagcaat ttccagacat caacaccatc aggaatgcta tccccgcaca ctgttttgag 120

gcatctctgg tgacttcagt tggttacttg gtgcgagatg tggccctcat caccgctctc 180

ggctgggccg ccttgaccta cattccccaa attccggatt cgactttgcg ctggaccgcc 240

tgggccgctt acggctttgt tcagggtctc tttggcaccg gtctctggat tctggcccac 300

gagtgcggcc acggtgcttt cagcaagcac acgcgcatta acaacattct tggctgggcc 360

gcccactcgg ccctgctggt accgtacttc agctggaagt tctctcacca ccgccaccac 420

aacttcaccg gccacatgga gaaggacatg gcctttgtgc ccccccaggc tgccgaccgc 480

gagtcccgcg ccagcttgct gtcccgcttc ggcatcgacc tcgaggtctt tgaggatacc 540

cccatctttc agcttgctcg cctcgtgagc caccagctct tcggctggca gacttacctg 600

ctcttcaacg ccacctgcgg caaggagtct ctgcagaaca agggtgccgc gtggttccgc 660

cagagccact ttgagcccac ctctgccgtc ttccgctcca gcgaggccct ctacatcgcc 720

atctctgaca ttggcctggc catcgttgcc gccgccatct actggggctc caccaaggtc 780

ggcgccggca ccatgttcct cctctacgcc gttccctaca tgtgggttca ccactggctc 840

gtcgccatca cctaccttca ccacaccaac aaggaggtgc accactacga ggccgacagc 900

tggacctttg tcaagggtgc cgtcgccact gtcgaccgtg actttggttt cattgaccgc 960

cacctgttcc acggtatcat tggaacccac gtcgcccacc atctgttccc tcgcattccc 1020

ttttacaagg cagaggaggc caccgaggcc atcaagcctg tcctcggaga cctttaccac 1080

agcgacaatc gccccttcat gcaggctctg tggagcaact tcaccacctg caagtacgtc 1140

gagaaggacc ccaaggttcc cggcgccatg aggtgggccg attga 1185

<210> 56
<211> 394
<212> PRT
<213> Magnaporthe grisea

<400> 56

```
Met Ser Thr Thr Val Thr Gln Arg Pro Gly Ala Ala Ser Arg Ala Glu
1               5                  10                  15

Ala Lys Pro Lys Glu Gln Gln Phe Pro Asp Ile Asn Thr Ile Arg Asn
            20                  25                  30

Ala Ile Pro Ala His Cys Phe Glu Ala Ser Leu Val Thr Ser Val Gly
        35                  40                  45

Tyr Leu Val Arg Asp Val Ala Leu Ile Thr Ala Leu Gly Trp Ala Ala
    50                  55                  60

Leu Thr Tyr Ile Pro Gln Ile Pro Asp Ser Thr Leu Arg Trp Thr Ala
65                  70                  75                  80

Trp Ala Ala Tyr Gly Phe Val Gln Gly Leu Phe Gly Thr Gly Leu Trp
                85                  90                  95

Ile Leu Ala His Glu Cys Gly His Gly Ala Phe Ser Lys His Thr Arg
            100                 105                 110

Ile Asn Asn Ile Leu Gly Trp Ala Ala His Ser Ala Leu Leu Val Pro
        115                 120                 125

Tyr Phe Ser Trp Lys Phe Ser His His Arg His His Asn Phe Thr Gly
    130                 135                 140

His Met Glu Lys Asp Met Ala Phe Val Pro Pro Gln Ala Ala Asp Arg
145                 150                 155                 160

Glu Ser Arg Ala Ser Leu Leu Ser Arg Phe Gly Ile Asp Leu Glu Val
                165                 170                 175
```

```
Phe Glu Asp Thr Pro Ile Phe Gln Leu Ala Arg Leu Val Ser His Gln
            180                 185                 190

Leu Phe Gly Trp Gln Thr Tyr Leu Leu Phe Asn Ala Thr Cys Gly Lys
        195                 200                 205

Glu Ser Leu Gln Asn Lys Gly Ala Ala Trp Phe Arg Gln Ser His Phe
    210                 215                 220

Glu Pro Thr Ser Ala Val Phe Arg Ser Ser Glu Ala Leu Tyr Ile Ala
225                 230                 235                 240

Ile Ser Asp Ile Gly Leu Ala Ile Val Ala Ala Ala Ile Tyr Trp Gly
                245                 250                 255

Ser Thr Lys Val Gly Ala Gly Thr Met Phe Leu Leu Tyr Ala Val Pro
            260                 265                 270

Tyr Met Trp Val His His Trp Leu Val Ala Ile Thr Tyr Leu His His
        275                 280                 285

Thr Asn Lys Glu Val His His Tyr Glu Ala Asp Ser Trp Thr Phe Val
    290                 295                 300

Lys Gly Ala Val Ala Thr Val Asp Arg Asp Phe Gly Phe Ile Asp Arg
305                 310                 315                 320

His Leu Phe His Gly Ile Ile Gly Thr His Val Ala His His Leu Phe
                325                 330                 335

Pro Arg Ile Pro Phe Tyr Lys Ala Glu Glu Ala Thr Glu Ala Ile Lys
            340                 345                 350

Pro Val Leu Gly Asp Leu Tyr His Ser Asp Asn Arg Pro Phe Met Gln
        355                 360                 365

Ala Leu Trp Ser Asn Phe Thr Thr Cys Lys Tyr Val Glu Lys Asp Pro
    370                 375                 380
```

Lys Val Pro Gly Ala Met Arg Trp Ala Asp
385 390

<210> 57
<211> 1290
<212> DNA
<213> Neurospora crassa

<220>
<221> misc_feature
<223> delta-15 desaturase

<400> 57
atgacggtca ccacccgcag ccacaaggcc gcggccgcca ccgagcccga ggttgtcagc 60

accggcgttg acgccgtctc tgctgctgct ccctcctcct cctcctcctc ttccagccaa 120

aagtcggccg agcccatcga ataccccgac atcaagacca tccgcgacgc catccccgac 180

cactgcttcc gcccgcgcgt ctggatctcc atggcctact tcatccgcga cttcgccatg 240

gcctttggcc tcggctacct cgcctggcag tacatccccc tgatcgcctc caccccgctc 300

cgctacggcg cctgggctct gtacggctac ctccagggtc tcgtctgcac gggcatctgg 360

attctggcgc acgagtgcgg ccacggcgcc ttctcgaggc acacgtggtt caacaacgtc 420

atggggtgga ttggccactc cttcctcttg gtcccttact tcagctggaa gttcagccac 480

catcgccacc atcgcttcac cggccacatg gagaaggaca tggcgtttgt gcctgccacc 540

gaggctgatc gcaaccagag gaagctggcc aacttgtaca tggacaagga gacggccgag 600

atgtttgagg atgtgcccat tgtccagctc gtcaagctca tcgcccacca gctggccggc 660

tggcagatgt acctcctctt caacgtctcc gccggtaagg gcagcaagca gtgggagact 720

ggcaagggcg gcatgggctg gttgagggtt agccactttg agccttcctc tgctgtgttc 780

cgcaactccg aggccatcta cattgccctg tccgatcttg gtctcatgat catgggctat 840

atcctctacc aggccgcgca ggttgttggc tggcagatgg taggtctgct gtacttccag 900

cagtacttct gggttcacca ttggttggtc gccatcactt acctccacca cacccacgag 960

gaagtccacc actttgacgc cgactcgtgg accttcgtca agggcgctct cgccaccgtc 1020

gaccgcgatt ttggcttcat tggcaagcac ctcttccaca acattatcga ccaccacgtc 1080

gtccaccact tgttccctcg catccccttc tactacgccg aagaagccac caactcgatc 1140

cgccccatgc tcggcccccc ctaccaccgc gacgaccgct ccttcatggg ccagctgtgg 1200

tacaacttca cccactgcaa gtgggtcgtt ccggaccccc aggtccccgg cgcgcttatt 1260

tgggcgcaca ccgttcagag cacccagtaa 1290

<210> 58
<211> 429
<212> PRT
<213> Neurospora crassa

<400> 58

Met Thr Val Thr Thr Arg Ser His Lys Ala Ala Ala Ala Thr Glu Pro
1 5 10 15

Glu Val Val Ser Thr Gly Val Asp Ala Val Ser Ala Ala Ala Pro Ser
20 25 30

Ser Ser Ser Ser Ser Ser Ser Gln Lys Ser Ala Glu Pro Ile Glu Tyr
35 40 45

Pro Asp Ile Lys Thr Ile Arg Asp Ala Ile Pro Asp His Cys Phe Arg
50 55 60

Pro Arg Val Trp Ile Ser Met Ala Tyr Phe Ile Arg Asp Phe Ala Met
65 70 75 80

Ala Phe Gly Leu Gly Tyr Leu Ala Trp Gln Tyr Ile Pro Leu Ile Ala
85 90 95

Ser Thr Pro Leu Arg Tyr Gly Ala Trp Ala Leu Tyr Gly Tyr Leu Gln
100 105 110

Gly Leu Val Cys Thr Gly Ile Trp Ile Leu Ala His Glu Cys Gly His
115 120 125

Gly Ala Phe Ser Arg His Thr Trp Phe Asn Asn Val Met Gly Trp Ile
130 135 140

Gly His Ser Phe Leu Leu Val Pro Tyr Phe Ser Trp Lys Phe Ser His
145 150 155 160

```
His Arg His His Arg Phe Thr Gly His Met Glu Lys Asp Met Ala Phe
               165                 170                 175

Val Pro Ala Thr Glu Ala Asp Arg Asn Gln Arg Lys Leu Ala Asn Leu
            180                 185                 190

Tyr Met Asp Lys Glu Thr Ala Glu Met Phe Glu Asp Val Pro Ile Val
        195                 200                 205

Gln Leu Val Lys Leu Ile Ala His Gln Leu Ala Gly Trp Gln Met Tyr
    210                 215                 220

Leu Leu Phe Asn Val Ser Ala Gly Lys Gly Ser Lys Gln Trp Glu Thr
225                 230                 235                 240

Gly Lys Gly Gly Met Gly Trp Leu Arg Val Ser His Phe Glu Pro Ser
               245                 250                 255

Ser Ala Val Phe Arg Asn Ser Glu Ala Ile Tyr Ile Ala Leu Ser Asp
            260                 265                 270

Leu Gly Leu Met Ile Met Gly Tyr Ile Leu Tyr Gln Ala Ala Gln Val
        275                 280                 285

Val Gly Trp Gln Met Val Gly Leu Leu Tyr Phe Gln Gln Tyr Phe Trp
    290                 295                 300

Val His His Trp Leu Val Ala Ile Thr Tyr Leu His His Thr His Glu
305                 310                 315                 320

Glu Val His His Phe Asp Ala Asp Ser Trp Thr Phe Val Lys Gly Ala
               325                 330                 335

Leu Ala Thr Val Asp Arg Asp Phe Gly Phe Ile Gly Lys His Leu Phe
            340                 345                 350

His Asn Ile Ile Asp His His Val Val His His Leu Phe Pro Arg Ile
        355                 360                 365
```

```
Pro Phe Tyr Tyr Ala Glu Glu Ala Thr Asn Ser Ile Arg Pro Met Leu
    370                 375                 380

Gly Pro Leu Tyr His Arg Asp Asp Arg Ser Phe Met Gly Gln Leu Trp
385                 390                 395                 400

Tyr Asn Phe Thr His Cys Lys Trp Val Val Pro Asp Pro Gln Val Pro
                405                 410                 415

Gly Ala Leu Ile Trp Ala His Thr Val Gln Ser Thr Gln
            420                 425
```

<210> 59
<211> 1212
<212> DNA
<213> Fusarium graminearium

<220>
<221> misc_feature
<223> delta-15 desaturase

<400> 59

```
atggccacca gacagcgaac tgccaccact gttgtggtcg agaaggacct gcccaaggtc     60

actctcgagg ccacttctca gcctcaattc cccgacatca agaccatcaa ggatgccatc    120

cccgcccact gcttccagcc ctcgctcatc acctcatact actatgtcgt ccgcgacttc    180

gccatggtcg gctccctcgt ctgggccgcc ctcacctaca tccccggcat tgaggaccag    240

tacctccgcg tcgccgcctg gatggcctac ggcttcctcc agggtctctt ctgcaccgga    300

atctggattc tcggtcatga gtgcggccac ggtgccttct ctacccacag caagctcaac    360

aatgtgaccg gctggttcct ccactcgttc ctcatggtcc cctatttcag ctggaagtac    420

tctcaccacc gtcaccaccg cttcaccggc cacatggatc tcgacatggc ctttgtcccc    480

cgcacttcgc ccaagccttc tttgtctttc cgcattgctg gtatggacgt cgctgagctg    540

attgaggaca cccccattgc ccaggccgtc aagctcatct tccaccagct cttcggatgg    600

caggtgtaca ccttcttcaa cgccagctct ggcaagggta gcaagcagtg ggagcccaag    660

agcggcttgg ccagctggtt ccgcgtcagc cacttcgagc ccaccagcgc tgtcttccgc    720
```

cccgccgagg ctcctttcat cctcatctcc gacattggtc tcgccctcac tggaactgct 780

ctgtactttg cttccaagga ggtcggcgtt tccaccgttc tctacctcta cctcgtcccc 840

tacctctggg tccaccactg gctcgtcgcc atcacctacc tccaccacca ccacaccgag 900

cttccccact acaccgccga gggctggacc tacgtcaagg gtgctctcgc tactgttgac 960

cgcgagtttg gcttcattgg caagcacctt ttccacggca tcattgagaa gcacgtcatt 1020

caccacctgt tccctaagat ccccttctac aaggctgacg aggccaccga ggccatcaag 1080

cccatcatcg gcgaccacta ctgccacgac gaccgcagct tccttggcca gctctggacc 1140

atctttggca gcctcaagta cgtcgagcac gaccccgccg tccctggtgc catgcgctgg 1200

gccaaggagt ag 1212

<210> 60
<211> 403
<212> PRT
<213> Fusarium graminearium

<400> 60

Met Ala Thr Arg Gln Arg Thr Ala Thr Thr Val Val Val Glu Lys Asp
1 5 10 15

Leu Pro Lys Val Thr Leu Glu Ala Thr Ser Gln Pro Gln Phe Pro Asp
20 25 30

Ile Lys Thr Ile Lys Asp Ala Ile Pro Ala His Cys Phe Gln Pro Ser
35 40 45

Leu Ile Thr Ser Tyr Tyr Tyr Val Val Arg Asp Phe Ala Met Val Gly
50 55 60

Ser Leu Val Trp Ala Ala Leu Thr Tyr Ile Pro Gly Ile Glu Asp Gln
65 70 75 80

Tyr Leu Arg Val Ala Ala Trp Met Ala Tyr Gly Phe Leu Gln Gly Leu
85 90 95

Phe Cys Thr Gly Ile Trp Ile Leu Gly His Glu Cys Gly His Gly Ala
100 105 110

```
Phe Ser Thr His Ser Lys Leu Asn Asn Val Thr Gly Trp Phe Leu His
        115                 120                 125

Ser Phe Leu Met Val Pro Tyr Phe Ser Trp Lys Tyr Ser His His Arg
    130                 135                 140

His His Arg Phe Thr Gly His Met Asp Leu Asp Met Ala Phe Val Pro
145                 150                 155                 160

Arg Thr Ser Pro Lys Pro Ser Leu Ser Phe Arg Ile Ala Gly Met Asp
                165                 170                 175

Val Ala Glu Leu Ile Glu Asp Thr Pro Ile Ala Gln Ala Val Lys Leu
            180                 185                 190

Ile Phe His Gln Leu Phe Gly Trp Gln Val Tyr Thr Phe Phe Asn Ala
        195                 200                 205

Ser Ser Gly Lys Gly Ser Lys Gln Trp Glu Pro Lys Ser Gly Leu Ala
    210                 215                 220

Ser Trp Phe Arg Val Ser His Phe Glu Pro Thr Ser Ala Val Phe Arg
225                 230                 235                 240

Pro Ala Glu Ala Pro Phe Ile Leu Ile Ser Asp Ile Gly Leu Ala Leu
                245                 250                 255

Thr Gly Thr Ala Leu Tyr Phe Ala Ser Lys Glu Val Gly Val Ser Thr
            260                 265                 270

Val Leu Tyr Leu Tyr Leu Val Pro Tyr Leu Trp Val His His Trp Leu
        275                 280                 285

Val Ala Ile Thr Tyr Leu His His His His Thr Glu Leu Pro His Tyr
    290                 295                 300

Thr Ala Glu Gly Trp Thr Tyr Val Lys Gly Ala Leu Ala Thr Val Asp
305                 310                 315                 320
```

Arg Glu Phe Gly Phe Ile Gly Lys His Leu Phe His Gly Ile Ile Glu
325 330 335

Lys His Val Ile His His Leu Phe Pro Lys Ile Pro Phe Tyr Lys Ala
340 345 350

Asp Glu Ala Thr Glu Ala Ile Lys Pro Ile Ile Gly Asp His Tyr Cys
355 360 365

His Asp Asp Arg Ser Phe Leu Gly Gln Leu Trp Thr Ile Phe Gly Ser
370 375 380

Leu Lys Tyr Val Glu His Asp Pro Ala Val Pro Gly Ala Met Arg Trp
385 390 395 400

Ala Lys Glu

<210> 61
<211> 1353
<212> DNA
<213> Mortierella alpina (GenBank Accession No. AB182163)

<220>
<221> misc_feature
<223> delta-15 desaturase

<400> 61
atggcccccc ctcacgttgt cgacgaacaa gttcgacgca ggatcgtcgt tgaggacgag 60

atcaagtcta agaagcaatt tgagcgcaac tatgtgccca tggactttac gattaaggag 120

attcgagatg cgatccctgc ccacctcttc atccgtgata ccacaaagtc gatcctgcat 180

gtcgtcaagg atctggtcac catcgccatc gtcttttact gtgcaacctt cattgagact 240

ctgccctcgc tcgctctgcg agttcctgcc tggatcacct actggatcat ccaaggaact 300

gtcatggtcg gcccctggat cttggctcat ggtaaggaaa cgaaaaatcc catgtgtatt 360

tctgtactac agaaggcgaa gtttgtacct gaaaagatca gcgtcgtccc ttgatttaga 420

atgtaactaa ccttgcaatc gtatgaccta aattttcttg tgtcaacgac agagtgcggc 480

cacggagctt tctcggatag caagacgatc aacaccatct ttggatgggt cctccactct 540

gctcttttgg tgccctacca ggcctgggct atgtcacact ccaagcatca caagggtact 600

ggatcgatga ccaaagatgt cgttttcatc cctgccactc gttcctacaa gggcctccca 660

gcactggaga agcctgccgt cgaagaggag gtttcggagc aggaacacca ccaccacgag 720

gagtccatct ttgccgaaac tcccatctac acgctcggag cgcttttgtt cgtcttgacc 780

ttcggatggc ccttgtactt gatcgtcaac ttttcaggac acgaggcccc tcactgggtc 840

aaccatttcc agactgtcgc tcctctctat gagcctcacc agcgcaagaa catcttctac 900

tccaactgcg gcattgtcgc catgggttcg atcttgactt acctttcgat ggtcttctcg 960

cccttgactg tcttcatgta ctatggcatc ccttacctcg gagtcaacgc ctggatcgtc 1020

tgcattacct atctccagca caccgatccc aaggtgcctc acttccgtga taacgagtgg 1080

aacttccagc gcggtgctgc ctgcactatc gaccgatcct tcggtaccat cgtgaaccac 1140

ctgcaccacc acattggcga ctctcaccag tgccaccata tgttctcgca gatgcccttc 1200

tacaatgctg tggaggctac aaagtacttg aaggccaaac ttggcaagta ctacatattt 1260

gacgacacgc ccattgccaa agccctctac cgcaattgga gagagtgcaa attcgtggag 1320

gacgagggag atgtagtgtt ttacaagcat taa 1353

<210> 62
<211> 403
<212> PRT
<213> Mortierella alpina (GenBank Accession No. AB182163)

<220>
<221> MISC_FEATURE
<223> delta-15 desaturase

<400> 62

Met Ala Pro Pro His Val Val Asp Glu Gln Val Arg Arg Arg Ile Val
1 5 10 15

Val Glu Asp Glu Ile Lys Ser Lys Lys Gln Phe Glu Arg Asn Tyr Val
20 25 30

Pro Met Asp Phe Thr Ile Lys Glu Ile Arg Asp Ala Ile Pro Ala His
35 40 45

```
Leu Phe Ile Arg Asp Thr Thr Lys Ser Ile Leu His Val Val Lys Asp
    50                  55                  60

Leu Val Thr Ile Ala Ile Val Phe Tyr Cys Ala Thr Phe Ile Glu Thr
65                  70                  75                  80

Leu Pro Ser Leu Ala Leu Arg Val Pro Ala Trp Ile Thr Tyr Trp Ile
                85                  90                  95

Ile Gln Gly Thr Val Met Val Gly Pro Trp Ile Leu Ala His Glu Cys
            100                 105                 110

Gly His Gly Ala Phe Ser Asp Ser Lys Thr Ile Asn Thr Ile Phe Gly
        115                 120                 125

Trp Val Leu His Ser Ala Leu Leu Val Pro Tyr Gln Ala Trp Ala Met
    130                 135                 140

Ser His Ser Lys His His Lys Gly Thr Gly Ser Met Thr Lys Asp Val
145                 150                 155                 160

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

Lys Pro Ala Val Glu Glu Glu Val Ser Glu Gln Glu His His His His
            180                 185                 190

Glu Glu Ser Ile Phe Ala Glu Thr Pro Ile Tyr Thr Leu Gly Ala Leu
        195                 200                 205

Leu Phe Val Leu Thr Phe Gly Trp Pro Leu Tyr Leu Ile Val Asn Phe
    210                 215                 220

Ser Gly His Glu Ala Pro His Trp Val Asn His Phe Gln Thr Val Ala
225                 230                 235                 240

Pro Leu Tyr Glu Pro His Gln Arg Lys Asn Ile Phe Tyr Ser Asn Cys
                245                 250                 255
```

```
Gly Ile Val Ala Met Gly Ser Ile Leu Thr Tyr Leu Ser Met Val Phe
            260                 265                 270

Ser Pro Leu Thr Val Phe Met Tyr Tyr Gly Ile Pro Tyr Leu Gly Val
        275                 280                 285

Asn Ala Trp Ile Val Cys Ile Thr Tyr Leu Gln His Thr Asp Pro Lys
    290                 295                 300

Val Pro His Phe Arg Asp Asn Glu Trp Asn Phe Gln Arg Gly Ala Ala
305                 310                 315                 320

Cys Thr Ile Asp Arg Ser Phe Gly Thr Ile Val Asn His Leu His His
                325                 330                 335

His Ile Gly Asp Ser His Gln Cys His His Met Phe Ser Gln Met Pro
            340                 345                 350

Phe Tyr Asn Ala Val Glu Ala Thr Lys Tyr Leu Lys Ala Lys Leu Gly
        355                 360                 365

Lys Tyr Tyr Ile Phe Asp Asp Thr Pro Ile Ala Lys Ala Leu Tyr Arg
    370                 375                 380

Asn Trp Arg Glu Cys Lys Phe Val Glu Asp Glu Gly Asp Val Val Phe
385                 390                 395                 400

Tyr Lys His
```

<210> 63
<211> 1248
<212> DNA
<213> Kluyveromyces lactis (GenBank Accession No. XM_451551)

<220>
<221> misc_feature
<223> delta-15 desaturase

<400> 63

atgagcaaaa gcaccggcgt cgagcatcat atcagtggtg ttgctactac agaaacggca 60

acagagacag ttactgttcc tcctgccaag acagctattg acacacatgg taacatcttc 120

aaggttccag actatactat caaagacatc ctaggtgcca tcccaaagga gtgttacaag 180

agagacactc tttggtcatt gcattatgtg gtaagagata ttattgccat ctgtattatt 240

gggtatgttg gtaccaacta cattccagtt tggtttccaa atagtggtct cctaagattt 300

gttgcctata tggtgcaatc atacttaatt gggttatttg gttttgggtt gtggattttg 360

gctcacgaat gcggtcatgg tgccttctct gattcaagat taatcaacga taccgttggg 420

tgggttttgc actcatggtg gatggtccca tacttttctt ggaaattctc tcactctaag 480

catcataaag ctaccggtca tttgactaga gatatggttt ttgttccata cacgaaaaaa 540

gaatacttgg aaatgaaggg taaatcaaaa ctaagggaga tcaccgaaga agcaccaatt 600

gtaactttat taactttgat tggccaacaa attggtggtt tgcagttgta tttagcaaca 660

aatgccactg gccaatctta tcctggtgtt cctaaattct tcaaatccca ttactggcct 720

acttctccgg tgttcgatac caaggacttc tggtacatca ttctgagtga cattggtatt 780

atttccacac ttacaatcaa ttatttgtgg gcaaagacct atggctccca tgtcatgttg 840

atcaactggt ttgttccatg gttatgggtt aaccactggt tagtatttgt cacatttttg 900

caacacactg acccaaccat gcctcattac gaagccagtg aatggacctt cgctaaaggt 960

gctgctgcca ccattgatag aaactttggc tttgttggtc aacatatttt ccacgatatt 1020

attgaaacgc atgttttgca tcattactgt tcaagaattc ctttctacaa cgctcgtgta 1080

gccactgaag ctattaagaa agttatgggt gaacattatc gttatgaagg cgaaaatatg 1140

tggcaatcat tatggaaggt ggcaagatcc tgtcagtttg tagatggtga caacggtgtt 1200

ttgatgttca gaaacacaaa cggagtcgga gcaccatgtc aagaatga 1248

<210> 64
<211> 415
<212> PRT
<213> Kluyveromyces lactis (GenBank Accession No. XM_451551)

<220>
<221> MISC_FEATURE
<223> delta-15 desaturase

<400> 64

Met Ser Lys Ser Thr Gly Val Glu His His Ile Ser Gly Val Ala Thr
1 5 10 15

Thr Glu Thr Ala Thr Glu Thr Val Thr Val Pro Pro Ala Lys Thr Ala
20 25 30

Ile Asp Thr His Gly Asn Ile Phe Lys Val Pro Asp Tyr Thr Ile Lys
35 40 45

Asp Ile Leu Gly Ala Ile Pro Lys Glu Cys Tyr Lys Arg Asp Thr Leu
50 55 60

Trp Ser Leu His Tyr Val Val Arg Asp Ile Ile Ala Ile Cys Ile Ile
65 70 75 80

Gly Tyr Val Gly Thr Asn Tyr Ile Pro Val Trp Phe Pro Asn Ser Gly
85 90 95

Leu Leu Arg Phe Val Ala Tyr Met Val Gln Ser Tyr Leu Ile Gly Leu
100 105 110

Phe Gly Phe Gly Leu Trp Ile Leu Ala His Glu Cys Gly His Gly Ala
115 120 125

Phe Ser Asp Ser Arg Leu Ile Asn Asp Thr Val Gly Trp Val Leu His
130 135 140

Ser Trp Trp Met Val Pro Tyr Phe Ser Trp Lys Phe Ser His Ser Lys
145 150 155 160

His His Lys Ala Thr Gly His Leu Thr Arg Asp Met Val Phe Val Pro
165 170 175

Tyr Thr Lys Lys Glu Tyr Leu Glu Met Lys Gly Lys Ser Lys Leu Arg
180 185 190

Glu Ile Thr Glu Glu Ala Pro Ile Val Thr Leu Leu Thr Leu Ile Gly
195 200 205

```
Gln Gln Ile Gly Gly Leu Gln Leu Tyr Leu Ala Thr Asn Ala Thr Gly
    210                 215                 220

Gln Ser Tyr Pro Gly Val Pro Lys Phe Phe Lys Ser His Tyr Trp Pro
225                 230                 235                 240

Thr Ser Pro Val Phe Asp Thr Lys Asp Phe Trp Tyr Ile Ile Leu Ser
                245                 250                 255

Asp Ile Gly Ile Ile Ser Thr Leu Thr Ile Asn Tyr Leu Trp Ala Lys
            260                 265                 270

Thr Tyr Gly Ser His Val Met Leu Ile Asn Trp Phe Val Pro Trp Leu
        275                 280                 285

Trp Val Asn His Trp Leu Val Phe Val Thr Phe Leu Gln His Thr Asp
    290                 295                 300

Pro Thr Met Pro His Tyr Glu Ala Ser Glu Trp Thr Phe Ala Lys Gly
305                 310                 315                 320

Ala Ala Ala Thr Ile Asp Arg Asn Phe Gly Phe Val Gly Gln His Ile
                325                 330                 335

Phe His Asp Ile Ile Glu Thr His Val Leu His His Tyr Cys Ser Arg
            340                 345                 350

Ile Pro Phe Tyr Asn Ala Arg Val Ala Thr Glu Ala Ile Lys Lys Val
        355                 360                 365

Met Gly Glu His Tyr Arg Tyr Glu Gly Glu Asn Met Trp Gln Ser Leu
    370                 375                 380

Trp Lys Val Ala Arg Ser Cys Gln Phe Val Asp Gly Asp Asn Gly Val
385                 390                 395                 400

Leu Met Phe Arg Asn Thr Asn Gly Val Gly Ala Pro Cys Gln Glu
                405                 410                 415
```

<210> 65
<211> 433
<212> PRT
<213> Candida albicans SC5314 (GenBank Accession No. EAL03493)

<220>
<221> MISC_FEATURE
<223> delta-15 desaturase

<400> 65

Met Ser Val Val Glu Ala Ser Ser Ser Ser Val Val Glu Asp Ser Thr
1 5 10 15

Ala Ser Asn Val Val Gln Arg Gly Asn Ile Ser Ser Phe Ala Ser Thr
20 25 30

Thr Ala Ser Ser Asn Leu Thr Thr Ile Asp Thr Asn Gly Lys Val Phe
35 40 45

Lys Val Pro Asp Tyr Ser Ile Lys Asp Ile Leu Gln Ala Ile Pro Lys
50 55 60

His Cys Tyr Glu Arg Ser Leu Ile Arg Ser Leu Gly Tyr Val Val Arg
65 70 75 80

Asp Ile Thr Met Met Val Ile Ile Gly Tyr Val Gly His Thr Phe Ile
85 90 95

Pro Met Val Gln Ile Pro Glu Tyr Pro Ser Leu Ala Tyr Gly Leu Arg
100 105 110

Gly Ala Leu Trp Met Val Gln Ser Tyr Cys Ile Gly Leu Phe Gly Phe
115 120 125

Gly Leu Trp Ile Leu Ala His Glu Cys Gly His Gly Ala Phe Ser Asp
130 135 140

Tyr Gln Asn Ile Asn Asp Phe Ile Gly Trp Val Leu His Ser Tyr Leu
145 150 155 160

```
Ile Val Pro Tyr Phe Ser Trp Lys Phe Ser His Ala Lys His His Lys
                165                 170                 175

Ala Thr Gly His Leu Thr Lys Asp Met Val Phe Ile Pro Tyr Thr Lys
            180                 185                 190

Glu Glu Tyr Leu Glu Lys Asn Lys Val Glu Lys Val Ala Asp Leu Met
        195                 200                 205

Glu Glu Ser Pro Ile Tyr Ser Phe Leu Val Leu Val Phe Gln Gln Leu
    210                 215                 220

Gly Gly Leu Gln Leu Tyr Leu Ala Thr Asn Ala Thr Gly Gln Val Tyr
225                 230                 235                 240

Pro Gly Tyr Ser Lys Ile Ala Lys Ser His Tyr Thr Pro Thr Ser Pro
                245                 250                 255

Val Phe Asp Lys His Gln Tyr Trp Tyr Ile Val Leu Ser Asp Ile Gly
            260                 265                 270

Ile Ile Leu Ala Phe Thr Thr Val Tyr Gln Trp Tyr Lys Asn Phe Gly
        275                 280                 285

Leu Phe Asn Met Met Ile Asn Trp Phe Val Pro Trp Leu Trp Val Asn
    290                 295                 300

His Trp Leu Val Phe Val Thr Phe Leu Gln His Thr Asp Pro Thr Met
305                 310                 315                 320

Pro His Tyr Thr Ser Lys Glu Trp Thr Phe Ala Arg Gly Ala Ala Ala
                325                 330                 335

Thr Ile Asp Arg Asn Phe Gly Phe Val Gly Gln His Ile Phe His Asp
            340                 345                 350

Ile Ile Glu Thr His Val Leu His His Tyr Val Ser Arg Ile Pro Phe
        355                 360                 365
```

Tyr Asn Ala Arg Glu Ala Thr Asp Ala Ile Arg Lys Val Met Gly Glu
370 375 380

His Tyr Arg Tyr Glu Gly Glu Ser Met Trp Tyr Ser Leu Trp Lys Cys
385 390 395 400

Met Arg Met Cys Gln Phe Val Asp Asp Asp Lys Glu Asp Ala Lys Gly
405 410 415

Val Met Met Phe Arg Asn Val Asn Gly Trp Gly Pro Val Lys Pro Lys
420 425 430

Asp

<210> 66
<211> 419
<212> PRT
<213> Saccharomyces kluyveri (GenBank Accession No. BAD11952)

<220>
<221> MISC_FEATURE
<223> delta-15 desaturase

<400> 66

Met Ser Ile Glu Thr Val Gly Ser Ser Ser Gly Val Ala Ile Asn Ser
1 5 10 15

Lys Ala Val Ser Ser Thr Ala Thr Thr Val Val Gln Pro Lys Thr Ala
20 25 30

Ile Asp Thr Asn Gly Asn Val Phe Lys Val Pro Asp Tyr Thr Ile Lys
35 40 45

Asp Ile Leu Ser Ala Ile Pro Lys Glu Cys Tyr Lys Arg Asp Thr Leu
50 55 60

Trp Ser Leu His Tyr Val Val Arg Asp Ile Ala Ala Ile Leu Val Ile
65 70 75 80

```
Gly Tyr Leu Gly Thr Asn Tyr Ile Pro Val Leu Phe Pro Asn Ser Ala
                85                  90                  95

Leu Leu Arg Gly Ile Ala Tyr Ala Ile Gln Ser Tyr Leu Ile Gly Leu
            100                 105                 110

Phe Gly Phe Gly Leu Trp Ile Leu Ala His Glu Cys Gly His Ser Ala
        115                 120                 125

Phe Ser Glu Ser Asn Ala Val Asn Asp Thr Val Gly Trp Val Leu His
    130                 135                 140

Ser Trp Trp Met Val Pro Tyr Phe Pro Trp Lys Phe Ser His Ser Lys
145                 150                 155                 160

His His Lys Ala Thr Gly His Met Thr Arg Asp Met Val Phe Ile Pro
                165                 170                 175

Tyr Thr Lys Asp Glu Phe Ile Thr Met Lys Lys Lys Ser Lys Phe Ala
            180                 185                 190

Glu Ile Thr Glu Glu Ala Pro Val Met Thr Leu Phe Asn Leu Ile Ala
        195                 200                 205

Gln Gln Val Gly Gly Leu Gln Leu Tyr Leu Ala Thr Asn Ala Thr Gly
    210                 215                 220

Gln Pro Tyr Pro Gly Val Lys Lys Phe Phe Lys Ser His Tyr Trp Pro
225                 230                 235                 240

Thr Ser Pro Val Phe Asp Ala Lys Asp Phe Trp Trp Ile Ile Met Ser
                245                 250                 255

Asp Ile Gly Ile Val Ser Thr Leu Leu Ile Asn Tyr Leu Trp Tyr Arg
            260                 265                 270

Ala Tyr Gly Ala His Val Val Leu Ile Asn Trp Phe Ile Pro Trp Leu
        275                 280                 285
```

```
Trp Val Asn His Trp Leu Val Phe Val Thr Phe Leu Gln His Thr Asp
    290                 295                 300

Pro Thr Met Pro His Tyr Asp Ala Glu Glu Trp Thr Phe Ala Lys Gly
305                 310                 315                 320

Ala Ala Ala Thr Ile Asp Arg Asn Phe Gly Phe Val Gly Gln His Ile
                325                 330                 335

Phe His Asp Ile Ile Glu Thr His Val Leu His His Tyr Cys Ser Arg
            340                 345                 350

Ile Pro Phe Tyr Asn Ala Arg Lys Ala Thr Ser Ala Ile Lys Glu Val
        355                 360                 365

Met Gly Gln His Tyr Arg Tyr Glu Gly Glu Asn Met Trp Lys Ser Leu
    370                 375                 380

Trp Lys Val Ala Arg Ser Cys Gln Tyr Val Glu Gly Asp Asn Gly Val
385                 390                 395                 400

Arg Met Phe Arg Asn Thr Asn Gly Val Gly Val Lys Pro Glu Asp Gly
                405                 410                 415

Ser Ser Gln
```

<210> 67
<211> 435
<212> PRT
<213> Debaryomyces hansenii CBS767 (GenBank Accession No. CAG88182)

<220>
<221> MISC_FEATURE
<223> delta-15 desaturase

<400> 67

```
Met Ser Val Val Asp Leu Thr Ser Thr Thr Ser Gly Ser Ala Ile Asn
1               5                   10                  15
```

```
Ser Ser Asn Ile Ser Gln Arg Gly Asn Gly Ser Thr Ile Val Glu Thr
            20                  25                  30

Lys Lys Gly Pro Ser Ser Asn Leu Lys Ala Ile Asp Thr Phe Gly Asn
        35                  40                  45

Glu Phe Lys Val Pro Asp Tyr Thr Ile Lys Gln Ile Leu Ser Ala Ile
    50                  55                  60

Pro Lys His Cys Tyr Glu Arg Ser Leu Val Arg Ser Leu Gly Tyr Val
65                  70                  75                  80

Ala Arg Asp Ile Thr Met Met Cys Leu Ile Gly Tyr Val Gly Gln Lys
                85                  90                  95

Thr Ile Pro Met Val Gln Ile Ala Asp Gln Glu Gly Leu Ser Thr Ala
            100                 105                 110

Ile Arg Gly Gly Leu Trp Cys Val Tyr Ser Tyr Leu Leu Gly Leu Phe
        115                 120                 125

Gly Phe Gly Leu Trp Ile Leu Ala His Glu Cys Gly His Gly Ala Phe
    130                 135                 140

Ser Asp Tyr Gln Asn Val Asn Asp Val Val Gly Trp Ile Leu His Ser
145                 150                 155                 160

Tyr Leu Ile Val Pro Tyr Phe Ser Trp Lys Phe Ser His Ser Lys His
                165                 170                 175

His Lys Ala Thr Gly His Leu Thr Lys Asp Met Val Phe Ile Pro Tyr
            180                 185                 190

Thr Lys Asp Glu Phe Val Glu Lys Ser Gly Val Ser Lys Val Ser Glu
        195                 200                 205

Val Met Glu Asp Ser Pro Ile Trp Ser Leu Met Val Leu Ile Phe Gln
    210                 215                 220
```

```
Gln Ile Gly Gly Leu Gln Leu Tyr Leu Ala Thr Asn Ala Thr Gly Gln
225                 230                 235                 240

Ser Tyr Gln Gly His Ser Lys Ile Ala Lys Ser His Tyr Ala Pro Ala
                245                 250                 255

Ser Pro Val Phe Asp Lys Glu His Tyr Trp Tyr Ile Ile Leu Ser Asp
            260                 265                 270

Ile Gly Ile Ile Thr Thr Ile Thr Val Val Tyr Gln Trp Tyr Lys Asn
        275                 280                 285

Phe Gly Phe Phe Asn Met Phe Val Asn Trp Phe Met Pro Trp Leu Trp
    290                 295                 300

Val Asn His Trp Leu Val Phe Val Thr Phe Leu Gln His Thr Asp Pro
305                 310                 315                 320

Thr Met Pro His Tyr Arg Asp Asn Glu Trp Thr Phe Ala Arg Gly Ala
                325                 330                 335

Ala Ala Thr Ile Asp Arg Asn Phe Gly Phe Ile Gly Gln His Ile Phe
            340                 345                 350

His Asp Ile Ile Glu Thr His Val Leu His His Tyr Val Ser Arg Ile
        355                 360                 365

Pro Phe Tyr Asn Ala Arg Glu Ala Thr Asp Ala Ile Arg Lys Val Met
    370                 375                 380

Gly Glu His Tyr Arg Tyr Glu Gly Glu Ser Met Trp Tyr Ser Leu Trp
385                 390                 395                 400

Lys Cys Met Arg Met Cys Gln Tyr Val Asp Asp Ala Asp Thr Asp Ala
                405                 410                 415

Lys Gly Val Leu Met Tyr Arg Asn Val Asn Gly Ala Gly Pro Val Lys
            420                 425                 430
```

Pro Ile Asp
435

<210> 68
<211> 396
<212> PRT
<213> Aspergillus fumigatus (GenBank Accession No. EAL85733)

<220>
<221> MISC_FEATURE
<223> delta-15 desaturase

<400> 68

Met Ala Gly Lys Met Ala Glu Val Arg Gln Arg Asp Val Gln Thr Glu
1 5 10 15

Thr Glu Val Val Gln Asp Ser Ile Pro Ser Leu Lys Ser Leu Lys Asp
20 25 30

Ala Ile Pro Lys Glu Cys Phe Glu Ser Ser Leu Ala Ile Ser Phe Leu
35 40 45

Tyr Leu Ala Arg Asp Ile Leu Tyr Cys Ala Ile Leu Thr Tyr Gly Ala
50 55 60

Phe His Ile His Leu Leu Pro Ser Leu Pro Leu Arg Val Leu Ala Trp
65 70 75 80

Ala Thr Tyr Gly Phe Phe Gln Gly Cys Val Gly Thr Gly Met Trp Ile
85 90 95

Leu Ala His Glu Cys Gly His Gly Ala Phe Ser Pro Tyr Gln Gly Ile
100 105 110

Asn Asp Phe Ile Gly Trp Ala Thr His Ser Phe Leu Leu Val Pro Tyr
115 120 125

Phe Ser Trp Lys Ile Thr His Ala Arg His His Arg Tyr Thr Gly His
130 135 140

```
Met Glu Lys Asp Thr Val Phe Val Pro Trp Thr Asp Glu Gln Leu Ala
145                 150                 155                 160

Lys Lys Arg Asn Val Arg Ile Glu Gln Leu Lys His Phe Ala Glu Glu
                165                 170                 175

Thr Pro Ile Val Ser Phe Leu Gln Leu Ile Gly His Gln Leu Gly Gly
            180                 185                 190

Trp Gln Leu Tyr Leu Leu Thr Asn Ala Thr Ala Gly Ala Gln Ser Trp
        195                 200                 205

Pro Glu Gly Lys Pro Lys Thr Gly Pro Ala Ser His Phe Asn Pro Val
    210                 215                 220

Gly Ala Leu Trp Thr Pro Ser Gln Arg Leu Ser Ile Ala Ile Ser Asp
225                 230                 235                 240

Leu Gly Leu Leu Ile Met Ala Ala Val Leu Tyr Tyr Ala Ser Thr Gln
                245                 250                 255

Ile Gly Ala Trp Asn Val Val Leu Leu Tyr Phe Val Pro Tyr Leu Trp
            260                 265                 270

Val His His Trp Leu Ile Ala Ile Thr Tyr Leu Gln His Thr His Pro
        275                 280                 285

Ser Val Pro His Tyr Thr Pro Glu Ala Trp Thr Tyr Thr Lys Gly Ala
    290                 295                 300

Leu Ala Thr Val Asp Arg Thr Met Gly Phe Ile Gly Arg His Phe Phe
305                 310                 315                 320

His Glu Ile Ile Asp Tyr His Val Val His His Leu Phe Ser Arg Ile
                325                 330                 335

Pro Phe Tyr Lys Ala Glu Gln Ala Thr Trp Ala Ile Gln Pro Leu Leu
            340                 345                 350
```

Gly Ala Gln Tyr His Glu Glu Lys Glu Gln Ser Phe Leu Gly Ser Leu
355 360 365

Val Thr Thr Phe Arg Lys Cys Ile Tyr Val Ser Ala Thr Gly Gln Pro
370 375 380

Gly Val Leu His Phe Val Lys Ala Asp Glu Gly Asn
385 390 395

<210> 69
<211> 792
<212> DNA
<213> Isochrysis galbana

<220>
<221> misc_feature
<223> delta-9 elongase

<400> 69
atggccctcg caaacgacgc gggagagcgc atctgggcgg ctgtgaccga cccggaaatc 60

ctcattggca ccttctcgta cttgctactc aaaccgctgc tccgcaattc cgggctggtg 120

gatgagaaga agggcgcata caggacgtcc atgatctggt acaacgttct gctggcgctc 180

ttctctgcgc tgagcttcta cgtgacggcg accgccctcg gctgggacta tggtacgggc 240

gcgtggctgc gcaggcaaac cggcgacaca ccgcagccgc tcttccagtg cccgtccccg 300

gtttgggact cgaagctctt cacatggacc gccaaggcat tctattactc caagtacgtg 360

gagtacctcg acacggcctg gctggtgctc aagggcaaga gggtctcctt tctccaggcc 420

ttccaccact ttggcgcgcc gtgggatgtg tacctcggca ttcggctgca caacgagggc 480

gtatggatct tcatgttttt caactcgttc attcacacca tcatgtacac ctactacggc 540

ctcaccgccg ccgggtataa gttcaaggcc aagccgctca tcaccgcgat gcagatctgc 600

cagttcgtgg gcggcttcct gttggtctgg gactacatca acgtcccctg cttcaactcg 660

gacaaaggga agttgttcag ctgggctttc aactatgcat acgtcggctc ggtcttcttg 720

ctcttctgcc acttttttcta ccaggacaac ttggcaacga agaaatcggc caaggcgggc 780

aagcagctct ag 792

<210> 70
<211> 263
<212> PRT
<213> Isochrysis galbana

<400> 70

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> 71
<211> 792
<212> DNA
<213> Artificial Sequence

<220>
<223> IgD9e: synthetic delta-9 elongase (codon-optimized for Yarrowia lipolytica)

<400> 71
atggctctgg ccaacgacgc tggcgagcga atctgggctg ccgtcaccga tcccgaaatc 60

ctcattggca ccttctccta cctgctcctg aagcctctcc tgcgaaactc tggtctcgtg 120

gacgagaaga aaggagccta ccgaacctcc atgatctggt acaacgtcct cctggctctc 180

ttctctgccc tgtccttcta cgtgactgcc accgctctcg gctgggacta cggtactgga 240

gcctggctgc gaagacagac cggtgatact ccccagcctc tctttcagtg tccctctcct 300

gtctgggact ccaagctgtt cacctggact gccaaggcct tctactattc taagtacgtg 360

gagtacctcg acaccgcttg gctggtcctc aagggcaagc gagtgtcctt tctgcaggcc 420

ttccatcact ttggagctcc ctgggacgtc tacctcggca ttcgactgca caacgagggt 480

gtgtggatct tcatgttctt taactcgttc attcacacca tcatgtacac ctactatgga 540

ctgactgccg ctggctacaa gttcaaggcc aagcctctga tcactgccat gcagatttgc 600

cagttcgtcg gtggctttct cctggtctgg gactacatca acgttccctg cttcaactct 660

gacaagggca agctgttctc ctgggctttc aactacgcct acgtcggatc tgtctttctc 720

ctgttctgtc acttctttta ccaggacaac ctggccacca agaaatccgc taaggctggt 780

aagcagcttt ag 792

<210> 72
<211> 1275
<212> DNA
<213> Euglena gracilis (GenBank Accession Nos. AF139720 and AAD45877)

<220>
<221> misc_feature
<222> (14)..(1273)
<223> non-functional delta-8 desaturase

<400> 72
attttttttc gaaatgaagt caaagcgcca agcgctatcc cccttacaat tgatggaaca 60

aacatatgat gtggtcaatt tccaccctgg tggtgcggaa attatagaga attaccaagg 120

aagggatgcc actgatgcct tcatggttat gcactttcaa gaagccttcg acaagctcaa 180

gcgcatgccc aaaatcaatc ccagttttga gttgccaccc caggctgcag tgaatgaagc 240

tcaagaggat ttccggaagc tccgagaaga gttgatcgca actggcatgt ttgatgcctc 300

ccccctctgg tactcataca aaatcagcac cacactgggc cttggagtgc tgggttattt 360

cctgatggtt cagtatcaga tgtatttcat tggggcagtg ttgcttggga tgcactatca 420

acagatgggc tggctttctc atgacatttg ccaccaccag actttcaaga accggaactg 480

gaacaacctc gtgggactgg tatttggcaa tggtctgcaa ggtttttccg tgacatgttg 540

gaaggacaga cacaatgcac atcattcggc aaccaatgtt caagggcacg accctgatat 600

tgacaacctc ccccccttag cctggtctga ggatgacgtc acacgggcgt caccgatttc 660

ccgcaagctc attcagttcc agcagtacta tttcttggtc atctgtatct tgttgcggtt 720

catttggtgt ttccagtgcg tgttgaccgt gcgcagtttg aaggacagag ataaccaatt 780

ctatcgctct cagtataaga aggaggccat tggcctcgcc ctgcactgga ccttgaaggc 840

cctgttccac ttattcttta tgcccagcat cctcacatcg ctgttggtgt ttttcgtttc 900

ggagctggtt ggcggcttcg gcattgcgat cgtggtgttc atgaaccact acccactgga 960

gaagatcggg gacccagtct gggatggcca tggattctcg gttggccaga tccatgagac 1020

catgaacatt cggcgaggga ttatcacaga ttggtttttc ggaggcttga attaccagat 1080

tgagcaccat ttgtggccga ccctccctcg ccacaacctg acagcggtta gctaccaggt 1140

ggaacagctg tgccagaagc acaacctgcc gtatcggaac ccgctgcccc atgaagggtt 1200

ggtcatcctg ctgcgctatc tggcggtgtt cgcccggatg gcggagaagc aacccgcggg 1260

gaaggctcta taagg 1275

<210> 73
<211> 419
<212> PRT
<213> Euglena gracilis (GenBank Accession No. AF139720 and AAD45877)

<400> 73

Met Lys Ser Lys Arg Gln Ala Leu Ser Pro Leu Gln Leu Met Glu Gln
1 5 10 15

Thr Tyr Asp Val Val Asn Phe His Pro Gly Gly Ala Glu Ile Ile Glu
20 25 30

Asn Tyr Gln Gly Arg Asp Ala Thr Asp Ala Phe Met Val Met His Phe
35 40 45

Gln Glu Ala Phe Asp Lys Leu Lys Arg Met Pro Lys Ile Asn Pro Ser
50 55 60

Phe Glu Leu Pro Pro Gln Ala Ala Val Asn Glu Ala Gln Glu Asp Phe
65 70 75 80

Arg Lys Leu Arg Glu Glu Leu Ile Ala Thr Gly Met Phe Asp Ala Ser
85 90 95

Pro Leu Trp Tyr Ser Tyr Lys Ile Ser Thr Thr Leu Gly Leu Gly Val
100 105 110

Leu Gly Tyr Phe Leu Met Val Gln Tyr Gln Met Tyr Phe Ile Gly Ala
115 120 125

```
Val Leu Leu Gly Met His Tyr Gln Gln Met Gly Trp Leu Ser His Asp
    130                 135                 140

Ile Cys His His Gln Thr Phe Lys Asn Arg Asn Trp Asn Asn Leu Val
145                 150                 155                 160

Gly Leu Val Phe Gly Asn Gly Leu Gln Gly Phe Ser Val Thr Cys Trp
                165                 170                 175

Lys Asp Arg His Asn Ala His His Ser Ala Thr Asn Val Gln Gly His
            180                 185                 190

Asp Pro Asp Ile Asp Asn Leu Pro Pro Leu Ala Trp Ser Glu Asp Asp
        195                 200                 205

Val Thr Arg Ala Ser Pro Ile Ser Arg Lys Leu Ile Gln Phe Gln Gln
    210                 215                 220

Tyr Tyr Phe Leu Val Ile Cys Ile Leu Leu Arg Phe Ile Trp Cys Phe
225                 230                 235                 240

Gln Cys Val Leu Thr Val Arg Ser Leu Lys Asp Arg Asp Asn Gln Phe
                245                 250                 255

Tyr Arg Ser Gln Tyr Lys Lys Glu Ala Ile Gly Leu Ala Leu His Trp
            260                 265                 270

Thr Leu Lys Ala Leu Phe His Leu Phe Phe Met Pro Ser Ile Leu Thr
        275                 280                 285

Ser Leu Leu Val Phe Phe Val Ser Glu Leu Val Gly Gly Phe Gly Ile
    290                 295                 300

Ala Ile Val Val Phe Met Asn His Tyr Pro Leu Glu Lys Ile Gly Asp
305                 310                 315                 320

Pro Val Trp Asp Gly His Gly Phe Ser Val Gly Gln Ile His Glu Thr
                325                 330                 335
```

```
Met Asn Ile Arg Arg Gly Ile Ile Thr Asp Trp Phe Phe Gly Gly Leu
            340                 345                 350

Asn Tyr Gln Ile Glu His His Leu Trp Pro Thr Leu Pro Arg His Asn
        355                 360                 365

Leu Thr Ala Val Ser Tyr Gln Val Glu Gln Leu Cys Gln Lys His Asn
    370                 375                 380

Leu Pro Tyr Arg Asn Pro Leu Pro His Glu Gly Leu Val Ile Leu Leu
385                 390                 395                 400

Arg Tyr Leu Ala Val Phe Ala Arg Met Ala Glu Lys Gln Pro Ala Gly
                405                 410                 415

Lys Ala Leu
```

<210> 74
<211> 422
<212> PRT
<213> Euglena gracilis

<400> 74

```
Met Lys Ser Lys Arg Gln Ala Leu Ser Pro Leu Gln Leu Met Glu Gln
1               5                  10                  15

Thr 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 Phe Gln Glu Ala Phe Asp Lys Leu Lys Arg Met Pro Lys Ile
    50                  55                  60

Asn Pro Ser Phe 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 Ser Thr Thr Leu Gly
            100                 105                 110

Leu Gly Val Leu Gly Tyr Phe Leu Met Val Gln Tyr Gln Met Tyr Phe
        115                 120                 125

Ile Gly Ala Val Leu 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 Leu Val Gly Leu Val Phe Gly Asn Gly Leu Gln Gly Phe Ser Val
                165                 170                 175

Thr Cys 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 Pro 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 Cys 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 Ala Leu Phe His Leu Phe Phe Met Pro Ser
        275                 280                 285
```

```
Ile Leu Thr Ser Leu 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 Pro 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 Ala Val Phe Ala Arg Met Ala Glu Lys Gln
                405                 410                 415

Pro Ala Gly Lys Ala Leu
            420
```

<210> 75
<211> 1270
<212> DNA
<213> Artificial Sequence

<220>
<223> D8S-1: Synthetic gene codon-optimized for expression in Yarrowia lipolytica

<400> 75

```
ccatggagtc caagcgacag gctctgtctc ccctccagct gatggaacag acctacgacg     60

tcgtgaactt ccaccctggt ggagctgaaa tcattgagaa ctaccaggga cgagatgcta    120

ctgacgcctt catggttatg cactttcagg aagccttcga caagctcaag cgaatgccca    180
```

agatcaaccc ctcctttgag ctgcctcccc aggctgccgt caacgaagct caggaggatt 240

tccgaaagct ccgagaagag ctgatcgcca ctggcatgtt tgacgcctct cccctctggt 300

actcgtacaa gatctccacc accctgggtc ttggcgtgct tggatacttc ctgatggtcc 360

agtaccagat gtacttcatt ggtgctgtgc tgctcggtat gcactaccag caaatgggat 420

ggctgtctca tgacatctgc caccaccaga ccttcaagaa ccgaaactgg aataacctcg 480

tgggtctggt ctttggcaac ggactccagg gcttctccgt gacctgttgg aaggacagac 540

acaacgccca tcattctgct accaacgttc agggtcacga tcccgacatt gataacctgc 600

ctcccctcgc ctggtccgag gacgatgtca ctcgagcttc tcccatctcc cgaaagctca 660

ttcagttcca acagtactat ttcctggtca tctgtattct cctgcgattc atctggtgtt 720

tccagtgcgt gctgaccgtt cgatccctca aggaccgaga caaccagttc taccgatctc 780

agtacaagaa agaggccatt ggactcgctc tgcactggac tctcaaggct ctgttccacc 840

tcttctttat gccctccatc ctgacctcgc tcctggtgtt ctttgtttcc gagctcgtcg 900

gtggcttcgg aattgccatc gtggtcttca tgaaccacta ccctctggag aagatcggtg 960

atcccgtctg ggacggacat ggcttctctg tgggtcagat ccatgagacc atgaacattc 1020

gacgaggcat cattactgac tggttctttg gaggcctgaa ctaccagatc gagcaccatc 1080

tctggcccac cctgcctcga cacaacctca ctgccgtttc ctaccaggtg gaacagctgt 1140

gccagaagca caacctcccc taccgaaacc ctctgcccca tgaaggtctc gtcatcctgc 1200

tccgatacct ggccgtgttc gctcgaatgg ccgagaagca gcccgctggc aaggctctct 1260

aagcggccgc 1270

<210> 76
<211> 1269
<212> DNA
<213> Artificial Sequence

<220>
<223> D8S-3: synthetic delta 8-desaturase gene codon-optimized for Yarrowia lipolytica in pDMW261

<400> 76
atgaagtcca agcgacaggc tctgtctccc ctccagctga tggaacagac ctacgacgtc 60

tccgcttggg tgaacttcca ccctggtgga gctgaaatca ttgagaacta ccagggacga 120

gatgctactg acgccttcat ggttatgcac tttcaggaag ccttcgacaa gctcaagcga 180

atgcccaaga tcaacccctc ctttgagctg cctccccagg ctgccgtcaa cgaagctcag 240

gaggatttcc gaaagctccg agaagagctg atcgccactg gcatgtttga cgcctctccc 300

ctctggtact cgtacaagat ctccaccacc ctgggtcttg gcgtgcttgg atacttcctg 360

atggtccagt accagatgta cttcattggt gctgtgctgc tcggtatgca ctaccagcaa 420

atgggatggc tgtctcatga catctgccac caccagacct tcaagaaccg aaactggaat 480

aacctcgtgg gtctggtctt tggcaacgga ctccagggct tctccgtgac ctgttggaag 540

gacagacaca acgcccatca ttctgctacc aacgttcagg gtcacgatcc cgacattgat 600

aacctgcctc ccctcgcctg gtccgaggac gatgtcactc gagcttctcc catctcccga 660

aagctcattc agttccaaca gtactatttc ctggtcatct gtattctcct gcgattcatc 720

tggtgtttcc agtgcgtgct gaccgttcga tccctcaagg accgagacaa ccagttctac 780

cgatctcagt acaagaaaga ggccattgga ctcgctctgc actggactct caaggctctg 840

ttccacctct tctttatgcc ctccatcctg acctcgctcc tggtgttctt tgtttccgag 900

ctcgtcggtg gcttcggaat tgccatcgtg gtcttcatga accactaccc tctggagaag 960

atcggtgatc ccgtctggga cggacatggc ttctctgtgg gtcagatcca tgagaccatg 1020

aacattcgac gaggcatcat tactgactgg ttctttggag gcctgaacta ccagatcgag 1080

caccatctct ggcccaccct gcctcgacac aacctcactg ccgtttccta ccaggtggaa 1140

cagctgtgcc agaagcacaa cctcccctac cgaaaccctc tgccccatga aggtctcgtc 1200

atcctgctcc gatacctggc cgtgttcgct cgaatggccg agaagcagcc cgctggcaag 1260

gctctctaa 1269

<210> 77
<211> 1271
<212> DNA
<213> Euglena gracilis

<220>
<221> misc_feature
<222> (4)..(1269)
<223> Eg5: delta-8 desaturase

<400> 77

gaaatgaagt caaagcgcca agcgcttccc cttacaattg atggaacaac atatgatgtg 60

tctgcctggg tcaatttcca ccctggtggt gcggaaatta tagagaatta ccaaggaagg 120

gatgccactg atgccttcat ggttatgcac tctcaagaag ccttcgacaa gctcaagcgc 180

atgcccaaaa tcaatcccag ttctgagttg ccaccccagg ctgcagtgaa tgaagctcaa 240

gaggatttcc ggaagctccg agaagagttg atcgcaactg gcatgtttga tgcctccccc 300

ctctggtact catacaaaat cagcaccaca ctgggccttg gagtgctggg ttatttcctg 360

atggttcagt atcagatgta tttcattggg gcagtgttgc ttgggatgca ctatcaacag 420

atgggctggc tttctcatga catttgccac caccagactt tcaagaaccg gaactggaac 480

aacctcgtgg gactggtatt tggcaatggt ctgcaaggtt tttccgtgac atggtggaag 540

gacagacaca atgcacatca ttcggcaacc aatgttcaag ggcacgaccc tgatattgac 600

aacctccccc tcttagcctg gtctgaggat gacgtcacac gggcgtcacc gatttcccgc 660

aagctcattc agttccagca gtactatttc ttggtcatct gtatcttgtt gcggttcatt 720

tggtgtttcc agagcgtgtt gaccgtgcgc agtttgaagg acagagataa ccaattctat 780

cgctctcagt ataagaagga ggccattggc ctcgccctgc actggacctt gaagaccctg 840

ttccacttat tctttatgcc cagcatcctc acatcgctgt tggtgttttt cgtttcggag 900

ctggttggcg gcttcggcat tgcgatcgtg gtgttcatga accactaccc actggagaag 960

atcggggact cagtctggga tggccatgga ttctcggttg gccagatcca tgagaccatg 1020

aacattcggc gagggattat cacagattgg tttttcggag gcttgaatta ccagattgag 1080

caccatttgt ggccgaccct ccctcgccac aacctgacag cggttagcta ccaggtggaa 1140

cagctgtgcc agaagcacaa cctgccgtat cggaacccgc tgccccatga agggttggtc 1200

atcctgctgc gctatctggc ggtgttcgcc cggatggcgg agaagcaacc cgcggggaag 1260

gctctataag g 1271

<210> 78
<211> 421
<212> PRT
<213> Euglena gracilis

<400> 78

Met Lys Ser Lys Arg Gln Ala Leu Pro Leu Thr Ile Asp Gly Thr Thr
1 5 10 15

```
Tyr Asp Val Ser Ala Trp Val Asn Phe His Pro Gly Gly Ala Glu Ile
            20                  25                  30

Ile Glu Asn Tyr Gln Gly Arg Asp Ala Thr Asp Ala Phe Met Val Met
        35                  40                  45

His Ser Gln Glu Ala Phe Asp Lys Leu Lys Arg Met Pro Lys Ile Asn
    50                  55                  60

Pro Ser Ser Glu Leu Pro Pro Gln Ala Ala Val Asn Glu Ala Gln Glu
65                  70                  75                  80

Asp Phe Arg Lys Leu Arg Glu Glu Leu Ile Ala Thr Gly Met Phe Asp
                85                  90                  95

Ala Ser Pro Leu Trp Tyr Ser Tyr Lys Ile Ser Thr Thr Leu Gly Leu
            100                 105                 110

Gly Val Leu Gly Tyr Phe Leu Met Val Gln Tyr Gln Met Tyr Phe Ile
        115                 120                 125

Gly Ala Val Leu Leu Gly Met His Tyr Gln Gln Met Gly Trp Leu Ser
    130                 135                 140

His Asp Ile Cys His His Gln Thr Phe Lys Asn Arg Asn Trp Asn Asn
145                 150                 155                 160

Leu Val Gly Leu Val Phe Gly Asn Gly Leu Gln Gly Phe Ser Val Thr
                165                 170                 175

Trp Trp Lys Asp Arg His Asn Ala His His Ser Ala Thr Asn Val Gln
            180                 185                 190

Gly His Asp Pro Asp Ile Asp Asn Leu Pro Leu Leu Ala Trp Ser Glu
        195                 200                 205

Asp Asp Val Thr Arg Ala Ser Pro Ile Ser Arg Lys Leu Ile Gln Phe
    210                 215                 220
```

```
Gln Gln Tyr Tyr Phe Leu Val Ile Cys Ile Leu Leu Arg Phe Ile Trp
225                 230                 235                 240

Cys Phe Gln Ser Val Leu Thr Val Arg Ser Leu Lys Asp Arg Asp Asn
                245                 250                 255

Gln Phe Tyr Arg Ser Gln Tyr Lys Lys Glu Ala Ile Gly Leu Ala Leu
            260                 265                 270

His Trp Thr Leu Lys Thr Leu Phe His Leu Phe Phe Met Pro Ser Ile
        275                 280                 285

Leu Thr Ser Leu Leu Val Phe Phe Val Ser Glu Leu Val Gly Gly Phe
    290                 295                 300

Gly Ile Ala Ile Val Val Phe Met Asn His Tyr Pro Leu Glu Lys Ile
305                 310                 315                 320

Gly Asp Ser Val Trp Asp Gly His Gly Phe Ser Val Gly Gln Ile His
                325                 330                 335

Glu Thr Met Asn Ile Arg Arg Gly Ile Ile Thr Asp Trp Phe Phe Gly
            340                 345                 350

Gly Leu Asn Tyr Gln Ile Glu His His Leu Trp Pro Thr Leu Pro Arg
        355                 360                 365

His Asn Leu Thr Ala Val Ser Tyr Gln Val Glu Gln Leu Cys Gln Lys
    370                 375                 380

His Asn Leu Pro Tyr Arg Asn Pro Leu Pro His Glu Gly Leu Val Ile
385                 390                 395                 400

Leu Leu Arg Tyr Leu Ala Val Phe Ala Arg Met Ala Glu Lys Gln Pro
                405                 410                 415

Ala Gly Lys Ala Leu
            420
```

<210> 79
<211> 1271
<212> DNA
<213> Euglena gracilis

<220>
<221> misc_feature
<222> (4)..(1269)
<223> Eg12: delta-8 desaturase

<400> 79
gaaatgaagt caaagcgcca agcgcttccc cttacaattg atggaacaac atatgatgtg 60

tctgcctggg tcaatttcca ccctggtggt gcggaaatta tagagaatta ccaaggaagg 120

gatgccactg atgccttcat ggttatgcac tctcaagaag ccttcgacaa gctcaagcgc 180

atgcccaaaa tcaatcccag ttctgagttg ccaccccagg ctgcagtgaa tgaagctcaa 240

gaggatttcc ggaagctccg agaagagttg atcgcaactg gcatgtttga tgcctccccc 300

ctctggtact catacaaaat cagcaccaca ctgggccttg gagtgctggg ttatttcctg 360

atggttcagt atcagatgta tttcattggg gcagtgttgc ttgggatgca ctatcaacag 420

atgggctggc tttctcatga catttgccac caccagactt tcaagaaccg gaactggaac 480

aacctcgtgg gactggtatt tggcaatggt ctgcaaggtt tttccgtgac atggtggaag 540

gacagacaca atgcacatca ttcggcaacc aatgttcaag ggcacgaccc tgatattgac 600

aacctccccc tcttagcctg gtctgaggat gacgtcacac gggcgtcacc gatttcccgc 660

aagctcattc agttccagca gtactatttc ttggtcatct gtatcttgtt gcggttcatt 720

tggtgtttcc agagcgtgtt gaccgtgcgc agtttgaagg acagagataa ccaattctat 780

cgctctcagt ataagaagga ggccattggc ctcgccctgc actggacctt gaaggccctg 840

ttccacttat tctttatgcc cagcatcctc acatcgctgt tggtgttttt cgtttcggag 900

ctggttggcg gcttcggcat tgcgatcgtg gtgttcatga accactaccc actggagaag 960

atcggggact cagtctggga tggccatgga ttctcggttg gccagatcca tgagaccatg 1020

aacattcggc gagggattat cacagattgg tttttcggag gcttgaatta ccagattgag 1080

caccatttgt ggccgaccct ccctcgccac aacctgacag cggttagcta ccaggtggaa 1140

cagctgtgcc agaagcacaa cctgccgtat cggaacccgc tgccccatga agggttggtc 1200

atcctgctgc gctatctggc ggtgttcgcc cggatggcgg agaagcaacc cgcggggaag 1260

gctctataag g 1271

<210> 80
<211> 421
<212> PRT
<213> Euglena gracilis

<400> 80

Met Lys Ser Lys Arg Gln Ala Leu Pro Leu Thr Ile Asp Gly Thr Thr
1 5 10 15

Tyr Asp Val Ser Ala Trp Val Asn Phe His Pro Gly Gly Ala Glu Ile
20 25 30

Ile Glu Asn Tyr Gln Gly Arg Asp Ala Thr Asp Ala Phe Met Val Met
35 40 45

His Ser Gln Glu Ala Phe Asp Lys Leu Lys Arg Met Pro Lys Ile Asn
50 55 60

Pro Ser Ser Glu Leu Pro Pro Gln Ala Ala Val Asn Glu Ala Gln Glu
65 70 75 80

Asp Phe Arg Lys Leu Arg Glu Glu Leu Ile Ala Thr Gly Met Phe Asp
85 90 95

Ala Ser Pro Leu Trp Tyr Ser Tyr Lys Ile Ser Thr Thr Leu Gly Leu
100 105 110

Gly Val Leu Gly Tyr Phe Leu Met Val Gln Tyr Gln Met Tyr Phe Ile
115 120 125

Gly Ala Val Leu Leu Gly Met His Tyr Gln Gln Met Gly Trp Leu Ser
130 135 140

His Asp Ile Cys His His Gln Thr Phe Lys Asn Arg Asn Trp Asn Asn
145 150 155 160

```
Leu Val Gly Leu Val Phe Gly Asn Gly Leu Gln Gly Phe Ser Val Thr
                165                 170                 175

Trp Trp Lys Asp Arg His Asn Ala His His Ser Ala Thr Asn Val Gln
            180                 185                 190

Gly His Asp Pro Asp Ile Asp Asn Leu Pro Leu Leu Ala Trp Ser Glu
        195                 200                 205

Asp Asp Val Thr Arg Ala Ser Pro Ile Ser Arg Lys Leu Ile Gln Phe
    210                 215                 220

Gln Gln Tyr Tyr Phe Leu Val Ile Cys Ile Leu Leu Arg Phe Ile Trp
225                 230                 235                 240

Cys Phe Gln Ser Val Leu Thr Val Arg Ser Leu Lys Asp Arg Asp Asn
                245                 250                 255

Gln Phe Tyr Arg Ser Gln Tyr Lys Lys Glu Ala Ile Gly Leu Ala Leu
            260                 265                 270

His Trp Thr Leu Lys Ala Leu Phe His Leu Phe Phe Met Pro Ser Ile
        275                 280                 285

Leu Thr Ser Leu Leu Val Phe Phe Val Ser Glu Leu Val Gly Gly Phe
    290                 295                 300

Gly Ile Ala Ile Val Val Phe Met Asn His Tyr Pro Leu Glu Lys Ile
305                 310                 315                 320

Gly Asp Ser Val Trp Asp Gly His Gly Phe Ser Val Gly Gln Ile His
                325                 330                 335

Glu Thr Met Asn Ile Arg Arg Gly Ile Ile Thr Asp Trp Phe Phe Gly
            340                 345                 350

Gly Leu Asn Tyr Gln Ile Glu His His Leu Trp Pro Thr Leu Pro Arg
        355                 360                 365
```

His Asn Leu Thr Ala Val Ser Tyr Gln Val Glu Gln Leu Cys Gln Lys
370 375 380

His Asn Leu Pro Tyr Arg Asn Pro Leu Pro His Glu Gly Leu Val Ile
385 390 395 400

Leu Leu Arg Tyr Leu Ala Val Phe Ala Arg Met Ala Glu Lys Gln Pro
405 410 415

Ala Gly Lys Ala Leu
420

<210> 81
<211> 1272
<212> DNA
<213> Artificial Sequence

<220>
<223> D8SF: synthetic delta-8 desaturase (codon-optimized for Yarrowia lipolytica)

<400> 81
catggtgaag tccaagcgac aggctctgcc cctcaccatc gacggaacta cctacgacgt 60

ctccgcttgg gtgaacttcc accctggtgg agctgaaatc attgagaact accagggacg 120

agatgctact gacgccttca tggttatgca ctctcaggaa gccttcgaca agctcaagcg 180

aatgcccaag atcaacccct cctccgagct gcctccccag gctgccgtca acgaagctca 240

ggaggatttc cgaaagctcc gagaagagct gatcgccact ggcatgtttg acgcctctcc 300

cctctggtac tcgtacaaga tctccaccac cctgggtctt ggcgtgcttg gatacttcct 360

gatggtccag taccagatgt acttcattgg tgctgtgctg ctcggtatgc actaccagca 420

aatgggatgg ctgtctcatg acatctgcca ccaccagacc ttcaagaacc gaaactggaa 480

taacctcgtg ggtctggtct ttggcaacgg actccagggc ttctccgtga cctggtggaa 540

ggacagacac aacgcccatc attctgctac caacgttcag ggtcacgatc ccgacattga 600

taacctgcct ctgctcgcct ggtccgagga cgatgtcact cgagcttctc ccatctcccg 660

aaagctcatt cagttccaac agtactattt cctggtcatc tgtattctcc tgcgattcat 720

ctggtgtttc cagtctgtgc tgaccgttcg atccctcaag gaccgagaca accagttcta 780

ccgatctcag tacaagaaag aggccattgg actcgctctg cactggactc tcaagaccct 840

gttccacctc ttctttatgc cctccatcct gacctcgctc ctggtgttct ttgtttccga 900

gctcgtcggt ggcttcggaa ttgccatcgt ggtcttcatg aaccactacc ctctggagaa 960

gatcggtgat tccgtctggg acggacatgg cttctctgtg ggtcagatcc atgagaccat 1020

gaacattcga cgaggcatca ttactgactg gttctttgga ggcctgaact accagatcga 1080

gcaccatctc tggcccaccc tgcctcgaca caacctcact gccgtttcct accaggtgga 1140

acagctgtgc cagaagcaca acctccccta ccgaaaccct ctgccccatg aaggtctcgt 1200

catcctgctc cgatacctgg ccgtgttcgc tcgaatggcc gagaagcagc ccgctggcaa 1260

ggctctctaa gc 1272

<210> 82
<211> 422
<212> PRT
<213> Artificial Sequence

<220>
<223> D8SF: synthetic delta-8 desaturase (codon-optimized for Yarrowia lipolytica)

<400> 82

Met Val Lys Ser Lys Arg Gln Ala Leu Pro Leu Thr Ile Asp Gly Thr
1 5 10 15

Thr 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 Pro Ser 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 Ser Thr Thr Leu Gly
            100                 105                 110

Leu Gly Val Leu Gly Tyr Phe Leu Met Val Gln Tyr Gln Met Tyr Phe
        115                 120                 125

Ile Gly Ala Val Leu 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 Leu Val 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 Leu 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 Ala Val Phe Ala Arg Met Ala Glu Lys Gln
                405                 410                 415

Pro Ala Gly Lys Ala Leu
            420
```

<210> 83
<211> 2628
<212> DNA
<213> Rattus norvegicus (GenBank Accession No. AB071986)

<220>
<221> CDS
<222> (90)..(893)
<223> C16/18 elongase

<400> 83

```
tggatgcgga cgctgggagg agagcccctg agctaggagc tgggagccga ggcgcagaga     60

acacgtagcg actccgaaga tcagcccca atg aac atg tca gtg ttg act tta      113
                                Met Asn Met Ser Val Leu Thr Leu
                                1               5
```

```
caa gaa tat gaa ttc gag aag cag ttc aac gag aat gaa gcc atc caa      161
Gln Glu Tyr Glu Phe Glu Lys Gln Phe Asn Glu Asn Glu Ala Ile Gln
    10                  15                  20

tgg atg cag gaa aac tgg aag aaa tct ttc ctg ttt tct gcg ctg tac      209
Trp Met Gln Glu Asn Trp Lys Lys Ser Phe Leu Phe Ser Ala Leu Tyr
25                  30                  35                  40

gct gcc ttt ata ttt ggt ggt cgg cat ctg atg aac aag cgg gcc aag      257
Ala Ala Phe Ile Phe Gly Gly Arg His Leu Met Asn Lys Arg Ala Lys
                45                  50                  55

ttt gaa ctg cgg aag ccg ctc gtg ctc tgg tca ctg act ctt gcg gtc      305
Phe Glu Leu Arg Lys Pro Leu Val Leu Trp Ser Leu Thr Leu Ala Val
            60                  65                  70

ttc agt ata ttc ggt gct ctt cga act ggt gct tac atg ctg tac att      353
Phe Ser Ile Phe Gly Ala Leu Arg Thr Gly Ala Tyr Met Leu Tyr Ile
        75                  80                  85

ttg atg acc aaa ggc ctg aag cag tca gtt tgt gac cag agt ttt tac      401
Leu Met Thr Lys Gly Leu Lys Gln Ser Val Cys Asp Gln Ser Phe Tyr
    90                  95                  100

aac gga cct gtc agc aaa ttc tgg gct tac gcg ttt gtg ctc agc aaa      449
Asn Gly Pro Val Ser Lys Phe Trp Ala Tyr Ala Phe Val Leu Ser Lys
105                 110                 115                 120

gca ccc gaa cta ggt gat acg ata ttc atc att ctg agg aag cag aag      497
Ala Pro Glu Leu Gly Asp Thr Ile Phe Ile Ile Leu Arg Lys Gln Lys
                125                 130                 135

ctg atc ttc ctg cac tgg tac cac cac atc act gtg ctc ctg tac tct      545
Leu Ile Phe Leu His Trp Tyr His His Ile Thr Val Leu Leu Tyr Ser
            140                 145                 150

tgg tac tcc tac aaa gac atg gta gct ggg ggt ggt tgg ttc atg act      593
Trp Tyr Ser Tyr Lys Asp Met Val Ala Gly Gly Gly Trp Phe Met Thr
        155                 160                 165

atg aac tat ggc gta cac gcc gtc atg tac tct tac tac gcc ttg cgg      641
Met Asn Tyr Gly Val His Ala Val Met Tyr Ser Tyr Tyr Ala Leu Arg
    170                 175                 180

gct gcg ggt ttc cgg gtc tcc cgg aag ttt gcc atg ttc atc acg ttg      689
Ala Ala Gly Phe Arg Val Ser Arg Lys Phe Ala Met Phe Ile Thr Leu
185                 190                 195                 200

tcc cag atc act cag atg ctg atg ggc tgt gtc att aac tac ctg gtc      737
Ser Gln Ile Thr Gln Met Leu Met Gly Cys Val Ile Asn Tyr Leu Val
                205                 210                 215
```

```
ttc aac tgg atg cag cat gac aat gac cag tgc tac tcc cac ttt cag      785
Phe Asn Trp Met Gln His Asp Asn Asp Gln Cys Tyr Ser His Phe Gln
            220                 225                 230

aac atc ttc tgg tcc tca ctc atg tac ctc agc tac ctt ctg ctc ttc      833
Asn Ile Phe Trp Ser Ser Leu Met Tyr Leu Ser Tyr Leu Leu Leu Phe
        235                 240                 245

tgc cat ttc ttc ttt gag gcc tac atc ggc aaa gtg aag aaa gcg acg      881
Cys His Phe Phe Phe Glu Ala Tyr Ile Gly Lys Val Lys Lys Ala Thr
    250                 255                 260

aag gcc gag tag tgtcagagct gaggaggaag acatagctca gggtcatcac          933
Lys Ala Glu
265

gaaaaataat agacaaaaag aaaatggcac aaggaatcac atatggtgca gctaaaacaa    993

aacaaaacat tatgagcaga cgctaagccc aaggcagctt gggagtgaag attaggttgt   1053

aagtttatga tccttttttgg gtgaggactc actgagaaca ctgctgctga gggacccccct  1113

tccctcttac ctgtcaactc tagaacacac tagaagccaa ggcagccatg ggcaaggaga   1173

ttagtggaca gcaagcaaaa cactgcagga agaggggggga gatctattca gagtttttttg  1233

ttttgttttg ttttgttttt ctctaaggat aaaggagttt ccccttttca aactgtgtga   1293

gcacacccac gcgcatgcgg acacacacac acacacacat acacacacac acacacatac   1353

acacatacac acacacacac acacacacac acacaatctt ttccacagga aaccagagct   1413

ggtagaaaag ataaacggta agcgacaggg tttctatcta ggacagcaat gcttttgcaa   1473

aaacctaagc cttttaaaga agttagcttg taactccttg acaaaagatg tcttaattct   1533

ttttactgta actgaaagtc aaaaaggtag ataccttccc cttcactgca cagcctcggg   1593

cttgttcgtt tgctacaacg gagcagagga cagttcttcc gtgatacttt atttctggga   1653

gaaagaaaac catgcagccc aaatcccgag aaggcggccc atagctaacc ctgcagttga   1713

agcatcacac tgacagctgt tatttctgct cttcggtgag aattgaagca gccgttgttc   1773

aattacccca aactttagga tgggggagta aatacggaat tgaaaagaaa gaagctcgac   1833

tggttggctt gaaaatggaa tcttgtacta tctataacaa aactcagccc atcgtccctg   1893

gagctggaag gactatcaga agagaccaga actgaagatt catccaatac acagatggca   1953

tgttcgcctc cttccccgtt tgacctcaca catagtcctg gctttctaaa tgaggtccta   2013
```

tgacccagtc tgtgttttct atatttttgt gactttcaaa aacagatccg cagggctctg 2073

catttggggt aaacactgtg tttctgcagc ctctgcattc gctcccttca gcaatgcaga 2133

ggcgtgagaa gtgccctctg ctggcttagt gagaagcttc aacaaacact tcataatagg 2193

ttgaaatagc tgaccacaaa gggcctgcgt agattaaacc ctaagttaag ttctaagtgc 2253

tgtcaaacac ctgacatata tttgaccaaa tcagaagaga gagaacctct atgcttcaag 2313

taagcttcat aaaaaatttt taaagtgact ttcacttggg aactcagaaa gtcaatgtat 2373

taagagccat attcttaaaa aaaaaaagaa agaaagaaaa gaagaagaag aaagctagac 2433

aatgttatct gtaatatttc agtcctttac aagccaaata aatgtgtcga tgttcctaaa 2493

aaaaaaaaaa aaaacggtcc gcggccgccc tttagtgagg gttaatttaa atcgtacgtc 2553

gcgattaatt aaccgcggta ccttctgagg cggaaagaac cagccggatc cctcgaggga 2613

tccagacatg ataaa 2628

<210> 84
<211> 267
<212> PRT
<213> Rattus norvegicus (GenBank Accession No. AB071986)

<400> 84

Met Asn Met Ser Val Leu Thr Leu Gln Glu Tyr Glu Phe Glu Lys Gln
1 5 10 15

Phe Asn Glu Asn Glu Ala Ile Gln Trp Met Gln Glu Asn Trp Lys Lys
20 25 30

Ser Phe Leu Phe Ser Ala Leu Tyr Ala Ala Phe Ile Phe Gly Gly Arg
35 40 45

His Leu Met Asn Lys Arg Ala Lys Phe Glu Leu Arg Lys Pro Leu Val
50 55 60

Leu Trp Ser Leu Thr Leu Ala Val Phe Ser Ile Phe Gly Ala Leu Arg
65 70 75 80

Thr Gly Ala Tyr Met Leu Tyr Ile Leu Met Thr Lys Gly Leu Lys Gln
85 90 95

```
Ser Val Cys Asp Gln Ser Phe Tyr Asn Gly Pro Val Ser Lys Phe Trp
            100                 105                 110

Ala Tyr Ala Phe Val Leu Ser Lys Ala Pro Glu Leu Gly Asp Thr Ile
        115                 120                 125

Phe Ile Ile Leu Arg Lys Gln Lys Leu Ile Phe Leu His Trp Tyr His
    130                 135                 140

His Ile Thr Val Leu Leu Tyr Ser Trp Tyr Ser Tyr Lys Asp Met Val
145                 150                 155                 160

Ala Gly Gly Gly Trp Phe Met Thr Met Asn Tyr Gly Val His Ala Val
                165                 170                 175

Met Tyr Ser Tyr Tyr Ala Leu Arg Ala Ala Gly Phe Arg Val Ser Arg
            180                 185                 190

Lys Phe Ala Met Phe Ile Thr Leu Ser Gln Ile Thr Gln Met Leu Met
        195                 200                 205

Gly Cys Val Ile Asn Tyr Leu Val Phe Asn Trp Met Gln His Asp Asn
    210                 215                 220

Asp Gln Cys Tyr Ser His Phe Gln Asn Ile Phe Trp Ser Ser Leu Met
225                 230                 235                 240

Tyr Leu Ser Tyr Leu Leu Leu Phe Cys His Phe Phe Phe Glu Ala Tyr
                245                 250                 255

Ile Gly Lys Val Lys Lys Ala Thr Lys Ala Glu
            260                 265
```

<210> 85
<211> 804
<212> DNA
<213> Rattus norvegicus

<220>
<221> misc_feature
<223> synthetic C16/18 elongase (codon-optimized)

<400> 85
atggacatgt ccgtcctgac tctccaagag tacgagttcg agaagcagtt caacgagaat 60

gaagccatcc aatggatgca ggaaaactgg aagaaatcct tcctgttttc tgccctctac 120

gctgccttta tctttggtgg acgacatctg atgaacaagc gagccaagtt tgagctgcga 180

aaacctctcg tgctctggtc cctgaccctc gctgtcttct ctatcttcgg tgctctgcga 240

actggagcct acatgctcta catcctgatg accaaaggcc tgaaacagtc tgtttgtgac 300

cagtcctttt acaacggacc cgtctcgaaa ttctgggctt acgcctttgt gctctccaaa 360

gctcccgaac ttggcgatac catcttcatc attctgcgaa agcagaaact catcttcctg 420

cactggtatc accacatcac cgtcctcctg tactcttggt actcctacaa ggacatggtg 480

gctggaggtg gctggttcat gactatgaac tacggtgtcc acgccgtgat gtactcctac 540

tacgccctcc gagctgccgg tttccgagtc tctcgaaagt ttgccatgtt catcaccctg 600

tcgcagatca ctcagatgct catgggctgt gtcattaact acctggtctt caactggatg 660

cagcatgaca atgaccagtg ctactcccac tttcagaaca tcttctggtc ctctctcatg 720

tacctctcct accttctgct cttctgccat ttcttctttg aggcctacat tggcaaagtg 780

aagaaagcca ccaaggctga gtaa 804

<210> 86
<211> 828
<212> DNA
<213> Mortierella alpina

<220>
<221> misc_feature
<223> C16/18 elongase

<400> 86
atggagtctg gacctatgcc tgccgggatc cccttccctg aatactatga ctttttcatg 60

gactggaaga cacccctggc aattgctgcc acctacaccg ccgctgttgg gctcttcaac 120

cccaaggttg gcaaagtctc gcgcgtggta gccaagtcgg ctaacgccaa gccggcagag 180

cgcacgcagt ccggcgccgc catgaccgcc tttgtctttg tccacaacct tatcctctgc 240

gtgtactctg gaatcacctt ctactacatg ttcccagcca tggtcaagaa ctttagaaca 300

cataccctcc atgaggccta ctgcgatacg gatcagagcc tgtggaacaa cgcccttggc 360

tactggggct acctcttcta cctttcaaag ttttacgagg tcattgacac catcatcatc 420

atcttgaagg ggcgccgctc gtccctgctc cagacctacc accacgccgg cgctatgatc 480

accatgtggt ccggcatcaa ctaccaggca acgcccattt ggatttttgt cgtcttcaac 540

tcgttcatcc acaccatcat gtactgttac tatgccttca cctcaatcgg cttccacccc 600

ccaggcaaga agtacctcac ctccatgcag atcacccagt ttttggtcgg catcactatc 660

gccgtctctt atctcttcgt ccctggatgt atccgcacac ccggtgctca gatggctgtc 720

tggatcaacg tcggatacct ctttcccctc acttatctct ttgtggattt tgccaagcgt 780

acttactcca agcgtagtgc catcgccgct cagaagaagg cccagtaa 828

<210> 87
<211> 275
<212> PRT
<213> Mortierella alpina

<400> 87

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> 88
<211> 607
<212> DNA
<213> Mortierella alpina

<220>
<221> misc_feature
<222> (133)..(135)
<223> ATG translation initiation codon

<400> 88
cccgtcttca ccctccctct tccccccgtt cctacgtcta cagccgttgg ctcatcttgc 60

agttgcttgt ctactatttg gtgccgacct atattcctct gtcacccaac ctaccgcact 120

cacactcgca taatggagtc tggacctatg cctgccggga tccccttccc tgaatactat 180

gacttttca tggactggaa gacacccctg gcaattgctg ccacctacac cgccgctgtt 240

gggctcttca accccaaggt tggcaaagtc tcgcgcgtgg tagccaagtc ggctaacgcc 300

aagccggcag agcgcacgca gtccggcgcc gccatgaccg cctttgtctt tgtccacaac 360

cttatcctct gcgtgtactc tggaatcacc ttctactaca tgttcccagc catggtcaag 420

aactttagaa cacataccct ccatgaggcc tactgcgata cggatcagag cctgtggaac 480

aacgcccttg gctactgggg ctacctcttc tacctttcaa agttttacga ggtcattgac 540

accatcatca tcatcttgaa ggggcgccgc tcgtccctgc tccagaccta ccaccacgcc 600

ggcgcta 607

<210> 89
<211> 1042
<212> DNA
<213> Mortierella alpina

<220>
<221> misc_feature
<222> (402)..(404)
<223> Stop codon

<400> 89
tgaaggggcg ccgctcgtcc ctgctccaga cctaccacca cgccggcgct atgatcacca 60

tgtggtccgg catcaactac caggcaacgc ccatttggat ttttgtcgtc ttcaactcgt 120

tcatccacac catcatgtac tgttactatg ccttcacctc aatcggcttc cacccccag 180

gcaagaagta cctcacctcc atgcagatca cccagttttt ggtcggcatc actatcgccg 240

tctcttatct cttcgtccct ggatgtatcc gcacacccgg tgctcagatg gctgtctgga 300

tcaacgtcgg atacctcttt cccctcactt atctctttgt ggattttgcc aagcgtactt 360

actccaagcg tagtgccatc gccgctcaga agaaggccca gtaaagaggc cacttcttgt 420

cgtttactct actgtttcca acatttacat ctctcaatgg cagcctcctc cctctatcct 480

tcgcactttc attttagttc cttgtattca taaaargtca aaattcattg taactccaaa 540

agtgccaatg gctttatcat ggatgagtgc catgtttatg gcccattgca tcaatacagg 600

aagatatgag cggtcagcga tagctattcc tctgagttgt cctggccttg tatgcactct 660

cgaaggagga aatggatcgg agcatcagca ctatttcagg actaggagag atacatccaa 720

aattttaaag gcatctctgg tcagtacaag agcgacttgg caacaagtct ttagtcaaac 780

ttcctcacct ttgaacttcc tcagaggtcg ccgtggcact gtggacggag acccgtgaaa 840

agaactcatg cagcaggtta agcaaccttg gaaatccatc gatgaatgcc tgatgtaaaa 900

attgagaagc tgcaagagcc aaaagttcag ttcagcacat ataagagact ttggccaaga 960

gcgatgaagt cacatgggaa ccacgtaaac aaacctcgtg cggagsacca gcccgggccg 1020

tcgaccacgc gtgccctata gt 1042

<210> 90
<211> 2223
<212> DNA
<213> Mortierella alpina

<220>
<221> misc_feature
<222> (2092)..(2094)
<223> ATG translation initiation codon

<400> 90
actatagggc acgcgtggtc gacggcccgg gctggtccta cctctcaatc ttgttgccct 60

gtagtcctga aatgagctcg tcatgggtct ctgtcctctc ctcgttatgg tcaaacgcat 120

tggacccggt tttctttaga ggagcattct tcttgtttaa gccgactcct cttcccgtcc 180

cgaacgacaa cttgataggt ccaccatttc catttggtgg gtgaagcggg ctgtcctcag 240

cggacatgac tctgagctag aaggaagggt gtagaggacg aaggagtaag tgctaccgtt 300

cactggcagg tgaggagcgt cgggtctaag ccacatgcct gccaattgtc ctcgccaagt 360

gcttggaagg tgtattgatg ctcatgcagc tccactctgg acctggaaac gccaccgcat 420

cgcattctta atgagctctg ctacaaacga caatctgcga accagtttgc acagggagct 480

agtcgagtct ggaaagcggg accagtgagt gactcggcgg ggcttgcatc gctgaacatc 540

gccattggga tttcttagaa agcagctctt gctcttatac aaacaccttc ctgcgtgcct 600

caccttactt acgaggagta atgttgtttt tacaggctag tggaactctt gcgcgcgagg 660

ctgatcgagt ccggctggcg ggacgatctg aaggcattta caaaaggtac accactcggc 720

acttggcctg ttttgttgcc actagctacc aggccgtgac tgaccacgaa cccttcctct 780

ggagaataga aaggatccag tcccgggaga ctacaatgtc tgttgacgag gtcatcagag 840

aagtcggccc acatgctaga ggtaagcccc gcgagaacct tatatgctga ttagtgccat 900

ccgttgaaca tcgtctgatc tgcgacaagt tctggtaata cttgatagcc accgtcccag 960

acaaggtcaa gacagaactt cttgcgcaaa tcagtacttt tatcgaagac aacatgtaac 1020

gaaaaataaa aatgctgata tggaagggaa catcacgtac gccaggtacg tgagatgttt 1080

ttccttttt ttttatttt tttgttgggg cttgagatgt tttgggtctt tcatgaagac 1140

gttcatgaca aacaagaaat acggacagcg agtttaccgc aatggaaaag tttacctcaa 1200

ccgaactttg aacgtttcgt gaaaacctgc gtgagaattc tttgagcatg gccataccta 1260

attgtcgccg gactgatctg tgatttgatc catggtccat cgctcctgac tcgctgcacg 1320

tcctgacaga gataacctcg catagtgtcg tacagaaacc aaaagcacgc gctcggtcca 1380

gtacagtaat ggggaactgg gtactgagga cagcttgaaa gtatatcctg taggggacga 1440

taacatgggc aatcttgatg gtgtggttga gtacagggta tatgcgttcc aggtaagcaa 1500

atatcacctt tgtgagacga taggggggcag ctatagtaaa tggccaggcg cttttatacg 1560

ggaagcaatt gcagcgatgg aggtgtgtac gctgcagccg cttgtatagc agaacgtaac 1620

cccaatcctc caactccggc aaggacaatg gggaaagaag gccggatata taaaatgctt 1680

gtcagcagaa ttaaagactg tgcggactga tctgccttaa tcggaaactt cagcagggct 1740

tctcgcagtg catgaaccct cggaccatct cctgtcccaa caacacgcca atgggcttat 1800

tattaatttt ttttggctgc cgttcaaaaa aaaaaagaaa aaaactgcca cttcttattt 1860

gagggctgca ggactgcttg ttaaggcgga taacctcaat ccgcccgatg ccttttaacg 1920

ctggcctgtt cttcctcttt tctttcttct ctcccttcac ccgtcttcac cctccctctt 1980

ccccccgttc ctacgtctac agccgttggc tcatcttgca gttgcttgtc tactatttgg 2040

tgccgaccta tattcctctg tcacccaacc taccgcactc acactcgcat aatggagtct 2100

ggacctatgc ctgccgggat ccccttccct gaatactatg acttttttcat ggactggaag 2160

acacccctgg caattgctgc cacctacacc gccgctgttg ggctcttcaa ccccaaggtt 2220

ggc 2223

<210> 91
<211> 3557
<212> DNA
<213> Mortierella alpina

<220>
<221> misc_feature
<222> (2092)..(2094)
<223> ATG translation initiation codon

<220>
<221> misc_feature
<222> (2917)..(2919)
<223> 'TAA' stop codon

<400> 91
actatagggc acgcgtggtc gacggcccgg gctggtccta cctctcaatc ttgttgccct 60

gtagtcctga aatgagctcg tcatgggtct ctgtcctctc ctcgttatgg tcaaacgcat 120

tggacccggt tttctttaga ggagcattct tcttgtttaa gccgactcct cttcccgtcc 180

cgaacgacaa cttgataggt ccaccatttc catttggtgg gtgaagcggg ctgtcctcag 240

cggacatgac tctgagctag aaggaagggt gtagaggacg aaggagtaag tgctaccgtt 300

cactggcagg tgaggagcgt cgggtctaag ccacatgcct gccaattgtc ctcgccaagt 360

gcttggaagg tgtattgatg ctcatgcagc tccactctgg acctggaaac gccaccgcat 420

cgcattctta atgagctctg ctacaaacga caatctgcga accagtttgc acagggagct 480

agtcgagtct ggaaagcggg accagtgagt gactcggcgg ggcttgcatc gctgaacatc 540

gccattggga tttcttagaa agcagctctt gctcttatac aaacaccttc ctgcgtgcct 600

caccttactt acgaggagta atgttgtttt tacaggctag tggaactctt gcgcgcgagg 660

ctgatcgagt ccggctggcg ggacgatctg aaggcattta caaaaggtac accactcggc 720

acttggcctg ttttgttgcc actagctacc aggccgtgac tgaccacgaa cccttcctct 780

ggagaataga aaggatccag tcccgggaga ctacaatgtc tgttgacgag gtcatcagag 840

aagtcggccc acatgctaga ggtaagcccc gcgagaacct tatatgctga ttagtgccat 900

ccgttgaaca tcgtctgatc tgcgacaagt tctggtaata cttgatagcc accgtcccag 960

acaaggtcaa gacagaactt cttgcgcaaa tcagtacttt tatcgaagac aacatgtaac 1020

gaaaaataaa aatgctgata tggaagggaa catcacgtac gccaggtacg tgagatgttt 1080

ttccttttt ttttatttt tttgttgggg cttgagatgt tttgggtctt tcatgaagac 1140

gttcatgaca aacaagaaat acggacagcg agtttaccgc aatggaaaag tttacctcaa 1200

ccgaactttg aacgtttcgt gaaaacctgc gtgagaattc tttgagcatg gccataccta 1260

attgtcgccg gactgatctg tgatttgatc catggtccat cgctcctgac tcgctgcacg 1320

tcctgacaga gataacctcg catagtgtcg tacagaaacc aaaagcacgc gctcggtcca 1380

gtacagtaat ggggaactgg gtactgagga cagcttgaaa gtatatcctg taggggacga 1440

taacatgggc aatcttgatg gtgtggttga gtacagggta tatgcgttcc aggtaagcaa 1500

atatcacctt tgtgagacga taggggggcag ctatagtaaa tggccaggcg cttttatacg 1560

ggaagcaatt gcagcgatgg aggtgtgtac gctgcagccg cttgtatagc agaacgtaac 1620

cccaatcctc caactccggc aaggacaatg gggaaagaag gccggatata taaaatgctt 1680

gtcagcagaa ttaaagactg tgcggactga tctgccttaa tcggaaactt cagcagggct 1740

tctcgcagtg catgaaccct cggaccatct cctgtcccaa caacacgcca atgggcttat 1800

tattaatttt ttttggctgc cgttcaaaaa aaaaaagaaa aaaactgcca cttcttattt 1860

gagggctgca ggactgcttg ttaaggcgga taacctcaat ccgcccgatg ccttttaacg 1920

ctggcctgtt cttcctcttt tctttcttct ctcccttcac ccgtcttcac cctccctctt 1980

ccccccgttc ctacgtctac agccgttggc tcatcttgca gttgcttgtc tactatttgg 2040

tgccgaccta tattcctctg tcacccaacc taccgcactc acactcgcat aatggagtct 2100

ggacctatgc ctgccgggat ccccttccct gaatactatg acttttttcat ggactggaag 2160

acacccctgg caattgctgc cacctacacc gccgctgttg ggctcttcaa ccccaaggtt 2220

ggcaaagtct cgcgcgtggt agccaagtcg gctaacgcca agccggcaga gcgcacgcag 2280

tccggcgccg ccatgaccgc ctttgtcttt gtccacaacc ttatcctctg cgtgtactct 2340

ggaatcacct tctactacat gttcccagcc atggtcaaga actttagaac acataccctc 2400

catgaggcct actgcgatac ggatcagagc ctgtggaaca acgcccttgg ctactggggc 2460

tacctcttct acctttcaaa gttttacgag gtcattgaca ccatcatcat catcttgaag 2520

gggcgccgct cgtccctgct ccagacctac caccacgccg gcgctatgat caccatgtgg 2580

tccggcatca actaccaggc aacgcccatt tggatttttg tcgtcttcaa ctcgttcatc 2640

cacaccatca tgtactgtta ctatgccttc acctcaatcg gcttccaccc cccaggcaag 2700

aagtacctca cctccatgca gatcacccag tttttggtcg gcatcactat cgccgtctct 2760

tatctcttcg tccctggatg tatccgcaca cccggtgctc agatggctgt ctggatcaac 2820

gtcggatacc tctttccccct cacttatctc tttgtggatt ttgccaagcg tacttactcc 2880

aagcgtagtg ccatcgccgc tcagaagaag gcccagtaaa gaggccactt cttgtcgttt 2940

actctactgt ttccaacatt tacatctctc aatggcagcc tcctccctct atccttcgca 3000

ctttcatttt agttccttgt attcataaaa rgtcaaaatt cattgtaact ccaaaagtgc 3060

caatggcttt atcatggatg agtgccatgt ttatggccca ttgcatcaat acaggaagat 3120

atgagcggtc agcgatagct attcctctga gttgtcctgg ccttgtatgc actctcgaag 3180

gaggaaatgg atcggagcat cagcactatt tcaggactag gagagataca tccaaaattt 3240

taaaggcatc tctggtcagt acaagagcga cttggcaaca agtctttagt caaacttcct 3300

caccttttgaa cttcctcaga ggtcgccgtg gcactgtgga cggagacccg tgaaaagaac 3360

tcatgcagca ggttaagcaa ccttggaaat ccatcgatga atgcctgatg taaaaattga 3420

gaagctgcaa gagccaaaag ttcagttcag cacatataag agactttggc caagagcgat 3480

gaagtcacat gggaaccacg taaacaaacc tcgtgcggag saccagcccg ggccgtcgac 3540

cacgcgtgcc ctatagt 3557

<210> 92
<211> 542
<212> DNA
<213> Mortierella alpina

<220>
<221> Intron
<222> (1)..(542)

<400> 92
cgtaagttga caacggttca aatttcgaaa cttgcactgt gggtttgtcc tggttttatt 60

cggcagcttt tttctccgaa gtcttggccc agggaggagg caagaagagc aattcagggt 120

aggggcgtcc ggagattgga ccctgccttg gactagcgaa tggaatacga gagcgaagct 180

gtctgcatgg gctgatgata atatttttta ctggagatga ttgtaaggtg cggcgcgagg 240

gtgtagtaca tcagtcgaca tggattcaac caaatgttct ctagcagcga taaccctggc 300

tgtcttttgt taatcacggc aatgctatcc ccaaagaaaa aaaaaagggt gcggttggca 360

cacggttgca caatggctta gccagacaat taacccaaaa aagtcaagag attccttttt 420

ttttttttt ttttttgggg ggggaacggg aaagtggggg ctgtgtgaaa ttggggcgaa 480

gtgaaccatg caactttgag ctacaaacac ctaattgttt ttgtcttgcc tgacattatt 540

ag 542

<210> 93
<211> 4099
<212> DNA
<213> Mortierella alpina

<220>
<221> misc_feature
<222> (2092)..(2094)
<223> ATG translation initiation codon

<220>
<221> Intron
<222> (2410)..(2951)

<220>
<221> misc_feature
<222> (3459)..(3461)
<223> 'TAA' stop codon

<400> 93
actatagggc acgcgtggtc gacggcccgg gctggtccta cctctcaatc ttgttgccct 60

gtagtcctga aatgagctcg tcatgggtct ctgtcctctc ctcgttatgg tcaaacgcat 120

tggacccggt tttctttaga ggagcattct tcttgtttaa gccgactcct cttcccgtcc 180

cgaacgacaa cttgataggt ccaccatttc catttggtgg gtgaagcggg ctgtcctcag 240

cggacatgac tctgagctag aaggaagggt gtagaggacg aaggagtaag tgctaccgtt 300

cactggcagg tgaggagcgt cgggtctaag ccacatgcct gccaattgtc ctcgccaagt 360

gcttggaagg tgtattgatg ctcatgcagc tccactctgg acctggaaac gccaccgcat 420

cgcattctta atgagctctg ctacaaacga caatctgcga accagtttgc acagggagct 480

agtcgagtct ggaaagcggg accagtgagt gactcggcgg ggcttgcatc gctgaacatc 540

gccattggga tttcttagaa agcagctctt gctcttatac aaacaccttc ctgcgtgcct 600

caccttactt acgaggagta atgttgtttt tacaggctag tggaactctt gcgcgcgagg 660

ctgatcgagt ccggctggcg ggacgatctg aaggcattta caaaaggtac accactcggc 720

acttggcctg ttttgttgcc actagctacc aggccgtgac tgaccacgaa cccttcctct 780

ggagaataga aaggatccag tcccgggaga ctacaatgtc tgttgacgag gtcatcagag 840

aagtcggccc acatgctaga ggtaagcccc gcgagaacct tatatgctga ttagtgccat 900

ccgttgaaca tcgtctgatc tgcgacaagt tctggtaata cttgatagcc accgtcccag 960

acaaggtcaa gacagaactt cttgcgcaaa tcagtacttt tatcgaagac aacatgtaac 1020

gaaaaataaa aatgctgata tggaagggaa catcacgtac gccaggtacg tgagatgttt 1080

ttccttttt ttttatttt tttgttgggg cttgagatgt tttgggtctt tcatgaagac 1140

gttcatgaca aacaagaaat acggacagcg agtttaccgc aatggaaaag tttacctcaa 1200

ccgaactttg aacgtttcgt gaaaacctgc gtgagaattc tttgagcatg gccataccta 1260

attgtcgccg gactgatctg tgatttgatc catggtccat cgctcctgac tcgctgcacg 1320

tcctgacaga gataacctcg catagtgtcg tacagaaacc aaaagcacgc gctcggtcca 1380

gtacagtaat ggggaactgg gtactgagga cagcttgaaa gtatatcctg taggggacga 1440

taacatgggc aatcttgatg gtgtggttga gtacagggta tatgcgttcc aggtaagcaa 1500

atatcacctt tgtgagacga taggggggcag ctatagtaaa tggccaggcg cttttatacg 1560

ggaagcaatt gcagcgatgg aggtgtgtac gctgcagccg cttgtatagc agaacgtaac 1620

cccaatcctc caactccggc aaggacaatg gggaaagaag gccggatata taaaatgctt 1680

gtcagcagaa ttaaagactg tgcggactga tctgccttaa tcggaaactt cagcagggct 1740

tctcgcagtg catgaaccct cggaccatct cctgtcccaa caacacgcca atgggcttat 1800

tattaatttt ttttggctgc cgttcaaaaa aaaaaagaaa aaaactgcca cttcttattt 1860

gagggctgca ggactgcttg ttaaggcgga taacctcaat ccgcccgatg ccttttaacg 1920

ctggcctgtt cttcctcttt tctttcttct ctcccttcac ccgtcttcac cctccctctt 1980

ccccccgttc ctacgtctac agccgttggc tcatcttgca gttgcttgtc tactatttgg 2040

tgccgaccta tattcctctg tcacccaacc taccgcactc acactcgcat aatggagtct 2100

ggacctatgc ctgccgggat ccccttccct gaatactatg acttttcat ggactggaag 2160

acacccctgg caattgctgc cacctacacc gccgctgttg ggctcttcaa ccccaaggtt 2220

ggcaaagtct cgcgcgtggt agccaagtcg gctaacgcca agccggcaga gcgcacgcag 2280

tccggcgccg ccatgaccgc ctttgtcttt gtccacaacc ttatcctctg cgtgtactct 2340

ggaatcacct tctactacat gttcccagcc atggtcaaga actttagaac acataccctc 2400

catgaggccc gtaagttgac aacggttcaa atttcgaaac ttgcactgtg ggtttgtcct 2460

ggttttattc ggcagctttt ttctccgaag tcttggccca gggaggaggc aagaagagca 2520

attcagggta ggggcgtccg gagattggac cctgccttgg actagcgaat ggaatacgag 2580

agcgaagctg tctgcatggg ctgatgataa tatttttac tggagatgat tgtaaggtgc 2640

ggcgcgaggg tgtagtacat cagtcgacat ggattcaacc aaatgttctc tagcagcgat 2700

aaccctggct gtcttttgtt aatcacggca atgctatccc caaagaaaaa aaaaagggtg 2760

cggttggcac acggttgcac aatggcttag ccagacaatt aacccaaaaa agtcaagaga 2820

ttccttttt ttttttttt tttttgggggg gggaacggga aagtgggggc tgtgtgaaat 2880

tggggcgaag tgaaccatgc aactttgagc tacaaacacc taattgtttt tgtcttgcct 2940

gacattatta gtactgcgat acggatcaga gcctgtggaa caacgccctt ggctactggg 3000

gctacctctt ctacctttca aagttttacg aggtcattga caccatcatc atcatcttga 3060

aggggcgccg ctcgtccctg ctccagacct accaccacgc cggcgctatg atcaccatgt 3120

ggtccggcat caactaccag gcaacgccca tttggatttt tgtcgtcttc aactcgttca 3180

tccacaccat catgtactgt tactatgcct tcacctcaat cggcttccac cccccaggca 3240

agaagtacct cacctccatg cagatcaccc agtttttggt cggcatcact atcgccgtct 3300

cttatctctt cgtccctgga tgtatccgca cacccggtgc tcagatggct gtctggatca 3360

acgtcggata cctctttccc ctcacttatc tctttgtgga ttttgccaag cgtacttact 3420

ccaagcgtag tgccatcgcc gctcagaaga aggcccagta aagaggccac ttcttgtcgt 3480

ttactctact gtttccaaca tttacatctc tcaatggcag cctcctccct ctatccttcg 3540

cactttcatt ttagttcctt gtattcataa aargtcaaaa ttcattgtaa ctccaaaagt 3600

gccaatggct ttatcatgga tgagtgccat gtttatggcc cattgcatca atacaggaag 3660

atatgagcgg tcagcgatag ctattcctct gagttgtcct ggccttgtat gcactctcga 3720

aggaggaaat ggatcggagc atcagcacta tttcaggact aggagagata catccaaaat 3780

tttaaaggca tctctggtca gtacaagagc gacttggcaa caagtcttta gtcaaacttc 3840

ctcacctttg aacttcctca gaggtcgccg tggcactgtg gacggagacc cgtgaaaaga 3900

actcatgcag caggttaagc aaccttggaa atccatcgat gaatgcctga tgtaaaaatt 3960

gagaagctgc aagagccaaa agttcagttc agcacatata agagactttg gccaagagcg 4020

atgaagtcac atgggaacca cgtaaacaaa cctcgtgcgg agsaccagcc cgggccgtcg 4080

accacgcgtg ccctatagt 4099

<210> 94
<211> 915
<212> DNA
<213> Yarrowia lipolytica

<220>
<221> CDS
<222> (1)..(915)
<223> C16/18 elongase

<400> 94
atg ctc tcg tca atc tcg ccc gac cta tac tcg tcc ttc tcg ttc aaa 48
Met Leu Ser Ser Ile Ser Pro Asp Leu Tyr Ser Ser Phe Ser Phe Lys
1 5 10 15

aac tcg ctc gcc gag gcc atg ccc tcc gtg cca cac gaa ctc atc aac 96
Asn Ser Leu Ala Glu Ala Met Pro Ser Val Pro His Glu Leu Ile Asn
20 25 30

tca aaa aca ctc tca tgg atg tac aat gcc tct ctg gac att cgg gtt 144
Ser Lys Thr Leu Ser Trp Met Tyr Asn Ala Ser Leu Asp Ile Arg Val
35 40 45

cct ctg act atc gga acc atc tac gcc gtc tcc gtg cac ctg acc aac 192
Pro Leu Thr Ile Gly Thr Ile Tyr Ala Val Ser Val His Leu Thr Asn

```
    50                  55                  60

tca tct gaa cga atc aag aaa cgc cag ccc att gcc ttt gcc aag acc      240
Ser Ser Glu Arg Ile Lys Lys Arg Gln Pro Ile Ala Phe Ala Lys Thr
65                  70                  75                  80

gca ctc ttc aag tgg ctc tgt gtc ctc cac aat gca ggt ctg tgt ctc      288
Ala Leu Phe Lys Trp Leu Cys Val Leu His Asn Ala Gly Leu Cys Leu
                85                  90                  95

tac tca gca tgg acc ttt gtc ggt atc ctc aac gcc gtc aaa cac gcc      336
Tyr Ser Ala Trp Thr Phe Val Gly Ile Leu Asn Ala Val Lys His Ala
            100                 105                 110

tac caa atc aca gga gac agc tcc gcc ccc ttc tcc ttc aac acc ctc      384
Tyr Gln Ile Thr Gly Asp Ser Ser Ala Pro Phe Ser Phe Asn Thr Leu
        115                 120                 125

tgg gga tcg ttt tgt tca cgt gac tcc ctc tgg gtc acc ggc ctc aac      432
Trp Gly Ser Phe Cys Ser Arg Asp Ser Leu Trp Val Thr Gly Leu Asn
    130                 135                 140

tac tac gga tac tgg ttc tat ctg tcc aaa ttc tac gaa gtg gtg gac      480
Tyr Tyr Gly Tyr Trp Phe Tyr Leu Ser Lys Phe Tyr Glu Val Val Asp
145                 150                 155                 160

acc atg atc atc ctc gca aag gga aaa ccg tcc tca atg ctc cag aca      528
Thr Met Ile Ile Leu Ala Lys Gly Lys Pro Ser Ser Met Leu Gln Thr
                165                 170                 175

tac cac cac acc ggc gcc atg ttc tcc atg tgg gcc ggc atc cga ttc      576
Tyr His His Thr Gly Ala Met Phe Ser Met Trp Ala Gly Ile Arg Phe
            180                 185                 190

gcc tct ccc ccc atc tgg atc ttt gtg gtt ttc aac tcc ctc atc cac      624
Ala Ser Pro Pro Ile Trp Ile Phe Val Val Phe Asn Ser Leu Ile His
        195                 200                 205

aca atc atg tac ttt tac tac acc ctc acc acc ctc aag atc aag gtt      672
Thr Ile Met Tyr Phe Tyr Tyr Thr Leu Thr Thr Leu Lys Ile Lys Val
    210                 215                 220

ccc aag atc ctc aag gca tct ctg acc acc gcc cag atc acc cag att      720
Pro Lys Ile Leu Lys Ala Ser Leu Thr Thr Ala Gln Ile Thr Gln Ile
225                 230                 235                 240

gtc gga ggt ggc atc ctg gct gcc tcc cac gcc ttt att tat tac aag      768
Val Gly Gly Gly Ile Leu Ala Ala Ser His Ala Phe Ile Tyr Tyr Lys
                245                 250                 255

gac cac cag act gag acc gtc tgt tct tgt ctc act acc cag ggt cag      816
Asp His Gln Thr Glu Thr Val Cys Ser Cys Leu Thr Thr Gln Gly Gln
```

```
            260                 265                 270

ttt ttc gct ctc gcc gtc aat gtc atc tat ctg agt cct ctg gcc tat      864
Phe Phe Ala Leu Ala Val Asn Val Ile Tyr Leu Ser Pro Leu Ala Tyr
        275                 280                 285

ctc ttt att gcc ttc tgg att cga tct tac ttg aag gcc aag tcc aac      912
Leu Phe Ile Ala Phe Trp Ile Arg Ser Tyr Leu Lys Ala Lys Ser Asn
    290                 295                 300

tag                                                                  915


<210>  95
<211>  304
<212>  PRT
<213>  Yarrowia lipolytica

<400>  95

Met Leu Ser Ser Ile Ser Pro Asp Leu Tyr Ser Ser Phe Ser Phe Lys
1               5                   10                  15


Asn Ser Leu Ala Glu Ala Met Pro Ser Val Pro His Glu Leu Ile Asn
            20                  25                  30


Ser Lys Thr Leu Ser Trp Met Tyr Asn Ala Ser Leu Asp Ile Arg Val
        35                  40                  45


Pro Leu Thr Ile Gly Thr Ile Tyr Ala Val Ser Val His Leu Thr Asn
    50                  55                  60


Ser Ser Glu Arg Ile Lys Lys Arg Gln Pro Ile Ala Phe Ala Lys Thr
65                  70                  75                  80


Ala Leu Phe Lys Trp Leu Cys Val Leu His Asn Ala Gly Leu Cys Leu
                85                  90                  95


Tyr Ser Ala Trp Thr Phe Val Gly Ile Leu Asn Ala Val Lys His Ala
            100                 105                 110


Tyr Gln Ile Thr Gly Asp Ser Ser Ala Pro Phe Ser Phe Asn Thr Leu
        115                 120                 125


Trp Gly Ser Phe Cys Ser Arg Asp Ser Leu Trp Val Thr Gly Leu Asn
    130                 135                 140
```

```
Tyr Tyr Gly Tyr Trp Phe Tyr Leu Ser Lys Phe Tyr Glu Val Val Asp
145                 150                 155                 160

Thr Met Ile Ile Leu Ala Lys Gly Lys Pro Ser Ser Met Leu Gln Thr
                165                 170                 175

Tyr His His Thr Gly Ala Met Phe Ser Met Trp Ala Gly Ile Arg Phe
            180                 185                 190

Ala Ser Pro Pro Ile Trp Ile Phe Val Val Phe Asn Ser Leu Ile His
        195                 200                 205

Thr Ile Met Tyr Phe Tyr Tyr Thr Leu Thr Thr Leu Lys Ile Lys Val
    210                 215                 220

Pro Lys Ile Leu Lys Ala Ser Leu Thr Thr Ala Gln Ile Thr Gln Ile
225                 230                 235                 240

Val Gly Gly Gly Ile Leu Ala Ala Ser His Ala Phe Ile Tyr Tyr Lys
                245                 250                 255

Asp His Gln Thr Glu Thr Val Cys Ser Cys Leu Thr Thr Gln Gly Gln
            260                 265                 270

Phe Phe Ala Leu Ala Val Asn Val Ile Tyr Leu Ser Pro Leu Ala Tyr
        275                 280                 285

Leu Phe Ile Ala Phe Trp Ile Arg Ser Tyr Leu Lys Ala Lys Ser Asn
    290                 295                 300
```

<210> 96
<211> 353
<212> PRT
<213> Candida albicans (GenBank Accession No. EAL04510)

<400> 96

```
Met Leu Ala Lys Phe Lys Leu Pro Asn Ala Glu Val Trp Lys Thr Phe
1               5                   10                  15
```

```
Pro Pro Phe Asp Thr Pro Val Pro Pro Ser Pro Phe Glu Asn Glu Phe
            20                  25                  30

Leu Asn Lys Ile Tyr Tyr Ala Ser Met Asn Val Thr Thr Pro Leu Thr
        35                  40                  45

Ile Ala Ile Ile Tyr Phe Thr Ser Val His Phe Ile Asn Ser Ile Ile
    50                  55                  60

Arg Asn Lys Gln Ile Ala Lys Tyr Asn Ser Lys Thr Ser Ser Glu Ser
65                  70                  75                  80

Lys Val Asp Ile Thr Lys Leu Asp Asp Lys Lys Leu Gln Lys Ile Leu
                85                  90                  95

Pro Thr Val Pro Asn Ser Ile Ala Lys Thr Ser Ile Phe Lys Leu Phe
            100                 105                 110

Val Phe Leu His Asn Val Phe Leu Cys Leu Tyr Ser Ile Trp Thr Phe
        115                 120                 125

Leu Gly Met Ser His Thr Ile Ala Thr Thr Met Asn Leu Phe Glu Gly
    130                 135                 140

Asn Phe Leu Gln Ser Leu Val Asn Tyr Gln Pro Lys Lys Leu Asp Ile
145                 150                 155                 160

Phe Leu His Ser Val Cys Asp Pro Lys Ile Gly Ile Phe Ser Arg Leu
                165                 170                 175

Leu Thr Asn Glu Lys Gly Leu His Asn Leu Glu Val Phe Gly Trp Trp
            180                 185                 190

Phe Tyr Ile Ser Lys Phe Tyr Glu Val Leu Asp Thr Ala Ile Ile Leu
        195                 200                 205

Leu Lys Gly Arg Pro Ser Ser Leu Leu Gln Ser Tyr His His Ala Gly
    210                 215                 220
```

```
Ala Met Met Cys Met Trp Ala Gly Ile Arg Tyr Gln Ser Pro Pro Ile
225                 230                 235                 240

Trp Ile Phe Val Val Phe Asn Ser Phe Ile His Ser Leu Met Tyr Phe
                245                 250                 255

Tyr Phe Ser Leu Ser Cys Leu Lys Ile Arg Val Pro Asn Phe Phe Lys
            260                 265                 270

Arg Ile Leu Thr Thr Met Gln Ile Thr Gln Phe Ile Val Gly Gly Ser
        275                 280                 285

Ile Ala Ile Leu His Ser Phe Val Trp Ile Val Asp Thr Ser His Val
    290                 295                 300

Ile Ser Pro Asp Asn Leu Lys Trp Val Ser Cys Ile Ser Thr Pro Asp
305                 310                 315                 320

Gln Ala Leu Pro Ile Leu Ile Asn Val Leu Tyr Leu Leu Pro Leu Thr
                325                 330                 335

Ala Leu Phe Thr Ala Phe Tyr Ile Glu Ser Tyr Leu Lys Lys Lys Ser
            340                 345                 350

Ala
```

<210> 97
<211> 978
<212> DNA
<213> Yarrowia lipolytica

<220>
<221> CDS
<222> (1)..(978)
<223> C14/16 elongase

<400> 97

```
atg agc gcc gtc cct att gaa ttc aac gtc ccc tcc gtg gac cga ccc       48
Met Ser Ala Val Pro Ile Glu Phe Asn Val Pro Ser Val Asp Arg Pro
1               5                   10                  15
```

```
ttt ggt atc tac ctc tgg gcc atc ttt gac cag gcc tgg gag aag ctt       96
Phe Gly Ile Tyr Leu Trp Ala Ile Phe Asp Gln Ala Trp Glu Lys Leu
            20                  25                  30

ttc ggc tgg ccc gcg tcc tct ttc att ttc gtg cga aat gac ccc aac      144
Phe Gly Trp Pro Ala Ser Ser Phe Ile Phe Val Arg Asn Asp Pro Asn
        35                  40                  45

atc ccc ttt tcc tct acc cct ccc gtg atc att gcc atc att gtg tac      192
Ile Pro Phe Ser Ser Thr Pro Pro Val Ile Ile Ala Ile Ile Val Tyr
    50                  55                  60

tac att gtc atc ttt ggc ggc cga gag gtg atg cga aac ctg tct ccc      240
Tyr Ile Val Ile Phe Gly Gly Arg Glu Val Met Arg Asn Leu Ser Pro
65                  70                  75                  80

atc cga ctc aac tgg ctc ttc cag atc cac aac atc ttc ctc acc ctt      288
Ile Arg Leu Asn Trp Leu Phe Gln Ile His Asn Ile Phe Leu Thr Leu
                85                  90                  95

ctg tcc ggt atg ctc ctc ctc ctc ctc gtt gag cag ctc ttc ccc atc      336
Leu Ser Gly Met Leu Leu Leu Leu Leu Val Glu Gln Leu Phe Pro Ile
            100                 105                 110

att gtc cga cag ggt atc ctc tac gcc atc tgc gac tac gga tct tgg      384
Ile Val Arg Gln Gly Ile Leu Tyr Ala Ile Cys Asp Tyr Gly Ser Trp
        115                 120                 125

act cag ccc att gtc ttc tgc tac tac ctc aac tac ctg acc aag tac      432
Thr Gln Pro Ile Val Phe Cys Tyr Tyr Leu Asn Tyr Leu Thr Lys Tyr
    130                 135                 140

ttt gag ctg atc gac acc gtt ttc ctt gtg ctg cga aag aag aag ctg      480
Phe Glu Leu Ile Asp Thr Val Phe Leu Val Leu Arg Lys Lys Lys Leu
145                 150                 155                 160

act ttc ctc cac acc tac cac cat ggt gcc act gct ctt ctg tgc tac      528
Thr Phe Leu His Thr Tyr His His Gly Ala Thr Ala Leu Leu Cys Tyr
                165                 170                 175

acc cag ctc att ggt aag acc tcg gtc tct tgg gtc ccc atc acc ctt      576
Thr Gln Leu Ile Gly Lys Thr Ser Val Ser Trp Val Pro Ile Thr Leu
            180                 185                 190

aac ctg ttt gtc cac gtt gtc atg tac ttc tac tac ttc ctg gct gcg      624
Asn Leu Phe Val His Val Val Met Tyr Phe Tyr Tyr Phe Leu Ala Ala
        195                 200                 205

cga ggt atc cga gtg tgg tgg aag gag tgg gtc acc cgg ctc cag atc      672
Arg Gly Ile Arg Val Trp Trp Lys Glu Trp Val Thr Arg Leu Gln Ile
    210                 215                 220
```

```
atc cag ttc gtt atc gat ctt gga ttt gtc tac ttt gcc tct tac acc      720
Ile Gln Phe Val Ile Asp Leu Gly Phe Val Tyr Phe Ala Ser Tyr Thr
225                 230                 235                 240

tac ttc acc tct acc tac tgg ccc tgg atg ccc aac atg ggc tct tgt      768
Tyr Phe Thr Ser Thr Tyr Trp Pro Trp Met Pro Asn Met Gly Ser Cys
                245                 250                 255

gcc ggc gag gag ttt gct gct att tac ggc tgt ggt ctg ctg acc tct      816
Ala Gly Glu Glu Phe Ala Ala Ile Tyr Gly Cys Gly Leu Leu Thr Ser
            260                 265                 270

tac ctc ttc ctc ttc atc gcc ttc tac atc aac tct tac cga aag ccc      864
Tyr Leu Phe Leu Phe Ile Ala Phe Tyr Ile Asn Ser Tyr Arg Lys Pro
        275                 280                 285

tct tcc aag gga cct tcc aag cct gtt gtt gct gtc gat ggc cct gtt      912
Ser Ser Lys Gly Pro Ser Lys Pro Val Val Ala Val Asp Gly Pro Val
    290                 295                 300

ggc ggc gtc aac gcc cag act ggt gct tct cga ggc cag acc act acc      960
Gly Gly Val Asn Ala Gln Thr Gly Ala Ser Arg Gly Gln Thr Thr Thr
305                 310                 315                 320

cga tct cga cga gca taa                                              978
Arg Ser Arg Arg Ala
                325
```

<210> 98
<211> 325
<212> PRT
<213> Yarrowia lipolytica

<400> 98

```
Met Ser Ala Val Pro Ile Glu Phe Asn Val Pro Ser Val Asp Arg Pro
1               5                  10                  15

Phe Gly Ile Tyr Leu Trp Ala Ile Phe Asp Gln Ala Trp Glu Lys Leu
            20                  25                  30

Phe Gly Trp Pro Ala Ser Ser Phe Ile Phe Val Arg Asn Asp Pro Asn
        35                  40                  45

Ile Pro Phe Ser Ser Thr Pro Pro Val Ile Ile Ala Ile Ile Val Tyr
    50                  55                  60
```

```
Tyr Ile Val Ile Phe Gly Gly Arg Glu Val Met Arg Asn Leu Ser Pro
65                  70                  75                  80

Ile Arg Leu Asn Trp Leu Phe Gln Ile His Asn Ile Phe Leu Thr Leu
                85                  90                  95

Leu Ser Gly Met Leu Leu Leu Leu Leu Val Glu Gln Leu Phe Pro Ile
            100                 105                 110

Ile Val Arg Gln Gly Ile Leu Tyr Ala Ile Cys Asp Tyr Gly Ser Trp
        115                 120                 125

Thr Gln Pro Ile Val Phe Cys Tyr Tyr Leu Asn Tyr Leu Thr Lys Tyr
    130                 135                 140

Phe Glu Leu Ile Asp Thr Val Phe Leu Val Leu Arg Lys Lys Lys Leu
145                 150                 155                 160

Thr Phe Leu His Thr Tyr His His Gly Ala Thr Ala Leu Leu Cys Tyr
                165                 170                 175

Thr Gln Leu Ile Gly Lys Thr Ser Val Ser Trp Val Pro Ile Thr Leu
            180                 185                 190

Asn Leu Phe Val His Val Val Met Tyr Phe Tyr Tyr Phe Leu Ala Ala
        195                 200                 205

Arg Gly Ile Arg Val Trp Trp Lys Glu Trp Val Thr Arg Leu Gln Ile
    210                 215                 220

Ile Gln Phe Val Ile Asp Leu Gly Phe Val Tyr Phe Ala Ser Tyr Thr
225                 230                 235                 240

Tyr Phe Thr Ser Thr Tyr Trp Pro Trp Met Pro Asn Met Gly Ser Cys
                245                 250                 255

Ala Gly Glu Glu Phe Ala Ala Ile Tyr Gly Cys Gly Leu Leu Thr Ser
            260                 265                 270
```

```
Tyr Leu Phe Leu Phe Ile Ala Phe Tyr Ile Asn Ser Tyr Arg Lys Pro
        275                 280                 285

Ser Ser Lys Gly Pro Ser Lys Pro Val Val Ala Val Asp Gly Pro Val
    290                 295                 300

Gly Gly Val Asn Ala Gln Thr Gly Ala Ser Arg Gly Gln Thr Thr Thr
305                 310                 315                 320

Arg Ser Arg Arg Ala
                325
```

<210> 99
<211> 337
<212> PRT
<213> Neurospora crassa (GenBank Accession #CAD70918)

<400> 99

```
Met Ala Thr Ala Ser Pro Val Ser Phe Leu Asp Ser Leu Pro Arg Pro
1               5                   10                  15

Thr Leu Asp Arg Pro Phe Gly Ile His Leu Trp Pro Ile Phe Ser Lys
            20                  25                  30

Ala Phe Glu Ala Val Val Gly Tyr Pro Ala Glu Asp Phe Arg Phe Val
        35                  40                  45

Pro Gly Glu Thr Pro Met Ser Thr Leu Lys Glu Thr Ser Ile Phe Ile
    50                  55                  60

Val Ile Tyr Tyr Thr Ile Ile Phe Gly Gly Arg Glu Leu Met Arg Asn
65                  70                  75                  80

Arg Glu Pro Phe Lys Leu Arg Thr Leu Phe Leu Ile His Asn Phe Tyr
                85                  90                  95

Leu Thr Ala Ile Ser Ala Ile Leu Leu Ala Leu Phe Ile Glu Gln Leu
            100                 105                 110

Leu Pro Ile Val Val Arg Gly Gly Ile Phe His Ala Ile Cys His Ile
        115                 120                 125
```

```
Asp Gly Gly Trp Thr Gln Pro Leu Val Val Leu Tyr Tyr Leu Asn Tyr
    130                 135                 140

Leu Thr Lys Tyr Leu Glu Leu Leu Asp Thr Cys Phe Leu Phe Leu Lys
145                 150                 155                 160

Lys Lys Pro Leu Thr Phe Leu His Cys Tyr His His Gly Ala Thr Ala
                165                 170                 175

Leu Leu Cys Tyr Thr Gln Leu Ile Gly Ser Thr Ser Val Ser Trp Val
            180                 185                 190

Pro Ile Thr Leu Asn Leu Thr Val His Val Val Met Tyr Trp Tyr Tyr
        195                 200                 205

Phe Gln Ser Ala Arg Gly Ile Arg Ile Trp Trp Lys Glu Trp Ile Thr
    210                 215                 220

Arg Leu Gln Ile Ile Gln Phe Ile Ile Asp Leu Gly Phe Val Tyr Phe
225                 230                 235                 240

Ala Ser Tyr Thr Tyr Phe Thr Ser Thr Tyr Phe Pro Trp Met Pro Asn
                245                 250                 255

Ala Gly Lys Cys Ala Gly Glu Glu Phe Ala Ala Phe Ser Gly Ile Gly
            260                 265                 270

Ile Leu Ser Ser Tyr Leu Val Leu Phe Ile Ser Phe Tyr Phe Ala Thr
        275                 280                 285

Tyr Lys Lys Asp Gly Lys Arg Pro Thr Gly Arg Lys Ser Leu Arg Arg
    290                 295                 300

Met Ser Gln Ala Pro Leu Pro Asp Pro Lys Asp Ile Val Ser Ala Lys
305                 310                 315                 320

Ala Ser Gly Ala Arg Pro Ser Gly Ala Ala Thr Ala Arg Ser Arg Lys
                325                 330                 335
```

Ala

<210> 100
<211> 903
<212> DNA
<213> Ostreococcus tauri (GenBank Accession No. AY591336)

<400> 100
atgagcgcct ccggtgcgct gctgcccgcg atcgcgtccg ccgcgtacgc gtacgcgacg 60

tacgcctacg cctttgagtg gtcgcacgcg aatggcatcg acaacgtcga cgcgcgcgag 120

tggatcggtg cgctgtcgtt gaggctcccg gcgatcgcga cgacgatgta cctgttgttc 180

tgcctggtcg gaccgaggtt gatggcgaag cgcgaggcgt tcgacccgaa ggggttcatg 240

ctggcgtaca atgcgtatca gacggcgttc aacgtcgtcg tgctcgggat gttcgcgcga 300

gagatctcgg ggctggggca gcccgtgtgg gggtcaacca tgccgtggag cgatagaaaa 360

tcgtttaaga tcctcctcgg ggtgtggttg cactacaaca acaaatattt ggagctattg 420

gacactgtgt tcatggttgc gcgcaagaag acgaagcagt tgagcttctt gcacgtttat 480

catcacgccc tgttgatctg ggcgtggtgg ttggtgtgtc acttgatggc cacgaacgat 540

tgtatcgatg cctacttcgg cgcggcgtgc aactcgttca ttcacatcgt gatgtactcg 600

tattatctca tgtcggcgct cggcattcga tgcccgtgga agcgatacat cacccaggct 660

caaatgctcc aattcgtcat tgtcttcgcg cacgccgtgt tcgtgctgcg tcagaagcac 720

tgcccggtca cccttccttg ggcgcaaatg ttcgtcatga cgaacatgct cgtgctcttc 780

gggaacttct acctcaaggc gtactcgaac aagtcgcgcg gcgacggcgc gagttccgtg 840

aaaccagccg agaccacgcg cgcgcccagc gtgcgacgca cgcgatctcg aaaaattgac 900

taa 903

<210> 101
<211> 300
<212> PRT
<213> Ostreococcus tauri (GenBank Accession No. AY591336)

<400> 101

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> 102
<211> 903
<212> DNA
<213> Ostreococcus tauri

<220>
<221> misc_feature
<223> synthetic C20/22 fatty acid elongase (codon-optimized)

<400> 102

```
atggctgcct ctggtgctct gcttcctgcc attgcctccg ctgcctacgc ctacgctacc     60

tacgcctacg ccttcgagtg gtcgcacgcc aacggcatcg acaacgtgga cgcacgagag    120

tggatcggag ccctgtcctt gcgacttccc gctattgcca ctaccatgta cctgctcttc    180

tgcctggtcg gtcctcgact catggccaag cgagaggcct ttgatcccaa gggcttcatg    240

ctggcataca acgcctacca gactgctttc aacgttgtcg tgctcggaat gtttgctcga    300

gaaatctctg gactgggaca gcccgtctgg ggttccacca tgccttggag cgatcgaaag    360

tccttcaaga ttctgctcgg tgtgtggctg cactacaaca acaagtatct ggagctgctt    420

gacaccgtgt tcatggttgc tcggaagaag accaagcagc tgtcgtttct ccacgtttac    480

catcacgcac tgctcatttg ggcctggtgg ctggtgtgtc atctcatggc taccaacgat    540

tgcatcgacg cctactttgg tgctgcctgc aattccttca ttcacatcgt catgtactcg    600
```

tactatctca tgtctgctct tggcattcga tgtccctgga agcgatacat cactcaggca 660

cagatgctgc aattcgtcat tgtcttcgct catgccgtgt ttgtcttgag acagaagcac 720

tgtcccgtta cacttccttg ggctcagatg ttcgtcatga ccaacatgct cgtgctgttt 780

ggcaacttct acctcaaagc ctactccaac aagtctcgag gcgacggtgc cagctccgtc 840

aaacctgccg agactacccg agcaccctcc gttcgaagaa cacgatctcg aaagatcgac 900

taa 903

<210> 103
<211> 300
<212> PRT
<213> Ostreococcus tauri

<220>
<221> MISC_FEATURE
<223> synthetic C20/22 fatty acid elongase (codon-optimized)

<400> 103

Met Ala 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> 104
<211> 1548
<212> DNA

<213> Thraustochytrium aureum (GenBank Accession No. AAN75707)

<400> 104

atgacggtcg ggtttgacga aacggtgact atggacacgg tccgcaacca caacatgccg 60

gacgacgcct ggtgcgcgat ccacggcacc gtgtacgaca tcaccaagtt cagcaaggtg 120

caccccggcg gggacatcat catgctggcc gctggcaagg aggccaccat cctgttcgag 180

acctaccaca tcaagggcgt cccggacgcg gtgctgcgca agtacaaggt cggcaagctc 240

ccccagggca agaagggcga aacgagccac atgcccaccg ggctcgactc ggcctcctac 300

tactcgtggg acagcgagtt ttacagggtg ctccgcgagc gcgtcgccaa gaagctggcc 360

gagcccggcc tcatgcagcg cgcgcgcatg gagctctggg ccaaggcgat cttcctcctg 420

gcaggtttct ggggctccct ttacgccatg tgcgtgctag acccgcacgg cggtgccatg 480

gtagccgccg ttacgctcgg cgtgttcgct gcctttgtcg gaacttgcat ccagcacgac 540

ggcagccacg gcgccttctc caagtcgcga ttcatgaaca aggcggcggg ctggaccctc 600

gacatgatcg gcgcgagcgc gatgacctgg gagatgcagc acgttcttgg tcaccacccg 660

tacaccaacc tcatcgagat ggagaacggt ttggccaagg tcaagggcgc cgacgtcgac 720

ccgaagaagg tcgaccagga gagcgacccg gacgtcttca gtacgtaccc gatgcttcgc 780

ctgcacccgt ggcaccgcca gcggttttac cacaagttcc agcacctgta cgccccgctt 840

atctttgggt ttatgacgat taacaaggtg atttcccagg atgtcggggt tgtgctgcgc 900

aagcgcctgt tccagatcga cgccaactgc cggtatggca gcccctggaa cgtggcccgc 960

ttctggatca tgaagctcct caccacgctc tacatggtgg cgcttcccat gtacatgcag 1020

gggcctgctc agggcttgaa gcttttcttc atggcccact tcacctgcgg agaggtcctc 1080

gccaccatgt ttattgtcaa ccacatcatc gagggcgtca gctacgcttc caaggacgcg 1140

gtcaagggcg tcatggctcc gccgcgcact gtgcacggtg tcaccccgat gcaggtgacg 1200

caaaaggcgc tcagtgcggc cgagtcgacc aagtcggacg ccgacaagac gaccatgatc 1260

cccctcaacg actgggccgc tgtgcagtgc cagacctctg tgaactgggc tgtcgggtcg 1320

tggttttgga accacttttc gggcggccta aaccaccaga ttgagcacca ctgcttcccc 1380

caaaaccccc acacggtcaa cgtctacatc tcgggcatcg tcaaggagac ctgcgaagaa 1440

tacggcgtgc cgtaccaggc tgagatcagc ctcttctctg cctatttcaa gatgctgtcg 1500

cacctccgca cgctcggcaa cgaggacctc acggcctggt ccacgtga 1548

<210> 105
<211> 515
<212> PRT
<213> Thraustochytrium aureum (GenBank Accession No. AAN75707)

<400> 105

Met Thr Val Gly Phe Asp Glu Thr Val Thr Met Asp Thr Val Arg Asn
1 5 10 15

His Asn Met Pro Asp Asp Ala Trp Cys Ala Ile His Gly Thr Val Tyr
20 25 30

Asp Ile Thr Lys Phe Ser Lys Val His Pro Gly Gly Asp Ile Ile Met
35 40 45

Leu Ala Ala Gly Lys Glu Ala Thr Ile Leu Phe Glu Thr Tyr His Ile
50 55 60

Lys Gly Val Pro Asp Ala Val Leu Arg Lys Tyr Lys Val Gly Lys Leu
65 70 75 80

Pro Gln Gly Lys Lys Gly Glu Thr Ser His Met Pro Thr Gly Leu Asp
85 90 95

Ser Ala Ser Tyr Tyr Ser Trp Asp Ser Glu Phe Tyr Arg Val Leu Arg
100 105 110

Glu Arg Val Ala Lys Lys Leu Ala Glu Pro Gly Leu Met Gln Arg Ala
115 120 125

Arg Met Glu Leu Trp Ala Lys Ala Ile Phe Leu Leu Ala Gly Phe Trp
130 135 140

Gly Ser Leu Tyr Ala Met Cys Val Leu Asp Pro His Gly Gly Ala Met
145 150 155 160

Val Ala Ala Val Thr Leu Gly Val Phe Ala Ala Phe Val Gly Thr Cys
165 170 175

```
Ile Gln His Asp Gly Ser His Gly Ala Phe Ser Lys Ser Arg Phe Met
            180                 185                 190

Asn Lys Ala Ala Gly Trp Thr Leu Asp Met Ile Gly Ala Ser Ala Met
        195                 200                 205

Thr Trp Glu Met Gln His Val Leu Gly His His Pro Tyr Thr Asn Leu
    210                 215                 220

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

Pro Lys Lys Val Asp Gln Glu Ser Asp Pro Asp Val Phe Ser Thr Tyr
                245                 250                 255

Pro Met Leu Arg Leu His Pro Trp His Arg Gln Arg Phe Tyr His Lys
            260                 265                 270

Phe Gln His Leu Tyr Ala Pro Leu Ile Phe Gly Phe Met Thr Ile Asn
        275                 280                 285

Lys Val Ile Ser Gln Asp Val Gly Val Val Leu Arg Lys Arg Leu Phe
    290                 295                 300

Gln Ile Asp Ala Asn Cys Arg Tyr Gly Ser Pro Trp Asn Val Ala Arg
305                 310                 315                 320

Phe Trp Ile Met Lys Leu Leu Thr Thr Leu Tyr Met Val Ala Leu Pro
                325                 330                 335

Met Tyr Met Gln Gly Pro Ala Gln Gly Leu Lys Leu Phe Phe Met Ala
            340                 345                 350

His Phe Thr Cys Gly Glu Val Leu Ala Thr Met Phe Ile Val Asn His
        355                 360                 365

Ile Ile Glu Gly Val Ser Tyr Ala Ser Lys Asp Ala Val Lys Gly Val
    370                 375                 380
```

```
Met Ala Pro Pro Arg Thr Val His Gly Val Thr Pro Met Gln Val Thr
385                 390                 395                 400

Gln Lys Ala Leu Ser Ala Ala Glu Ser Thr Lys Ser Asp Ala Asp Lys
                405                 410                 415

Thr Thr Met Ile Pro Leu Asn Asp Trp Ala Ala Val Gln Cys Gln Thr
            420                 425                 430

Ser Val Asn Trp Ala Val Gly Ser Trp Phe Trp Asn His Phe Ser Gly
        435                 440                 445

Gly Leu Asn His Gln Ile Glu His His Cys Phe Pro Gln Asn Pro His
    450                 455                 460

Thr Val Asn Val Tyr Ile Ser Gly Ile Val Lys Glu Thr Cys Glu Glu
465                 470                 475                 480

Tyr Gly Val Pro Tyr Gln Ala Glu Ile Ser Leu Phe Ser Ala Tyr Phe
                485                 490                 495

Lys Met Leu Ser His Leu Arg Thr Leu Gly Asn Glu Asp Leu Thr Ala
            500                 505                 510

Trp Ser Thr
        515
```

<210> 106
<211> 1545
<212> DNA
<213> Thraustochytrium aureum

<220>
<221> misc_feature
<223> synthetic delta-4 desaturase (codon-optimized)

<400> 106
atggtcggct tcgacgagac cgtcactatg gacaccgtcc gaaaccacaa catgcccgac 60

gatgcctggt gcgctatcca cggcaccgtg tacgacatca ccaagttctc caaggtgcat 120

cctggtggag acattatcat gctggccgct ggcaaggagg ccaccatcct gttcgagacc 180

taccacatca agggtgttcc cgacgctgtc ctgcgaaagt acaaggtcgg caagctccct 240

cagggcaaga agggtgagac ctcccacatg cccaccggac tcgactctgc ctcctactac 300

tcctgggact ctgagttcta ccgagtcctc cgagagcgag tcgccaagaa gctggccgag 360

cccggtctca tgcagcgagc tcgaatggag ctctgggcca aggctatctt tctcctggct 420

ggtttctggg gatccctgta cgccatgtgc gtgctcgacc ctcacggagg tgctatggtc 480

gctgccgtca ccctcggtgt gtttgctgcc ttcgtcggaa cctgcatcca gcacgacggc 540

tcccacggag ccttctccaa gtctcgattc atgaacaagg ctgccggttg gaccctcgac 600

atgattggtg cctccgctat gacctgggag atgcagcacg tcctcggtca ccatccctac 660

accaacctga tcgagatgga gaacggtctg gccaaggtca agggtgccga cgtggacccc 720

aagaaggtcg accaggagtc cgaccctgac gtcttctcta cctaccctat gctgcgactg 780

catccctggc accgacagcg attctaccac aagtttcagc acctgtacgc tcccctgatc 840

ttcggcttca tgaccatcaa caaggtgatc tcccaggacg tcggtgtcgt gctgcgaaag 900

cgactgtttc agatcgacgc caactgccga tatggctctc cctggaacgt ggctcgattc 960

tggatcatga agctcctgac cactctctac atggtggctc tgcccatgta catgcaggga 1020

cccgctcagg gtctgaagct gttctttatg gctcacttca cctgcggaga ggtcctcgcc 1080

accatgttca ttgttaacca cattatcgag ggtgtctcct acgcctccaa ggacgctgtc 1140

aagggcgtca tggctcctcc ccgaaccgtg cacggtgtca ctcccatgca ggtgacccag 1200

aaggctctct ctgctgccga gtccaccaag tctgacgctg acaagaccac tatgatccct 1260

ctcaacgact gggctgccgt ccagtgccag acctccgtca actgggctgt cggatcctgg 1320

ttctggaacc acttctctgg aggtctcaac caccagattg agcaccattg ctttcctcag 1380

aacccccaca ccgtcaacgt ctacatctcc ggcatcgtca aggagacctg cgaggagtac 1440

ggcgtgccct accaggctga gatctccctc ttctctgcct atttcaagat gctgtcccac 1500

ctccgaaccc tcggcaacga ggacctcacc gcctggtcca cctaa 1545

<210> 107
<211> 514
<212> PRT
<213> Thraustochytrium aureum

<220>
<221> MISC_FEATURE
<223> synthetic delta-4 desaturase (codon-optimized)

<400> 107

```
Met Val Gly Phe Asp Glu Thr Val Thr Met Asp Thr Val Arg Asn His
1               5                  10                  15

Asn Met Pro Asp Asp Ala Trp Cys Ala Ile His Gly Thr Val Tyr Asp
            20                  25                  30

Ile Thr Lys Phe Ser Lys Val His Pro Gly Gly Asp Ile Ile Met Leu
        35                  40                  45

Ala Ala Gly Lys Glu Ala Thr Ile Leu Phe Glu Thr Tyr His Ile Lys
    50                  55                  60

Gly Val Pro Asp Ala Val Leu Arg Lys Tyr Lys Val Gly Lys Leu Pro
65                  70                  75                  80

Gln Gly Lys Lys Gly Glu Thr Ser His Met Pro Thr Gly Leu Asp Ser
                85                  90                  95

Ala Ser Tyr Tyr Ser Trp Asp Ser Glu Phe Tyr Arg Val Leu Arg Glu
            100                 105                 110

Arg Val Ala Lys Lys Leu Ala Glu Pro Gly Leu Met Gln Arg Ala Arg
        115                 120                 125

Met Glu Leu Trp Ala Lys Ala Ile Phe Leu Leu Ala Gly Phe Trp Gly
    130                 135                 140

Ser Leu Tyr Ala Met Cys Val Leu Asp Pro His Gly Gly Ala Met Val
145                 150                 155                 160

Ala Ala Val Thr Leu Gly Val Phe Ala Ala Phe Val Gly Thr Cys Ile
                165                 170                 175

Gln His Asp Gly Ser His Gly Ala Phe Ser Lys Ser Arg Phe Met Asn
            180                 185                 190
```

```
Lys Ala Ala Gly Trp Thr Leu Asp Met Ile Gly Ala Ser Ala Met Thr
        195                 200                 205

Trp Glu Met Gln His Val Leu Gly His His Pro Tyr Thr Asn Leu Ile
    210                 215                 220

Glu Met Glu Asn Gly Leu Ala Lys Val Lys Gly Ala Asp Val Asp Pro
225                 230                 235                 240

Lys Lys Val Asp Gln Glu Ser Asp Pro Asp Val Phe Ser Thr Tyr Pro
                245                 250                 255

Met Leu Arg Leu His Pro Trp His Arg Gln Arg Phe Tyr His Lys Phe
            260                 265                 270

Gln His Leu Tyr Ala Pro Leu Ile Phe Gly Phe Met Thr Ile Asn Lys
        275                 280                 285

Val Ile Ser Gln Asp Val Gly Val Val Leu Arg Lys Arg Leu Phe Gln
    290                 295                 300

Ile Asp Ala Asn Cys Arg Tyr Gly Ser Pro Trp Asn Val Ala Arg Phe
305                 310                 315                 320

Trp Ile Met Lys Leu Leu Thr Thr Leu Tyr Met Val Ala Leu Pro Met
                325                 330                 335

Tyr Met Gln Gly Pro Ala Gln Gly Leu Lys Leu Phe Phe Met Ala His
            340                 345                 350

Phe Thr Cys Gly Glu Val Leu Ala Thr Met Phe Ile Val Asn His Ile
        355                 360                 365

Ile Glu Gly Val Ser Tyr Ala Ser Lys Asp Ala Val Lys Gly Val Met
    370                 375                 380

Ala Pro Pro Arg Thr Val His Gly Val Thr Pro Met Gln Val Thr Gln
385                 390                 395                 400
```

Lys Ala Leu Ser Ala Ala Glu Ser Thr Lys Ser Asp Ala Asp Lys Thr
405 410 415

Thr Met Ile Pro Leu Asn Asp Trp Ala Ala Val Gln Cys Gln Thr Ser
420 425 430

Val Asn Trp Ala Val Gly Ser Trp Phe Trp Asn His Phe Ser Gly Gly
435 440 445

Leu Asn His Gln Ile Glu His His Cys Phe Pro Gln Asn Pro His Thr
450 455 460

Val Asn Val Tyr Ile Ser Gly Ile Val Lys Glu Thr Cys Glu Glu Tyr
465 470 475 480

Gly Val Pro Tyr Gln Ala Glu Ile Ser Leu Phe Ser Ala Tyr Phe Lys
485 490 495

Met Leu Ser His Leu Arg Thr Leu Gly Asn Glu Asp Leu Thr Ala Trp
500 505 510

Ser Thr

<210> 108
<211> 945
<212> DNA
<213> Mortierella alpina

<400> 108
atgtccatag gttcttccaa tcctgtcctg ctggcagcga tccccttcgt ctacctcttc 60

gtcctccctc gtgtcctcgc cttcctccct caaaaggccc agttcctcgc aaaatgcatc 120

gtggtcttga tcgccaccct tatcatgtcc gtcgcaggct gcttcatttc catcgtctgt 180

gcgctcctcg ataaacgcta tgtgatcaac tacgtcgtct caagactctt ctcattcctc 240

gctgcaagac cctgcggtgt cacctacaag atcgtcggcg aggaacatct ggacaagtac 300

cccgccattg tcgtctgcaa ccaccagagc tccatggaca tgatggtcct gggacgcgtc 360

ttcccaaagc actgtgtcgt catggcaaag aaggaacttc tttactttcc gttcctgggc 420

atgtttatga agctgagtaa cgccatcttc attgaccgca agaaccacaa gaaggcgatc 480

gagtccacca cccaagctgt cgccgacatg aagaagcaca actctggaat ctggattttc 540

cccgaaggaa cacgttcccg cttggacaag gccgatctct tgcccttcaa gaagggagcc 600

ttccacctcg ccattcaagc ccaactcccg atcctcccca tcatctcgca aggatactca 660

cacatctacg attcgtcaaa acgctacttc cccggtggag agctcgagat cagagtcctg 720

gaacctatcc ccaccacggg attgaccaca gacgatgtga acgacctgat ggacaagact 780

cgcaacctga tgctgaagca cctcaaggag atggattctc aatactcctc ctccaccgct 840

gaaaacggat cgacccatat tgacgccgat atcgcaaagt caactgccac atcgatcgga 900

aacacggacg atgctatcac aaagaggagg acaccaaaag agtag 945

<210> 109
<211> 314
<212> PRT
<213> Mortierella alpina

<400> 109

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> 110
<211> 927
<212> DNA
<213> Mortierella alpina

<220>
<221> misc_feature
<223> LPAAT2

<400> 110
atgctcgggt ccgtcacccg acccacaaag gccctgctct atggatcagc cctcttcagt 60

ttctgctcat tgctcaatgt ggtccaggtg ttctccatac tcctgcagcc gttctcgaag 120

cgtctcttct ttgaagtgaa cgctcgcgtg gccggctcca tgtggaaggt tatgcagctg 180

attatggaga aaaagcacaa ggccgccatc aycttctcag gagacaagat ccctcaccac 240

gagagtgcca tcgtctttgg caaccaccgg tcctttgtcg acttttacat gtttcacacc 300

gttgctgcty ggagaggcat gctcaactat atgaagtact ttgccaagga ctctctgaaa 360

tacattccat tctatggatg gggcatgtgg aycatgggaa tgctattcat caatcgcaac 420

tggcagcagg atcagctcaa gatcaacaag atgtttgcac ggatattgga catccaagcg 480

cccgtttggg tcgccagttt cttggagggc tctcggttga cgcccagcaa actgkctgcc 540

tctcaaaagt tcatgctggg acgcggattg cctctgctgt caaacgtcat gatgcccagg 600

accaagggat tcattgcctg tgtcaacaaa ttccggggaa ctcatgtgaa atgtgtttat 660

gatttcacgt tcgcctacta ccacaagacc aagggctttg gagtgyctcc agatctggtc 720

cgtgttcaca stggccagct cagccccgag tacaaattcc atgttcatgt gagacgctat 780

cagctcgacg atctgcccac ggatgaggag aagctgagcg agtgggtggt ccaaaagtat 840

gtggagaagg acgccttttt ggagcagatg aaggagaatt ggacagatgg tattgatggg 900

ggtgtgtggt cagagaactg gatgtga 927

<210> 111
<211> 308
<212> PRT
<213> Mortierella alpina

<220>
<221> misc_feature
<222> (71)..(71)
<223> Xaa can be any naturally occurring amino acid

<220>
<221> misc_feature
<222> (104)..(104)
<223> Xaa can be any naturally occurring amino acid

<220>
<221> misc_feature
<222> (131)..(131)
<223> Xaa can be any naturally occurring amino acid

<220>
<221> misc_feature
<222> (179)..(179)
<223> Xaa can be any naturally occurring amino acid

<220>
<221> misc_feature
<222> (236)..(236)
<223> Xaa can be any naturally occurring amino acid

<220>
<221> misc_feature
<222> (244)..(244)
<223> Xaa can be any naturally occurring amino acid

<400> 111

```
Met Leu Gly Ser Val Thr Arg Pro Thr Lys Ala Leu Leu Tyr Gly Ser
1               5                  10                  15

Ala Leu Phe Ser Phe Cys Ser Leu Leu Asn Val Val Gln Val Phe Ser
            20                  25                  30

Ile Leu Leu Gln Pro Phe Ser Lys Arg Leu Phe Phe Glu Val Asn Ala
        35                  40                  45

Arg Val Ala Gly Ser Met Trp Lys Val Met Gln Leu Ile Met Glu Lys
    50                  55                  60

Lys His Lys Ala Ala Ile Xaa Phe Ser Gly Asp Lys Ile Pro His His
65                  70                  75                  80

Glu Ser Ala Ile Val Phe Gly Asn His Arg Ser Phe Val Asp Phe Tyr
                85                  90                  95
```

```
Met Phe His Thr Val Ala Ala Xaa Arg Gly Met Leu Asn Tyr Met Lys
            100                 105                 110

Tyr Phe Ala Lys Asp Ser Leu Lys Tyr Ile Pro Phe Tyr Gly Trp Gly
        115                 120                 125

Met Trp Xaa Met Gly Met Leu Phe Ile Asn Arg Asn Trp Gln Gln Asp
    130                 135                 140

Gln Leu Lys Ile Asn Lys Met Phe Ala Arg Ile Leu Asp Ile Gln Ala
145                 150                 155                 160

Pro Val Trp Val Ala Ser Phe Leu Glu Gly Ser Arg Leu Thr Pro Ser
                165                 170                 175

Lys Leu Xaa Ala Ser Gln Lys Phe Met Leu Gly Arg Gly Leu Pro Leu
            180                 185                 190

Leu Ser Asn Val Met Met Pro Arg Thr Lys Gly Phe Ile Ala Cys Val
        195                 200                 205

Asn Lys Phe Arg Gly Thr His Val Lys Cys Val Tyr Asp Phe Thr Phe
    210                 215                 220

Ala Tyr Tyr His Lys Thr Lys Gly Phe Gly Val Xaa Pro Asp Leu Val
225                 230                 235                 240

Arg Val His Xaa Gly Gln Leu Ser Pro Glu Tyr Lys Phe His Val His
                245                 250                 255

Val Arg Arg Tyr Gln Leu Asp Asp Leu Pro Thr Asp Glu Glu Lys Leu
            260                 265                 270

Ser Glu Trp Val Val Gln Lys Tyr Val Glu Lys Asp Ala Phe Leu Glu
        275                 280                 285

Gln Met Lys Glu Asn Trp Thr Asp Gly Ile Asp Gly Gly Val Trp Ser
    290                 295                 300
```

Glu Asn Trp Met
305

<210> 112
<211> 1549
<212> DNA
<213> Yarrowia lipolytica

<220>
<221> misc_feature
<223> LPAAT1

<220>
<221> CDS
<222> (501)..(1349)

<400> 112
cacagcataa taccacggca tgaccccgct gactccaacc ttcatttcgg cacatgtagg 60

tgcacaaggg acttcaagag gggccaattt catgcggaca catggcgcaa aaaacgcccg 120

actttgatta cacagacacg taataacgac gaagccgaga tgagcacacg tggccaagtc 180

tgccaatggc cccctggacc cccctgacaa agtttcccaa caagcccagc cgtgcatggt 240

gtgtttttgt gcggagacac acgccaatta ggctcatttg agggtatgca gcgaaaaaaa 300

attagtgtgg gtagtttgtt tgcaggaatc aagtgggtgg ttgaaaaaca agaaagagcg 360

acgacaagag agagagaaaa agagagagag actccataaa gcgtgcatca aaattaaggt 420

gtgtgactat ccgaaaacca aacatgaaca gttggatata tgtcgctgtg attgcagttg 480

ctgccgttct cattgcccga atg tcc gtt gca tcc aag ctc gtc ttc tac gtc 533
Met Ser Val Ala Ser Lys Leu Val Phe Tyr Val
1 5 10

cgc gcc gcc atc gcc gtg gtc atc ttt gcc gcc tgt gcc acc tac ggc 581
Arg Ala Ala Ile Ala Val Val Ile Phe Ala Ala Cys Ala Thr Tyr Gly
15 20 25

gtg ctg gcg tcc acc att ctc acc gcc atc ggc aag cag ggc ctg gcc 629
Val Leu Ala Ser Thr Ile Leu Thr Ala Ile Gly Lys Gln Gly Leu Ala
30 35 40

caa tgg acc gtt gcc aga gcc ttc tac tac tcg gtg cgc atc ttc ctg 677
Gln Trp Thr Val Ala Arg Ala Phe Tyr Tyr Ser Val Arg Ile Phe Leu
45 50 55

```
ggt atc agc atc aag ctg cgt agc cgg cag gtg acc gga acc gcc ggt      725
Gly Ile Ser Ile Lys Leu Arg Ser Arg Gln Val Thr Gly Thr Ala Gly
60                  65                  70                  75

ctg gat gcc tcc aag atc cag gtc gcc aac acc acc aag ccc att gac      773
Leu Asp Ala Ser Lys Ile Gln Val Ala Asn Thr Thr Lys Pro Ile Asp
                80                  85                  90

gac atc acc aaa cac ctg ccc cga cca tgc att ctg att tcc aac cac      821
Asp Ile Thr Lys His Leu Pro Arg Pro Cys Ile Leu Ile Ser Asn His
            95                  100                 105

cag aac gaa atg gac att ctg gtg ctc ggt cgc atc ttc ccc cag tac      869
Gln Asn Glu Met Asp Ile Leu Val Leu Gly Arg Ile Phe Pro Gln Tyr
        110                 115                 120

tgc tcc gtc acc gcc aaa aag gcc ctc aag tgg tac cct ctg ctg ggc      917
Cys Ser Val Thr Ala Lys Lys Ala Leu Lys Trp Tyr Pro Leu Leu Gly
    125                 130                 135

cag ttc atg gcg ctg tcc ggc acc atc ttc ctg gac cga aag gac cga      965
Gln Phe Met Ala Leu Ser Gly Thr Ile Phe Leu Asp Arg Lys Asp Arg
140                 145                 150                 155

acc aag tcc gtg cag acc ctc ggc ggc gcc gtc aag acc atc cag agc     1013
Thr Lys Ser Val Gln Thr Leu Gly Gly Ala Val Lys Thr Ile Gln Ser
                160                 165                 170

ggc aac gga ggc aag ggc cag agc gtc ttc atg ttc ccc gag gga acc     1061
Gly Asn Gly Gly Lys Gly Gln Ser Val Phe Met Phe Pro Glu Gly Thr
            175                 180                 185

cga tcc tac tcc aag gac gtc ggc atc atg ccc ttc aag aag ggc tgt     1109
Arg Ser Tyr Ser Lys Asp Val Gly Ile Met Pro Phe Lys Lys Gly Cys
        190                 195                 200

ttc cac ctg gcg gtc cag tcg ggc gct ccc att gtc ccc gtg gtg gtc     1157
Phe His Leu Ala Val Gln Ser Gly Ala Pro Ile Val Pro Val Val Val
    205                 210                 215

cag aac acc tcc cga atg ttt tct ttc ggc cga ggc aag ctg gac gcc     1205
Gln Asn Thr Ser Arg Met Phe Ser Phe Gly Arg Gly Lys Leu Asp Ala
220                 225                 230                 235

gga gag atc ctt gtc gac gtc ctg agc ccc att gag acc aag ggt ctg     1253
Gly Glu Ile Leu Val Asp Val Leu Ser Pro Ile Glu Thr Lys Gly Leu
                240                 245                 250

gac gcc agc aac gtc gac gct ctc atg gcc acc act tat aag gcc atg     1301
Asp Ala Ser Asn Val Asp Ala Leu Met Ala Thr Thr Tyr Lys Ala Met
            255                 260                 265
```

tgc gag act gcc gac cag att ggc tac gct ggc cag aag act cag tag 1349
Cys Glu Thr Ala Asp Gln Ile Gly Tyr Ala Gly Gln Lys Thr Gln
270 275 280

agactgcagc acaagaagtg cttgtagcta ctttaggaga gagataggta atatgaaaca 1409

tttttcagat cgacacccac ggcgaaccat tggctgtgga gctatgggtg aatggattaa 1469

tatagcaacg aaatctacct cgattaccaa cgcaaaacga gcccactttc tctgtactgt 1529

gctatatcgt gtatacccca 1549

<210> 113
<211> 282
<212> PRT
<213> Yarrowia lipolytica

<400> 113

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> 114
<211> 1495
<212> DNA
<213> Yarrowia lipolytica

<220>
<221> misc_feature
<223> LPAAT2

<400> 114

tagagtactc atagctactg tattccataa tcaaggttca ctgtgcttcg atactgcctg 60

gagataagac agatgcgata atacacgaca tgttggatgt aagggcatgt aatacgtgtt 120

gatattaatt gatctgacgt attatataaa cgatgtatat ccaggactca ttcaacacta 180

gacgcaaccc gatgctctcg tcacaatgat tggattcttg gtctcggtct caggttgcgt 240

cttctttcgg ctcctctcac cacacttctt taacacccct caacccacca ccattctatt 300

ccacatcacc tcacaccact tcaagagaga taccgacaca atataatttc atcgatgaaa 360

ctggaaacca gaaaatacct gctggtgacc atcctggtgg tcatgttcac tatcggctcg 420

gtcatggacc tatggctcta cctcctgtgt cttagccgat cacccaacct caagcactac 480

agagctatgt tggctcgaag atggtggtcg tggacgatgt tctcggtggg tcaggtggtc 540

aacttcaaga tcaagatgca cggcgacaaa atgatacccg agggaaacat gctgataatc 600

agtgagatac ctatgatgtg gtcgacgaga cacgaaatac taacacaggt aatcatctat 660

caagatacga ctatatgatc tacgcactaa tcaacaacca gcattgcgcc aagtcgcacc 720

ttgcattcct gtcctggaaa cacatctttt ctctctttag cctcagcaag attctgccct 780

tctaccgact tgacgaaaac tggatgttcg aatcagcagc cgaccttgag aaagaactag 840

ccaagctaat caacccttac tgtctggtgc tgttcccaga agtctcggtg gcaaccccac 900

aacttatcag acgacataga gagctctgca gagcgtgttt tgcacctgaa ctcactcatg 960

tactctaccc tcgccacagc agctttgcag acttcatatt gggactgaat aagggccagg 1020

ctcttagtta catctacgac gccaccatct cgtatactga caagaaagat aagattctat 1080

gtaatcctgg aaacatagat actctgttga ctcaagtgga aacggttcat gttcacatac 1140

accgagaaca gtatcgtcgg ctgccgcgcc acagacgtgg tattcagaag tggctggaaa 1200

acacatgggt ccacaaggat aagagtatca gaaaggccta caagcagtct ggatcaatca 1260

tggatgatgg caagctcaag gagaaggata agtgatacta actaactgca tgtaaaaaca 1320

gataaatatg ttaatgaatc aatgaaatgg tatacgatga ctacagtaaa ctctggtata 1380

ccttacttga atgtttttga taaatattat tgacaacact atattgacaa tgatgctatc 1440

tagttgatgt gagccttctt ctccagctca tcggcaagat cgtcaacggt cttgt 1495

<210> 115

<211> 672
<212> DNA
<213> Yarrowia lipolytica

<400> 115
atggtggtcg tggacgatgt tctcggtaat catctatcaa gatacgacta tatgatctac 60

gcactaatca acaaccagca ttgcgccaag tcgcaccttg cattcctgtc ctggaaacac 120

atcttttctc tctttagcct cagcaagatt ctgcccttct accgacttga cgaaaactgg 180

atgttcgaat cagcagccga ccttgagaaa gaactagcca agctaatcaa cccttactgt 240

ctggtgctgt tcccagaagt ctcggtggca accccacaac ttatcagacg acatagagag 300

ctctgcagag cgtgttttgc acctgaactc actcatgtac tctaccctcg ccacagcagc 360

tttgcagact tcatattggg actgaataag ggccaggctc ttagttacat ctacgacgcc 420

accatctcgt atactgacaa gaaagataag attctatgta atcctggaaa catagatact 480

ctgttgactc aagtggaaac ggttcatgtt cacatacacc gagaacagta tcgtcggctg 540

ccgcgccaca gacgtggtat tcagaagtgg ctggaaaaca catgggtcca caaggataag 600

agtatcagaa aggcctacaa gcagtctgga tcaatcatgg atgatggcaa gctcaaggag 660

aaggataagt ga 672

<210> 116
<211> 223
<212> PRT
<213> Yarrowia lipolytica

<400> 116

Met Val Val Val Asp Asp Val Leu Gly Asn His Leu Ser Arg Tyr Asp
1 5 10 15

Tyr Met Ile Tyr Ala Leu Ile Asn Asn Gln His Cys Ala Lys Ser His
20 25 30

Leu Ala Phe Leu Ser Trp Lys His Ile Phe Ser Leu Phe Ser Leu Ser
35 40 45

Lys Ile Leu Pro Phe Tyr Arg Leu Asp Glu Asn Trp Met Phe Glu Ser
50 55 60

```
Ala Ala Asp Leu Glu Lys Glu Leu Ala Lys Leu Ile Asn Pro Tyr Cys
65                  70                  75                  80

Leu Val Leu Phe Pro Glu Val Ser Val Ala Thr Pro Gln Leu Ile Arg
                85                  90                  95

Arg His Arg Glu Leu Cys Arg Ala Cys Phe Ala Pro Glu Leu Thr His
            100                 105                 110

Val Leu Tyr Pro Arg His Ser Ser Phe Ala Asp Phe Ile Leu Gly Leu
        115                 120                 125

Asn Lys Gly Gln Ala Leu Ser Tyr Ile Tyr Asp Ala Thr Ile Ser Tyr
    130                 135                 140

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

Leu Leu Thr Gln Val Glu Thr Val His Val His Ile His Arg Glu Gln
                165                 170                 175

Tyr Arg Arg Leu Pro Arg His Arg Arg Gly Ile Gln Lys Trp Leu Glu
            180                 185                 190

Asn Thr Trp Val His Lys Asp Lys Ser Ile Arg Lys Ala Tyr Lys Gln
        195                 200                 205

Ser Gly Ser Ile Met Asp Asp Gly Lys Leu Lys Glu Lys Asp Lys
    210                 215                 220
```

<210> 117
<211> 2326
<212> DNA
<213> Yarrowia lipolytica

<220>
<221> misc_feature
<223> PDAT

<220>

<221> misc_feature
<222> (2271)..(2271)
<223> n is a, c, g, or t

<400> 117
tattaatatt atgctcttca tgcaccagca aaataaccga aacgcgcata tgatagtggg 60

attctcgatt tgcccggcag acaaacgccg ctaaaatcgc cacagtatcg aattttaatt 120

gaatacgaac gtcaattccg gcttatcctt ctagcagttg tctcccgcag ctcgctccat 180

gactaatcat tcacgcgaca tgtctcagct accccggtct ggctcatgta aaaaaagtgt 240

aatcggcttt tttccggttg atcacaacca tcaatgacac aacctgtgaa tcggaaggcg 300

actgtcgagc gggtcgagcc agcagtggag gtggctgact ccgagtccga ggccaagacc 360

gacgtccacg ttcaccacca tcatcaccac cacaagcgaa aatccgtcaa gggcaagatt 420

ctcaacttct tcacccgaag tcgacgtatc accttcgtcc tcggcgccgt ggtcggtgtg 480

atagccgcgg gatactacgc tgcgccaccg gagctcagca ttgatatcga tgctcttctc 540

ggcgacttgc cctcgttcga ctttgacgct ctatctctcg acaacttgtc catggacagt 600

gtgtcggact ttgtacaaga catgaaatcg cggtttccga ccaagattct gcaggaggcg 660

gccaagatcg agaagcacca gaaaagcgaa cagaaggctg ccccttttgc tgtgggcaag 720

gctatgaaga gcgagggact caacgccaag tacccggtgg tgctggtgcc cggcgtcatc 780

tccacgggac tggagagctg gtccctggag ggaaccgagg agtgtcccac cgagtcgcac 840

ttcagaaagc gaatgtgggg ctcctggtac atgatccgag tcatgctgct ggacaagtac 900

tgctggctgc agaacctgat gctggacaca gagaccggtc tagaccctcc ccatttcaag 960

ctgcgagccg cccagggatt tgcctccgcc gacttcttta tggcaggcta ctggctgtgg 1020

aacaagctgc tcgagaacct ggctgttatt ggatacgata cggatacaat gtctgctgcg 1080

gcgtacgact ggagactgtc ctaccctgat ttggagcacc gagacggata cttctccaag 1140

ctcaaagctt caatcgaaga gactaagcgt atgacaggtg agaagacagt tctgacgggc 1200

cattccatgg gctcccaggt catcttctac ttcatgaagt gggctgaggc cgagggatat 1260

ggaggaggag gtcccaactg ggtcaatgac catattgaat cctttgtcga catttccggc 1320

tccatgctgg gtactcccaa gaccctggtt gctcttctgt ctggagaaat gaaggatacc 1380

gtgcagctga acgcgatggc tgtgtatgga ctggagcagt tcttctctcg acgagagcga 1440

gccgatctgc tgcgaacatg gggaggaatt gcttccatga ttcccaaggg tggtaaggct 1500

atctggggtg atcattctgg agcccctgat gacgagcccg gccagaatgt caccttggc 1560

aacttcatca agttcaagga gtccttgacc gagtactctg ctaagaacct caccatggat 1620

gaaaccgttg acttcctgta ttctcagtct cccgagtggt ttgtgaaccg aaccgagggt 1680

gcttactcct ttggaattgc caagactcga aagcaggttg agcagaatga gaagcgacct 1740

tctacctgga gcaaccctct ggaagctgct ctccccaatg cccccgatct caagatctac 1800

tgcttctatg gagtcggtaa ggataccgag cgagcctact actaccagga tgagcccaat 1860

cccgagcaga ccaacttgaa cgtcagtatc gctggaaacg accctgatgg tgtgcttatg 1920

ggtcagggcg atggaaccgt ctcccttgtg acccatacca tgtgtcaccg atggaaggac 1980

gagaattcca agttcaaccc tggtaacgcc caggtcaagg ttgtggagat gttgcaccag 2040

cctgatcgac ttgatattcg aggcggtgct cagactgccg agcatgtgga cattctgggg 2100

cgttctgagt tgaacgagat ggttctgaag gtggctagtg gaaagggaaa tgagattgaa 2160

gagagagtca tctccaacat tgatgagtgg gtgtggaaga ttgatctcgg cagcaattag 2220

agagtccgtt ttgtagagta atatgttttg tatatcacac tgatggagaa nggcgttcga 2280

tttctcatga ttccatgtgg ttgtttaatg agcacgtaga acgacg 2326

<210> 118
<211> 648
<212> PRT
<213> Yarrowia lipolytica

<400> 118

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> 119
<211> 1632
<212> DNA
<213> Yarrowia lipolytica

<220>
<221> misc_feature
<223> ARE2

<400> 119
atggccacac tccaccccga agacgccgca ggacggcccg tgcgacgacg acctcgtccc 60

tccagttcgg gcggctccag atcgccgtcc accaaacgac actcgatagt gcgggagcat 120

ctcggagaag agctcaatgt gcccgacggc caggaaatgg acctgggcca ggtcaacaag 180

aacctcaatg ccgcatacgc caaggccgag aaggactcgg acgacgagaa ggaaaagaag 240

gaggagggcg tggtggacga gctgccagag aagtattcct accctcgatt ctcaaagaac 300

aaccgacgct acagattcac cgacatcaag ttcaagccaa caccgtcgat tctcgacaag 360

ttcgcccaca aggactcgga gttctttggc ttctacaccc tgctgtggat ggtgtttgcc 420

ttctgcgtct tccgaaccgg cctgctcaac tacacaaacg aaggcatcct gttccggggc 480

cagattttcg ccattctcag caaagatctc tggaaagtcg cattggtcga tctgggcatg 540

tacctgacca cctatctgtc tgtgtttctg caattggccg tcaagcacgg tctggtcgac 600

tggaactcgt ttggctggat catccagaac gtgcaccaga ccctgttcct cttcttctac 660

ctttgggtcg ccaagtcgag taacctgcct tggatcggta acatcttcat tgtgcttcat 720

gcctttgtca tgctcatgaa acaacactcg tacgccttct acaatggcta cctatggact 780

gtcgaggacg agctctccca cgcaaagcag cgtctcaccg aagacattcc tgtttcagag 840

aaggaggatc tcaagctgga catcgagttc tgcgagacag agctcaaggt ccaatccaga 900

cacacccctt tccccaccaa catcaccttt tctaactact tctggtactc catgttccca 960

acgctcgtct acgaaattga gttccctcga acccccccgaa tcaagtggac atacgtgctg 1020

gagaaggtcg ccgcagtctt tggcgtcttc ttccttatga tctgggtcgc agagtcgtac 1080

ctgtatcccc ctgtggtggc tgttattcaa atgcgagacg aacccttctg gaacaaggtc 1140

cgaatctatc ccattttcct gtcggacatt ctgctgccct ttgtcattga gtacatgctt 1200

gttttctaca tcatctggga cgccattctc aacggcattg ccgagctcac tcgcttcgcc 1260

gacagagact tttatggccc ctggtggaac tgtaccagct gggagcagtt tagccgagaa 1320

tggaacattc ctgtctacca gttcctcaag cgacacgtct accactcgtc catctctgct 1380

ttcaagttct ccaagggcgc agctaccctc accaccttct tgctgtcttc tcttgtccac 1440

gagctggtca tgtttgccat ctttaagaag ttccgaggat acctgctgtt gctgcagatg 1500

acccagctgc ccctggccat gctgcagaaa accaaatgga tccaggacag acccgttttt 1560

ggcaacgctt tcttctggtt ctcgctcatg atcggacctt ctctcatgtg ttccatgtac 1620

ctcctcttct aa 1632

<210> 120
<211> 543
<212> PRT
<213> Yarrowia lipolytica

<400> 120

Met Ala Thr Leu His Pro Glu Asp Ala Ala Gly Arg Pro Val Arg Arg
1 5 10 15

Arg Pro Arg Pro Ser Ser Ser Gly Gly Ser Arg Ser Pro Ser Thr Lys
20 25 30

Arg His Ser Ile Val Arg Glu His Leu Gly Glu Glu Leu Asn Val Pro
35 40 45

Asp Gly Gln Glu Met Asp Leu Gly Gln Val Asn Lys Asn Leu Asn Ala
50 55 60

Ala Tyr Ala Lys Ala Glu Lys Asp Ser Asp Asp Glu Lys Glu Lys Lys
65 70 75 80

Glu Glu Gly Val Val Asp Glu Leu Pro Glu Lys Tyr Ser Tyr Pro Arg
85 90 95

Phe Ser Lys Asn Asn Arg Arg Tyr Arg Phe Thr Asp Ile Lys Phe Lys
100 105 110

Pro Thr Pro Ser Ile Leu Asp Lys Phe Ala His Lys Asp Ser Glu Phe
115 120 125

Phe Gly Phe Tyr Thr Leu Leu Trp Met Val Phe Ala Phe Cys Val Phe
130 135 140

Arg Thr Gly Leu Leu Asn Tyr Thr Asn Glu Gly Ile Leu Phe Arg Gly
145 150 155 160

```
Gln Ile Phe Ala Ile Leu Ser Lys Asp Leu Trp Lys Val Ala Leu Val
               165                 170                 175

Asp Leu Gly Met Tyr Leu Thr Thr Tyr Leu Ser Val Phe Leu Gln Leu
            180                 185                 190

Ala Val Lys His Gly Leu Val Asp Trp Asn Ser Phe Gly Trp Ile Ile
        195                 200                 205

Gln Asn Val His Gln Thr Leu Phe Leu Phe Phe Tyr Leu Trp Val Ala
    210                 215                 220

Lys Ser Ser Asn Leu Pro Trp Ile Gly Asn Ile Phe Ile Val Leu His
225                 230                 235                 240

Ala Phe Val Met Leu Met Lys Gln His Ser Tyr Ala Phe Tyr Asn Gly
               245                 250                 255

Tyr Leu Trp Thr Val Glu Asp Glu Leu Ser His Ala Lys Gln Arg Leu
            260                 265                 270

Thr Glu Asp Ile Pro Val Ser Glu Lys Glu Asp Leu Lys Leu Asp Ile
        275                 280                 285

Glu Phe Cys Glu Thr Glu Leu Lys Val Gln Ser Arg His Thr Pro Phe
    290                 295                 300

Pro Thr Asn Ile Thr Phe Ser Asn Tyr Phe Trp Tyr Ser Met Phe Pro
305                 310                 315                 320

Thr Leu Val Tyr Glu Ile Glu Phe Pro Arg Thr Pro Arg Ile Lys Trp
               325                 330                 335

Thr Tyr Val Leu Glu Lys Val Ala Ala Val Phe Gly Val Phe Phe Leu
            340                 345                 350

Met Ile Trp Val Ala Glu Ser Tyr Leu Tyr Pro Pro Val Val Ala Val
        355                 360                 365
```

```
Ile Gln Met Arg Asp Glu Pro Phe Trp Asn Lys Val Arg Ile Tyr Pro
    370                 375                 380

Ile Phe Leu Ser Asp Ile Leu Leu Pro Phe Val Ile Glu Tyr Met Leu
385                 390                 395                 400

Val Phe Tyr Ile Ile Trp Asp Ala Ile Leu Asn Gly Ile Ala Glu Leu
                405                 410                 415

Thr Arg Phe Ala Asp Arg Asp Phe Tyr Gly Pro Trp Trp Asn Cys Thr
            420                 425                 430

Ser Trp Glu Gln Phe Ser Arg Glu Trp Asn Ile Pro Val Tyr Gln Phe
        435                 440                 445

Leu Lys Arg His Val Tyr His Ser Ser Ile Ser Ala Phe Lys Phe Ser
    450                 455                 460

Lys Gly Ala Ala Thr Leu Thr Thr Phe Leu Leu Ser Ser Leu Val His
465                 470                 475                 480

Glu Leu Val Met Phe Ala Ile Phe Lys Lys Phe Arg Gly Tyr Leu Leu
                485                 490                 495

Leu Leu Gln Met Thr Gln Leu Pro Leu Ala Met Leu Gln Lys Thr Lys
            500                 505                 510

Trp Ile Gln Asp Arg Pro Val Phe Gly Asn Ala Phe Phe Trp Phe Ser
        515                 520                 525

Leu Met Ile Gly Pro Ser Leu Met Cys Ser Met Tyr Leu Leu Phe
    530                 535                 540
```

<210> 121
<211> 282
<212> PRT
<213> Caenorhabditis elegans (clone TO6E8.1)

<220>
<221> misc_feature
<223> LPCAT

<300>
<310> WO 2004/076617A2
<311> 2004-01-29
<312> 2004-09-10

<400> 121

Met Glu Asn Phe Trp Ser Ile Val Val Phe Phe Leu Leu Ser Ile Leu
1 5 10 15

Phe Ile Leu Tyr Asn Ile Ser Thr Val Cys His Tyr Tyr Met Arg Ile
20 25 30

Ser Phe Tyr Tyr Phe Thr Ile Leu Leu His Gly Met Glu Val Cys Val
35 40 45

Thr Met Ile Pro Ser Trp Leu Asn Gly Lys Gly Ala Asp Tyr Val Phe
50 55 60

His Ser Phe Phe Tyr Trp Cys Lys Trp Thr Gly Val His Thr Thr Val
65 70 75 80

Tyr Gly Tyr Glu Lys Thr Gln Val Glu Gly Pro Ala Val Val Ile Cys
85 90 95

Asn His Gln Ser Ser Leu Asp Ile Leu Ser Met Ala Ser Ile Trp Pro
100 105 110

Lys Asn Cys Val Val Met Met Lys Arg Ile Leu Ala Tyr Val Pro Phe
115 120 125

Phe Asn Leu Gly Ala Tyr Phe Ser Asn Thr Ile Phe Ile Asp Arg Tyr
130 135 140

Asn Arg Glu Arg Ala Met Ala Ser Val Asp Tyr Cys Ala Ser Glu Met
145 150 155 160

Lys Asn Arg Asn Leu Lys Leu Trp Val Phe Pro Glu Gly Thr Arg Asn
165 170 175

```
Arg Glu Gly Gly Phe Ile Pro Phe Lys Lys Gly Ala Phe Asn Ile Ala
            180                 185                 190

Val Arg Ala Gln Ile Pro Ile Ile Pro Val Val Phe Ser Asp Tyr Arg
        195                 200                 205

Asp Phe Tyr Ser Lys Pro Gly Arg Tyr Phe Lys Asn Asp Gly Glu Val
    210                 215                 220

Val Ile Arg Val Leu Asp Ala Ile Pro Thr Lys Gly Leu Thr Leu Asp
225                 230                 235                 240

Asp Val Ser Glu Leu Ser Asp Met Cys Arg Asp Val Met Leu Ala Ala
                245                 250                 255

Tyr Lys Glu Val Thr Leu Glu Ala Gln Gln Arg Asn Ala Thr Arg Arg
            260                 265                 270

Gly Glu Thr Lys Asp Gly Lys Lys Ser Glu
        275                 280
```

<210> 122
<211> 1578
<212> DNA
<213> Yarrowia lipolytica

<220>
<221> misc_feature
<223> DGAT1

<400> 122

```
atggaggtcc gacgacgaaa gatagacgtg ctcaaggccc agaaaaacgg ctacgaatcg     60

ggcccaccat ctcgacaatc gtcgcagccc tcctcaagag catcgtccag aacccgcaac    120

aaacactcct cgtccaccct gtcgctcagc ggactgacca tgaaagtcca gaagaaacct    180

gcgggacccc cggcgaactc caaaacgcca ttcctacaca tcaagcccgt gcacacgtgc    240

tgctccacat caatgctttc gcgcgattat gacggctcca accccagctt caagggcttc    300

aaaaacatcg gcatgatcat tctcattgtg ggaaatctac ggctcgcatt cgaaaactac    360
```

ctcaaatacg gcatttccaa cccgttcttc gaccccaaaa ttactccttc cgagtggcag 420

ctctcaggct tgctcatagt cgtggcctac gcacatatcc tcatggccta cgctattgag 480

agcgctgcca agctgctgtt cctctctagc aaacaccact acatggccgt ggggcttctg 540

cataccatga acactttgtc gtccatctcg ttgctgtcct acgtcgtcta ctactacctg 600

cccaaccccg tggcaggcac aatagtcgag tttgtggccg ttattctgtc tctcaaactc 660

gcctcatacg ccctcactaa ctcggatctc cgaaaagccg caattcatgc ccagaagctc 720

gacaagacgc aagacgataa cgaaaaggaa tccacctcgt cttcctcttc ttcagatgac 780

gcagagactt tggcagacat tgacgtcatt cctgcatact acgcacagct gccctacccc 840

cagaatgtga cgctgtcgaa cctgctgtac ttctggtttg ctcccacact ggtctaccag 900

cccgtgtacc ccaagacgga gcgtattcga cccaagcacg tgatccgaaa cctgtttgag 960

ctcgtctctc tgtgcatgct tattcagttt ctcatcttcc agtacgccta ccccatcatg 1020

cagtcgtgtc tggctctgtt cttccagccc aagctcgatt atgccaacat ctccgagcgc 1080

ctcatgaagt tggcctccgt gtctatgatg gtctggctca ttggattcta cgctttcttc 1140

cagaacggtc tcaatcttat tgccgagctc acctgttttg gaaacagaac cttctaccag 1200

cagtggtgga attcccgctc cattggccag tactggactc tatggaacaa gccagtcaac 1260

cagtacttta gacaccacgt ctacgtgcct cttctcgctc ggggcatgtc gcggttcaat 1320

gcgtcggtgg tggttttctt tttctccgcc gtcatccatg aactgcttgt cggcatcccc 1380

actcacaaca tcatcggagc cgccttcttc ggcatgatgt cgcaggtgcc tctgatcatg 1440

gctactgaga accttcagca tattaactcc tctctgggcc ccttccttgg caactgtgca 1500

ttctggttca ccttttttcct gggacaaccc acttgtgcat tcctttacta tttggcctac 1560

aactacaagc agaaccag 1578

<210> 123
<211> 526
<212> PRT
<213> Yarrowia lipolytica

<220>
<221> misc_feature
<222> (518)..(518)

<223> Xaa can be any naturally occurring amino acid

<400> 123

```
Met Glu Val Arg Arg Arg Lys Ile Asp Val Leu Lys Ala Gln Lys Asn
1               5                  10                  15

Gly Tyr Glu Ser Gly Pro Pro Ser Arg Gln Ser Ser Gln Pro Ser Ser
            20                  25                  30

Arg Ala Ser Ser Arg Thr Arg Asn Lys His Ser Ser Ser Thr Leu Ser
        35                  40                  45

Leu Ser Gly Leu Thr Met Lys Val Gln Lys Lys Pro Ala Gly Pro Pro
    50                  55                  60

Ala Asn Ser Lys Thr Pro Phe Leu His Ile Lys Pro Val His Thr Cys
65                  70                  75                  80

Cys Ser Thr Ser Met Leu Ser Arg Asp Tyr Asp Gly Ser Asn Pro Ser
                85                  90                  95

Phe Lys Gly Phe Lys Asn Ile Gly Met Ile Ile Leu Ile Val Gly Asn
            100                 105                 110

Leu Arg Leu Ala Phe Glu Asn Tyr Leu Lys Tyr Gly Ile Ser Asn Pro
        115                 120                 125

Phe Phe Asp Pro Lys Ile Thr Pro Ser Glu Trp Gln Leu Ser Gly Leu
    130                 135                 140

Leu Ile Val Val Ala Tyr Ala His Ile Leu Met Ala Tyr Ala Ile Glu
145                 150                 155                 160

Ser Ala Ala Lys Leu Leu Phe Leu Ser Ser Lys His His Tyr Met Ala
                165                 170                 175

Val Gly Leu Leu His Thr Met Asn Thr Leu Ser Ser Ile Ser Leu Leu
            180                 185                 190

Ser Tyr Val Val Tyr Tyr Tyr Leu Pro Asn Pro Val Ala Gly Thr Ile
        195                 200                 205
```

```
Val Glu Phe Val Ala Val Ile Leu Ser Leu Lys Leu Ala Ser Tyr Ala
    210                 215                 220

Leu Thr Asn Ser Asp Leu Arg Lys Ala Ala Ile His Ala Gln Lys Leu
225                 230                 235                 240

Asp Lys Thr Gln Asp Asp Asn Glu Lys Glu Ser Thr Ser Ser Ser Ser
                245                 250                 255

Ser Ser Asp Asp Ala Glu Thr Leu Ala Asp Ile Asp Val Ile Pro Ala
            260                 265                 270

Tyr Tyr Ala Gln Leu Pro Tyr Pro Gln Asn Val Thr Leu Ser Asn Leu
        275                 280                 285

Leu Tyr Phe Trp Phe Ala Pro Thr Leu Val Tyr Gln Pro Val Tyr Pro
    290                 295                 300

Lys Thr Glu Arg Ile Arg Pro Lys His Val Ile Arg Asn Leu Phe Glu
305                 310                 315                 320

Leu Val Ser Leu Cys Met Leu Ile Gln Phe Leu Ile Phe Gln Tyr Ala
                325                 330                 335

Tyr Pro Ile Met Gln Ser Cys Leu Ala Leu Phe Phe Gln Pro Lys Leu
            340                 345                 350

Asp Tyr Ala Asn Ile Ser Glu Arg Leu Met Lys Leu Ala Ser Val Ser
        355                 360                 365

Met Met Val Trp Leu Ile Gly Phe Tyr Ala Phe Phe Gln Asn Gly Leu
    370                 375                 380

Asn Leu Ile Ala Glu Leu Thr Cys Phe Gly Asn Arg Thr Phe Tyr Gln
385                 390                 395                 400

Gln Trp Trp Asn Ser Arg Ser Ile Gly Gln Tyr Trp Thr Leu Trp Asn
                405                 410                 415
```

Lys Pro Val Asn Gln Tyr Phe Arg His His Val Tyr Val Pro Leu Leu
420 425 430

Ala Arg Gly Met Ser Arg Phe Asn Ala Ser Val Val Val Phe Phe Phe
435 440 445

Ser Ala Val Ile His Glu Leu Leu Val Gly Ile Pro Thr His Asn Ile
450 455 460

Ile Gly Ala Ala Phe Phe Gly Met Met Ser Gln Val Pro Leu Ile Met
465 470 475 480

Ala Thr Glu Asn Leu Gln His Ile Asn Ser Ser Leu Gly Pro Phe Leu
485 490 495

Gly Asn Cys Ala Phe Trp Phe Thr Phe Phe Leu Gly Gln Pro Thr Cys
500 505 510

Ala Phe Leu Tyr Tyr Xaa Ala Tyr Asn Tyr Lys Gln Asn Gln
515 520 525

<210> 124
<211> 1578
<212> DNA
<213> Mortierella alpina

<220>
<221> misc_feature
<223> DGAT1

<220>
<221> misc_feature
<222> (466)..(466)
<223> n is a, c, g, or t

<400> 124
atgacagagt cgacaacaac gacatgtgca aaggaggagg gcattgccaa cagcgctgct 60

ttgcctgaca ttcccccaaa gatggaagac ctcaagtcct ccaggaagac cggctcttct 120

tacaagcaca ccttccccgt ccatacaaaa accatcccca gcccattgtc taaagaggca 180

cctccagaga gctatcgtgg attcgtcaac ctcggcatgc tcctactttt cggcaacaac 240

atccgattga tcatcgagaa ttacctcaaa tacggcttcc tgctctcaat ccctggatca 300

agcgtctcga agcaggactg gatcctggct gccctcaccc acgccatcct acccgtcaac 360

ctcatcctgg cctacaagct tgagagctgg gccaaggaga gagccgtcgg ctatcgcaag 420

cgtcgatctg acgaacccat tgcccaggaa tcaaccaagg ccgtgncagc aggagataat 480

gacgctatca aaaccacaaa acccgccaag gcccaggatc tcacacccga ggcccttgca 540

aggaaggaac aatcgaccgt gggctggctc catgtcttca atctgttcac catcgttgcc 600

tggccctcct tcatgtccta ctttatgatc taccaccccc tcgtggccat gtcctgcctc 660

atgaacggac ttatcctctt cctcaaaatg acctcctttg cgcttgtgaa ccaggagctc 720

cgagcagcct acatctttgg aacacccgtg gacacgttcc agcacatggc taaagtgcac 780

gacatctctg gcaaggacct gacaaagaag gagatcttcc agtatgacat ccagtacccc 840

gacaacatca ccctcaagaa cattggctat ttctggctcg cccccacgct ctgctaccag 900

ccatcatacc caaggacgac cgtcttccgc aaatccttct tcctcaagcg tgtggccgag 960

atcgtgacct gtctgggcat gatgtacttt ttagtcgagc agtacgccac ccccaccctg 1020

cagaactcgg tccgagcatt cgatgagttg gcgttcggca ccattctgga gagagtgctg 1080

aagctgagca ccaccagtgt catcatctgg ctactcatgt tctacacctt tttccactcg 1140

ttctttaatg ctcttgcaga ggcactgtac tttggagacc gtcgcttcta tctcgcctgg 1200

tggaatgcca ctggtgtcgg catgtactgg aagacgtgga actcgcccgt ctacaccttc 1260

ttcaaacgcc acgtatacct gcccctgatc acctctggca cctctcccat ggtcgcctcg 1320

atcgtcatct tcctcatctc ggctgtcttg cacgagatct tgatcggctt ccccactcat 1380

atgatctatg gatacgcatt cgccggcatg ttcctccaga tcccgctgat cattctgacc 1440

cgacccctcg aaaaatggcg aggcaccgga tcgggtctcg gcaacatgat cttctgggtc 1500

tcgttcacca tcctgggcca gccagcgtgt gcgctgctct actactacca ctggaccaag 1560

cgccatatgg atgtttga 1578

<210> 125
<211> 525
<212> PRT
<213> Mortierella alpina

<220>
<221> misc_feature
<222> (156)..(156)
<223> Xaa can be any naturally occurring amino acid

<400> 125

Met Thr Glu Ser Thr Thr Thr Thr Cys Ala Lys Glu Glu Gly Ile Ala
1 5 10 15

Asn Ser Ala Ala Leu Pro Asp Ile Pro Pro Lys Met Glu Asp Leu Lys
20 25 30

Ser Ser Arg Lys Thr Gly Ser Ser Tyr Lys His Thr Phe Pro Val His
35 40 45

Thr Lys Thr Ile Pro Ser Pro Leu Ser Lys Glu Ala Pro Pro Glu Ser
50 55 60

Tyr Arg Gly Phe Val Asn Leu Gly Met Leu Leu Leu Phe Gly Asn Asn
65 70 75 80

Ile Arg Leu Ile Ile Glu Asn Tyr Leu Lys Tyr Gly Phe Leu Leu Ser
85 90 95

Ile Pro Gly Ser Ser Val Ser Lys Gln Asp Trp Ile Leu Ala Ala Leu
100 105 110

Thr His Ala Ile Leu Pro Val Asn Leu Ile Leu Ala Tyr Lys Leu Glu
115 120 125

Ser Trp Ala Lys Glu Arg Ala Val Gly Tyr Arg Lys Arg Arg Ser Asp
130 135 140

Glu Pro Ile Ala Gln Glu Ser Thr Lys Ala Val Xaa Ala Gly Asp Asn
145 150 155 160

Asp Ala Ile Lys Thr Thr Lys Pro Ala Lys Ala Gln Asp Leu Thr Pro
165 170 175

```
Glu Ala Leu Ala Arg Lys Glu Gln Ser Thr Val Gly Trp Leu His Val
            180                 185                 190

Phe Asn Leu Phe Thr Ile Val Ala Trp Pro Ser Phe Met Ser Tyr Phe
        195                 200                 205

Met Ile Tyr His Pro Phe Val Ala Met Ser Cys Leu Met Asn Gly Leu
    210                 215                 220

Ile Leu Phe Leu Lys Met Thr Ser Phe Ala Leu Val Asn Gln Glu Leu
225                 230                 235                 240

Arg Ala Ala Tyr Ile Phe Gly Thr Pro Val Asp Thr Phe Gln His Met
                245                 250                 255

Ala Lys Val His Asp Ile Ser Gly Lys Asp Leu Thr Lys Lys Glu Ile
            260                 265                 270

Phe Gln Tyr Asp Ile Gln Tyr Pro Asp Asn Ile Thr Leu Lys Asn Ile
        275                 280                 285

Gly Tyr Phe Trp Leu Ala Pro Thr Leu Cys Tyr Gln Pro Ser Tyr Pro
    290                 295                 300

Arg Thr Thr Val Phe Arg Lys Ser Phe Phe Leu Lys Arg Val Ala Glu
305                 310                 315                 320

Ile Val Thr Cys Leu Gly Met Met Tyr Phe Leu Val Glu Gln Tyr Ala
                325                 330                 335

Thr Pro Thr Leu Gln Asn Ser Val Arg Ala Phe Asp Glu Leu Ala Phe
            340                 345                 350

Gly Thr Ile Leu Glu Arg Val Leu Lys Leu Ser Thr Thr Ser Val Ile
        355                 360                 365

Ile Trp Leu Leu Met Phe Tyr Thr Phe Phe His Ser Phe Phe Asn Ala
    370                 375                 380
```

```
Leu Ala Glu Ala Leu Tyr Phe Gly Asp Arg Arg Phe Tyr Leu Ala Trp
385                 390                 395                 400

Trp Asn Ala Thr Gly Val Gly Met Tyr Trp Lys Thr Trp Asn Ser Pro
                405                 410                 415

Val Tyr Thr Phe Phe Lys Arg His Val Tyr Leu Pro Leu Ile Thr Ser
            420                 425                 430

Gly Thr Ser Pro Met Val Ala Ser Ile Val Ile Phe Leu Ile Ser Ala
        435                 440                 445

Val Leu His Glu Ile Leu Ile Gly Phe Pro Thr His Met Ile Tyr Gly
    450                 455                 460

Tyr Ala Phe Ala Gly Met Phe Leu Gln Ile Pro Leu Ile Ile Leu Thr
465                 470                 475                 480

Arg Pro Leu Glu Lys Trp Arg Gly Thr Gly Ser Gly Leu Gly Asn Met
                485                 490                 495

Ile Phe Trp Val Ser Phe Thr Ile Leu Gly Gln Pro Ala Cys Ala Leu
            500                 505                 510

Leu Tyr Tyr Tyr His Trp Thr Lys Arg His Met Asp Val
        515                 520                 525
```

<210> 126
<211> 533
<212> PRT
<213> Neurospora crassa strain OR74A [GenBank Accession No. XP_322121]

<220>
<221> misc_feature
<223> DGAT1

<400> 126

```
Met Ser Ser Ser Thr Ala Thr Thr Thr Gly Leu Asp Pro Ala Val His
1               5                   10                  15
```

```
Thr Ser Asn Asp Asn Val Ile Arg Arg Thr His Gly Thr Glu Asn Gly
            20                  25                  30

Ser Thr Pro Asn Asp Lys Ala Asn Ala Gly Gly Glu Pro Glu Thr Glu
        35                  40                  45

Thr Lys Arg His Ser Lys Lys Val Val Arg Ser Lys Tyr Arg His Val
    50                  55                  60

Glu Ala Val His Ser Gln Ser Arg Pro Ser Cys Leu Ser His Asp Thr
65                  70                  75                  80

Thr Glu Ser Pro Ser Phe Leu Gly Phe Arg Asn Leu Met Val Ile Val
                85                  90                  95

Leu Ala Asn Asn Ser His Gln Tyr Gly Val Leu Ile Cys Ile Gly Cys
            100                 105                 110

His Asp Phe Arg Lys Ser Asp Ile Asn Leu Gly Leu Leu Leu Tyr Phe
        115                 120                 125

Leu Ile Pro Cys His Leu Phe Ile Ala Tyr Ile Ile Glu Tyr Tyr Ala
    130                 135                 140

Ala Val Gln Ala Arg Ala Glu Arg Asn Val Ser Ala Ser Glu Gln Asn
145                 150                 155                 160

Ala Lys Glu His Gln His Gln Asp Gly Thr Asn Ser Pro Thr Glu Glu
                165                 170                 175

Gln His Arg Lys Phe Gln Ser Thr Trp Lys Leu Val Arg Leu Leu His
            180                 185                 190

Ala Ile Asn Val Thr Thr Ala Leu Val Leu Thr Ser Tyr Val Val Tyr
        195                 200                 205

Tyr His Ile His His Pro Leu Ile Gly Thr Leu Thr Glu Val His Ala
    210                 215                 220
```

```
Ile Val Val Trp Leu Lys Thr Ala Ser Tyr Ala Phe Thr Asn Arg Asp
225                 230                 235                 240

Leu Arg His Ala Tyr Leu His Pro Ala Arg Gly Glu Leu Asp Ala Leu
                245                 250                 255

Pro Gly Leu Tyr Ala Glu Cys Pro Tyr Pro Glu Asn Ile Thr Met Gly
            260                 265                 270

Asn Leu Cys Tyr Phe Trp Trp Ala Pro Thr Leu Val Tyr Gln Pro Val
        275                 280                 285

Tyr Pro Arg Thr Ala Lys Ile Arg Trp Ser Phe Val Ala Lys Arg Cys
    290                 295                 300

Gly Glu Val Ile Cys Leu Ser Val Phe Ile Trp Phe Leu Ser Ala Gln
305                 310                 315                 320

Tyr Ala Thr Pro Val Leu Arg Asn Ser Leu Asp Lys Ile Ala Ser Leu
                325                 330                 335

Asp Ile Pro Ser Ile Val Glu Arg Leu Leu Lys Leu Ser Thr Ile Ser
            340                 345                 350

Leu Ile Ile Trp Leu Ala Gly Phe Phe Ala Leu Phe Gln Ser Phe Leu
        355                 360                 365

Asn Ala Leu Ala Glu Val Thr Arg Phe Ala Asp Arg Ser Phe Tyr Asp
    370                 375                 380

Glu Trp Trp Asn Ser Glu Ser Leu Gly Val Tyr Trp Arg Thr Trp Asn
385                 390                 395                 400

Lys Pro Val Tyr Gln Tyr Phe Lys Arg His Val Tyr Ser Pro Met Arg
                405                 410                 415

Ser Arg Gly Trp Ser Asn Ala Thr Ala Ser Leu Ala Val Phe Phe Leu
            420                 425                 430
```

```
Ser Ala Val Leu His Glu Leu Leu Val Gly Val Pro Thr His Asn Leu
        435                 440                 445

Ile Gly Val Ala Phe Leu Gly Met Phe Leu Gln Leu Pro Leu Ile Gln
    450                 455                 460

Phe Thr Lys Pro Leu Glu Lys Lys Thr Ser Pro Asn Gly Lys Leu Leu
465                 470                 475                 480

Gly Asn Ile Ile Phe Trp Val Ser Phe Thr Ile Phe Gly Gln Pro Phe
                485                 490                 495

Ala Ala Leu Met Tyr Phe Tyr Ala Trp Gln Ala Lys Tyr Gly Ser Val
            500                 505                 510

Ser Lys Met Thr Thr Ser Gln Gln Leu Val Gln Gln Gly Gln Gly Thr
        515                 520                 525

Cys Pro Pro Leu Val
    530
```

<210> 127
<211> 499
<212> PRT
<213> Gibberella zeae PH-1 [GenBank Accession No. EAA77624]

<220>
<221> misc_feature
<223> DGAT1

<400> 127

```
Met Asn Ser Ala Thr Thr Thr Ser Thr Glu Thr Ser Asn Gly Ser Thr
1               5                   10                  15

Ser Val Ser Lys Arg Asn Gly His Asp Val Thr Arg Thr Asn Gly Asn
            20                  25                  30

Gly Thr Thr Thr Thr Ser Pro Pro Lys Lys Ala Gly Gln Lys Tyr Arg
        35                  40                  45
```

```
His Val Ala Ala Val His Lys Lys Thr Arg Pro Ser Cys Leu Ser His
    50                  55                  60

Asp Ser Asp Ala Ala Pro Ser Phe Ile Gly Phe Arg Asn Leu Met Val
65                  70                  75                  80

Ile Val Leu Gly Ile Tyr His Ile Gly Met Ser Gln Phe Asp Ser Glu
                85                  90                  95

Gln Pro Ile Asp Thr Ala Ser Tyr Arg Gln Asp Ile Phe Leu Gly Leu
            100                 105                 110

Leu Leu Tyr Phe Leu Ile Pro Cys His Leu Leu Ala Ala Tyr Leu Ile
        115                 120                 125

Glu Leu Ala Ala Ala Gln Gln Ala Arg Gly Ser Leu Lys Arg Tyr Asn
    130                 135                 140

Asp Ser Ala Ser Gly Gly Pro Ser Asp Gln Glu Arg Lys Lys Phe His
145                 150                 155                 160

Lys Thr Trp Val Ile Val Ala Trp Ala His Leu Phe Asn Ile Thr Leu
                165                 170                 175

Ala Leu Val Leu Thr Thr Trp Val Val Tyr Phe Lys Ile His His Pro
            180                 185                 190

Leu Ile Gly Thr Leu Thr Glu Met His Ala Ile Ala Val Trp Leu Lys
        195                 200                 205

Thr Ala Ser Tyr Ala Phe Thr Asn Arg Asp Leu Arg His Ala Tyr Leu
    210                 215                 220

His Pro Val Glu Gly Glu Arg Glu Leu Val Pro Glu Leu Tyr Thr Gln
225                 230                 235                 240

Cys Pro Tyr Pro Gln Asn Ile Thr Phe Ser Asn Leu Ala Tyr Phe Trp
                245                 250                 255
```

```
Trp Ala Pro Thr Leu Val Tyr Gln Pro Val Tyr Pro Arg Thr Asp Lys
            260                 265                 270

Ile Arg Trp Gly Phe Val Ala Lys Arg Val Gly Glu Ile Phe Gly Leu
        275                 280                 285

Ser Val Phe Ile Trp Val Ala Ser Ala Gln Tyr Ala Ala Pro Val Leu
    290                 295                 300

Arg Asn Ser Leu Asp Lys Ile Ala Ser Leu Asp Leu Met Ser Ile Leu
305                 310                 315                 320

Glu Arg Leu Leu Lys Leu Ser Thr Ile Ser Leu Ala Ile Trp Leu Ala
                325                 330                 335

Gly Phe Phe Ala Leu Phe Gln Ser Phe Leu Asn Ala Leu Ala Glu Val
            340                 345                 350

Leu Arg Phe Gly Asp Arg Ser Phe Tyr Asp Asp Trp Trp Asn Ser Glu
        355                 360                 365

Ser Leu Gly Ala Tyr Trp Arg Thr Trp Asn Lys Pro Val Tyr Thr Tyr
    370                 375                 380

Phe Lys Arg His Leu Tyr Met Pro Met Ile Gly Arg Gly Trp Ser Pro
385                 390                 395                 400

Gln Ala Ala Ser Phe Phe Val Phe Leu Val Ser Ala Ile Leu His Glu
                405                 410                 415

Ile Leu Val Gly Val Pro Thr His Asn Ile Ile Gly Val Ala Phe Leu
            420                 425                 430

Gly Met Phe Leu Gln Leu Pro Leu Ile His Leu Thr Lys Pro Leu Glu
        435                 440                 445

Asn Met Lys Leu Gly His Thr Gly Lys Ile Val Gly Asn Thr Ile Phe
    450                 455                 460
```

```
Trp Val Ser Phe Thr Ile Phe Gly Gln Pro Phe Ala Ala Leu Met Tyr
465                 470                 475                 480

Phe Tyr Ala Trp Gln Ala Lys Tyr Gly Ser Val Thr Asp Ser Gly Phe
                485                 490                 495

Ser Ile Ser
```

<210> 128
<211> 503
<212> PRT
<213> Magnaporthe grisea 70-15 [GenBank Accession No. EAA52634]

<220>
<221> misc_feature
<223> DGAT1

<400> 128

```
Met Ala Ala Ala Thr Ala Thr Gly Leu Asp Leu Ala Ala Gln Glu Gly
1               5                   10                  15

Ala Gln Gln Arg Arg Ser Thr Ala Thr Asn Gln Ser Ala Asp Asp Asp
            20                  25                  30

Val Thr Thr Asn Ala Asp Gly Ala Ala Ala Ala Pro Ser Leu Lys Gly
        35                  40                  45

Thr Thr Ala Asp Thr Asn Gly Thr Ser Asn Gly Asn Gly Asn Gly Asn
    50                  55                  60

Gly Asn Val Asp Glu Asp Glu Gln Thr Lys Ala Leu Arg Lys Ala Phe
65                  70                  75                  80

Thr Arg Lys Tyr Arg His Val Ala Ala Leu His Ser Gln Ala Arg Pro
                85                  90                  95

Ser Thr Leu Ser His Asp Ser Glu Ala Ser Pro Ser Phe Val Gly Phe
            100                 105                 110
```

```
Arg Asn Leu Met Val Ile Val Leu Glu Leu Leu Ala Ala Gln Gln Ala
        115                 120                 125

Arg Asn Ser Arg Gly Tyr Phe Asn Arg Gly Arg Thr Gly Ser Ser Arg
    130                 135                 140

Asp Gly Ser Thr Ser Pro Thr Glu Asp Glu Ser Arg Arg Phe Val Ser
145                 150                 155                 160

Thr Trp Lys Leu Ile Ala Leu Val His Gly Ile Asn Val Asn Ser Ala
                165                 170                 175

Leu Leu Ile Thr Thr Tyr Thr Val Tyr Phe His Ile His His Pro Leu
            180                 185                 190

Ile Gly Thr Leu Thr Glu Met His Ala Val Ile Val Trp Leu Lys Thr
        195                 200                 205

Ala Ser Tyr Ala Phe Thr Asn Arg Asp Leu Arg His Ala Tyr Leu His
    210                 215                 220

Pro Val Lys Gly Glu Leu Asp Ala Leu Pro Glu Leu Tyr Lys Gln Cys
225                 230                 235                 240

Pro Tyr Pro Asn Asn Ile Thr Met Lys Asn Leu Cys Tyr Phe Trp Trp
                245                 250                 255

Ala Pro Thr Leu Ile Tyr Gln Pro Val Tyr Pro Arg Ser Gly Arg Ile
            260                 265                 270

Arg Trp Val Phe Phe Phe Lys Arg Val Ala Glu Val Phe Cys Leu Ser
        275                 280                 285

Val Cys Ile Trp Phe Leu Ser Ala Gln Tyr Ala Thr Pro Val Leu Val
    290                 295                 300

Asn Ser Leu Asp Lys Ile Ala Ser Leu Asp Met Pro Ala Ile Leu Glu
305                 310                 315                 320
```

```
Arg Leu Leu Lys Leu Ser Thr Ile Ser Leu Ala Ile Trp Leu Ala Gly
                325                 330                 335

Phe Phe Ala Leu Phe Gln Ser Phe Leu Asn Ala Leu Ala Glu Ile Thr
            340                 345                 350

Arg Phe Gly Asp Arg Ser Phe Tyr Glu Ala Trp Trp Asn Ser Glu Ser
        355                 360                 365

Leu Gly Val Tyr Trp Arg Thr Trp Asn Lys Pro Val Tyr Gln Tyr Phe
    370                 375                 380

Lys Arg His Val Tyr Ser Pro Met Leu Gly Arg Gly Trp Ala Pro Arg
385                 390                 395                 400

Thr Ala Ser Ala Ser Val Phe Leu Ile Ser Ala Val Leu His Glu Ile
                405                 410                 415

Leu Val Gly Val Pro Thr His Asn Ile Ile Gly Val Ala Phe Met Gly
            420                 425                 430

Met Phe Leu Gln Val Pro Leu Ile Ile Leu Thr Ala Pro Leu Glu Lys
        435                 440                 445

Arg Lys Ser Pro Thr Gly Lys Leu Ile Gly Asn Ser Ile Phe Trp Val
    450                 455                 460

Ser Phe Thr Ile Phe Gly Gln Pro Leu Ala Ala Leu Met Tyr Phe Tyr
465                 470                 475                 480

Ala Trp Gln Ala Lys Tyr Gly Ser Val Ser Lys Met Gly Tyr Ala Thr
                485                 490                 495

Ser Lys Ala Ala Leu Thr Asn
            500
```

<210> 129
<211> 458
<212> PRT
<213> Aspergillus nidulans FGSC A4 [GenBank Accession No. EAA57945]

<220>
<221> misc_feature
<223> DGAT1

<400> 129

```
Met Ala Thr Arg Lys Thr Ala Ile Tyr Arg His Ala Val Ala Val His
1               5                  10                  15

Ser Gln Val Gln His Ser Cys Leu Ser Arg Asp Ser Thr Lys Ala Thr
            20                  25                  30

Ser Phe Ile Gly Phe Arg Asn Leu Met Val Val Val Leu Val Ala Met
        35                  40                  45

Asn Leu Arg Leu Val Ile Glu Asn Phe Leu Lys Tyr Gly Val Leu Ile
    50                  55                  60

Cys Ile Arg Cys His Asp Tyr Arg Lys Gln Asp Val Val Ile Gly Ala
65                  70                  75                  80

Ile Leu Phe Ala Leu Val Pro Cys Gln Leu Leu Cys Ser Tyr Phe Ile
                85                  90                  95

Glu Leu Ala Ala Ser Arg His Ala Gln Arg Val Ile Gly Arg Ala Lys
            100                 105                 110

Lys Gln Asp Lys Asp Arg Ile Leu Asn Glu Ser Lys Arg Thr Trp Phe
        115                 120                 125

Ala Ile Ala Leu Leu His Ser Ile Ile Ser Phe Phe Gly Leu Ala Ala
    130                 135                 140

Thr Ser Tyr Val Ile Phe Tyr Tyr Val Asn His Pro Gly Ile Gly Thr
145                 150                 155                 160

Val Cys Glu Val Gln Val Ile Ile Val Ser Leu Lys Ser Tyr Ser Tyr
                165                 170                 175

Ala Leu Thr Asn Arg Asp Leu Arg Arg Ala Met Leu Gly Ser Pro Ser
            180                 185                 190
```

```
Ala Asp Ser Asp Ile Pro Glu Leu Tyr Arg Ser Cys Pro Tyr Pro Arg
        195                 200                 205

Asn Ile Thr Leu Gly Asn Leu Ala Tyr Phe Leu Trp Ala Pro Thr Leu
    210                 215                 220

Val Tyr Gln Pro Val Tyr Pro Arg Thr Pro Arg Ile Arg Trp Ser Phe
225                 230                 235                 240

Val Gly Lys Arg Leu Phe Glu Phe Val Cys Leu Ser Val Val Met Trp
                245                 250                 255

Leu Leu Ser Ala Gln Tyr Ala Ala Pro Leu Leu Arg Asn Ala Thr Gln
            260                 265                 270

Lys Ile Ala Thr Leu Asp Ile Ala Ser Ile Leu Glu Arg Gly Leu Lys
        275                 280                 285

Leu Ser Thr Ile Ser Leu Val Ile Trp Leu Ala Gly Phe Tyr Ala Leu
    290                 295                 300

Phe Gln Ser Leu Leu Asn Gly Leu Ala Glu Ile Met Arg Phe Gly Asp
305                 310                 315                 320

Arg Glu Phe Tyr Thr Asp Trp Trp Asn Ser Pro Ser Phe Gly Val Tyr
                325                 330                 335

Trp Arg Ser Trp Asn Arg Pro Val Tyr Ile Phe Met Lys Arg His Val
            340                 345                 350

Tyr Met Pro Leu Val Thr Arg Gly Trp Asn Pro Thr Leu Ala Gly Thr
        355                 360                 365

Val Val Phe Ala Val Ser Ala Val Leu His Glu Ile Leu Val Gly Val
    370                 375                 380

Pro Thr His Asn Leu Ile Gly Val Ala Ser Ile Ala Met Met Phe Gln
385                 390                 395                 400
```

Leu Pro Leu Ile Leu Leu Thr Ala Pro Phe Glu Arg Phe Lys Ser Pro
405 410 415

Leu Gly Lys Ala Ile Gly Asn Ser Phe Phe Trp Val Thr Phe Cys Val
420 425 430

Val Gly Gln Pro Leu Gly Ala Leu Leu Tyr Phe Phe Ala Trp Gln Ala
435 440 445

Lys Tyr Gly Ser Val Ser Gln Thr His Pro
450 455

<210> 130
<211> 2119
<212> DNA
<213> Yarrowia lipolytica

<220>
<221> CDS
<222> (291)..(1835)
<223> DGAT2 opening reading frame, comprising 2 smaller internal opening reading frames

<220>
<221> misc_feature
<222> (291)..(293)
<223> initiation codon ('ATG')

<220>
<221> misc_feature
<222> (456)..(458)
<223> initiation codon ('ATG')

<220>
<221> misc_feature
<222> (768)..(770)
<223> initiation codon ('ATG')

<400> 130
aaacgcaccc actgctcgtc ctccttgctc ctcgaaaccg actcctctac acacgtcaaa 60

tccgaggttg aaatcttccc cacatttggc agccaaacca gcacatccca gcaacctcgc 120

acagcgccga aatcgacctg tcgacttggc cacaaaaaaa agcaccggct ctgcaacagt 180

```
tctcacgacc aattacgtac aagtacgaaa tcgttcgtgg accgtgactg ataagctccc    240

acttttttctt ctaacaacag gcaacagaca agtcacacaa aacaaaagct atg act       296
                                                          Met Thr
                                                          1

atc gac tca caa tac tac aag tcg cga gac aaa aac gac acg gca ccc      344
Ile Asp Ser Gln Tyr Tyr Lys Ser Arg Asp Lys Asn Asp Thr Ala Pro
        5                  10                  15

aaa atc gcg gga atc cga tat gcc ccg cta tcg aca cca tta ctc aac      392
Lys Ile Ala Gly Ile Arg Tyr Ala Pro Leu Ser Thr Pro Leu Leu Asn
    20                  25                  30

cga tgt gag acc ttc tct ctg gtc tgg cac att ttc agc att ccc act      440
Arg Cys Glu Thr Phe Ser Leu Val Trp His Ile Phe Ser Ile Pro Thr
35                  40                  45                  50

ttc ctc aca att ttc atg cta tgc tgc gca att cca ctg ctc tgg cca      488
Phe Leu Thr Ile Phe Met Leu Cys Cys Ala Ile Pro Leu Leu Trp Pro
                55                  60                  65

ttt gtg att gcg tat gta gtg tac gct gtt aaa gac gac tcc ccg tcc      536
Phe Val Ile Ala Tyr Val Val Tyr Ala Val Lys Asp Asp Ser Pro Ser
            70                  75                  80

aac gga gga gtg gtc aag cga tac tcg cct att tca aga aac ttc ttc      584
Asn Gly Gly Val Val Lys Arg Tyr Ser Pro Ile Ser Arg Asn Phe Phe
        85                  90                  95

atc tgg aag ctc ttt ggc cgc tac ttc ccc ata act ctg cac aag acg      632
Ile Trp Lys Leu Phe Gly Arg Tyr Phe Pro Ile Thr Leu His Lys Thr
    100                 105                 110

gtg gat ctg gag ccc acg cac aca tac tac cct ctg gac gtc cag gag      680
Val Asp Leu Glu Pro Thr His Thr Tyr Tyr Pro Leu Asp Val Gln Glu
115                 120                 125                 130

tat cac ctg att gct gag aga tac tgg ccg cag aac aag tac ctc cga      728
Tyr His Leu Ile Ala Glu Arg Tyr Trp Pro Gln Asn Lys Tyr Leu Arg
                135                 140                 145

gca atc atc tcc acc atc gag tac ttt ctg ccc gcc ttc atg aaa cgg      776
Ala Ile Ile Ser Thr Ile Glu Tyr Phe Leu Pro Ala Phe Met Lys Arg
            150                 155                 160

tct ctt tct atc aac gag cag gag cag cct gcc gag cga gat cct ctc      824
Ser Leu Ser Ile Asn Glu Gln Glu Gln Pro Ala Glu Arg Asp Pro Leu
        165                 170                 175

ctg tct ccc gtt tct ccc agc tct ccg ggt tct caa cct gac aag tgg      872
Leu Ser Pro Val Ser Pro Ser Ser Pro Gly Ser Gln Pro Asp Lys Trp
```

```
    180                 185                 190

att aac cac gac agc aga tat agc cgt gga gaa tca tct ggc tcc aac      920
Ile Asn His Asp Ser Arg Tyr Ser Arg Gly Glu Ser Ser Gly Ser Asn
195                 200                 205                 210

ggc cac gcc tcg ggc tcc gaa ctt aac ggc aac ggc aac aat ggc acc      968
Gly His Ala Ser Gly Ser Glu Leu Asn Gly Asn Gly Asn Asn Gly Thr
                215                 220                 225

act aac cga cga cct ttg tcg tcc gcc tct gct ggc tcc act gca tct     1016
Thr Asn Arg Arg Pro Leu Ser Ser Ala Ser Ala Gly Ser Thr Ala Ser
            230                 235                 240

gat tcc acg ctt ctt aac ggg tcc ctc aac tcc tac gcc aac cag atc     1064
Asp Ser Thr Leu Leu Asn Gly Ser Leu Asn Ser Tyr Ala Asn Gln Ile
        245                 250                 255

att ggc gaa aac gac cca cag ctg tcg ccc aca aaa ctc aag ccc act     1112
Ile Gly Glu Asn Asp Pro Gln Leu Ser Pro Thr Lys Leu Lys Pro Thr
    260                 265                 270

ggc aga aaa tac atc ttc ggc tac cac ccc cac ggc att atc ggc atg     1160
Gly Arg Lys Tyr Ile Phe Gly Tyr His Pro His Gly Ile Ile Gly Met
275                 280                 285                 290

gga gcc ttt ggt gga att gcc acc gag gga gct gga tgg tcc aag ctc     1208
Gly Ala Phe Gly Gly Ile Ala Thr Glu Gly Ala Gly Trp Ser Lys Leu
                295                 300                 305

ttt ccg ggc atc cct gtt tct ctt atg act ctc acc aac aac ttc cga     1256
Phe Pro Gly Ile Pro Val Ser Leu Met Thr Leu Thr Asn Asn Phe Arg
            310                 315                 320

gtg cct ctc tac aga gag tac ctc atg agt ctg gga gtc gct tct gtc     1304
Val Pro Leu Tyr Arg Glu Tyr Leu Met Ser Leu Gly Val Ala Ser Val
        325                 330                 335

tcc aag aag tcc tgc aag gcc ctc ctc aag cga aac cag tct atc tgc     1352
Ser Lys Lys Ser Cys Lys Ala Leu Leu Lys Arg Asn Gln Ser Ile Cys
    340                 345                 350

att gtc gtt ggt gga gca cag gaa agt ctt ctg gcc aga ccc ggt gtc     1400
Ile Val Val Gly Gly Ala Gln Glu Ser Leu Leu Ala Arg Pro Gly Val
355                 360                 365                 370

atg gac ctg gtg cta ctc aag cga aag ggt ttt gtt cga ctt ggt atg     1448
Met Asp Leu Val Leu Leu Lys Arg Lys Gly Phe Val Arg Leu Gly Met
                375                 380                 385

gag gtc gga aat gtc gcc ctt gtt ccc atc atg gcc ttt ggt gag aac     1496
Glu Val Gly Asn Val Ala Leu Val Pro Ile Met Ala Phe Gly Glu Asn
```

```
            390                 395                 400

gac ctc tat gac cag gtt agc aac gac aag tcg tcc aag ctg tac cga     1544
Asp Leu Tyr Asp Gln Val Ser Asn Asp Lys Ser Ser Lys Leu Tyr Arg
        405                 410                 415

ttc cag cag ttt gtc aag aac ttc ctt gga ttc acc ctt cct ttg atg     1592
Phe Gln Gln Phe Val Lys Asn Phe Leu Gly Phe Thr Leu Pro Leu Met
    420                 425                 430

cat gcc cga ggc gtc ttc aac tac gat gtc ggt ctt gtc ccc tac agg     1640
His Ala Arg Gly Val Phe Asn Tyr Asp Val Gly Leu Val Pro Tyr Arg
435                 440                 445                 450

cga ccc gtc aac att gtg gtt ggt tcc ccc att gac ttg cct tat ctc     1688
Arg Pro Val Asn Ile Val Val Gly Ser Pro Ile Asp Leu Pro Tyr Leu
                455                 460                 465

cca cac ccc acc gac gaa gaa gtg tcc gaa tac cac gac cga tac atc     1736
Pro His Pro Thr Asp Glu Glu Val Ser Glu Tyr His Asp Arg Tyr Ile
            470                 475                 480

gcc gag ctg cag cga atc tac aac gag cac aag gat gaa tat ttc atc     1784
Ala Glu Leu Gln Arg Ile Tyr Asn Glu His Lys Asp Glu Tyr Phe Ile
        485                 490                 495

gat tgg acc gag gag ggc aaa gga gcc cca gag ttc cga atg att gag     1832
Asp Trp Thr Glu Glu Gly Lys Gly Ala Pro Glu Phe Arg Met Ile Glu
    500                 505                 510

taa ggaaaactgc ctgggttagg caaatagcta atgagtattt ttttgatggc          1885

aaccaaatgt agaaagaaaa aaaaaaaaaa agaaaaaaaa aagagaatat tatatctatg   1945

taattctatt aaaagctctg ttgagtgagc ggaataaata ctgttgaaga ggggattgtg   2005

tagagatctg tttactcaat ggcaaactca tctgggggag atccttccac tgtgggaagc   2065

tcctggatag cctttgcatc ggggttcaag aagaccattg tgaacagccc ttga         2119


<210>  131
<211>  514
<212>  PRT
<213>  Yarrowia lipolytica

<400>  131

Met Thr Ile Asp Ser Gln Tyr Tyr Lys Ser Arg Asp Lys Asn Asp Thr
1               5                   10                  15


Ala Pro Lys Ile Ala Gly Ile Arg Tyr Ala Pro Leu Ser Thr Pro Leu
            20                  25                  30
```

```
Leu Asn Arg Cys Glu Thr Phe Ser Leu Val Trp His Ile Phe Ser Ile
        35                  40                  45

Pro Thr Phe Leu Thr Ile Phe Met Leu Cys Cys Ala Ile Pro Leu Leu
    50                  55                  60

Trp Pro Phe Val Ile Ala Tyr Val Val Tyr Ala Val Lys Asp Asp Ser
65                  70                  75                  80

Pro Ser Asn Gly Gly Val Val Lys Arg Tyr Ser Pro Ile Ser Arg Asn
                85                  90                  95

Phe Phe Ile Trp Lys Leu Phe Gly Arg Tyr Phe Pro Ile Thr Leu His
            100                 105                 110

Lys Thr Val Asp Leu Glu Pro Thr His Thr Tyr Tyr Pro Leu Asp Val
        115                 120                 125

Gln Glu Tyr His Leu Ile Ala Glu Arg Tyr Trp Pro Gln Asn Lys Tyr
    130                 135                 140

Leu Arg Ala Ile Ile Ser Thr Ile Glu Tyr Phe Leu Pro Ala Phe Met
145                 150                 155                 160

Lys Arg Ser Leu Ser Ile Asn Glu Gln Glu Gln Pro Ala Glu Arg Asp
                165                 170                 175

Pro Leu Leu Ser Pro Val Ser Pro Ser Ser Pro Gly Ser Gln Pro Asp
            180                 185                 190

Lys Trp Ile Asn His Asp Ser Arg Tyr Ser Arg Gly Glu Ser Ser Gly
        195                 200                 205

Ser Asn Gly His Ala Ser Gly Ser Glu Leu Asn Gly Asn Gly Asn Asn
    210                 215                 220

Gly Thr Thr Asn Arg Arg Pro Leu Ser Ser Ala Ser Ala Gly Ser Thr
225                 230                 235                 240
```

```
Ala Ser Asp Ser Thr Leu Leu Asn Gly Ser Leu Asn Ser Tyr Ala Asn
               245                 250                 255

Gln Ile Ile Gly Glu Asn Asp Pro Gln Leu Ser Pro Thr Lys Leu Lys
            260                 265                 270

Pro Thr Gly Arg Lys Tyr Ile Phe Gly Tyr His Pro His Gly Ile Ile
        275                 280                 285

Gly Met Gly Ala Phe Gly Gly Ile Ala Thr Glu Gly Ala Gly Trp Ser
    290                 295                 300

Lys Leu Phe Pro Gly Ile Pro Val Ser Leu Met Thr Leu Thr Asn Asn
305                 310                 315                 320

Phe Arg Val Pro Leu Tyr Arg Glu Tyr Leu Met Ser Leu Gly Val Ala
               325                 330                 335

Ser Val Ser Lys Lys Ser Cys Lys Ala Leu Leu Lys Arg Asn Gln Ser
            340                 345                 350

Ile Cys Ile Val Val Gly Gly Ala Gln Glu Ser Leu Leu Ala Arg Pro
        355                 360                 365

Gly Val Met Asp Leu Val Leu Leu Lys Arg Lys Gly Phe Val Arg Leu
    370                 375                 380

Gly Met Glu Val Gly Asn Val Ala Leu Val Pro Ile Met Ala Phe Gly
385                 390                 395                 400

Glu Asn Asp Leu Tyr Asp Gln Val Ser Asn Asp Lys Ser Ser Lys Leu
               405                 410                 415

Tyr Arg Phe Gln Gln Phe Val Lys Asn Phe Leu Gly Phe Thr Leu Pro
            420                 425                 430

Leu Met His Ala Arg Gly Val Phe Asn Tyr Asp Val Gly Leu Val Pro
        435                 440                 445
```

Tyr Arg Arg Pro Val Asn Ile Val Val Gly Ser Pro Ile Asp Leu Pro
450 455 460

Tyr Leu Pro His Pro Thr Asp Glu Glu Val Ser Glu Tyr His Asp Arg
465 470 475 480

Tyr Ile Ala Glu Leu Gln Arg Ile Tyr Asn Glu His Lys Asp Glu Tyr
485 490 495

Phe Ile Asp Trp Thr Glu Glu Gly Lys Gly Ala Pro Glu Phe Arg Met
500 505 510

Ile Glu

<210> 132
<211> 1380
<212> DNA
<213> Yarrowia lipolytica

<220>
<221> misc_feature
<223> DGAT2

<400> 132
atgctatgct gcgcaattcc actgctctgg ccatttgtga ttgcgtatgt agtgtacgct 60

gttaaagacg actccccgtc caacggagga gtggtcaagc gatactcgcc tatttcaaga 120

aacttcttca tctggaagct ctttggccgc tacttcccca taactctgca caagacggtg 180

gatctggagc ccacgcacac atactaccct ctggacgtcc aggagtatca cctgattgct 240

gagagatact ggccgcagaa caagtacctc cgagcaatca tctccaccat cgagtacttt 300

ctgcccgcct tcatgaaacg gtctctttct atcaacgagc aggagcagcc tgccgagcga 360

gatcctctcc tgtctcccgt ttctcccagc tctccgggtt ctcaacctga caagtggatt 420

aaccacgaca gcagatatag ccgtggagaa tcatctggct ccaacggcca cgcctcgggc 480

tccgaactta acggcaacgg caacaatggc accactaacc gacgaccttt gtcgtccgcc 540

tctgctggct ccactgcatc tgattccacg cttcttaacg ggtccctcaa ctcctacgcc 600

aaccagatca ttggcgaaaa cgacccacag ctgtcgccca caaaactcaa gcccactggc 660

agaaaatata tcttcggcta ccacccccac ggcattatcg gcatgggagc ctttggtgga 720

attgccaccg agggagctgg atggtccaag ctctttccgg gcatccctgt ttctcttatg 780

actctcacca acaacttccg agtgcctctc tacagagagt acctcatgag tctgggagtc 840

gcttctgtct ccaagaagtc ctgcaaggcc ctcctcaagc gaaaccagtc tatctgcatt 900

gtcgttggtg gagcacagga aagtcttctg gccagacccg gtgtcatgga cctggtgcta 960

ctcaagcgaa agggttttgt tcgacttggt atggaggtcg gaaatgtcgc ccttgttccc 1020

atcatggcct ttggtgagaa cgacctctat gaccaggtta gcaacgacaa gtcgtccaag 1080

ctgtaccgat tccagcagtt tgtcaagaac ttccttggat tcacccttcc tttgatgcat 1140

gcccgaggcg tcttcaacta cgatgtcggt cttgtcccct acaggcgacc cgtcaacatt 1200

gtggttggtt cccccattga cttgccttat ctcccacacc ccaccgacga agaagtgtcc 1260

gaataccacg accgatacat cgccgagctg cagcgaatct acaacgagca caaggatgaa 1320

tatttcatcg attggaccga ggagggcaaa ggagccccag agttccgaat gattgagtaa 1380

<210> 133
<211> 459
<212> PRT
<213> Yarrowia lipolytica

<400> 133

Met Leu Cys Cys Ala Ile Pro Leu Leu Trp Pro Phe Val Ile Ala Tyr
1 5 10 15

Val Val Tyr Ala Val Lys Asp Asp Ser Pro Ser Asn Gly Gly Val Val
20 25 30

Lys Arg Tyr Ser Pro Ile Ser Arg Asn Phe Phe Ile Trp Lys Leu Phe
35 40 45

Gly Arg Tyr Phe Pro Ile Thr Leu His Lys Thr Val Asp Leu Glu Pro
50 55 60

Thr His Thr Tyr Tyr Pro Leu Asp Val Gln Glu Tyr His Leu Ile Ala
65 70 75 80

```
Glu Arg Tyr Trp Pro Gln Asn Lys Tyr Leu Arg Ala Ile Ile Ser Thr
                85                  90                  95

Ile Glu Tyr Phe Leu Pro Ala Phe Met Lys Arg Ser Leu Ser Ile Asn
            100                 105                 110

Glu Gln Glu Gln Pro Ala Glu Arg Asp Pro Leu Leu Ser Pro Val Ser
        115                 120                 125

Pro Ser Ser Pro Gly Ser Gln Pro Asp Lys Trp Ile Asn His Asp Ser
    130                 135                 140

Arg Tyr Ser Arg Gly Glu Ser Ser Gly Ser Asn Gly His Ala Ser Gly
145                 150                 155                 160

Ser Glu Leu Asn Gly Asn Gly Asn Asn Gly Thr Thr Asn Arg Arg Pro
                165                 170                 175

Leu Ser Ser Ala Ser Ala Gly Ser Thr Ala Ser Asp Ser Thr Leu Leu
            180                 185                 190

Asn Gly Ser Leu Asn Ser Tyr Ala Asn Gln Ile Ile Gly Glu Asn Asp
        195                 200                 205

Pro Gln Leu Ser Pro Thr Lys Leu Lys Pro Thr Gly Arg Lys Tyr Ile
    210                 215                 220

Phe Gly Tyr His Pro His Gly Ile Ile Gly Met Gly Ala Phe Gly Gly
225                 230                 235                 240

Ile Ala Thr Glu Gly Ala Gly Trp Ser Lys Leu Phe Pro Gly Ile Pro
                245                 250                 255

Val Ser Leu Met Thr Leu Thr Asn Asn Phe Arg Val Pro Leu Tyr Arg
            260                 265                 270

Glu Tyr Leu Met Ser Leu Gly Val Ala Ser Val Ser Lys Lys Ser Cys
        275                 280                 285
```

```
Lys Ala Leu Leu Lys Arg Asn Gln Ser Ile Cys Ile Val Val Gly Gly
    290                 295                 300

Ala Gln Glu Ser Leu Leu Ala Arg Pro Gly Val Met Asp Leu Val Leu
305                 310                 315                 320

Leu Lys Arg Lys Gly Phe Val Arg Leu Gly Met Glu Val Gly Asn Val
                325                 330                 335

Ala Leu Val Pro Ile Met Ala Phe Gly Glu Asn Asp Leu Tyr Asp Gln
            340                 345                 350

Val Ser Asn Asp Lys Ser Ser Lys Leu Tyr Arg Phe Gln Gln Phe Val
        355                 360                 365

Lys Asn Phe Leu Gly Phe Thr Leu Pro Leu Met His Ala Arg Gly Val
    370                 375                 380

Phe Asn Tyr Asp Val Gly Leu Val Pro Tyr Arg Arg Pro Val Asn Ile
385                 390                 395                 400

Val Val Gly Ser Pro Ile Asp Leu Pro Tyr Leu Pro His Pro Thr Asp
                405                 410                 415

Glu Glu Val Ser Glu Tyr His Asp Arg Tyr Ile Ala Glu Leu Gln Arg
            420                 425                 430

Ile Tyr Asn Glu His Lys Asp Glu Tyr Phe Ile Asp Trp Thr Glu Glu
        435                 440                 445

Gly Lys Gly Ala Pro Glu Phe Arg Met Ile Glu
    450                 455
```

<210> 134
<211> 1068
<212> DNA
<213> Yarrowia lipolytica

<220>
<221> misc_feature
<223> DGAT2

<400> 134
atgaaacggt ctctttctat caacgagcag gagcagcctg ccgagcgaga tcctctcctg 60

tctcccgttt ctcccagctc tccgggttct caacctgaca agtggattaa ccacgacagc 120

agatatagcc gtggagaatc atctggctcc aacggccacg cctcgggctc cgaacttaac 180

ggcaacggca acaatggcac cactaaccga cgacctttgt cgtccgcctc tgctggctcc 240

actgcatctg attccacgct tcttaacggg tccctcaact cctacgccaa ccagatcatt 300

ggcgaaaacg acccacagct gtcgcccaca aaactcaagc ccactggcag aaaatacatc 360

ttcggctacc acccccacgg cattatcggc atgggagcct ttggtggaat tgccaccgag 420

ggagctggat ggtccaagct ctttccgggc atccctgttt ctcttatgac tctcaccaac 480

aacttccgag tgcctctcta cagagagtac ctcatgagtc tgggagtcgc ttctgtctcc 540

aagaagtcct gcaaggccct cctcaagcga aaccagtcta tctgcattgt cgttggtgga 600

gcacaggaaa gtcttctggc cagacccggt gtcatggacc tggtgctact caagcgaaag 660

ggttttgttc gacttggtat ggaggtcgga aatgtcgccc ttgttcccat catggccttt 720

ggtgagaacg acctctatga ccaggttagc aacgacaagt cgtccaagct gtaccgattc 780

cagcagtttg tcaagaactt ccttggattc acccttcctt tgatgcatgc ccgaggcgtc 840

ttcaactacg atgtcggtct tgtcccctac aggcgacccg tcaacattgt ggttggttcc 900

cccattgact tgccttatct cccacacccc accgacgaag aagtgtccga ataccacgac 960

cgatacatcg ccgagctgca gcgaatctac aacgagcaca aggatgaata tttcatcgat 1020

tggaccgagg agggcaaagg agccccagag ttccgaatga ttgagtaa 1068

<210> 135
<211> 355
<212> PRT
<213> Yarrowia lipolytica

<400> 135

Met Lys Arg Ser Leu Ser Ile Asn Glu Gln Glu Gln Pro Ala Glu Arg
1 5 10 15

```
Asp Pro Leu Leu Ser Pro Val Ser Pro Ser Ser Pro Gly Ser Gln Pro
            20                  25                  30

Asp Lys Trp Ile Asn His Asp Ser Arg Tyr Ser Arg Gly Glu Ser Ser
        35                  40                  45

Gly Ser Asn Gly His Ala Ser Gly Ser Glu Leu Asn Gly Asn Gly Asn
    50                  55                  60

Asn Gly Thr Thr Asn Arg Arg Pro Leu Ser Ser Ala Ser Ala Gly Ser
65                  70                  75                  80

Thr Ala Ser Asp Ser Thr Leu Leu Asn Gly Ser Leu Asn Ser Tyr Ala
                85                  90                  95

Asn Gln Ile Ile Gly Glu Asn Asp Pro Gln Leu Ser Pro Thr Lys Leu
            100                 105                 110

Lys Pro Thr Gly Arg Lys Tyr Ile Phe Gly Tyr His Pro His Gly Ile
        115                 120                 125

Ile Gly Met Gly Ala Phe Gly Gly Ile Ala Thr Glu Gly Ala Gly Trp
    130                 135                 140

Ser Lys Leu Phe Pro Gly Ile Pro Val Ser Leu Met Thr Leu Thr Asn
145                 150                 155                 160

Asn Phe Arg Val Pro Leu Tyr Arg Glu Tyr Leu Met Ser Leu Gly Val
                165                 170                 175

Ala Ser Val Ser Lys Lys Ser Cys Lys Ala Leu Leu Lys Arg Asn Gln
            180                 185                 190

Ser Ile Cys Ile Val Val Gly Gly Ala Gln Glu Ser Leu Leu Ala Arg
        195                 200                 205

Pro Gly Val Met Asp Leu Val Leu Leu Lys Arg Lys Gly Phe Val Arg
    210                 215                 220
```

```
Leu Gly Met Glu Val Gly Asn Val Ala Leu Val Pro Ile Met Ala Phe
225                 230                 235                 240

Gly Glu Asn Asp Leu Tyr Asp Gln Val Ser Asn Asp Lys Ser Ser Lys
                245                 250                 255

Leu Tyr Arg Phe Gln Gln Phe Val Lys Asn Phe Leu Gly Phe Thr Leu
            260                 265                 270

Pro Leu Met His Ala Arg Gly Val Phe Asn Tyr Asp Val Gly Leu Val
        275                 280                 285

Pro Tyr Arg Arg Pro Val Asn Ile Val Val Gly Ser Pro Ile Asp Leu
    290                 295                 300

Pro Tyr Leu Pro His Pro Thr Asp Glu Glu Val Ser Glu Tyr His Asp
305                 310                 315                 320

Arg Tyr Ile Ala Glu Leu Gln Arg Ile Tyr Asn Glu His Lys Asp Glu
                325                 330                 335

Tyr Phe Ile Asp Trp Thr Glu Glu Gly Lys Gly Ala Pro Glu Phe Arg
            340                 345                 350

Met Ile Glu
        355


<210>  136
<211>  996
<212>  DNA
<213>  Mortierella alpina


<220>
<221>  misc_feature
<223>  DGAT2

<400>  136
atgccgctct ttgcgccttt acggatgccc gtcaagcgtc gtatgcagac aggagctgtc     60

ctatactgga ttgcggggat gatttactgc attggcatct ttgccttcct ctgcacgttc    120
```

aagatccttc gacccttgat catcatctat gtcctgtggg cctacatgct cgaccgagca 180

ccagagcggg gtgcacgcac agtccaatgg tattgtaact ggatcggatg gaaacacttt 240

gcacagtact ttcctatgac ccttgtcaag gagggagagc tggacccatc caagaactac 300

atctttgggt atcacccaca cggaatcatt tccttgggtg ccctctgcgc ctttgggacc 360

gagggccttc atttctccaa acgcttcccg ggtatcaagc ctcatctgct caccattcac 420

gccaactttc agatcccact ctatcgcgat atgatcatgg cccacggctg tgcttccgtg 480

tcgagggcct cttgtgaaca catcctgcgg tctggcgaag gatcctcggt cgtgatcgtt 540

gtcgggggtg cacaagaaag tttgtcgact caacctggca cgttaaatct gacactcaag 600

aaaagactgg gattttgcaa gctggccttt gtcaatggcg caagtctggt acctacgttg 660

gcctttggtg agaacgagct ctatgaggtg taccacacca agcccacaag cctgatatac 720

aagctccagc agttgactaa acgcacgatc ggcttcacaa tgcccgtgtt caacggacga 780

ggaatcttca attatgagtt tggactgctg ccaaggagga agcctgtcta tatcgttata 840

ggaaacccca ttcatgtaga caaggtcgag aacccaacga ttgaacagat gcagaaactg 900

cagtcaattt acattgatga ggtgctaaac atttgggaaa gatacaagga caagtatgcc 960

gcaggacgaa ctcaggaact gtgcatcatc gaatag 996

<210> 137
<211> 331
<212> PRT
<213> Mortierella alpina

<400> 137

Met Pro Leu Phe Ala Pro Leu Arg Met Pro Val Lys Arg Arg Met Gln
1 5 10 15

Thr Gly Ala Val Leu Tyr Trp Ile Ala Gly Met Ile Tyr Cys Ile Gly
20 25 30

Ile Phe Ala Phe Leu Cys Thr Phe Lys Ile Leu Arg Pro Leu Ile Ile
35 40 45

Ile Tyr Val Leu Trp Ala Tyr Met Leu Asp Arg Ala Pro Glu Arg Gly
50 55 60

```
Ala Arg Thr Val Gln Trp Tyr Cys Asn Trp Ile Gly Trp Lys His Phe
65                  70                  75                  80

Ala Gln Tyr Phe Pro Met Thr Leu Val Lys Glu Gly Glu Leu Asp Pro
                85                  90                  95

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

Gly Ala Leu Cys Ala Phe Gly Thr Glu Gly Leu His Phe Ser Lys Arg
        115                 120                 125

Phe Pro Gly Ile Lys Pro His Leu Leu Thr Ile His Ala Asn Phe Gln
    130                 135                 140

Ile Pro Leu Tyr Arg Asp Met Ile Met Ala His Gly Cys Ala Ser Val
145                 150                 155                 160

Ser Arg Ala Ser Cys Glu His Ile Leu Arg Ser Gly Glu Gly Ser Ser
                165                 170                 175

Val Val Ile Val Val Gly Gly Ala Gln Glu Ser Leu Ser Thr Gln Pro
            180                 185                 190

Gly Thr Leu Asn Leu Thr Leu Lys Lys Arg Leu Gly Phe Cys Lys Leu
        195                 200                 205

Ala Phe Val Asn Gly Ala Ser Leu Val Pro Thr Leu Ala Phe Gly Glu
    210                 215                 220

Asn Glu Leu Tyr Glu Val Tyr His Thr Lys Pro Thr Ser Leu Ile Tyr
225                 230                 235                 240

Lys Leu Gln Gln Leu Thr Lys Arg Thr Ile Gly Phe Thr Met Pro Val
                245                 250                 255

Phe Asn Gly Arg Gly Ile Phe Asn Tyr Glu Phe Gly Leu Leu Pro Arg
            260                 265                 270
```

Arg Lys Pro Val Tyr Ile Val Ile Gly Asn Pro Ile His Val Asp Lys
275 280 285

Val Glu Asn Pro Thr Ile Glu Gln Met Gln Lys Leu Gln Ser Ile Tyr
290 295 300

Ile Asp Glu Val Leu Asn Ile Trp Glu Arg Tyr Lys Asp Lys Tyr Ala
305 310 315 320

Ala Gly Arg Thr Gln Glu Leu Cys Ile Ile Glu
325 330

<210> 138
<211> 2151
<212> DNA
<213> Mortierella alpina

<400> 138
atggcccttc agatctacga ctttgtgtcg ttcttcttca ctatcctgct cgacatcttc 60

ttcagggaga ttcgtcccag aggcgcacac aagattccac aaaaaggccc cgtgatcttt 120

gtcgccgctc ctcatgccaa tcagtttgtc gatcctctcg tcttgatgcg agagtgcggc 180

cgcagagtct cattccttgc agccaaaaag tccatggacc gccggtggat tggtgcaatg 240

gcacgctcga tgaatgcgat tcctgttgag cgcccccagg accttgctaa agccggctcg 300

ggaatcatca aacttctgga tcgttatggc gaccctcttc gggtcaccgg tgtcggcact 360

aaattcacaa aggagctgct tgtgggcgac cagatctccc ttccaaagga cgtcggtgtc 420

ycagctgtgg gcgagatcat atctgatacc gagctgattg tcaagaagga attcaaagag 480

ctcaaggccc ttgagttgct gaccagtgct gaaggaacca agtacaaatg cctaccccat 540

atggaccaga cgaacgtcta caaaactgtc tttgagcgcc ttaacgcagg acattgcgtt 600

ggcattttcc ccgagggagg ctcccacgat cgtgctgaaa tgctgccatt gaaagctgga 660

gtcaccatca tggccctggg cgcattagcc gccaaccctt ccttggatct caagattgtc 720

acctgcggcc tcaactactt tcacccgcat cgcttccgct cgcgtgcagt agtcgagttt 780

ggcgagccat tgacggtttc gcctgagctg gtcgaaatgt acaagcgagg cggggcggaa 840

aagcgagagg cttgcggaaa actgctcgac acgatctatg aagctctccg cggtgtcact 900

ctcaacgcgc ctgattacga gacattgatg gtcattcaag cggcccgtcg cctttacaag 960

cctactcatc gcaagctgca gatctcgcag gtcgttgagt tgaaccgtag gttcgtcgca 1020

ggatacatgc acttcaagga caatccaaaa gtcattgaag ccaaggataa agtcatgcat 1080

tacaacactc agctgcggta ccatggattg cgggatcacc aagtgaatat tcgcaccacc 1140

aggaagcacg ctatcggcat gctcatctct cggctgatcc agatgatctt tttgagttgt 1200

ctggcgctac ctggaactct gatgaatctt ccggtcgcca ttgtcgctcg tgtcatcagc 1260

aacaaaaagg ccaaagaggc gctggctgcc tcgacagtca aaattgctgg aagggacgtc 1320

ttggctacgt ggaagttgct ggtcgctcta ggattgatgc ctgttctgta cttcacgtac 1380

tccgtcatgg tcttcatcta ttgcagccgc ttcgacctat cgttcaagtc gcgtcttttg 1440

gttgcttggg cagcatgggc gcttattcct tttgtaacct acgcaagcat ccgctttggt 1500

gaagttggta tcgatatctt caagtctatt cgcccattgt tcctgtccat catcccgggc 1560

gaggagagca cgatcaacga cctgcgcaag gcgcgcgcag aacttcaaaa gaccatcacc 1620

aatctcatca atgagctagc gccgcagata tatcccgact ttgattcaaa gcgcatcctc 1680

gatccatccc ctgcagatcg tcccagccgc tcagcatcag gcaccaacct tgcacagacg 1740

attttcaaca cggctgctca gcccttgaac caatggctag gcaaggatgg ccgctttgaa 1800

tgggagcgta ccgaggactc tgatgcagat gacgtgttct tctttttgga cccagcgaga 1860

ggaattatgg ggcggtctag ggcgtcgtct tggggaggtg gagcatttac gcctgctgtt 1920

gatgggtcgc gatcccggaa tcggagcagg acaagcagct tcacgtcggg ccagatccag 1980

ctgggcgagg gcttcaaact cgaggcactg acggagctgc cgcgggacaa ccctttgca 2040

gaagtgacca ggcggttgag tgtgagccga atgcagagat atgggctgga gggtatgacg 2100

cgctcggata cggatgaaaa cgaaggcccc gccaagtcaa aagacattta g 2151

<210> 139
<211> 716
<212> PRT
<213> Mortierella alpina

<220>

<221> misc_feature
<222> (141)..(141)
<223> Xaa can be any naturally occurring amino acid

<400> 139

Met Ala Leu Gln Ile Tyr Asp Phe Val Ser Phe Phe Phe Thr Ile Leu
1 5 10 15

Leu Asp Ile Phe Phe Arg Glu Ile Arg Pro Arg Gly Ala His Lys Ile
20 25 30

Pro Gln Lys Gly Pro Val Ile Phe Val Ala Ala Pro His Ala Asn Gln
35 40 45

Phe Val Asp Pro Leu Val Leu Met Arg Glu Cys Gly Arg Arg Val Ser
50 55 60

Phe Leu Ala Ala Lys Lys Ser Met Asp Arg Arg Trp Ile Gly Ala Met
65 70 75 80

Ala Arg Ser Met Asn Ala Ile Pro Val Glu Arg Pro Gln Asp Leu Ala
85 90 95

Lys Ala Gly Ser Gly Ile Ile Lys Leu Leu Asp Arg Tyr Gly Asp Pro
100 105 110

Leu Arg Val Thr Gly Val Gly Thr Lys Phe Thr Lys Glu Leu Leu Val
115 120 125

Gly Asp Gln Ile Ser Leu Pro Lys Asp Val Gly Val Xaa Ala Val Gly
130 135 140

Glu Ile Ile Ser Asp Thr Glu Leu Ile Val Lys Lys Glu Phe Lys Glu
145 150 155 160

Leu Lys Ala Leu Glu Leu Leu Thr Ser Ala Glu Gly Thr Lys Tyr Lys
165 170 175

Cys Leu Pro His Met Asp Gln Thr Asn Val Tyr Lys Thr Val Phe Glu
180 185 190

```
Arg Leu Asn Ala Gly His Cys Val Gly Ile Phe Pro Glu Gly Gly Ser
        195                 200                 205

His Asp Arg Ala Glu Met Leu Pro Leu Lys Ala Gly Val Thr Ile Met
    210                 215                 220

Ala Leu Gly Ala Leu Ala Ala Asn Pro Ser Leu Asp Leu Lys Ile Val
225                 230                 235                 240

Thr Cys Gly Leu Asn Tyr Phe His Pro His Arg Phe Arg Ser Arg Ala
                245                 250                 255

Val Val Glu Phe Gly Glu Pro Leu Thr Val Ser Pro Glu Leu Val Glu
            260                 265                 270

Met Tyr Lys Arg Gly Gly Ala Glu Lys Arg Glu Ala Cys Gly Lys Leu
        275                 280                 285

Leu Asp Thr Ile Tyr Glu Ala Leu Arg Gly Val Thr Leu Asn Ala Pro
    290                 295                 300

Asp Tyr Glu Thr Leu Met Val Ile Gln Ala Ala Arg Arg Leu Tyr Lys
305                 310                 315                 320

Pro Thr His Arg Lys Leu Gln Ile Ser Gln Val Val Glu Leu Asn Arg
                325                 330                 335

Arg Phe Val Ala Gly Tyr Met His Phe Lys Asp Asn Pro Lys Val Ile
            340                 345                 350

Glu Ala Lys Asp Lys Val Met His Tyr Asn Thr Gln Leu Arg Tyr His
        355                 360                 365

Gly Leu Arg Asp His Gln Val Asn Ile Arg Thr Thr Arg Lys His Ala
    370                 375                 380

Ile Gly Met Leu Ile Ser Arg Leu Ile Gln Met Ile Phe Leu Ser Cys
385                 390                 395                 400
```

```
Leu Ala Leu Pro Gly Thr Leu Met Asn Leu Pro Val Ala Ile Val Ala
                405                 410                 415

Arg Val Ile Ser Asn Lys Lys Ala Lys Glu Ala Leu Ala Ala Ser Thr
            420                 425                 430

Val Lys Ile Ala Gly Arg Asp Val Leu Ala Thr Trp Lys Leu Leu Val
        435                 440                 445

Ala Leu Gly Leu Met Pro Val Leu Tyr Phe Thr Tyr Ser Val Met Val
    450                 455                 460

Phe Ile Tyr Cys Ser Arg Phe Asp Leu Ser Phe Lys Ser Arg Leu Leu
465                 470                 475                 480

Val Ala Trp Ala Ala Trp Ala Leu Ile Pro Phe Val Thr Tyr Ala Ser
                485                 490                 495

Ile Arg Phe Gly Glu Val Gly Ile Asp Ile Phe Lys Ser Ile Arg Pro
            500                 505                 510

Leu Phe Leu Ser Ile Ile Pro Gly Glu Glu Ser Thr Ile Asn Asp Leu
        515                 520                 525

Arg Lys Ala Arg Ala Glu Leu Gln Lys Thr Ile Thr Asn Leu Ile Asn
    530                 535                 540

Glu Leu Ala Pro Gln Ile Tyr Pro Asp Phe Asp Ser Lys Arg Ile Leu
545                 550                 555                 560

Asp Pro Ser Pro Ala Asp Arg Pro Ser Arg Ser Ala Ser Gly Thr Asn
                565                 570                 575

Leu Ala Gln Thr Ile Phe Asn Thr Ala Ala Gln Pro Leu Asn Gln Trp
            580                 585                 590

Leu Gly Lys Asp Gly Arg Phe Glu Trp Glu Arg Thr Glu Asp Ser Asp
        595                 600                 605
```

Ala Gly Asp Val Phe Phe Phe Leu Asp Pro Ala Arg Gly Ile Met Gly
610 615 620

Arg Ser Arg Ala Ser Ser Trp Gly Gly Gly Ala Phe Thr Pro Ala Val
625 630 635 640

Asp Gly Ser Arg Ser Arg Asn Arg Ser Arg Thr Ser Ser Phe Thr Ser
645 650 655

Gly Gln Ile Gln Leu Gly Glu Gly Phe Lys Leu Glu Ala Leu Thr Glu
660 665 670

Leu Pro Arg Asp Asn Pro Phe Ala Glu Val Thr Arg Arg Leu Ser Val
675 680 685

Ser Arg Met Gln Arg Tyr Gly Leu Glu Gly Met Thr Arg Ser Asp Thr
690 695 700

Asp Glu Asn Glu Gly Pro Ala Lys Ser Lys Asp Ile
705 710 715

<210> 140
<211> 1212
<212> DNA
<213> Mortierella alpina

<400> 140
ccgcatgcca atcagtttgt cgatcctctc gtcttgatgc gagagtgcgg ccgcagagtc 60

tcattccttg cagccaaaaa gtccatggac cgccggtgga ttggtgcaat ggcacgctcg 120

atgaatgcga ttcctgttga gcgtccccag gaccttgcta aagccggctc gggaatcatc 180

aaacttctgg atcgttatgg cgacctctt cgggtcaccg gtgtcggcac taaattcaca 240

aaggagctgc ttgtgggcga ccagatctcc cttccaaagg acgtcggtgt cycagctgtg 300

ggcgagatca tatctgatac cgagctgatt gtcaagaagg aattcaaaga gctcaaggcc 360

cttgagttgc tgaccagtgc tgaaggaacc aagtacaaat gcctacccca tatggaccag 420

acgaacgtct acaaaactgt ctttgagcgc cttaacgcag gacattgcgt tggcattttc 480

cccgagggag gctcccacga tcgtgctgaa atgctgccat tgaaagctgg agtcaccatc 540

atggccctgg gcgcattagc cgccaaccct tccttggatc tcaagattgt cacctgcggc 600

ctcaactact ttcacccgca tcgcttccgc tcgcgtgcag tagtcgagtt tggcgagcca 660

ttgacggttt cgcctgagct ggtcgaaatg tacaagcgag gcggggcgga aaagcgagag 720

gcttgcggaa aactgctcga cacgatctat gaagctctcc gcggtgtcac tctcaacgcg 780

cctgattacg agacattgat ggtcattcaa gcggcccgtc gcctttacaa gcctactcat 840

cgcaagctgc agatctcgca ggtcgttgag ttgaaccgta ggttcgtcgc aggatacatg 900

cacttcaagg acaatccaaa agtcattgaa gccaaggata aagtcatgca ttacaacact 960

cagctgcggt accatggatt gcgggatcac caagtgaata ttcgcaccac caggaagcac 1020

gctatcggca tgctcatctc tcggctgatc cagatgatct ttttgagttg tctggcgcta 1080

cctggaacyc tgatgaatct tccggtcgcc attgtcgctc gtgtcatcag caacaaaaag 1140

gccaaagagg cgctggctgc ctcgacagtc aaaattgctg gaagggacga catggcyaca 1200

tggaaaaggg cg 1212

<210> 141
<211> 3935
<212> DNA
<213> Mortierella alpina

<220>
<221> misc_feature
<222> (1051)..(1053)
<223> ATG translation initiation codon

<220>
<221> Intron
<222> (1195)..(1469)

<220>
<221> misc_feature
<222> (1379)..(1379)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (1381)..(1381)
<223> n is a, c, g, or t

<220>
<221> Intron
<222> (1585)..(1839)

<220>
<221> Intron
<222> (2795)..(2877)

<220>
<221> Intron
<222> (2940)..(3038)

<220>
<221> misc_feature
<222> (3911)..(3913)
<223> TAG stop codon

<400> 141
actatagggc acgcgtggtc gacggcccgg gctggtcctt gaccctcgtt ctatgaatat 60

gcattacgac aaaagaaaat gaaaaaaaaa gagcacgcga gcaggagcgc ggttgtcaat 120

tttcaagcct acctcactcc tgtcatgtcg cgccattttc tttctttttc tcgtctcacg 180

atctctttga tcgaccagag tcgaatagaa tatgaacatg atcaacctaa ttgcgcgaga 240

gtgagccatt attgactatc agaccgggga gcatgcagcc tgatgcagca aaaatttgag 300

ctgtgcacaa tatttgtcat acataaaagt cggcccacat gtccaacagg gcggggtggc 360

ctgaacaccc aacgtgagaa tatggacctg actgctcaca ccatcgatac cgaggattgc 420

cgcatgcact cagctgccgt agacacgcac catataccat gtgataacgg atctggctgt 480

ggtgtaccca ggggcgatat acgcgcgcgt ccctgcggac ttgacgtgtt aacgactctg 540

tatgtggctc tgatgatgca tcgggtgcat tgaggctacg gccgaggcgg caatcgagac 600

gatgccattc tctcttttgg acggagaaag agaaggagga aggaaagcga ggccgaggga 660

cggaggaacg aggaagagag agcgagaggw agcgagagcg aaggcgacgg gctgtccaat 720

ggagcgcgta gagggtgatt tgggttcgct ctgccatact ttattcgcct gcatattcac 780

cacccgctgc atctttgcag tgtttgccct cttccttttc cttttccttt tctttttttct 840

tttttcgtct ttctgcccct catccgcatc tcccctgtcg cccattacct gctcgctttc 900

caccctgtt tgcacaccta ccccagccca atcctgtcct agcgcactct ctctgcccgt 960

ccgcacgttt cttctatggc cctctaggcg ccttttttgct gcccattctc tattgtccac 1020

ttacccacgc aggcacggcc gagtccagtc atggcccttc agatctacga ctttgtgtcg 1080

ttcttcttca ctatcctgct cgacatcttc ttcagggaga ttcgtcccag aggcgcacac 1140

aagattccac aaaaaggccc cgtgatcttt gtcgccgctc ctcatgccaa tcaggtacgt 1200

gcacgcaggc gttgttttct cggaccctac ctcgtcaaat agctgtacac gctggaatga 1260

gagcgcataa tcgcctggac acgggtgcag agttcatgca cgataagccg aacagcggga 1320

gccgacactc caaaagacct caggactagg ggaaaaaaaa aaaaaaaaaa aaaaaaaana 1380

nctggactcg attagcccta ttcccaaggg tgggggttcg tcattggttg tcgcaaacat 1440

cccaacttac agcactcttg tcctcatagt ttgtcgatcc tctcgtcttg atgcgagagt 1500

gcggccgcag agtctcattc cttgcagcca aaaagtccat ggaccgccgg tggattggtg 1560

caatggcacg ctcgatgaat gcgagtaagt tgctaggatt cccaccctcc cggctttatt 1620

ttgcagctct ctcatacacg tacacacaca cacacaaaga agattctcga tggtcactgg 1680

atgctttggt ggtgctgtcg aaagggctgc agactcttgc tgtgtggtga tagagtgccg 1740

gctttgatcc cccatcgacc gtttgggtcc ctcaccaaga catcagccct gagagcagat 1800

taattgatct gatactgttg aatcgttttc tacccatagt tcctgttgag cgcccccagg 1860

accttgctaa agccggctcg ggaatcatca aacttctgga tcgttatggc gaccctcttc 1920

gggtcaccgg tgtcggcact aaattcacaa aggagctgct tgtgggcgac cagatctccc 1980

ttccaaagga cgtcggtgtc ycagctgtgg gcgagatcat atctgatacc gagctgattg 2040

tcaagaagga attcaaagag ctcaaggccc ttgagttgct gaccagtgct gaaggaacca 2100

agtacaaatg cctaccccat atggaccaga cgaacgtcta caaaactgtc tttgagcgcc 2160

ttaacgcagg acattgcgtt ggcattttcc ccgagggagg ctcccacgat cgtgctgaaa 2220

tgctgccatt gaaagctgga gtcaccatca tggccctggg cgcattagcc gccaaccctt 2280

ccttggatct caagattgtc acctgcggcc tcaactactt tcacccgcat cgcttccgct 2340

cgcgtgcagt agtcgagttt ggcgagccat tgacggtttc gcctgagctg gtcgaaatgt 2400

acaagcgagg cggggcggaa aagcgagagg cttgcggaaa actgctcgac acgatctatg 2460

aagctctccg cggtgtcact ctcaacgcgc ctgattacga gacattgatg gtcattcaag 2520

cggcccgtcg cctttacaag cctactcatc gcaagctgca gatctcgcag gtcgttgagt 2580

tgaaccgtag gttcgtcgca ggatacatgc acttcaagga caatccaaaa gtcattgaag 2640

ccaaggataa agtcatgcat tacaacactc agctgcggta ccatggattg cgggatcacc 2700

aagtgaatat tcgcaccacc aggaagcacg ctatcggcat gctcatctct cggctgatcc 2760

agatgatctt tttgagttgt ctggcgctac ctgggtaaga gcatttttct atatcgacaa 2820

gggtctctag tgagttggct acggaatggt cactaacgca tgcttttgcc aatacagaac 2880

tctgatgaat cttccggtcg ccattgtcgc tcgtgtcatc agcaacaaaa aggccaaagg 2940

tatacttttt ttttttttt ttttttttt tcgcctttgt gttttgtgct cgacctgtga 3000

aaactaattt atttcttcct tcgttctgcc accggtagag gcgctggctg cctcgacagt 3060

caaaattgct ggaagggacg tcttggctac gtggaagttg ctggtcgctc taggattgat 3120

gcctgttctg tacttcacgt actccgtcat ggtcttcatc tattgcagcc gcttcgacct 3180

atcgttcaag tcgcgtcttt tggttgcttg ggcagcatgg gcgcttattc cttttgtaac 3240

ctacgcaagc atccgctttg gtgaagttgg tatcgatatc ttcaagtcta ttcgcccatt 3300

gttcctgtcc atcatcccgg gcgaggagag cacgatcaac gacctgcgca aggcgcgcgc 3360

agaacttcaa aagaccatca ccaatctcat caatgagcta gcgccgcaga tatatcccga 3420

ctttgattca aagcgcatcc tcgatccatc ccctgcagat cgtcccagcc gctcagcatc 3480

aggcaccaac cttgcacaga cgattttcaa cacggctgct cagcccttga accaatggct 3540

aggcaaggat ggccgctttg aatgggagcg taccgaggac tctgatgcag atgacgtgtt 3600

cttctttttg gacccagcga gaggaattat ggggcggtct agggcgtcgt cttggggagg 3660

tggagcattt acgcctgctg ttgatgggtc gcgatcccgg aatcggagca ggacaagcag 3720

cttcacgtcg ggccagatcc agctgggcga gggcttcaaa ctcgaggcac tgacggagct 3780

gccgcgggac aacccttttg cagaagtgac caggcggttg agtgtgagcc gaatgcagag 3840

atatgggctg gagggtatga cgcgctcgga tacggatgaa aacgaaggcc ccgccaagtc 3900

aaaagacatt taggaataaa cgcccgtatt tcccc 3935

<210> 142
<211> 965
<212> DNA
<213> Mortierella alpina

<400> 142

cattgtcgct cgtgtcatca gcaacaaaaa ggccaaagag gcgctggctg cctcgacagt 60

caaaattgct ggaagggacg tcttggctac gtggaagttg ctggtcgctc taggattgat 120

gcctgttctg tacttcacgt actccgtcat ggtcttcatc tattgcagcc gcttcgacct 180

atcgttcaag tcgcgtcttt tggttgcttg ggcagcatgg gcgcttattc cttttgtaac 240

ctacgcaagc atccgctttg gtgaagttgg tatcgatatc ttcaagtcta ttcgcccatt 300

gttcctgtcc atcatcccgg gcgaggagag cacgattaac gacctgcgca aggcgcgcgc 360

agaacttcaa aagaccatca ccaatctcat caatgagcta gcgccgcaga tatatcccga 420

ctttgattca aagcgcatcc tcgatccatc ccctgcagat cgtcccagcc gctcagcatc 480

aggcaccaac cttgcacaga cgattttcaa cacggctgct cagcccttga accaatggct 540

aggcaaggat ggccgctttg aatgggagcg taccgaggac tctgatgcag gtgacgtgtt 600

cttctttttg gacccagcga gaggaattat ggggcggtct agggcgtcgt cttggggagg 660

tggagcattt acccctgctg ttgatgggtc gcgatcccgg aatcggagca ggacaagcag 720

cttcacgtcg ggccagatcc agctgggcga gggcttcaaa ctcgaggcac tgacggagct 780

gccgcgggac aacccttttg cagaagtgac caggcggttg agtgtgagcc gaatgcagag 840

atatgggctg gagggtatga cgcgctcgga tacggatgaa aacgaaggcc ccgccaagtc 900

aaaagacatt taggaataaa cgcccgtatt tccccaaaaa aaaaaaaaaa aaaaaaaaaa 960

aaaaa 965

<210> 143
<211> 1908
<212> DNA
<213> Mortierella alpina

<220>
<221> misc_feature
<222> (1051)..(1053)
<223> ATG translation initiation codon

<220>
<221> Intron
<222> (1195)..(1469)

<220>
<221> misc_feature
<222> (1379)..(1379)

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

<220>
<221> misc_feature
<222> (1381)..(1381)
<223> n is a, c, g, or t

<220>
<221> Intron
<222> (1585)..(1839)

<400> 143
actatagggc acgcgtggtc gacggcccgg gctggtcctt gaccctcgtt ctatgaatat 60

gcattacgac aaaagaaaat gaaaaaaaaa gagcacgcga gcaggagcgc ggttgtcaat 120

tttcaagcct acctcactcc tgtcatgtcg cgccattttc tttctttttc tcgtctcacg 180

atctctttga tcgaccagag tcgaatagaa tatgaacatg atcaacctaa ttgcgcgaga 240

gtgagccatt attgactatc agaccgggga gcatgcagcc tgatgcagca aaaatttgag 300

ctgtgcacaa tatttgtcat acataaaagt cggcccacat gtccaacagg gcggggtggc 360

ctgaacaccc aacgtgagaa tatggacctg actgctcaca ccatcgatac cgaggattgc 420

cgcatgcact cagctgccgt agacacgcac catataccat gtgataacgg atctggctgt 480

ggtgtaccca ggggcgatat acgcgcgcgt ccctgcggac ttgacgtgtt aacgactctg 540

tatgtggctc tgatgatgca tcgggtgcat tgaggctacg gccgaggcgg caatcgagac 600

gatgccattc tctcttttgg acggagaaag agaaggagga aggaaagcga ggccgaggga 660

cggaggaacg aggaagagag agcgagaggw agcgagagcg aaggcgacgg gctgtccaat 720

ggagcgcgta gagggtgatt tgggttcgct ctgccatact ttattcgcct gcatattcac 780

cacccgctgc atctttgcag tgtttgccct cttccttttc cttttccttt tctttttttct 840

tttttcgtct ttctgcccct catccgcatc tcccctgtcg cccattacct gctcgctttc 900

caccccctgtt tgcacaccta ccccagccca atcctgtcct agcgcactct ctctgcccgt 960

ccgcacgttt cttctatggc cctctaggcg ccttttttgct gcccattctc tattgtccac 1020

ttacccacgc aggcacggcc gagtccagtc atggcccttc agatctacga ctttgtgtcg 1080

ttcttcttca ctatcctgct cgacatcttc ttcagggaga ttcgtcccag aggcgcacac 1140

aagattccac aaaaaggccc cgtgatcttt gtcgccgctc ctcatgccaa tcaggtacgt 1200

gcacgcaggc gttgttttct cggaccctac ctcgtcaaat agctgtacac gctggaatga 1260

gagcgcataa tcgcctggac acgggtgcag agttcatgca cgataagccg aacagcggga 1320

gccgacactc caaaagacct caggactagg ggaaaaaaaa aaaaaaaaaa aaaaaaaana 1380

nctggactcg attagcccta ttcccaaggg tgggggttcg tcattggttg tcgcaaacat 1440

cccaacttac agcactcttg tcctcatagt ttgtcgatcc tctcgtcttg atgcgagagt 1500

gcggccgcag agtctcattc cttgcagcca aaaagtccat ggaccgccgg tggattggtg 1560

caatggcacg ctcgatgaat gcgagtaagt tgctaggatt cccaccctcc cggctttatt 1620

ttgcagctct ctcatacacg tacacacaca cacacaaaga agattctcga tggtcactgg 1680

atgctttggt ggtgctgtcg aaagggctgc agactcttgc tgtgtggtga tagagtgccg 1740

gctttgatcc cccatcgacc gtttgggtcc ctcaccaaga catcagccct gagagcagat 1800

taattgatct gatactgttg aatcgttttc tacccatagt tcctgttgag cgccccccagg 1860

accttgctaa agccggctcg ggaatcatca aacttctgga tcgttatg 1908

<210> 144
<211> 966
<212> DNA
<213> Mortierella alpina

<220>
<221> Intron
<222> (19)..(101)

<220>
<221> Intron
<222> (164)..(262)

<400> 144
ttgtctggcg ctacctgggt aagagcattt ttctatatcg acaagggtct ctagtgagtt 60

ggctacggaa tggtcactaa cgcatgcttt tgccaataca gaactctgat gaatcttccg 120

gtcgccattg tcgctcgtgt catcagcaac aaaaaggcca aaggtatact ttttttttttt 180

ttttttttttt tttttcgcct ttgtgttttg tgctcgacct gtgaaaacta atttatttct 240

tccttcgttc tgccaccggt agaggcgctg gctgcctcga cagtcaaaat tgctggaagg 300

gacgtcttgg ctacgtggaa gttgctggtc gctctaggat tgatgcctgt tctgtacttc 360

acgtactccg tcatggtctt catctattgc agccgcttcg acctatcgtt caagtcgcgt 420

cttttggttg cttgggcagc atgggcgctt attccttttg taacctacgc aagcatccgc 480

tttggtgaag ttggtatcga tatcttcaag tctattcgcc cattgttcct gtccatcatc 540

ccgggcgagg agagcacgat caacgacctg cgcaaggcgc gcgcagaact tcaaaagacc 600

atcaccaatc tcatcaatga gctagcgccg cagatatatc ccgactttga ttcaaagcgc 660

atcctcgatc catcccctgc agatcgtccc agccgctcag catcaggcac caaccttgca 720

cagacgattt tcaacacggc tgctcagccc ttgaaccaat ggctaggcaa ggatggccgc 780

tttgaatggg agcgtaccga ggactctgat gcagatgacg tgttcttctt tttggaccca 840

gcgagaggaa ttatggggcg gtctagggcg tcgtcttggg gaggtggagc atttacgcct 900

gctgttgatg ggtcgcgatc ccggaatcgg agcaggacaa gcagcttcac gtcgggccag 960

atccag 966

<210> 145
<211> 275
<212> DNA
<213> Mortierella alpina

<220>
<221> Intron
<222> (1)..(275)

<220>
<221> Intron
<222> (1)..(275)

<220>
<221> misc_feature
<222> (185)..(185)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (187)..(187)
<223> n is a, c, g, or t

<400> 145
gtacgtgcac gcaggcgttg ttttctcgga ccctacctcg tcaaatagct gtacacgctg 60

gaatgagagc gcataatcgc ctggacacgg gtgcagagtt catgcacgat aagccgaaca 120

gcgggagccg acactccaaa agacctcagg actaggggaa aaaaaaaaaa aaaaaaaaaa 180

aaaananctg gactcgatta gccctattcc caagggtggg ggttcgtcat tggttgtcgc 240

aaacatccca acttacagca ctcttgtcct catag 275

<210> 146
<211> 255
<212> DNA
<213> Mortierella alpina

<220>
<221> Intron
<222> (1)..(255)

<220>
<221> Intron
<222> (1)..(255)

<400> 146
gtaagttgct aggattccca ccctcccggc tttattttgc agctctctca tacacgtaca 60

cacacacaca caaagaagat tctcgatggt cactggatgc tttggtggtg ctgtcgaaag 120

ggctgcagac tcttgctgtg tggtgataga gtgccggctt tgatccccca tcgaccgttt 180

gggtccctca ccaagacatc agccctgaga gcagattaat tgatctgata ctgttgaatc 240

gttttctacc catag 255

<210> 147
<211> 83
<212> DNA
<213> Mortierella alpina

<220>
<221> Intron
<222> (1)..(83)

<220>
<221> Intron
<222> (1)..(83)

<400> 147
gtaagagcat ttttctatat cgacaagggt ctctagtgag ttggctacgg aatggtcact 60

aacgcatgct tttgccaata cag 83

<210> 148
<211> 99
<212> DNA
<213> Mortierella alpina

<220>
<221> Intron
<222> (1)..(99)

<220>
<221> Intron
<222> (1)..(99)

<400> 148
gtatactttt ttttttttt ttttttttt ttcgcctttg tgttttgtgc tcgacctgtg 60

aaaactaatt tatttcttcc ttcgttctgc caccggtag 99

<210> 149
<211> 2133
<212> DNA
<213> Yarrowia lipolytica

<220>
<221> misc_feature
<223> CPT1

<400> 149
attctgttgc tcttgttgcc agctttcgtc tggtatattt ggatctcgat cttcatgtgt 60

gatgactcag ctggcaggag gtgaacaatt ttggagcggc caaatgcgcc acactgaatg 120

ccgcaagaat cggatgagaa ccgtagacat actggcatgt caattgaaac tgtgggtgtc 180

tactcgtgta ctcttctgat atataattat cgattttctt cttccagact ggaaatagct 240

cagattccca actgtccctc cacctacctc cctggttgtc tcacaaggct gaaacttgaa 300

ctgggccatc tttatcaagt ataaaaccca ttgtcacctt ctctgtcccc acagaccttc 360

caacggcttc tttttggccc cgcaagttca tcctgaagcc ccccacccac ctatcacgag 420

aataagcttt ctggcaacaa cccaaataaa tttatttggc aagcgaaaaa agtaaaatca 480

aacttgacca ttcaaccatc atgggcgtat tcattaaaca ggagcagctt ccggctctca 540

agaagtacaa gtgagtatgg atgtggcgaa gttggaagaa cggtcacgtg agctcgacga 600

gtggtgttga cgttcacttt gacagtcctg aggtggagat aatgacgtac taccaatagt 660

gtaggggatg aggacgtgtt ggggttgtgg cagcagatac atggtcataa tatatggctg 720

tctgttatct aagacaagtg ccggatcgat cgcgtggatg tatatcctaa tctacgccag 780

tggttcatgc taacccaggt actccgccga ggatcactcg ttcatctcca acaacattct 840

gcgccccttc tggcgacagt ttgtcaaaat cttccctctg tggatggccc ccaacatggt 900

gactctgttg ggcttcttct ttgtcattgt gaacttcatc accatgctca ttgttgatcc 960

cacccacgac cgcgagcctc ccagatgggt ctacctcacc tacgctctgg gtctgttcct 1020

ttaccagaca tttgatgcct gtgacggatc ccatgcccga cgaactggcc agagtggacc 1080

ccttggagag ctgtttgacc actgtgtcga cgccatgaat acctctctga ttctcacggt 1140

ggtggtgtcc accacccata tgggatataa catgaagctg ctgattgtgc agattgccgc 1200

tctcggaaac ttctacctgt cgacctggga gacctaccat accggaactc tgtacctttc 1260

tggcttctct ggtcctgttg aaggtatctt gattctggtg gctcttttcg tcctcacctt 1320

cttcactggt cccaacgtgt acgctctgac cgtctacgag gctcttcccg aatccatcac 1380

ttcgctgctg cctgccagct tcctggacgt caccatcacc cagatctaca ttggattcgg 1440

agtgctgggc atggtgttca acatctacgg cgcctgcgga aacgtgatca agtactacaa 1500

caacaagggc aagagcgctc tccccgccat tctcggaatc gccccctttg gcatcttcta 1560

cgtcggcgtc tttgcctggg cccatgttgc tcctctgctt ctctccaagt acgccatcgt 1620

ctatctgttt gccattgggg ctgcctttgc catgcaagtc ggccagatga ttcttgccca 1680

tctcgtgctt gctcccttcc cccactggaa cgtgctgctc ttcttccccc ttgtgggact 1740

ggcagtgcac tacattgcac ccgtgtttgg ctgggacgcc gatatcgtgt cggttaacac 1800

tctcttcacc tgttttggcg ccaccctctc catttacgcc ttctttgtgc ttgagatcat 1860

cgacgagatc accaactacc tcgatatctg gtgtctgcga atcaagtacc ctcaggagaa 1920

gaagactgag taatcgagtc ccgacacggt ccaaaagaca ttatatatag acagaaataa 1980

atttatacat tgactaacga gtgagtacag cagtgaatgg caagtcacca gccacaagac 2040

aagttcaagt aagtacaaga acagtgcgta tcagtagatg ctgtacttgc aacgtgcttc 2100

tctgtcaatc tcagtctctc attttgtctt gtc 2133

<210> 150
<211> 1185
<212> DNA
<213> Yarrowia lipolytica

<400> 150
atgggcgtat tcattaaaca ggagcagctt ccggctctca agaagtacaa gtactccgcc 60

gaggatcact cgttcatctc caacaacatt ctgcgcccct tctggcgaca gtttgtcaaa 120

atcttccctc tgtggatggc ccccaacatg gtgactctgt tgggcttctt ctttgtcatt 180

gtgaacttca tcaccatgct cattgttgat cccacccacg accgcgagcc tcccagatgg 240

gtctacctca cctacgctct gggtctgttc ctttaccaga catttgatgc ctgtgacgga 300

tcccatgccc gacgaactgg ccagagtgga ccccttggag agctgtttga ccactgtgtc 360

gacgccatga atacctctct gattctcacg gtggtggtgt ccaccaccca tatgggatat 420

aacatgaagc tgctgattgt gcagattgcc gctctcggaa acttctacct gtcgacctgg 480

gagacctacc ataccggaac tctgtacctt tctggcttct ctggtcctgt tgaaggtatc 540

ttgattctgg tggctctttt cgtcctcacc ttcttcactg gtcccaacgt gtacgctctg 600

accgtctacg aggctcttcc cgaatccatc acttcgctgc tgcctgccag cttcctggac 660

gtcaccatca cccagatcta cattggattc ggagtgctgg gcatggtgtt caacatctac 720

ggcgcctgcg gaaacgtgat caagtactac aacaacaagg gcaagagcgc tctccccgcc 780

attctcggaa tcgccccctt tggcatcttc tacgtcggcg tctttgcctg ggcccatgtt 840

gctcctctgc ttctctccaa gtacgccatc gtctatctgt ttgccattgg ggctgccttt 900

gccatgcaag tcggccagat gattcttgcc catctcgtgc ttgctccctt cccccactgg 960

aacgtgctgc tcttcttccc ctttgtggga ctggcagtgc actacattgc acccgtgttt 1020

ggctgggacg ccgatatcgt gtcggttaac actctcttca cctgttttgg cgccaccctc 1080

tccatttacg ccttctttgt gcttgagatc atcgacgaga tcaccaacta cctcgatatc 1140

tggtgtctgc gaatcaagta ccctcaggag aagaagactg agtaa 1185

<210> 151
<211> 394
<212> PRT
<213> Yarrowia lipolytica

<400> 151

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> 152
<211> 1434
<212> DNA

<213> Saccharomyces cerevisiae

<400> 152

atgtacaaca gaaaagacag agatgttcac gagaggaagg aagatggtca atctgagttt 60

gaagcactga acgggaccaa cgcaattatg tccgataata gtaaagcgta ttccataaag 120

tttctgacct tcaatacatg ggggttaaaa tacgtctcca aacaccgtaa agaaagactc 180

agagcaattg ctgataaatt ggcgggccac tcaatgctta cgccaatatc tgacgagttg 240

ttgcccaatg gtggagatag taatgaaaac gaagattacg acgtgattgc cttacaagaa 300

atctggtgtg tggaagactg gaagtatcta gcttctgcgt gtgcctccaa gtatccgtat 360

cagcgtttgt tccattctgg tattctgacg gggcctgggt tggccatact gtccaaggtc 420

ccgatagagt cgacctttct ttaccggttc ccgataaacg gtagaccgag tgcggtgttc 480

cgtggcgact ggtacgtagg gaaatctata gcaatcaccg tattgaacac aggaacccgc 540

cccattgcaa taatgaacag tcacatgcac gccccatacg ccaagcaggg tgatgccgcc 600

tacttgtgcc acagatcttg tcaggcctgg gatttcagca ggctcattaa gctttacagg 660

caggccggtt atgcggtgat tgtggtgggt gacttaaact ccagaccggg ctcactgccc 720

cacaaatttc tcacgcagga ggccggcctg gtcgactcct gggagcaatt gcatgggaag 780

caagacttgg cggtgatcgc tcgtctgtct ccattgcaac aattgcttaa gggctgtacc 840

acgtgcgatt cgctgctcaa cacatggagg gcccaaagac aacccgatga ggcatgcagg 900

ttggattatg ctcttatcga ccctgatttc ttgcaaacag tagacgcagg tgtcaggttc 960

actgaacgga tccctcacct ggactgcagt gtctctgacc attttgcata ctcatgcacc 1020

cttaacatcg tcccacaggg cacagagtcc cgtccatcca cctccgttaa gcgtgcgaag 1080

actcatgata gagagctgat cttgcagaga tactccaact acgaaaccat gatagaatgc 1140

atccacacgt acttgaagac agcccaaaga cagaaatttt tccgtggcct acatttctgg 1200

gcctcaatac ttctcctaat agcgtcgttg gtcgtgacaa cgtttactgc aaacaaggca 1260

ggctggtcct ccatcttctg ggtccttttc gctattgctg tctccatctc gggcaccatc 1320

gacggtgcca tctccttctt gtttggcagg tctgaaatca gagccctcat cgaagtcgaa 1380

caagaggttc tggacgcgga gcaccacctg caaactttct tgagcgagaa atga 1434

<210> 153
<211> 477
<212> PRT
<213> Saccharomyces cerevisiae (GenBank Accession No. NP_010935)

<400> 153

Met Tyr Asn Arg Lys Asp Arg Asp Val His Glu Arg Lys Glu Asp Gly
1 5 10 15

Gln Ser Glu Phe Glu Ala Leu Asn Gly Thr Asn Ala Ile Met Ser Asp
20 25 30

Asn Ser Lys Ala Tyr Ser Ile Lys Phe Leu Thr Phe Asn Thr Trp Gly
35 40 45

Leu Lys Tyr Val Ser Lys His Arg Lys Glu Arg Leu Arg Ala Ile Ala
50 55 60

Asp Lys Leu Ala Gly His Ser Met Leu Thr Pro Ile Ser Asp Glu Leu
65 70 75 80

Leu Pro Asn Gly Gly Asp Ser Asn Glu Asn Glu Asp Tyr Asp Val Ile
85 90 95

Ala Leu Gln Glu Ile Trp Cys Val Glu Asp Trp Lys Tyr Leu Ala Ser
100 105 110

Ala Cys Ala Ser Lys Tyr Pro Tyr Gln Arg Leu Phe His Ser Gly Ile
115 120 125

Leu Thr Gly Pro Gly Leu Ala Ile Leu Ser Lys Val Pro Ile Glu Ser
130 135 140

Thr Phe Leu Tyr Arg Phe Pro Ile Asn Gly Arg Pro Ser Ala Val Phe
145 150 155 160

Arg Gly Asp Trp Tyr Val Gly Lys Ser Ile Ala Ile Thr Val Leu Asn
165 170 175

Thr Gly Thr Arg Pro Ile Ala Ile Met Asn Ser His Met His Ala Pro
180 185 190

```
Tyr Ala Lys Gln Gly Asp Ala Ala Tyr Leu Cys His Arg Ser Cys Gln
        195                 200                 205

Ala Trp Asp Phe Ser Arg Leu Ile Lys Leu Tyr Arg Gln Ala Gly Tyr
    210                 215                 220

Ala Val Ile Val Val Gly Asp Leu Asn Ser Arg Pro Gly Ser Leu Pro
225                 230                 235                 240

His Lys Phe Leu Thr Gln Glu Ala Gly Leu Val Asp Ser Trp Glu Gln
                245                 250                 255

Leu His Gly Lys Gln Asp Leu Ala Val Ile Ala Arg Leu Ser Pro Leu
            260                 265                 270

Gln Gln Leu Leu Lys Gly Cys Thr Thr Cys Asp Ser Leu Leu Asn Thr
        275                 280                 285

Trp Arg Ala Gln Arg Gln Pro Asp Glu Ala Cys Arg Leu Asp Tyr Ala
    290                 295                 300

Leu Ile Asp Pro Asp Phe Leu Gln Thr Val Asp Ala Gly Val Arg Phe
305                 310                 315                 320

Thr Glu Arg Ile Pro His Leu Asp Cys Ser Val Ser Asp His Phe Ala
                325                 330                 335

Tyr Ser Cys Thr Leu Asn Ile Val Pro Gln Gly Thr Glu Ser Arg Pro
            340                 345                 350

Ser Thr Ser Val Lys Arg Ala Lys Thr His Asp Arg Glu Leu Ile Leu
        355                 360                 365

Gln Arg Tyr Ser Asn Tyr Glu Thr Met Ile Glu Cys Ile His Thr Tyr
    370                 375                 380

Leu Lys Thr Ala Gln Arg Gln Lys Phe Phe Arg Gly Leu His Phe Trp
385                 390                 395                 400
```

Ala Ser Ile Leu Leu Leu Ile Ala Ser Leu Val Val Thr Thr Phe Thr
405 410 415

Ala Asn Lys Ala Gly Trp Ser Ser Ile Phe Trp Val Leu Phe Ala Ile
420 425 430

Ala Val Ser Ile Ser Gly Thr Ile Asp Gly Ala Ile Ser Phe Leu Phe
435 440 445

Gly Arg Ser Glu Ile Arg Ala Leu Ile Glu Val Glu Gln Glu Val Leu
450 455 460

Asp Ala Glu His His Leu Gln Thr Phe Leu Ser Glu Lys
465 470 475

<210> 154
<211> 8953
<212> DNA
<213> Artificial Sequence

<220>
<223> Plasmid pY5-30

<400> 154
ggtggagctc cagcttttgt tccctttagt gagggttaat ttcgagcttg gcgtaatcat 60

ggtcatagct gtttcctgtg tgaaattgtt atccgctcac aattccacac aacatacgag 120

ccggaagcat aaagtgtaaa gcctggggtg cctaatgagt gagctaactc acattaattg 180

cgttgcgctc actgcccgct ttccagtcgg gaaacctgtc gtgccagctg cattaatgaa 240

tcggccaacg cgcggggaga ggcggtttgc gtattgggcg ctcttccgct tcctcgctca 300

ctgactcgct gcgctcggtc gttcggctgc ggcgagcggt atcagctcac tcaaaggcgg 360

taatacggtt atccacagaa tcaggggata acgcaggaaa gaacatgtga gcaaaaggcc 420

agcaaaaggc caggaaccgt aaaaaggccg cgttgctggc gtttttccat aggctccgcc 480

cccctgacga gcatcacaaa aatcgacgct caagtcagag gtggcgaaac ccgacaggac 540

tataaagata ccaggcgttt ccccctggaa gctccctcgt gcgctctcct gttccgaccc 600

tgccgcttac cggatacctg tccgcctttc tcccttcggg aagcgtggcg ctttctcata 660

gctcacgctg taggtatctc agttcggtgt aggtcgttcg ctccaagctg ggctgtgtgc 720

acgaaccccc cgttcagccc gaccgctgcg ccttatccgg taactatcgt cttgagtcca 780

acccggtaag acacgactta tcgccactgg cagcagccac tggtaacagg attagcagag 840

cgaggtatgt aggcggtgct acagagttct tgaagtggtg gcctaactac ggctacacta 900

gaaggacagt atttggtatc tgcgctctgc tgaagccagt taccttcgga aaaagagttg 960

gtagctcttg atccggcaaa caaaccaccg ctggtagcgg tggtttttt gtttgcaagc 1020

agcagattac gcgcagaaaa aaaggatctc aagaagatcc tttgatcttt tctacggggt 1080

ctgacgctca gtggaacgaa aactcacgtt aagggatttt ggtcatgaga ttatcaaaaa 1140

ggatcttcac ctagatcctt ttaaattaaa aatgaagttt taaatcaatc taaagtatat 1200

atgagtaaac ttggtctgac agttaccaat gcttaatcag tgaggcacct atctcagcga 1260

tctgtctatt tcgttcatcc atagttgcct gactccccgt cgtgtagata actacgatac 1320

gggagggctt accatctggc cccagtgctg caatgatacc gcgagaccca cgctcaccgg 1380

ctccagattt atcagcaata aaccagccag ccggaagggc cgagcgcaga agtggtcctg 1440

caactttatc cgcctccatc cagtctatta attgttgccg ggaagctaga gtaagtagtt 1500

cgccagttaa tagtttgcgc aacgttgttg ccattgctac aggcatcgtg gtgtcacgct 1560

cgtcgtttgg tatggcttca ttcagctccg gttcccaacg atcaaggcga gttacatgat 1620

cccccatgtt gtgcaaaaaa gcggttagct ccttcggtcc tccgatcgtt gtcagaagta 1680

agttggccgc agtgttatca ctcatggtta tggcagcact gcataattct cttactgtca 1740

tgccatccgt aagatgcttt tctgtgactg gtgagtactc aaccaagtca ttctgagaat 1800

agtgtatgcg gcgaccgagt tgctcttgcc cggcgtcaat acgggataat accgcgccac 1860

atagcagaac tttaaaagtg ctcatcattg gaaaacgttc ttcggggcga aaactctcaa 1920

ggatcttacc gctgttgaga tccagttcga tgtaacccac tcgtgcaccc aactgatctt 1980

cagcatcttt tactttcacc agcgtttctg ggtgagcaaa aacaggaagg caaaatgccg 2040

caaaaaaggg aataagggcg acacggaaat gttgaatact catactcttc ctttttcaat 2100

attattgaag catttatcag ggttattgtc tcatgagcgg atacatattt gaatgtattt 2160

agaaaaataa acaaataggg gttccgcgca catttccccg aaaagtgcca cctgacgcgc 2220

cctgtagcgg cgcattaagc gcggcgggtg tggtggttac gcgcagcgtg accgctacac 2280

ttgccagcgc cctagcgccc gctcctttcg ctttcttccc ttcctttctc gccacgttcg 2340

ccggctttcc ccgtcaagct ctaaatcggg ggctcccttt agggttccga tttagtgctt 2400

tacggcacct cgaccccaaa aaacttgatt agggtgatgg ttcacgtagt gggccatcgc 2460

cctgatagac ggtttttcgc cctttgacgt tggagtccac gttctttaat agtggactct 2520

tgttccaaac tggaacaaca ctcaacccta tctcggtcta ttcttttgat ttataaggga 2580

ttttgccgat ttcggcctat tggttaaaaa atgagctgat ttaacaaaaa tttaacgcga 2640

attttaacaa aatattaacg cttacaattt ccattcgcca ttcaggctgc gcaactgttg 2700

ggaagggcga tcggtgcggg cctcttcgct attacgccag ctggcgaaag ggggatgtgc 2760

tgcaaggcga ttaagttggg taacgccagg gttttcccag tcacgacgtt gtaaaacgac 2820

ggccagtgaa ttgtaatacg actcactata gggcgaattg ggtaccgggc ccccccctcga 2880

ggtcgatggt gtcgataagc ttgatatcga attcatgtca cacaaaccga tcttcgcctc 2940

aaggaaacct aattctacat ccgagagact gccgagatcc agtctacact gattaatttt 3000

cgggccaata atttaaaaaa atcgtgttat ataatattat atgtattata tatatacatc 3060

atgatgatac tgacagtcat gtcccattgc taaatagaca gactccatct gccgcctcca 3120

actgatgttc tcaatattta aggggtcatc tcgcattgtt taataataaa cagactccat 3180

ctaccgcctc caaatgatgt tctcaaaata tattgtatga acttatttt attacttagt 3240

attattagac aacttacttg ctttatgaaa aacacttcct atttaggaaa caatttataa 3300

tggcagttcg ttcatttaac aatttatgta gaataaatgt tataaatgcg tatgggaaat 3360

cttaaatatg gatagcataa atgatatctg cattgcctaa ttcgaaatca acagcaacga 3420

aaaaaatccc ttgtacaaca taaatagtca tcgagaaata tcaactatca aagaacagct 3480

attcacacgt tactattgag attattattg gacgagaatc acacactcaa ctgtctttct 3540

ctcttctaga aatacaggta caagtatgta ctattctcat tgttcatact tctagtcatt 3600

tcatcccaca tattccttgg atttctctcc aatgaatgac attctatctt gcaaattcaa 3660

caattataat aagatatacc aaagtagcgg tatagtggca atcaaaaagc ttctctggtg 3720

tgcttctcgt atttatttt attctaatga tccattaaag gtatatattt atttcttgtt 3780

atataatcct tttgtttatt acatgggctg gatacataaa ggtattttga tttaattttt 3840

tgcttaaatt caatcccccc tcgttcagtg tcaactgtaa tggtaggaaa ttaccatact 3900

tttgaagaag caaaaaaaat gaaagaaaaa aaaaatcgta tttccaggtt agacgttccg 3960

cagaatctag aatgcggtat gcggtacatt gttcttcgaa cgtaaaagtt gcgctccctg 4020

agatattgta cattttgct tttacaagta caagtacatc gtacaactat gtactactgt 4080

tgatgcatcc acaacagttt gttttgtttt tttttgtttt tttttttct aatgattcat 4140

taccgctatg tatacctact tgtacttgta gtaagccggg ttattggcgt tcaattaatc 4200

atagacttat gaatctgcac ggtgtgcgct gcgagttact tttagcttat gcatgctact 4260

tgggtgtaat attgggatct gttcggaaat caacggatgc tcaaccgatt tcgacagtaa 4320

taatttgaat cgaatcggag cctaaaatga acccgagtat atctcataaa attctcggtg 4380

agaggtctgt gactgtcagt acaaggtgcc ttcattatgc cctcaacctt accatacctc 4440

actgaatgta gtgtacctct aaaaatgaaa tacagtgcca aaagccaagg cactgagctc 4500

gtctaacgga cttgatatac aaccaattaa aacaaatgaa aagaaataca gttctttgta 4560

tcatttgtaa caattaccct gtacaaacta aggtattgaa atcccacaat attcccaaag 4620

tccacccctt tccaaattgt catgcctaca actcatatac caagcactaa cctaccaaac 4680

accactaaaa ccccacaaaa tatatcttac cgaatataca gtaacaagct accaccacac 4740

tcgttgggtg cagtcgccag cttaaagata tctatccaca tcagccacaa ctcccttcct 4800

ttaataaacc gactacaccc ttggctattg aggttatgag tgaatatact gtagacaaga 4860

cactttcaag aagactgttt ccaaaacgta ccactgtcct ccactacaaa cacacccaat 4920

ctgcttcttc tagtcaaggt tgctacaccg gtaaattata aatcatcatt tcattagcag 4980

ggcagggccc tttttataga gtcttataca ctagcggacc ctgccggtag accaacccgc 5040

aggcgcgtca gtttgctcct tccatcaatg cgtcgtagaa acgacttact ccttcttgag 5100

cagctccttg accttgttgg caacaagtct ccgacctcgg aggtggagga agagcctccg 5160

atatcggcgg tagtgatacc agcctcgacg gactccttga cggcagcctc aacagcgtca 5220

ccggcgggct tcatgttaag agagaacttg agcatcatgg cggcagacag aatggtggca 5280

atggggttga ccttctgctt gccgagatcg ggggcagatc cgtgacaggg ctcgtacaga 5340

ccgaacgcct cgttggtgtc gggcagagaa gccagagagg cggagggcag cagacccaga 5400

gaaccgggga tgacggaggc ctcgtcggag atgatatcgc caaacatgtt ggtggtgatg 5460

atgataccat tcatcttgga gggctgcttg atgaggatca tggcggccga gtcgatcagc 5520

tggtggttga gctcgagctg ggggaattcg tccttgagga ctcgagtgac agtctttcgc 5580

caaagtcgag aggaggccag cacgttggcc ttgtcaagag accacacggg aagaggggag 5640

ttgtgctgaa gggccaggaa ggcggccatt cgggcaattc gctcaacctc aggaacggag 5700

taggtctcgg tgtcggaagc gacgccagat ccgtcatcct cctttcgctc tccaaagtag 5760

atacctccga cgagctctcg gacaatgatg aagtcggtgc cctcaacgtt tcggatgggg 5820

gagagatcgg cgagcttggg cgacagcagc tggcagggtc gcaggttggc gtacaggttc 5880

aggtcctttc gcagcttgag gagaccctgc tcgggtcgca cgtcggttcg tccgtcggga 5940

gtggtccata cggtgttggc agcgcctccg acagcaccga gcataatagа gtcagccttt 6000

cggcagatgt cgagagtagc gtcggtgatg ggctcgccct ccttctcaat ggcagctcct 6060

ccaatgagtc ggtcctcaaa cacaaactcg gtgccggagg cctcagcaac agacttgagc 6120

accttgacgg cctcggcaat cacctcgggg ccacagaagt cgccgccgag aagaacaatc 6180

ttcttggagt cagtcttggt cttcttagtt tcgggttcca ttgtggatgt gtgtggttgt 6240

atgtgtgatg tggtgtgtgg agtgaaaatc tgtggctggc aaacgctctt gtatatatac 6300

gcacttttgc ccgtgctatg tggaagacta aacctccgaa gattgtgact caggtagtgc 6360

ggtatcggct agggacccaa accttgtcga tgccgatagc gctatcgaac gtaccccagc 6420

cggccgggag tatgtcggag gggacatacg agatcgtcaa gggtttgtgg ccaactggta 6480

aataaatgat gtcgactcag gcgacgacgg aattcctgca gcccatctgc agaattcagg 6540

agagaccggg ttggcggcgt atttgtgtcc caaaaaacag ccccaattgc cccaattgac 6600

cccaaattga cccagtagcg ggcccaaccc cggcgagagc ccccttcacc ccacatatca 6660

aacctccccc ggttcccaca cttgccgtta agggcgtagg gtactgcagt ctggaatcta 6720

cgcttgttca gactttgtac tagtttcttt gtctggccat ccgggtaacc catgccggac 6780

gcaaaataga ctactgaaaa ttttttttgct ttgtggttgg gactttagcc aagggtataa 6840

aagaccaccg tccccgaatt acctttcctc ttctttttctc tctctccttg tcaactcaca 6900

cccgaaatcg ttaagcattt ccttctgagt ataagaatca ttcaccatgg atggtacgtc 6960

ctgtagaaac cccaacccgt gaaatcaaaa aactcgacgg cctgtgggca ttcagtctgg 7020

atcgcgaaaa ctgtggaatt gatcagcgtt ggtgggaaag cgcgttacaa gaaagccggg 7080

caattgctgt gccaggcagt tttaacgatc agttcgccga tgcagatatt cgtaattatg 7140

cgggcaacgt ctggtatcag cgcgaagtct ttataccgaa aggttgggca ggccagcgta 7200

tcgtgctgcg tttcgatgcg gtcactcatt acggcaaagt gtgggtcaat aatcaggaag 7260

tgatggagca tcagggcggc tatacgccat ttgaagccga tgtcacgccg tatgttattg 7320

ccgggaaaag tgtacgtatc accgtttgtg tgaacaacga actgaactgg cagactatcc 7380

cgccgggaat ggtgattacc gacgaaaacg gcaagaaaaa gcagtcttac ttccatgatt 7440

tctttaacta tgccgggatc catcgcagcg taatgctcta caccacgccg aacacctggg 7500

tggacgatat caccgtggtg acgcatgtcg cgcaagactg taaccacgcg tctgttgact 7560

ggcaggtggt ggccaatggt gatgtcagcg ttgaactgcg tgatgcggat caacaggtgg 7620

ttgcaactgg acaaggcact agcgggactt tgcaagtggt gaatccgcac ctctggcaac 7680

cgggtgaagg ttatctctat gaactgtgcg tcacagccaa aagccagaca gagtgtgata 7740

tctacccgct tcgcgtcggc atccggtcag tggcagtgaa gggcgaacag ttcctgatta 7800

accacaaacc gttctacttt actggctttg gtcgtcatga agatgcggac ttacgtggca 7860

aaggattcga taacgtgctg atggtgcacg accacgcatt aatggactgg attggggcca 7920

actcctaccg tacctcgcat taccccttacg ctgaagagat gctcgactgg gcagatgaac 7980

atggcatcgt ggtgattgat gaaactgctg ctgtcggctt taacctctct ttaggcattg 8040

gtttcgaagc gggcaacaag ccgaaagaac tgtacagcga agaggcagtc aacggggaaa 8100

ctcagcaagc gcacttacag gcgattaaag agctgatagc gcgtgacaaa aaccacccaa 8160

gcgtggtgat gtggagtatt gccaacgaac cggatacccg tccgcaagtg cacgggaata 8220

tttcgccact ggcggaagca acgcgtaaac tcgacccgac gcgtccgatc acctgcgtca 8280

atgtaatgtt ctgcgacgct cacaccgata ccatcagcga tctctttgat gtgctgtgcc 8340

tgaaccgtta ttacggatgg tatgtccaaa gcggcgattt ggaaacggca gagaaggtac 8400

tggaaaaaga acttctggcc tggcaggaga aactgcatca gccgattatc atcaccgaat 8460

acggcgtgga tacgttagcc gggctgcact caatgtacac cgacatgtgg agtgaagagt 8520

atcagtgtgc atggctggat atgtatcacc gcgtctttga tcgcgtcagc gccgtcgtcg 8580

gtgaacaggt atggaatttc gccgattttg cgacctcgca aggcatattg cgcgttggcg 8640

gtaacaagaa agggatcttc actcgcgacc gcaaaccgaa gtcggcggct tttctgctgc 8700

aaaaacgctg gactggcatg aacttcggtg aaaaaccgca gcagggaggc aaacaatgat 8760

taattaacta gagcggccgc caccgcggcc cgagattccg gcctcttcgg ccgccaagcg 8820

acccgggtgg acgtctagag gtacctagca attaacagat agtttgccgg tgataattct 8880

cttaacctcc cacactcctt tgacataacg atttatgtaa cgaaactgaa atttgaccag 8940

atattgtgtc cgc 8953

<210> 155
<211> 10
<212> DNA
<213> Yarrowia lipolytica

<220>
<221> misc_feature
<222> (8)..(8)
<223> n is a, c, g, or t

<400> 155
mammatgnhs 10

<210> 156
<211> 12649
<212> DNA
<213> Artificial Sequence

<220>
<223> Plasmid pKUNF12T6E

<220>
<221> misc_feature
<222> (2507)..(2507)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (2512)..(2515)
<223> n is a, c, g, or t

<400> 156

taaccctcac taaagggaac aaaagctgga gctccaccgc ggacacaata tctggtcaaa 60

tttcagtttc gttacataaa tcgttatgtc aaaggagtgt gggaggttaa gagaattatc 120

accggcaaac tatctgttaa ttgctaggta cctctagacg tccacccggg tcgcttggcg 180

gccgaagagg ccggaatctc gggccgcggt ggcggccgct tagttggtct tggacttctt 240

gggcttcttc aggtaggact ggacaaagaa gttgccgaac agagcgagca gggtgatcat 300

gtacacgccg agcagctgga ccagagcctg agggtagtcg caggggaaga ggtagtcgta 360

cagggactgc accagcatag ccatgaactg ggtcatctgc agagtggtga tgtagggctt 420

gatgggcttg acgaagccga agccctgaga ggaaaagaag tagtaggcgt acatgacggt 480

gtggacgaag gagttgagga tgacggagaa gtaggcgtcg ccaccaggag cgtacttggc 540

aatagcccac cagatggcga agatggtggc atggtggtac acgtgcagga aggagacctg 600

gttgaacttc ttgcacagga tcatgatagc ggtgtccagg aactcgtagg ccttggagac 660

gtagaacacg tagacgattc gggacatgcc ctgagcgtgg gactcgttgc ccttctccat 720

gtcgttgccg aagaccttgt agccacccag gatagcctgt cggatggtct cgacgcacat 780

gtagagggac agtccgaaga ggaacaggtt gtggagcagc ttgatggtct tcagctcgaa 840

gggcttctcc atctgcttca tgatgggaat gccgaagagc agcatggcca tgtagccgac 900

ctcgaaggcg agcatggtgg agacgtccat catgggcaga ccgtcggtca gagcgtaggg 960

cttagctccg tccatccact ggtcgacacc ggtctcgact cgtccgacca cgtcgtccca 1020

gacagaggag ttggccatgg tgaatgattc ttatactcag aaggaaatgc ttaacgattt 1080

cgggtgtgag ttgacaagga gagagagaaa agaagaggaa aggtaattcg gggacggtgg 1140

tcttttatac ccttggctaa agtcccaacc acaaagcaaa aaaattttca gtagtctatt 1200

ttgcgtccgg catgggttac ccggatggcc agacaaagaa actagtacaa agtctgaaca 1260

agcgtagatt ccagactgca gtaccctacg cccttaacgg caagtgtggg aaccgggggа 1320

ggtttgatat gtggggtgaa gggggctctc gccggggttg ggcccgctac tgggtcaatt 1380

tggggtcaat tggggcaatt ggggctgttt tttgggacac aaatacgccg ccaacccggt 1440

ctctcctgaa ttctgcatcg atcgaggaag aggacaagcg gctgcttctt aagtttgtga 1500

catcagtatc caaggcacca ttgcaaggat tcaaggcttt gaacccgtca tttgccattc 1560

gtaacgctgg tagacaggtt gatcggttcc ctacggcctc cacctgtgtc aatcttctca 1620

agctgcctga ctatcaggac attgatcaac ttcggaagaa acttttgtat gccattcgat 1680

cacatgctgg tttcgatttg tcttagagga acgcatatac agtaatcata gagaataaac 1740

gatattcatt tattaaagta gatagttgag gtagaagttg taaagagtga taaatagcgg 1800

ccgcgcctac ttaagcaacg ggcttgataa cagcgggggg ggtgcccacg ttgttgcggt 1860

tgcggaagaa cagaacaccc ttaccagcac cctcggcacc agcgctgggc tcaacccact 1920

ggcacatacg cgcactgcgg tacatggcgc ggatgaagcc acgaggacca tcctggacat 1980

cagcccggta gtgcttgccc atgatgggct taatggcctc ggtggcctcg tccgcgttgt 2040

agaaggggat gctgctgacg tagtggtgga ggacatgagt ctcgatgatg ccgtggagaa 2100

ggtggcggcc gatgaagccc atctcacggt caatggtagc agcggcacca cggacgaagt 2160

tccactcgtc gttggtgtag tggggaaggg tagggtcggt gtgctggagg aaggtgatgg 2220

caacgagcca gtggttaacc cagaggtagg gaacaaagta ccagatggcc atgttgtaga 2280

aaccgaactt ctgaacgagg aagtacagag cagtggccat cagaccgata ccaatatcgc 2340

tgaggacgat gagcttagcg tcactgttct cgtacagagg gctgcgggga tcgaagtggt 2400

taacaccacc gccgaggccg ttatgcttgc ccttgccgcg accctcacgc tggcgctcgt 2460

ggtagttgtg gccggtaaca ttggtgatga ggtagttggg ccagccnacg annnnctcag 2520

taagatgagc gagctcgtgg gtcatctttc cgagacgagt agcctgctgc tcgcgggttc 2580

ggggaacgaa gaccatgtca cgctccatgt tgccagtggc cttgtggtgc tttcggtggg 2640

agatttgcca gctgaagtag gggacaagga gggaagagtg aagaacccag ccagtaatgt 2700

cgttgatgat gcgagaatcg gagaaagcac cgtgaccgca ctcatgggca ataacccaga 2760

gaccagtacc gaaaagaccc tgaagaacgg tgtacacggc ccacagacca gcgcgggcgg 2820

gggtggaggg gatatattcg ggggtcacaa agttgtacca gatgctgaaa gtggtagtca 2880

ggaggacaat gtcgcggagg atataaccgt atcccttgag agcggagcgc ttgaagcagt 2940

gcttagggat ggcattgtag atgtccttga tggtaaagtc gggaacctcg aactggttgc 3000

cgtaggtgtc gagcatgaca ccatactcgg acttgggctt ggcgatatca acctcggaca 3060

tggacgagag cgatgtggaa gaggccgagt ggcggggaga gtctgaagga gagacggcgg 3120

cagactcaga atccgtcaca gtagttgagg tgacggtgcg tctaagcgca gggttctgct 3180

tgggcagagc cgaagtggac gccatggaga gctgggttag tttgtgtaga gagtgtgtgt 3240

tgctagcgac tttcggattg tgtcattaca caaaacgcgt cgtctcgaca ctgatcttgt 3300

cgtggatact cacggctcgg acatcgtcgc cgacgatgac accggacttt cgcttaagga 3360

cgtcagtaac aggcattgtg tgatgtgtag tttagatttc gaatctgtgg ggaaagaaag 3420

gaaaaaagag actggcaacc gattgggaga gccactgttt atatataccc tagacaagcc 3480

ccccgcttgt aagatgttgg tcaatgtaaa ccagtattaa ggttggcaag tgcaggagaa 3540

gcaaggtgtg ggtaccgagc aatggaaatg tgcggaaggc aaaaaaatga ggccacggcc 3600

tattgtcggg gctatatcca gggggcgatt gaagtacact aacatgacat gtgtccacag 3660

accctcaatc tggcctgatg agccaaatcc atacgcgctt tcgcagctct aaaggctata 3720

acaagtcaca ccaccctgct cgacctcagc gccctcactt tttgttaaga caaactgtac 3780

acgctgttcc agcgttttct gcctgcacct ggtgggacat ttggtgcaac ctaaagtgct 3840

cggaacctct gtggtgtcca gatcagcgca gcagttccga ggtagttttg aggcccttag 3900

atgatgcaat ggtgtcagtc gctggatcac gagtcttaat ggcagtattc gttcttattt 3960

gtgccattga gccccgttat cctcgtatct tctacccccc atcccatccc tttgttggtg 4020

caaccctacc catttattgt tgggtgcagc ccaaccgacg tggagagctt ggcttggcca 4080

tataaaaagg ccccccccta gtggcaatgg cagaaagtca gctgtgagtt gttgaatttg 4140

tcatctaggc ggcctggccg tcttctccgg ggcaattgtt cctctatagt actgcgtaca 4200

ctgtttaaac agtgtacgca gatctgcgac gacggaattc ctgcagccca tctgcagaat 4260

tcaggagaga ccgggttggc ggcgtatttg tgtcccaaaa aacagcccca attgccccaa 4320

ttgaccccaa attgacccag tagcgggccc aaccccggcg agagcccccc tcaccccaca 4380

tatcaaacct cccccggttc ccacacttgc cgttaagggc gtagggtact gcagtctgga 4440

atctacgctt gttcagactt tgtactagtt tctttgtctg gccatccggg taacccatgc 4500

cggacgcaaa atagactact gaaaattttt ttgctttgtg gttgggactt tagccaaggg 4560

tataaaagac caccgtcccc gaattacctt tcctcttctt ttctctctct ccttgtcaac 4620

tcacacccga aatcgttaag catttccttc tgagtataag aatcattcac catggctgcc 4680

gctccctctg tgcgaacctt tacccgagcc gaggttctga acgctgaggc tctgaacgag 4740

ggcaagaagg acgctgaggc tcccttcctg atgatcatcg acaacaaggt gtacgacgtc 4800

cgagagttcg tccctgacca tcctggaggc tccgtgattc tcacccacgt tggcaaggac 4860

ggcaccgacg tctttgacac ctttcatccc gaggctgctt gggagactct cgccaacttc 4920

tacgttggag acattgacga gtccgaccga gacatcaaga acgatgactt tgccgctgag 4980

gtccgaaagc tgcgaaccct gttccagtct ctcggctact acgactcctc taaggcctac 5040

tacgccttca aggtctcctt caacctctgc atctggggac tgtccaccgt cattgtggcc 5100

aagtggggtc agacctccac cctcgccaac gtgctctctg ctgccctgct cggcctgttc 5160

tggcagcagt gcggatggct ggctcacgac tttctgcacc accaggtctt ccaggaccga 5220

ttctggggtg atctcttcgg agccttcctg ggaggtgtct gccagggctt ctcctcttcc 5280

tggtggaagg acaagcacaa cactcaccat gccgctccca acgtgcatgg cgaggatcct 5340

gacattgaca cccaccctct cctgacctgg tccgagcacg ctctggagat gttctccgac 5400

gtccccgatg aggagctgac ccgaatgtgg tctcgattca tggtcctgaa ccagacctgg 5460

ttctacttcc ccattctctc cttcgctcga ctgtcttggt gcctccagtc cattctcttt 5520

gtgctgccca acggtcaggc tcacaagccc tccggagctc gagtgcccat ctccctggtc 5580

gagcagctgt ccctcgccat gcactggacc tggtacctcg ctaccatgtt cctgttcatc 5640

aaggatcctg tcaacatgct cgtgtacttc ctggtgtctc aggctgtgtg cggaaacctg 5700

ctcgccatcg tgttctccct caaccacaac ggtatgcctg tgatctccaa ggaggaggct 5760

gtcgacatgg atttctttac caagcagatc atcactggtc gagatgtcca tcctggactg 5820

ttcgccaact ggttcaccgg tggcctgaac taccagatcg agcatcacct gttcccttcc 5880

atgcctcgac acaacttctc caagatccag cctgccgtcg agaccctgtg caagaagtac 5940

aacgtccgat accacaccac tggtatgatc gagggaactg ccgaggtctt ctcccgactg 6000

aacgaggtct ccaaggccac ctccaagatg ggcaaggctc agtaagcggc cgcatgagaa 6060

gataaatata taaatacatt gagatattaa atgcgctaga ttagagagcc tcatactgct 6120

cggagagaag ccaagacgag tactcaaagg ggattacacc atccatatcc acagacacaa 6180

gctggggaaa ggttctatat acactttccg gaataccgta gtttccgatg ttatcaatgg 6240

gggcagccag gatttcaggc acttcggtgt ctcggggtga aatggcgttc ttggcctcca 6300

tcaagtcgta ccatgtcttc atttgcctgt caaagtaaaa cagaagcaga tgaagaatga 6360

acttgaagtg aaggaattta aattgccccg gagaagacgg ccaggccgcc tagatgacaa 6420

attcaacaac tcacagctga ctttctgcca ttgccactag ggggggggcct ttttatatgg 6480

ccaagccaag ctctccacgt cggttgggct gcacccaaca ataaatgggt agggttgcac 6540

caacaaaggg atgggatggg gggtagaaga tacgaggata acggggctca atggcacaaa 6600

taagaacgaa tactgccatt aagactcgtg atccagcgac tgacaccatt gcatcatcta 6660

agggcctcaa aactacctcg gaactgctgc gctgatctgg acaccacaga ggttccgagc 6720

actttaggtt gcaccaaatg tcccaccagg tgcaggcaga aaacgctgga acagcgtgta 6780

cagtttgtct taacaaaaag tgagggcgct gaggtcgagc agggtggtgt gacttgttat 6840

agcctttaga gctgcgaaag cgcgtatgga tttggctcat caggccagat tgagggtctg 6900

tggacacatg tcatgttagt gtacttcaat cgccccctgg atatagcccc gacaataggc 6960

cgtggcctca ttttttgcc ttccgcacat ttccattgct cggtacccac accttgcttc 7020

tcctgcactt gccaacctta atactggttt acattgacca acatcttaca agcgggggc 7080

ttgtctaggg tatatataaa cagtggctct cccaatcggt tgccagtctc ttttttcctt 7140

tctttcccca cagattcgaa atctaaacta cacatcacac aatgcctgtt actgacgtcc 7200

ttaagcgaaa gtccggtgtc atcgtcggcg acgatgtccg agccgtgagt atccacgaca 7260

agatcagtgt cgagacgacg cgttttgtgt aatgacacaa tccgaaagtc gctagcaaca 7320

cacactctct acacaaacta acccagctct ccatggagtc cattgctccc ttcctgccct 7380

ccaagatgcc tcaggacctg ttcatggacc tcgccagcgc tatcggtgtc cgagctgctc 7440

cctacgtcga tcccctggag gctgccctgg ttgcccaggc cgagaagtac attcccacca 7500

ttgtccatca cactcgaggc ttcctggttg ccgtggagtc tcccctggct cgagagctgc 7560

ctctgatgaa ccccttccac gtgctcctga tcgtgctcgc ctacctggtc accgtgtttg 7620

tgggtatgca gatcatgaag aactttgaac gattcgaggt caagaccttc tccctcctgc 7680

acaacttctg tctggtctcc atctccgcct acatgtgcgg tggcatcctg tacgaggctt 7740

atcaggccaa ctatggactg tttgagaacg ctgccgatca caccttcaag ggtctcccta 7800

tggctaagat gatctggctc ttctacttct ccaagatcat ggagtttgtc gacaccatga 7860

tcatggtcct caagaagaac aaccgacaga tttccttttct gcacgtgtac caccactctt 7920

ccatcttcac catctggtgg ctggtcacct tcgttgctcc caacggtgaa gcctacttct 7980

ctgctgccct gaactccttc atccacgtca tcatgtacgg ctactacttt ctgtctgccc 8040

tgggcttcaa gcaggtgtcg ttcatcaagt tctacatcac tcgatcccag atgacccagt 8100

tctgcatgat gtctgtccag tcttcctggg acatgtacgc catgaaggtc cttggccgac 8160

ctggataccc cttcttcatc accgctctgc tctggttcta catgtggacc atgctcggtc 8220

tcttctacaa cttttaccga aagaacgcca agctcgccaa gcaggccaag gctgacgctg 8280

ccaaggagaa ggccagaaag ctccagtaag cggccgcaag tgtggatggg gaagtgagtg 8340

cccggttctg tgtgcacaat tggcaatcca agatggatgg attcaacaca gggatatagc 8400

gagctacgtg gtggtgcgag gatatagcaa cggatattta tgtttgacac ttgagaatgt 8460

acgatacaag cactgtccaa gtacaatact aaacatactg tacatactca tactcgtacc 8520

cgggcaacgg tttcacttga gtgcagtggc tagtgctctt actcgtacag tgtgcaatac 8580

tgcgtatcat agtctttgat gtatatcgta ttcattcatg ttagttgcgt acgaagtcgt 8640

caatgatgtc gatatgggtt ttgatcatgc acacataagg tccgacctta tcggcaagct 8700

caatgagctc cttggtggtg gtaacatcca gagaagcaca caggttggtt ttcttggctg 8760

ccacgagctt gagcactcga gcggcaaagg cggacttgtg gacgttagct cgagcttcgt 8820

aggagggcat tttggtggtg aagaggagac tgaaataaat ttagtctgca gaacttttta 8880

tcggaacctt atctggggca gtgaagtata tgttatggta atagttacga gttagttgaa 8940

cttatagata gactggacta tacggctatc ggtccaaatt agaaagaacg tcaatggctc 9000

tctgggcgtc gcctttgccg acaaaaatgt gatcatgatg aaagccagca atgacgttgc 9060

agctgatatt gttgtcggcc aaccgcgccg aaaacgcagc tgtcagaccc acagcctcca 9120

acgaagaatg tatcgtcaaa gtgatccaag cacactcata gttggagtcg tactccaaag 9180

gcggcaatga cgagtcagac agatactcgt cgacctttc cttgggaacc accaccgtca 9240

gcccttctga ctcacgtatt gtagccaccg acacaggcaa cagtccgtgg atagcagaat 9300

atgtcttgtc ggtccatttc tcaccaactt taggcgtcaa gtgaatgttg cagaagaagt 9360

atgtgccttc attgagaatc ggtgttgctg atttcaataa agtcttgaga tcagtttggc 9420

gcgccagctg cattaatgaa tcggccaacg cgcggggaga ggcggtttgc gtattgggcg 9480

ctcttccgct tcctcgctca ctgactcgct gcgctcggtc gttcggctgc ggcgagcggt 9540

atcagctcac tcaaaggcgg taatacggtt atccacagaa tcaggggata acgcaggaaa 9600

gaacatgtga gcaaaaggcc agcaaaaggc caggaaccgt aaaaaggccg cgttgctggc 9660

gtttttccat aggctccgcc cccctgacga gcatcacaaa aatcgacgct caagtcagag 9720

gtggcgaaac ccgacaggac tataaagata ccaggcgttt ccccctggaa gctccctcgt 9780

gcgctctcct gttccgaccc tgccgcttac cggatacctg tccgcctttc tcccttcggg 9840

aagcgtggcg ctttctcata gctcacgctg taggtatctc agttcggtgt aggtcgttcg 9900

ctccaagctg ggctgtgtgc acgaaccccc cgttcagccc gaccgctgcg ccttatccgg 9960

taactatcgt cttgagtcca acccggtaag acacgactta tcgccactgg cagcagccac 10020

tggtaacagg attagcagag cgaggtatgt aggcggtgct acagagttct tgaagtggtg 10080

gcctaactac ggctacacta gaagaacagt atttggtatc tgcgctctgc tgaagccagt 10140

taccttcgga aaaagagttg gtagctcttg atccggcaaa caaaccaccg ctggtagcgg 10200

tggtttttt gtttgcaagc agcagattac gcgcagaaaa aaaggatctc aagaagatcc 10260

tttgatcttt tctacggggt ctgacgctca gtggaacgaa aactcacgtt aagggatttt 10320

ggtcatgaga ttatcaaaaa ggatcttcac ctagatcctt ttaaattaaa aatgaagttt 10380

taaatcaatc taaagtatat atgagtaaac ttggtctgac agttaccaat gcttaatcag 10440

tgaggcacct atctcagcga tctgtctatt tcgttcatcc atagttgcct gactccccgt 10500

cgtgtagata actacgatac gggagggctt accatctggc cccagtgctg caatgatacc 10560

gcgagaccca cgctcaccgg ctccagattt atcagcaata aaccagccag ccggaagggc 10620

cgagcgcaga agtggtcctg caactttatc cgcctccatc cagtctatta attgttgccg 10680

ggaagctaga gtaagtagtt cgccagttaa tagtttgcgc aacgttgttg ccattgctac 10740

aggcatcgtg gtgtcacgct cgtcgtttgg tatggcttca ttcagctccg gttcccaacg 10800

atcaaggcga gttacatgat cccccatgtt gtgcaaaaaa gcggttagct ccttcggtcc 10860

tccgatcgtt gtcagaagta agttggccgc agtgttatca ctcatggtta tggcagcact 10920

gcataattct cttactgtca tgccatccgt aagatgcttt tctgtgactg gtgagtactc 10980

aaccaagtca ttctgagaat agtgtatgcg gcgaccgagt tgctcttgcc cggcgtcaat 11040

acgggataat accgcgccac atagcagaac tttaaaagtg ctcatcattg gaaaacgttc 11100

ttcggggcga aaactctcaa ggatcttacc gctgttgaga tccagttcga tgtaacccac 11160

tcgtgcaccc aactgatctt cagcatcttt tactttcacc agcgtttctg ggtgagcaaa 11220

aacaggaagg caaaatgccg caaaaaaggg aataagggcg acacggaaat gttgaatact 11280

catactcttc ctttttcaat attattgaag catttatcag ggttattgtc tcatgagcgg 11340

atacatattt gaatgtattt agaaaaataa acaaataggg gttccgcgca catttccccg 11400

aaaagtgcca cctgatgcgg tgtgaaatac cgcacagatg cgtaaggaga aaataccgca 11460

tcaggaaatt gtaagcgtta atattttgtt aaaattcgcg ttaaattttt gttaaatcag 11520

ctcatttttt aaccaatagg ccgaaatcgg caaaatccct tataaatcaa aagaatagac 11580

cgagataggg ttgagtgttg ttccagtttg gaacaagagt ccactattaa agaacgtgga 11640

ctccaacgtc aaagggcgaa aaaccgtcta tcagggcgat ggcccactac gtgaaccatc 11700

accctaatca agtttttttgg ggtcgaggtg ccgtaaagca ctaaatcgga accctaaagg 11760

gagcccccga tttagagctt gacggggaaa gccggcgaac gtggcgagaa aggaagggaa 11820

gaaagcgaaa ggagcgggcg ctagggcgct ggcaagtgta gcggtcacgc tgcgcgtaac 11880

caccacaccc gccgcgctta atgcgccgct acagggcgcg tccattcgcc attcaggctg 11940

cgcaactgtt gggaagggcg atcggtgcgg gcctcttcgc tattacgcca gctggcgaaa 12000

gggggatgtg ctgcaaggcg attaagttgg gtaacgccag ggttttccca gtcacgacgt 12060

tgtaaaacga cggccagtga attgtaatac gactcactat agggcgaatt gggcccgacg 12120

tcgcatgcag tggtggtatt gtgactgggg atgtagttga gaataagtca tacacaagtc 12180

agctttcttc gagcctcata taagtataag tagttcaacg tattagcact gtacccagca 12240

tctccgtatc gagaaacaca acaacatgcc ccattggaca gatcatgcgg atacacaggt 12300

tgtgcagtat catacatact cgatcagaca ggtcgtctga ccatcataca agctgaacaa 12360

gcgctccata cttgcacgct ctctatatac acagttaaat tacatatcca tagtctaacc 12420

tctaacagtt aatcttctgg taagcctccc agccagcctt ctggtatcgc ttggcctcct 12480

caataggatc tcggttctgg ccgtacagac ctcggccgac aattatgata tccgttccgg 12540

tagacatgac atcctcaaca gttcggtact gctgtccgag agcgtctccc ttgtcgtcaa 12600

gacccacccc gggggtcaga ataagccagt cctcagagtc gcccttaat 12649

<210> 157
<211> 13034
<212> DNA
<213> Artificial Sequence

<220>
<223> Plasmid pDMW271

<400> 157
cgatgcagaa ttcaggagag accgggttgg cggcgtattt gtgtcccaaa aaacagcccc 60

aattgcccca attgacccca aattgaccca gtagcgggcc caaccccggc gagagccccc 120

ttcaccccac atatcaaacc tcccccggtt cccacacttg ccgttaaggg cgtagggtac 180

tgcagtctgg aatctacgct tgttcagact ttgtactagt ttctttgtct ggccatccgg 240

gtaacccatg ccggacgcaa aatagactac tgaaaatttt tttgctttgt ggttgggact 300

ttagccaagg gtataaaaga ccaccgtccc cgaattacct ttcctcttct tttctctctc 360

tccttgtcaa ctcacacccg aaatcgttaa gcatttcctt ctgagtataa gaatcattca 420

ccatggatgg ctcccgaccc tgtcgctgcc gagaccgctg cccagggtcc cactccccga 480

tacttcacct gggacgaggt cgcccagcga tccggttgcg aggaacgatg gctggtcatc 540

gaccgaaagg tgtacaacat ctctgagttc acccgacgac atcccggtgg ctcccgagtg 600

atctcgcact acgctggaca ggacgccact gacccccttcg ttgcctttca cattaacaag 660

ggcctggtta agaagtacat gaactccctg ctcattggag agctgtctcc cgaacagcct 720

tcgtttgagc ctaccaagaa caaggagctg accgacgagt ttcgagagct ccgagccacc 780

gttgagcgaa tgggactgat gaaggccaac catgtcttct ttctgctcta cctgctccac 840

attcttctcc ttgacggagc tgcctggctt accctgtggg tcttcggcac ttcctttctg 900

ccctttcttc tctgcgccgt cctgctctct gccgtgcagg ctcaggctgg ttggcttcag 960

catgactttg gtcacctttc cgtgttctct acctccaagt ggaaccacct gctccatcac 1020

ttcgtgatcg gccacctcaa gggtgctcct gcctcgtggt ggaaccacat gcatttccag 1080

caccatgcca agcccaactg ttttcgaaag gatcccgaca tcaacatgca ccccttcttt 1140

ttcgctcttg gcaagatcct gtccgtcgag ctcggaaagc agaagaagaa gtacatgccc 1200

tacaaccacc agcacaagta cttcttcctg attggacctc ccgctctcct gcctctttac 1260

tttcagtggt acatctttta ctttgttatt cagcgaaaga agtgggttga tcttgcctgg 1320

atgatcacct tctacgtccg attcttcctg acctacgtcc ctctccttgg actgaaggcc 1380

tttctcggtc tgttctttat cgtccgattc ctggagtcca actggttcgt gtgggtgacc 1440

cagatgaacc acattcccat gcacattgac catgatcgaa acatggactg ggtgtcgact 1500

cagctgcagg ccacctgcaa cgttcacaag tctgctttca acgactggtt ttccggtcac 1560

ctcaactttc agattgagca ccatctgttt cccaccatgc ctcgacacaa ctaccacaag 1620

gttgctcccc tggtccagtc gctctgtgcc aagcatggca tcgagtacca gtccaagccc 1680

ctgctctctg ccttcgctga catcattcac tcgctgaagg aatctggcca gctctggctc 1740

gatgcctacc tgcaccagta agcggccgca ttgatgattg gaaacacaca catgggttat 1800

atctaggtga gagttagttg gacagttata tattaaatca gctatgccaa cggtaacttc 1860

attcatgtca acgaggaacc agtgactgca agtaatatag aatttgacca ccttgccatt 1920

ctcttgcact cctttactat atctcattta tttcttatat acaaatcact tcttcttccc 1980

agcatcgagc tcggaaacct catgagcaat aacatcgtgg atctcgtcaa tagagggctt 2040

tttggactcc ttgctgttgg ccaccttgtc cttgctgtct ggctcattct gtttcaacgc 2100

cttttaatta acggagtagg tctcggtgtc ggaagcgacg ccagatccgt catcctcctt 2160

tcgctctcca aagtagatac ctccgacgag ctctcggaca atgatgaagt cggtgccctc 2220

aacgtttcgg atgggggaga gatcggcgag cttgggcgac agcagctggc agggtcgcag 2280

gttggcgtac aggttcaggt cctttcgcag cttgaggaga ccctgctcgg gtcgcacgtc 2340

ggttcgtccg tcgggagtgg tccatacggt gttggcagcg cctccgacag caccgagcat 2400

aatagagtca gcctttcggc agatgtcgag agtagcgtcg gtgatgggct cgccctcctt 2460

ctcaatggca gctcctccaa tgagtcggtc ctcaaacaca aactcggtgc cggaggcctc 2520

agcaacagac ttgagcacct tgacggcctc ggcaatcacc tcggggccac agaagtcgcc 2580

gccgagaaga acaatcttct tggagtcagt cttggtcttc ttagtttcgg gttccattgt 2640

ggatgtgtgt ggttgtatgt gtgatgtggt gtgtggagtg aaaatctgtg gctggcaaac 2700

gctcttgtat atatacgcac ttttgcccgt gctatgtgga agactaaacc tccgaagatt 2760

gtgactcagg tagtgcggta tcggctaggg acccaaacct tgtcgatgcc gatagcatgc 2820

gacgtcgggc ccaattcgcc ctatagtgag tcgtattaca attcactggc cgtcgtttta 2880

caacgtcgtg actgggaaaa ccctggcgtt acccaactta atcgccttgc agcacatccc 2940

cctttcgcca gctggcgtaa tagcgaagag gcccgcaccg atcgcccttc ccaacagttg 3000

cgcagcctga atggcgaatg gacgcgccct gtagcggcgc attaagcgcg gcgggtgtgg 3060

tggttacgcg cagcgtgacc gctacacttg ccagcgccct agcgcccgct cctttcgctt 3120

tcttcccttc ctttctcgcc acgttcgccg gctttccccg tcaagctcta aatcgggggc 3180

tccctttagg gttccgattt agtgctttac ggcacctcga ccccaaaaaa cttgattagg 3240

gtgatggttc acgtagtggg ccatcgccct gatagacggt ttttcgccct ttgacgttgg 3300

agtccacgtt ctttaatagt ggactcttgt tccaaactgg aacaacactc aaccctatct 3360

cggtctattc ttttgattta taagggattt tgccgatttc ggcctattgg ttaaaaaatg 3420

agctgattta acaaaaattt aacgcgaatt ttaacaaaat attaacgctt acaatttcct 3480

gatgcggtat tttctcctta cgcatctgtg cggtatttca caccgcatca ggtggcactt 3540

ttcggggaaa tgtgcgcgga acccctattt gtttatttttt ctaaatacat tcaaatatgt 3600

atccgctcat gagacaataa ccctgataaa tgcttcaata atattgaaaa aggaagagta 3660

tgagtattca acatttccgt gtcgccctta ttccctttttt tgcggcattt tgccttcctg 3720

tttttgctca cccagaaacg ctggtgaaag taaaagatgc tgaagatcag ttgggtgcac 3780

gagtgggtta catcgaactg gatctcaaca gcggtaagat ccttgagagt tttcgccccg 3840

aagaacgttt tccaatgatg agcactttta aagttctgct atgtggcgcg gtattatccc 3900

gtattgacgc cgggcaagag caactcggtc gccgcataca ctattctcag aatgacttgg 3960

ttgagtactc accagtcaca gaaaagcatc ttacggatgg catgacagta agagaattat 4020

gcagtgctgc cataaccatg agtgataaca ctgcggccaa cttacttctg acaacgatcg 4080

gaggaccgaa ggagctaacc gcttttttgc acaacatggg ggatcatgta actcgccttg 4140

atcgttggga accggagctg aatgaagcca taccaaacga cgagcgtgac accacgatgc 4200

ctgtagcaat ggcaacaacg ttgcgcaaac tattaactgg cgaactactt actctagctt 4260

cccggcaaca attaatagac tggatggagg cggataaagt tgcaggacca cttctgcgct 4320

cggcccttcc ggctggctgg tttattgctg ataaatctgg agccggtgag cgtgggtctc 4380

gcggtatcat tgcagcactg gggccagatg gtaagccctc ccgtatcgta gttatctaca 4440

cgacggggag tcaggcaact atggatgaac gaaatagaca gatcgctgag ataggtgcct 4500

cactgattaa gcattggtaa ctgtcagacc aagtttactc atatatactt tagattgatt 4560

taaaacttca tttttaattt aaaaggatct aggtgaagat ccttttttgat aatctcatga 4620

ccaaaatccc ttaacgtgag ttttcgttcc actgagcgtc agaccccgta gaaaagatca 4680

aaggatcttc ttgagatcct ttttttctgc gcgtaatctg ctgcttgcaa acaaaaaaac 4740

caccgctacc agcggtggtt tgtttgccgg atcaagagct accaactctt tttccgaagg 4800

taactggctt cagcagagcg cagataccaa atactgttct tctagtgtag ccgtagttag 4860

gccaccactt caagaactct gtagcaccgc ctacatacct cgctctgcta atcctgttac 4920

cagtggctgc tgccagtggc gataagtcgt gtcttaccgg gttggactca agacgatagt 4980

taccggataa ggcgcagcgg tcgggctgaa cgggggggttc gtgcacacag cccagcttgg 5040

agcgaacgac ctacaccgaa ctgagatacc tacagcgtga gctatgagaa agcgccacgc 5100

ttcccgaagg gagaaaggcg gacaggtatc cggtaagcgg cagggtcgga acaggagagc 5160

gcacgaggga gcttccaggg ggaaacgcct ggtatcttta tagtcctgtc gggtttcgcc 5220

acctctgact tgagcgtcga tttttgtgat gctcgtcagg ggggcggagc ctatggaaaa 5280

acgccagcaa cgcggccttt ttacggttcc tggccttttg ctggcctttt gctcacatgt 5340

tctttcctgc gttatcccct gattctgtgg ataaccgtat taccgccttt gagtgagctg 5400

ataccgctcg ccgcagccga acgaccgagc gcagcgagtc agtgagcgag gaagcggaag 5460

agcgcccaat acgcaaaccg cctctccccg cgcgttggcc gattcattaa tgcagctggc 5520

gcgcccactg agctcgtcta acggacttga tatacaacca attaaaacaa atgaaaagaa 5580

atacagttct ttgtatcatt tgtaacaatt accctgtaca aactaaggta ttgaaatccc 5640

acaatattcc caaagtccac ccctttccaa attgtcatgc ctacaactca tataccaagc 5700

actaacctac caaacaccac taaaacccca caaaatatat cttaccgaat atacagtaac 5760

aagctaccac cacactcgtt gggtgcagtc gccagcttaa agatatctat ccacatcagc 5820

cacaactccc ttcctttaat aaaccgacta cacccttggc tattgaggtt atgagtgaat 5880

atactgtaga caagacactt tcaagaagac tgtttccaaa acgtaccact gtcctccact 5940

acaaacacac ccaatctgct tcttctagtc aaggttgcta caccggtaaa ttataaatca 6000

tcatttcatt agcagggcag ggccctttt atagagtctt atacactagc ggaccctgcc 6060

ggtagaccaa cccgcaggcg cgtcagtttg ctccttccat caatgcgtcg tagaaacgac 6120

ttactccttc ttgagcagct ccttgacctt gttggcaaca agtctccgac ctcggaggtg 6180

gaggaagagc ctccgatatc ggcggtagtg ataccagcct cgacggactc cttgacggca 6240

gcctcaacag cgtcaccggc gggcttcatg ttaagagaga acttgagcat catggcggca 6300

gacagaatgg tggcgtacgc aactaacatg aatgaatacg atatacatca aagactatga 6360

tacgcagtat tgcacactgt acgagtaaga gcactagcca ctgcactcaa gtgaaaccgt 6420

tgcccgggta cgagtatgag tatgtacagt atgtttagta ttgtacttgg acagtgcttg 6480

tatcgtacat tctcaagtgt caaacataaa tatccgttgc tatatcctcg caccaccacg 6540

tagctcgcta tatccctgtg ttgaatccat ccatcttgga ttgccaattg tgcacacaga 6600

accgggcact cacttcccca tccacacttg cggccgctta gctgcctact cttccttggg 6660

acggagtcca agaacacgca agtgctccaa atgtgaagca aatgcttgcc aaaacgtatc 6720

cttgacaagg tatggaacct tgtactcgct gcaggtgttc ttgatgatgg ccagaatatc 6780

gggataatgg tgctgcgaca cgttggggaa cagatggtgc acagcctggt agttcaagct 6840

gccagtgatg ctggtccaga ggtgcgaatc gtgtgcgtaa tcctgcgtag tctcgacctg 6900

catagctgcc cagtcctttt ggatgatccc gttctcgtca ggcaacggcc actgaacttc 6960

ctcaacaacg tggttcgcct ggaaggtcag cgccagccag taagacgaca ccatgtccgc 7020

gaccgtgaac aagagcagca ccttgcccag gggcagatac tgcaggggaa caatcaggcg 7080

ataccagaca aagaaagcct tgccgcccca gaacatcaca gtgtgccatg tcgagatggg 7140

attgacacga atagcgtcat tggtcttgac aaagtacaaa atgttgatgt cctgaatgcg 7200

caccttgaac gccagcagtc cgtacaggaa aggaacaaac atgtgctggt tgatgtggtt 7260

gacaaaccac ttttggttgg gcttgatacg acgaacatcg ggctcagacg tcgacacgtc 7320

gggatctgct ccagcaatgt tggtgtaggg gtgatggccg agcatatgtt ggtacatcca 7380

caccaggtac gatgctccgt tgaaaaagtc gtgcgtggct cccagaatct tccagacagt 7440

ggggttgtgg gtcactgaaa agtgagacgc atcatgaaga gggttgagtc cgacttgtgc 7500

gcacgcaaat cccatgatga ttgcaaacac cacctgaagc catgtgcgtt cgacaacgaa 7560

aggcacaaag agctgcgcgt agtaggaagc gatcaaggat ccaaagataa gagcgtatcg 7620

tccccagatc tctggtctat tcttgggatc aatgttccga tccgtaaagt agccctcgac 7680

tctcgtcttg atggttttgt ggaacaccgt tggctccggg aagatgggca gctcattcga 7740

gaccagtgta ccgacatagt acttcttcat aatggcatct gcagccccaa acgcgtgata 7800

catctcaaag accggagtaa catctcggcc agctccgagc aggagagtgt ccactccacc 7860

aggatggcgg ctcaagaact ttgtgacatc gtacaccctg ccgcggatgg ccaagagtag 7920

gtcgtccttg gtgttatggg ccgccagctc ttcccaggtg aaggtttttc cttggtccgt 7980

tcccatggag agctgggtta gtttgtgtag agagtgtgtg ttgctagcga ctttcggatt 8040

gtgtcattac acaaaacgcg tcgtctcgac actgatcttg tcgtggatac tcacggctcg 8100

gacatcgtcg ccgacgatga caccggactt tcgcttaagg acgtcagtaa caggcattgt 8160

gtgatgtgta gtttagattt cgaatctgtg gggaaagaaa ggaaaaaaga gactggcaac 8220

cgattgggag agccactgtt tatatatacc ctagacaagc cccccgcttg taagatgttg 8280

gtcaatgtaa accagtatta aggttggcaa gtgcaggaga agcaaggtgt gggtaccgag 8340

caatggaaat gtgcggaagg caaaaaaatg aggccacggc ctattgtcgg ggctatatcc 8400

agggggcgat tgaagtacac taacatgaca tgtgtccaca gaccctcaat ctggcctgat 8460

gagccaaatc catacgcgct ttcgcagctc taaaggctat aacaagtcac accaccctgc 8520

tcgacctcag cgccctcact ttttgttaag acaaactgta cacgctgttc cagcgttttc 8580

tgcctgcacc tggtgggaca tttggtgcaa cctaaagtgc tcggaacctc tgtggtgtcc 8640

agatcagcgc agcagttccg aggtagtttt gaggcccttа gatgatgcaa tggtgtcagt 8700

cgctggatca cgagtcttaa tggcagtatt cgttcttatt tgtgccattg agccccgtta 8760

tcctcgtatc ttctaccccc catcccatcc ctttgttggt gcaaccctac ccatttattg 8820

ttgggtgcag cccaaccgac gtggagagct tggcttggcc atataaaaag gccccccccct 8880

agtggcaatg gcagaaagtc agctgtgagt tgttgaattt gtcatctagg cggcctggcc 8940

gtcttctccg gggcaattta aattccttca cttcaagttc attcttcatc tgcttctgtt 9000

ttactttgac aggcaaatga agacatggta cgacttgatg gaggccaaga acgccatttc 9060

accccgagac accgaagtgc ctgaaatcct ggctgccccc attgataaca tcggaaacta 9120

cggtattccg gaaagtgtat atagaacctt tccccagctt gtgtctgtgg atatggatgg 9180

tgtaatcccc tttgagtact cgtcttggct tctctccgag cagtatgagg ctctctaatc 9240

tagcgcattt aatatctcaa tgtatttata tatttatctt ctcatgcggc cgcttagctg 9300

cctactcttc cttgggacgg agtccaagaa cacgcaagtg ctccaaatgt gaagcaaatg 9360

cttgccaaaa cgtatccttg acaaggtatg gaaccttgta ctcgctgcag gtgttcttga 9420

tgatggccag aatatcggga taatggtgct gcgacacgtt ggggaacaga tggtgcacag 9480

cctggtagtt caagctgcca gtgatgctgg tccagaggtg cgaatcgtgt gcgtaatcct 9540

gcgtagtctc gacctgcata gctgcccagt ccttttggat gatcccgttc tcgtcaggca 9600

acggccactg aacttcctca acaacgtggt tcgcctggaa ggtcagcgcc agccagtaag 9660

acgacaccat gtccgcgacc gtgaacaaga gcagcacctt gcccaggggc agatactgca 9720

ggggaacaat caggcgatac cagacaaaga aagccttgcc gccccagaac atcacagtgt 9780

gccatgtcga gatgggattg acacgaatag cgtcattggt cttgacaaag tacaaaatgt 9840

tgatgtcctg aatgcgcacc ttgaacgcca gcagtccgta caggaaagga acaaacatgt 9900

gctggttgat gtggttgaca aaccactttt ggttgggctt gatacgacga acatcgggct 9960

cagacgtcga cacgtcggga tctgctccag caatgttggt gtaggggtga tggccgagca 10020

tatgttggta catccacacc aggtacgatg ctccgttgaa aaagtcgtgc gtggctccca 10080

gaatcttcca gacagtgggg ttgtgggtca ctgaaaagtg agacgcatca tgaagagggt 10140

tgagtccgac ttgtgcgcac gcaaatccca tgatgattgc aaacaccacc tgaagccatg 10200

tgcgttcgac aacgaaaggc acaaagagct gcgcgtagta ggaagcgatc aaggatccaa 10260

agataagagc gtatcgtccc cagatctctg gtctattctt gggatcaatg ttccgatccg 10320

taaagtagcc ctcgactctc gtcttgatgg ttttgtggaa caccgttggc tccgggaaga 10380

tgggcagctc attcgagacc agtgtaccga catagtactt cttcataatg gcatctgcag 10440

ccccaaacgc gtgatacatc tcaaagaccg gagtaacatc tcggccagct ccgagcagga 10500

gagtgtccac tccaccagga tggcggctca agaactttgt gacatcgtac accctgccgc 10560

ggatggccaa gagtaggtcg tccttggtgt tatgggccgc cagctcttcc caggtgaagg 10620

tttttccttg gtccgttccc atggtgaatg attcttatac tcagaaggaa atgcttaacg 10680

atttcgggtg tgagttgaca aggagagaga gaaaagaaga ggaaaggtaa ttcggggacg 10740

gtggtctttt ataccccttgg ctaaagtccc aaccacaaag caaaaaaatt ttcagtagtc 10800

tattttgcgt ccggcatggg ttacccggat ggccagacaa agaaactagt acaaagtctg 10860

aacaagcgta gattccagac tgcagtaccc tacgccctta acggcaagtg tgggaaccgg 10920

gggaggtttg atatgtgggg tgaaggggggc tctcgccggg gttgggcccg ctactgggtc 10980

aatttggggt caattggggc aattggggct gttttttggg acacaaatac gccgccaacc 11040

cggtctctcc tgatcgatgg gctgcaggaa ttctacaata cgtgagtcag aagggctgac 11100

ggtggtggtt cccaaggaaa aggtcgacga gtatctgtct gactcgtcat tgccgccttt 11160

ggagtacgac tccaactatg agtgtgcttg gatcactttg acgatacatt cttcgttgga 11220

ggctgtgggt ctgacagctg cgttttcggc gcggttggcc gacaacaata tcagctgcaa 11280

cgtcattgct ggctttcatc atgatcacat ttttgtcggc aaaggcgacg cccagagagc 11340

cattgacgtt ctttctaatt tggaccgata gccgtatagt ccagtctatc tataagttca 11400

actaactcgt aactattacc ataacatata cttcactgcc ccagataagg ttccgataaa 11460

aagttctgca gactaaattt atttcagtct cctcttcacc accaaaatgc cctcctacga 11520

agctcgagct aacgtccaca agtccgcctt tgccgctcga gtgctcaagc tcgtggcagc 11580

caagaaaacc aacctgtgtg cttctctgga tgttaccacc accaaggagc tcattgagct 11640

tgccgataag gtcggacctt atgtgtgcat gatcaaaacc catatcgaca tcattgacga 11700

cttcacctac gccggcactg tgctcccccc caaggaactt gctcttaagc acggtttctt 11760

cctgttcgag gacagaaagt tcgcagatat tggcaacact gtcaagcacc agtaccggtg 11820

tcaccgaatc gccgagtggt ccgatatcac caacgcccac ggtgtacccg gaaccggaat 11880

cattgctggc ctgcgagctg gtgccgagga aactgtctct gaacagaaga aggaggacgt 11940

ctctgactac gagaactccc agtacaagga gttcctagtc ccctctccca acgagaagct 12000

ggccagaggt ctgctcatgc tggccgagct gtcttgcaag ggctctctgg ccactggcga 12060

gtactccaag cagaccattg agcttgcccg atccgacccc gagtttgtgg ttggcttcat 12120

tgcccagaac cgacctaagg gcgactctga ggactggctt attctgaccc ccggggtggg 12180

tcttgacgac aagggagacg ctctcggaca gcagtaccga actgttgagg atgtcatgtc 12240

taccggaacg gatatcataa ttgtcggccg aggtctgtac ggccagaacc gagatcctat 12300

tgaggaggcc aagcgatacc agaaggctgg ctgggaggct taccagaaga ttaactgtta 12360

gaggttagac tatggatatg taatttaact gtgtatatag agagcgtgca agtatggagc 12420

gcttgttcag cttgtatgat ggtcagacga cctgtctgat cgagtatgta tgatactgca 12480

caacctgtgt atccgcatga tctgtccaat ggggcatgtt gttgtgtttc tcgatacgga 12540

gatgctgggt acagtgctaa tacgttgaac tacttatact tatatgaggc tcgaagaaag 12600

ctgacttgtg tatgacttat tctcaactac atccccagtc acaataccac cactgcacta 12660

ccactacacc agatctgcgt acactgttta aacggtaggt tagtgcttgg tatatgagtt 12720

gtaggcatga caatttggaa aggggtggac tttgggaata ttgtgggatt tcaatacctt 12780

agtttgtaca gggtaattgt tacaaatgat acaaagaact gtatttcttt tcatttgttt 12840

taattggttg tatatcaagt ccgttagacg agctcagtgc cttggctttt ggcactgtat 12900

ttcattttta gaggtacact acattcagtg aggtatggta aggttgaggg cataatgaag 12960

gcaccttgta ctgacagtca cagacctctc accgagaatt ttatgagata tactcgggtt 13020

cattttaggc tcat 13034

<210> 158
<211> 12690
<212> DNA
<213> Artificial Sequence

<220>
<223> Plasmid pZP3L37

<400> 158
aaataccagt tggccacaaa cccttgacga tctcgtatgt cccctccgac atactcccgg 60

ccggctgggg tacgttcgat agcgctatcg gcatcgacaa ggtttgggtc cctagccgat 120

accgcactac ctgagtcaca atcttcggag gtttagtctt ccacatagca cgggcaaaag 180

tgcgtatata tacaagagcg tttgccagcc acagattttc actccacaca ccacatcaca 240

catacaacca cacacatcca caatggaacc cgaaactaag aagaccaaga ctgactccaa 300

gaagattgtt cttctcggcg gcgacttctg tggccccgag gtgattgccg aggccgtcaa 360

ggtgctcaag tctgttgctg aggcctccgg caccgagttt gtgtttgagg accgactcat 420

tggaggagct gccattgaga aggagggcga gcccatcacc gacgctactc tcgacatctg 480

ccgaaaggct gactctatta tgctcggtgc tgtcggaggc gctgccaaca ccgtatggac 540

cactcccgac ggacgaaccg acgtgcgacc cgagcagggt ctcctcaagc tgcgaaagga 600

cctgaacctg tacgccaacc tgcgaccctg ccagctgctg tcgcccaagc tcgccgatct 660

ctcccccatc cgaaacgttg agggcaccga cttcatcatt gtccgagagc tcgtcggagg 720

tatctacttt ggagagcgaa aggaggatga cggatctggc gtcgcttccg acaccgagac 780

ctactccgtt cctgaggttg agcgaattgc ccgaatggcc gccttcctgg cccttcagca 840

caacccccct cttcccgtgt ggtctcttga caaggccaac gtgctggcct cctctcgact 900

ttggcgaaag actgtcactc gagtcctcaa ggacgaattc ccccagctcg agctcaacca 960

ccagctgatc gactcggccg ccatgatcct catcaagcag ccctccaaga tgaatggtat 1020

catcatcacc accaacatgt ttggcgatat catctccgac gaggcctccg tcatccccgg 1080

ttctctgggt ctgctgccct ccgcctctct ggcttctctg cccgacacca acgaggcgtt 1140

cggtctgtac gagccctgtc acggatctgc ccccgatctc ggcaagcaga aggtcaaccc 1200

cattgccacc attctgtctg ccgccatgat gctcaagttc tctcttaaca tgaagcccgc 1260

cggtgacgct gttgaggctg ccgtcaagga gtccgtcgag gctggtatca ctaccgccga 1320

tatcggaggc tcttcctcca cctccgaggt cggagacttg ttgccaacaa ggtcaaggag 1380

ctgctcaaga aggagtaagt cgtttctacg acgcattgat ggaaggagca aactgacgcg 1440

cctgcgggtt ggtctaccgg cagggtccgc tagtgtataa gactctataa aaagggccct 1500

gccctgctaa tgaaatgatg atttataatt taccggtgta gcaaccttga ctagaagaag 1560

cagattgggt gtgtttgtag tggaggacag tggtacgttt tggaaacagt cttcttgaaa 1620

gtgtcttgtc tacagtatat tcactcataa cctcaatagc caagggtgta gtcggtttat 1680

taaaggaagg gagttgtggc tgatgtggat agatatcttt aagctggcga ctgcacccaa 1740

cgagtgtggt ggtagcttgt tagatctgta tattcggtaa gatatatttt gtggggtttt 1800

agtggtgttt aaacagtgta cgcagtacta tagaggaaca attgccccgg agaagacggc 1860

caggccgcct agatgacaaa ttcaacaact cacagctgac tttctgccat tgccactagg 1920

ggggggcctt tttatatggc caagccaagc tctccacgtc ggttgggctg cacccaacaa 1980

taaatgggta gggttgcacc aacaaaggga tgggatgggg ggtagaagat acgaggataa 2040

cggggctcaa tggcacaaat aagaacgaat actgccatta agactcgtga tccagcgact 2100

gacaccattg catcatctaa gggcctcaaa actacctcgg aactgctgcg ctgatctgga 2160

caccacagag gttccgagca ctttaggttg caccaaatgt cccaccaggt gcaggcagaa 2220

aacgctggaa cagcgtgtac agtttgtctt aacaaaaagt gagggcgctg aggtcgagca 2280

gggtggtgtg acttgttata gcctttagag ctgcgaaagc gcgtatggat ttggctcatc 2340

aggccagatt gagggtctgt ggacacatgt catgttagtg tacttcaatc gccccctgga 2400

tatagccccg acaataggcc gtggcctcat ttttttgcct tccgcacatt tccattgctc 2460

ggtacccaca ccttgcttct cctgcacttg ccaaccttaa tactggttta cattgaccaa 2520

catcttacaa gcgggggggct tgtctagggt atatataaac agtggctctc ccaatcggtt 2580

gccagtctct tttttccttt ctttccccac agattcgaaa tctaaactac acatcacaca 2640

atgcctgtta ctgacgtcct taagcgaaag tccggtgtca tcgtcggcga cgatgtccga 2700

gccgtgagta tccacgacaa gatcagtgtc gagacgacgc gttttgtgta atgacacaat 2760

ccgaaagtcg ctagcaacac acactctcta cacaaactaa cccagctctc catggctgag 2820

gataagacca aggtcgagtt ccctaccctg actgagctga agcactctat ccctaacgct 2880

tgctttgagt ccaacctcgg actctcgctc tactacactg cccgagcgat cttcaacgca 2940

tctgcctctg ctgctctgct ctacgctgcc cgatctactc ccttcattgc cgataacgtt 3000

ctgctccacg ctctggtttg cgccacctac atctacgtgc agggtgtcat cttctggggt 3060

ttctttaccg tcggtcacga ctgtggtcac tctgccttct cccgatacca ctccgtcaac 3120

ttcatcattg gctgcatcat gcactctgcc attctgactc ccttcgagtc ctggcgagtg 3180

acccaccgac accatcacaa gaacactggc aacattgata aggacgagat cttctaccct 3240

catcggtccg tcaaggacct ccaggacgtg cgacaatggg tctacaccct cggaggtgct 3300

tggtttgtct acctgaaggt cggatatgct cctcgaacca tgtcccactt tgacccctgg 3360

gaccctctcc tgcttcgacg agcctccgct gtcatcgtgt ccctcggagt ctgggctgcc 3420

ttcttcgctg cctacgccta cctcacatac tcgctcggct ttgccgtcat gggcctctac 3480

tactatgctc ctctctttgt ctttgcttcg ttcctcgtca ttactacctt cttgcatcac 3540

aacgacgaag ctactccctg gtacggtgac tcggagtgga cctacgtcaa gggcaacctg 3600

agctccgtcg accgatcgta cggagctttc gtggacaacc tgtctcacca cattggcacc 3660

caccaggtcc atcacttgtt ccctatcatt ccccactaca agctcaacga agccaccaag 3720

cactttgctg ccgcttaccc tcacctcgtg agacgtaacg acgagcccat cattactgcc 3780

ttcttcaaga ccgctcacct ctttgtcaac tacggagctg tgcccgagac tgctcagatt 3840

ttcaccctca aagagtctgc cgctgcagcc aaggccaaga gcgactaagc ggccgctatt 3900

tatcactctt tacaacttct acctcaacta tctactttaa taaatgaata tcgtttattc 3960

tctatgatta ctgtatatgc gttcctctaa gacaaatcga aaccagcatg tgatcgaatg 4020

gcatacaaaa gtttcttccg aagttgatca atgtcctgat agtcaggcag cttgagaaga 4080

ttgacacagg tggaggccgt agggaaccga tcaacctgtc taccagcgtt acgaatggca 4140

aatgacgggt tcaaagcctt gaatccttgc aatggtgcct tggatactga tgtcacaaac 4200

ttaagaagca gccgcttgtc ctcttcctcg atcgatcagg agagaccggg ttggcggcgt 4260

atttgtgtcc caaaaaacag ccccaattgc cccaattgac cccaaattga cccagtagcg 4320

ggcccaaccc cggcgagagc ccccttcacc ccacatatca aacctccccc ggttcccaca 4380

cttgccgtta agggcgtagg gtactgcagt ctggaatcta cgcttgttca gactttgtac 4440

tagtttcttt gtctggccat ccgggtaacc catgccggac gcaaaataga ctactgaaaa 4500

ttttttgct ttgtggttgg gactttagcc aagggtataa aagaccaccg tccccgaatt 4560

acctttcctc ttcttttctc tctctccttg tcaactcaca cccgaaatcg ttaagcattt 4620

ccttctgagt ataagaatca ttcaccatgg ctgaggataa gaccaaggtc gagttcccta 4680

ccctgactga gctgaagcac tctatcccta acgcttgctt tgagtccaac ctcggactct 4740

cgctctacta cactgcccga gcgatcttca acgcatctgc ctctgctgct ctgctctacg 4800

ctgcccgatc tactcccttc attgccgata acgttctgct ccacgctctg gtttgcgcca 4860

cctacatcta cgtgcagggt gtcatcttct ggggtttctt taccgtcggt cacgactgtg 4920

gtcactctgc cttctcccga taccactccg tcaacttcat cattggctgc atcatgcact 4980

ctgccattct gactcccttc gagtcctggc gagtgaccca ccgacaccat cacaagaaca 5040

ctggcaacat tgataaggac gagatcttct accctcatcg gtccgtcaag gacctccagg 5100

acgtgcgaca atgggtctac accctcggag gtgcttggtt tgtctacctg aaggtcggat 5160

atgctcctcg aaccatgtcc cactttgacc cctgggaccc tctcctgctt cgacgagcct 5220

ccgctgtcat cgtgtccctc ggagtctggg ctgccttctt cgctgcctac gcctacctca 5280

catactcgct cggctttgcc gtcatgggcc tctactacta tgctcctctc tttgtctttg 5340

cttcgttcct cgtcattact accttcttgc atcacaacga cgaagctact ccctggtacg 5400

gtgactcgga gtggacctac gtcaagggca acctgagctc cgtcgaccga tcgtacggag 5460

ctttcgtgga caacctgtct caccacattg gcacccacca ggtccatcac ttgttcccta 5520

tcattcccca ctacaagctc aacgaagcca ccaagcactt tgctgccgct taccctcacc 5580

tcgtgagacg taacgacgag cccatcatta ctgccttctt caagaccgct cacctctttg 5640

tcaactacgg agctgtgccc gagactgctc agattttcac cctcaaagag tctgccgctg 5700

cagccaaggc caagagcgac taagcggccg caagtgtgga tggggaagtg agtgcccggt 5760

tctgtgtgca caattggcaa tccaagatgg atggattcaa cacagggata tagcgagcta 5820

cgtggtggtg cgaggatata gcaacggata tttatgtttg acacttgaga atgtacgata 5880

caagcactgt ccaagtacaa tactaaacat actgtacata ctcatactcg tacccgggca 5940

acggtttcac ttgagtgcag tggctagtgc tcttactcgt acagtgtgca atactgcgta 6000

tcatagtctt tgatgtatat cgtattcatt catgttagtt gcgtacggtg tgtatcgtag 6060

aggtagtgac gtgttgtcca cagggcgact gtgtccgtgt atatatatat tcctcggccc 6120

gagcttattt gtgtggggtt gaggaaatca aaccaaatcg gtagtcagag aaataaaaca 6180

aaaagaaata aaaagaaata gaggacgcac aacgccatca ccgtcggaga gacaggagaa 6240

gggaaaatgg gcaaaaatgc ccttatcaca cccgcccgct ttgtgctctc attcggctcc 6300

cacaagagcc tcttgtcctg gttccccccc cccacatttt aacaccccac acgacgttgc 6360

tgcacgtgga attttcggcc gaaaacctgt ggggtactta cttttggcac tggagagaag 6420

catctgggat tttgggaacc taggcagaag atgaggaaaa aaataagagg aaccgttgtg 6480

agcttgctta tcagtgtcat atactccccc ctccttgcgt ttttgcgtct tttcccccta 6540

tttttcaaat tttgcgattt tttttctctt tttttccgct tttttccgct ttttttttgg 6600

ccggctttta tccatttctc caagccgagg atcacatcta tgcagcccag tccgttggag 6660

catatctgcg gtagagtttc ggaacggcgt taagcactgt gtccgggtcg gtctggaacg 6720

agattgagcg ggaaattcgg gggaataaga ccaccgttgg actccccgca atgaggagat 6780

caagatgtgc ttttcagaat tctgattggt ggcgcgccag ctgcattaat gaatcggcca 6840

acgcgcgggg agaggcggtt tgcgtattgg gcgctcttcc gcttcctcgc tcactgactc 6900

gctgcgctcg gtcgttcggc tgcggcgagc ggtatcagct cactcaaagg cggtaatacg 6960

gttatccaca gaatcagggg ataacgcagg aaagaacatg tgagcaaaag gccagcaaaa 7020

ggccaggaac cgtaaaaagg ccgcgttgct ggcgtttttc cataggctcc gccccccctga 7080

cgagcatcac aaaaatcgac gctcaagtca gaggtggcga aacccgacag gactataaag 7140

ataccaggcg tttccccctg gaagctccct cgtgcgctct cctgttccga ccctgccgct 7200

taccggatac ctgtccgcct ttctcccttc gggaagcgtg gcgctttctc atagctcacg 7260

ctgtaggtat ctcagttcgg tgtaggtcgt tcgctccaag ctgggctgtg tgcacgaacc 7320

ccccgttcag cccgaccgct gcgccttatc cggtaactat cgtcttgagt ccaacccggt 7380

aagacacgac ttatcgccac tggcagcagc cactggtaac aggattagca gagcgaggta 7440

tgtaggcggt gctacagagt tcttgaagtg gtggcctaac tacggctaca ctagaagaac 7500

agtatttggt atctgcgctc tgctgaagcc agttaccttc ggaaaaagag ttggtagctc 7560

ttgatccggc aaacaaacca ccgctggtag cggtggtttt tttgtttgca agcagcagat 7620

tacgcgcaga aaaaaaggat ctcaagaaga tcctttgatc ttttctacgg ggtctgacgc 7680

tcagtggaac gaaaactcac gttaagggat tttggtcatg agattatcaa aaaggatctt 7740

cacctagatc cttttaaatt aaaaatgaag ttttaaatca atctaaagta tatatgagta 7800

aacttggtct gacagttacc aatgcttaat cagtgaggca cctatctcag cgatctgtct 7860

atttcgttca tccatagttg cctgactccc cgtcgtgtag ataactacga tacgggaggg 7920

cttaccatct ggccccagtg ctgcaatgat accgcgagac ccacgctcac cggctccaga 7980

tttatcagca ataaaccagc cagccggaag ggccgagcgc agaagtggtc ctgcaacttt 8040

atccgcctcc atccagtcta ttaattgttg ccgggaagct agagtaagta gttcgccagt 8100

taatagtttg cgcaacgttg ttgccattgc tacaggcatc gtggtgtcac gctcgtcgtt 8160

tggtatggct tcattcagct ccggttccca acgatcaagg cgagttacat gatcccccat 8220

gttgtgcaaa aaagcggtta gctccttcgg tcctccgatc gttgtcagaa gtaagttggc 8280

cgcagtgtta tcactcatgg ttatggcagc actgcataat tctcttactg tcatgccatc 8340

cgtaagatgc ttttctgtga ctggtgagta ctcaaccaag tcattctgag aatagtgtat 8400

gcggcgaccg agttgctctt gcccggcgtc aatacgggat aataccgcgc cacatagcag 8460

aactttaaaa gtgctcatca ttggaaaacg ttcttcgggg cgaaaactct caaggatctt 8520

accgctgttg agatccagtt cgatgtaacc cactcgtgca cccaactgat cttcagcatc 8580

ttttactttc accagcgttt ctgggtgagc aaaaacagga aggcaaaatg ccgcaaaaaa 8640

gggaataagg gcgacacgga aatgttgaat actcatactc ttccttttc aatattattg 8700

aagcatttat cagggttatt gtctcatgag cggatacata tttgaatgta tttagaaaaa 8760

taaacaaata ggggttccgc gcacatttcc ccgaaaagtg ccacctgatg cggtgtgaaa 8820

taccgcacag atgcgtaagg agaaaatacc gcatcaggaa attgtaagcg ttaatatttt 8880

gttaaaattc gcgttaaatt tttgttaaat cagctcattt tttaaccaat aggccgaaat 8940

cggcaaaatc ccttataaat caaaagaata gaccgagata gggttgagtg ttgttccagt 9000

ttggaacaag agtccactat taaagaacgt ggactccaac gtcaaagggc gaaaaaccgt 9060

ctatcagggc gatggcccac tacgtgaacc atcaccctaa tcaagttttt tggggtcgag 9120

gtgccgtaaa gcactaaatc ggaaccctaa agggagcccc cgatttagag cttgacgggg 9180

aaagccggcg aacgtggcga gaaaggaagg gaagaaagcg aaaggagcgg gcgctagggc 9240

gctggcaagt gtagcggtca cgctgcgcgt aaccaccaca cccgccgcgc ttaatgcgcc 9300

gctacagggc gcgtccattc gccattcagg ctgcgcaact gttgggaagg gcgatcggtg 9360

cgggcctctt cgctattacg ccagctggcg aaaggggga gtgctgcaag gcgattaagt 9420

tgggtaacgc cagggttttc ccagtcacga cgttgtaaaa cgacggccag tgaattgtaa 9480

tacgactcac tatagggcga attgggcccg acgtcgcatg cgtcgagata tcgacattgt 9540

tccatctcca gtttaacccc aacttatcga gagtatttgt gagacacgca ataaatgaat 9600

ttataccaat caaatccata ttctacgctg tctacatata gatacttttt gtcatctctt 9660

gccctactat ttcgtcgata tatgaaggat acgccaaccg aacccatact ccacgctaca 9720

cacgcgcctt ttcacgcatt tctggggaaa atagacaccc ttggtgtcac ctgaagaata 9780

tgaaagaaga tattcattgt attgagctgt agatctgtgt atttcttgac ctcatcaatg 9840

acttctgggc tctttacctc gaatcatggt ggtactgtac cacatctcaa caccttgtag 9900

cacacctatg ggaaaattga gactatgaat ggattcccgt gcccgtatta ctctactaat 9960

ttgatcttgg aacgcgaaaa tacgtttcta ggactccaaa gaatctcaac tcttgtcctt 10020

actaaatata ctacccatag ttgatggttt acttgaacag agaggacatg ttcacttgac 10080

ccaaagtttc tcgcatctct tggatatttg aacaacggcg tccactgacc gtcagttatc 10140

cagtcacaaa acccccacat tcatacattc ccatgtacgt ttacaaagtt ctcaattcca 10200

tcgtgcaaat caaaatcaca tctattcatt catcatatat aaacccatca tgtctactaa 10260

cactcacaac tccatagaaa acatcgactc agaacacacg ctccatctat tcctcgtcca 10320

gctcgcaaat gtcgtcatct taattaaaag gcgttgaaac agaatgagcc agacagcaag 10380

gacaaggtgg ccaacagcaa ggagtccaaa aagccctcta ttgacgagat ccacgatgtt 10440

attgctcatg aggtttccga gctcgatgct gggaagaaga agtgatttgt atataagaaa 10500

taaatgagat atagtaaagg agtgcaagag aatggcaagg tggtcaaatt ctatattact 10560

tgcagtcact ggttcctcgt tgacatgaat gaagttaccg ttggcatagc tgatttaata 10620

tataactgtc caactaactc tcacctagat ataacccatg tgtgtgtttc caatcatcaa 10680

tgcggccgct tagtcgctct tggccttggc tgcagcggca gactctttga gggtgaaaat 10740

ctgagcagtc tcgggcacag ctccgtagtt gacaaagagg tgagcggtct tgaagaaggc 10800

agtaatgatg ggctcgtcgt tacgtctcac gaggtgaggg taagcggcag caaagtgctt 10860

ggtggcttcg ttgagcttgt agtggggaat gatagggaac aagtgatgga cctggtgggt 10920

gccaatgtgg tgagacaggt tgtccacgaa agctccgtac gatcggtcga cggagctcag 10980

gttgcccttg acgtaggtcc actccgagtc accgtaccag ggagtagctt cgtcgttgtg 11040

atgcaagaag gtagtaatga cgaggaacga agcaaagaca aagagaggag catagtagta 11100

gaggcccatg acggcaaagc cgagcgagta tgtgaggtag gcgtaggcag cgaagaaggc 11160

agcccagact ccgagggaca cgatgacagc ggaggctcgt cgaagcagga gagggtccca 11220

ggggtcaaag tgggacatgg ttcgaggagc atatccgacc ttcaggtaga caaaccaagc 11280

acctccgagg gtgtagaccc attgtcgcac gtcctggagg tccttgacgg accgatgagg 11340

gtagaagatc tcgtccttat caatgttgcc agtgttcttg tgatggtgtc ggtgggtcac 11400

tcgccaggac tcgaagggag tcagaatggc agagtgcatg atgcagccaa tgatgaagtt 11460

gacggagtgg tatcgggaga aggcagagtg accacagtcg tgaccgacgg taaagaaacc 11520

ccagaagatg acaccctgca cgtagatgta ggtggcgcaa accagagcgt ggagcagaac 11580

gttatcggca atgaagggag tagatcgggc agcgtagagc agagcagcag aggcagatgc 11640

gttgaagatc gctcgggcag tgtagtagag cgagagtccg aggttggact caaagcaagc 11700

gttagggata gagtgcttca gctcagtcag ggtagggaac tcgaccttgg tcttatcctc 11760

agccatggta ccagagctgg gttagtttgt gtagagagtg tgtgttgcta gcgactttcg 11820

gattgtgtca ttacacaaaa cgcgtcgtct cgacactgat cttgtcgtgg atactcacgg 11880

ctcggaattc tgtgatgtgt agtttagatt tcgaatctgt ggggaaagaa aggaaaaaag 11940

agactggcaa ccgattggga gagccactgt ttatatatac cctagacaag ccccccgctt 12000

gtaagatgtt ggtcaatgta aaccagtatt aaggttggca agtgcaggag aagcaaggtg 12060

tgggtaccga gcaatggaaa tgtgcggaag gcaaaaaaat gaggccacgg cctattgtcg 12120

gggctatatc caggggggcga ttgaagtaca ctaacatgac atgtgtccac agaccctcaa 12180

tctggcctga tgagccaaat ccatacgcgc tttcgcagct ctaaaggcta taacaagtca 12240

caccaccctg ctcgacctca gcgccctcac tttttgttaa gacaaactgt acacgctgtt 12300

ccagcgtttt ctgcctgcac ctggtgggac atttggtgca acctaaagtg ctcggaacct 12360

ctgtggtgtc cagatcagcg cagcagttcc gaggtagttt tgaggccctt agatgatgca 12420

atggtgtcag tcgctggatc acgagtctta atggcagtat tcgttcttat ttgtgccatt 12480

gagccccgtt atcctcgtat cttctacccc ccatcccatc cctttgttgg tgcaacccta 12540

cccatttatt gttgggtgca gcccaaccga cgtggagagc ttggcttggc catataaaaa 12600

ggccccccccc tagtggcaat ggcagaaagt cagctgtgag ttgttgaatt tgtcatctag 12660

gcggcctggc cgtcttctcc ggggcaattt 12690

<210> 159
<211> 5833
<212> DNA
<213> Artificial Sequence

<220>
<223> Plasmid pZKUT16

<400> 159
gtacgaggaa actgtctctg aacagaagaa ggaggacgtc tctgactacg agaactccca 60

gtacaaggag ttcctagtcc cctctcccaa cgagaagctg gccagaggtc tgctcatgct 120

ggccgagctg tcttgcaagg gctctctggc cactggcgag tactccaagc agaccattga 180

gcttgcccga tccgaccccg agtttgtggt tggcttcatt gcccagaacc gacctaaggg 240

cgactctgag gactggctta ttctgacccc cggggtgggt cttgacgaca agggagacgc 300

tctcggacag cagtaccgaa ctgttgagga tgtcatgtct accggaacgg atatcataat 360

tgtcggccga ggtctgtacg gccagaaccg agatcctatt gaggaggcca agcgatacca 420

gaaggctggc tgggaggctt accagaagat taactgttag aggttagact atggatatgt 480

aatttaactg tgtatataga gagcgtgcaa gtatggagcg cttgttcagc ttgtatgatg 540

gtcagacgac ctgtctgatc gagtatgtat gatactgcac aacctgtgta tccgcatgat 600

ctgtccaatg gggcatgttg ttgtgtttct cgatacggag atgctgggta cagtgctaat 660

acgttgaact acttatactt atatgaggct cgaagaaagc tgacttgtgt atgacttaat 720

taatcgagct tggcgtaatc atggtcatag ctgtttcctg tgtgaaattg ttatccgctc 780

acaattccac acaacatacg agccggaagc ataaagtgta aagcctgggg tgcctaatga 840

gtgagctaac tcacattaat tgcgttgcgc tcactgcccg ctttccagtc gggaaacctg 900

tcgtgccagc tgcattaatg aatcggccaa cgcgcgggga gaggcggttt gcgtattggg 960

cgctcttccg cttcctcgct cactgactcg ctgcgctcgg tcgttcggct gcggcgagcg 1020

gtatcagctc actcaaaggc ggtaatacgg ttatccacag aatcagggga taacgcagga 1080

aagaacatgt gagcaaaagg ccagcaaaag gccaggaacc gtaaaaaggc cgcgttgctg 1140

gcgtttttcc ataggctccg ccccccctgac gagcatcaca aaaatcgacg ctcaagtcag 1200

aggtggcgaa acccgacagg actataaaga taccaggcgt ttcccccctgg aagctccctc 1260

gtgcgctctc ctgttccgac cctgccgctt accggatacc tgtccgcctt tctcccttcg 1320

ggaagcgtgg cgctttctca tagctcacgc tgtaggtatc tcagttcggt gtaggtcgtt 1380

cgctccaagc tgggctgtgt gcacgaaccc cccgttcagc ccgaccgctg cgccttatcc 1440

ggtaactatc gtcttgagtc caacccggta agacacgact tatcgccact ggcagcagcc 1500

actggtaaca ggattagcag agcgaggtat gtaggcggtg ctacagagtt cttgaagtgg 1560

tggcctaact acggctacac tagaaggaca gtatttggta tctgcgctct gctgaagcca 1620

gttaccttcg gaaaaagagt tggtagctct tgatccggca aacaaaccac cgctggtagc 1680

ggtggttttt ttgtttgcaa gcagcagatt acgcgcagaa aaaaaggatc tcaagaagat 1740

cctttgatct tttctacggg gtctgacgct cagtggaacg aaaactcacg ttaagggatt 1800

ttggtcatga gattatcaaa aaggatcttc acctagatcc ttttaaatta aaaatgaagt 1860

tttaaatcaa tctaaagtat atatgagtaa acttggtctg acagttacca atgcttaatc 1920

agtgaggcac ctatctcagc gatctgtcta tttcgttcat ccatagttgc ctgactcccc 1980

gtcgtgtaga taactacgat acgggagggc ttaccatctg gccccagtgc tgcaatgata 2040

ccgcgagacc cacgctcacc ggctccagat ttatcagcaa taaaccagcc agccggaagg 2100

gccgagcgca gaagtggtcc tgcaacttta tccgcctcca tccagtctat taattgttgc 2160

cgggaagcta gagtaagtag ttcgccagtt aatagtttgc gcaacgttgt tgccattgct 2220

acaggcatcg tggtgtcacg ctcgtcgttt ggtatggctt cattcagctc cggttcccaa 2280

cgatcaaggc gagttacatg atcccccatg ttgtgcaaaa aagcggttag ctccttcggt 2340

cctccgatcg ttgtcagaag taagttggcc gcagtgttat cactcatggt tatggcagca 2400

ctgcataatt ctcttactgt catgccatcc gtaagatgct tttctgtgac tggtgagtac 2460

tcaaccaagt cattctgaga atagtgtatg cggcgaccga gttgctcttg cccggcgtca 2520

atacgggata ataccgcgcc acatagcaga actttaaaag tgctcatcat tggaaaacgt 2580

tcttcggggc gaaaactctc aaggatctta ccgctgttga gatccagttc gatgtaaccc 2640

actcgtgcac ccaactgatc ttcagcatct tttactttca ccagcgtttc tgggtgagca 2700

aaaacaggaa ggcaaaatgc cgcaaaaaag ggaataaggg cgacacggaa atgttgaata 2760

ctcatactct tccttttca atattattga agcatttatc agggttattg tctcatgagc 2820

ggatacatat ttgaatgtat ttagaaaaat aaacaaatag gggttccgcg cacatttccc 2880

cgaaaagtgc cacctgacgc gccctgtagc ggcgcattaa gcgcggcggg tgtggtggtt 2940

acgcgcagcg tgaccgctac acttgccagc gccctagcgc ccgctccttt cgctttcttc 3000

ccttcctttc tcgccacgtt cgccggcttt ccccgtcaag ctctaaatcg ggggctccct 3060

ttagggttcc gatttagtgc tttacggcac ctcgacccca aaaaacttga ttagggtgat 3120

ggttcacgta gtgggccatc gccctgatag acggtttttc gccctttgac gttggagtcc 3180

acgttcttta atagtggact cttgttccaa actggaacaa cactcaaccc tatctcggtc 3240

tattcttttg atttataagg gattttgccg atttcggcct attggttaaa aaatgagctg 3300

atttaacaaa aatttaacgc gaattttaac aaaatattaa cgcttacaat ttccattcgc 3360

cattcaggct gcgcaactgt tgggaagggc gatcggtgcg ggcctcttcg ctattacgcc 3420

agctggcgaa aggggatgt gctgcaaggc gattaagttg ggtaacgcca gggttttccc 3480

agtcacgacg ttgtaaaacg acggccagtg aattgtaata cgactcacta tagggcgaat 3540

tgggtaccgg gccccccctc gaggtcgacg agtatctgtc tgactcgtca ttgccgcctt 3600

tggagtacga ctccaactat gagtgtgctt ggatcacttt gacgatacat tcttcgttgg 3660

aggctgtggg tctgacagct gcgttttcgg cgcggttggc cgacaacaat atcagctgca 3720

acgtcattgc tggctttcat catgatcaca tttttgtcgg caaaggcgac gcccagagag 3780

ccattgacgt tctttctaat ttggaccgat agccgtatag tccagtctat ctataagttc 3840

aactaactcg taactattac cataacatat acttcactgc cccagataag gttccgataa 3900

aaagttctgc agactaaatt tatttcagtc tcctcttcac caccaaaatg ccctcctacg 3960

aagctcgagt gctcaagctc gtggcagcca agaaaaccaa cctgtgtgct tctctggatg 4020

ttaccaccac caaggagctc attgagcttg ccgataaggt cggaccttat gtgtgcatga 4080

tcaaaaccca tatcgacatc attgacgact tcacctacgc cggcactgtg ctccccctca 4140

aggaacttgc tcttaagcac ggtttcttcc tgttcgagga cagaaagttc gcagatattg 4200

gcaacactgt caagcaccag taccggtgtc accgaatcgc cgagtggtcc gatatcacca 4260

acgcccacgg tgtacccgga accggaatcg atgcagaatt caggagagac cgggttggcg 4320

gcgtatttgt gtcccaaaaa acagccccaa ttgccccaat tgaccccaaa ttgacccagt 4380

agcgggccca accccggcga gagccccctt caccccacat atcaaacctc ccccggttcc 4440

cacacttgcc gttaagggcg tagggtactg cagtctggaa tctacgcttg ttcagacttt 4500

gtactagttt ctttgtctgg ccatccgggt aacccatgcc ggacgcaaaa tagactactg 4560

aaaatttttt tgctttgtgg ttgggacttt agccaagggt ataaaagacc accgtccccg 4620

aattaccttt cctcttcttt tctctctctc cttgtcaact cacacccgaa atcgttaagc 4680

atttccttct gagtataaga atcattcacc atggacatgt ccgtcctgac tctccaagag 4740

tacgagttcg agaagcagtt caacgagaat gaagccatcc aatggatgca ggaaaactgg 4800

aagaaatcct tcctgttttc tgccctctac gctgccttta tctttggtgg acgacatctg 4860

atgaacaagc gagccaagtt tgagctgcga aaacctctcg tgctctggtc cctgaccctc 4920

gctgtcttct ctatcttcgg tgctctgcga actggagcct acatgctcta catcctgatg 4980

accaaaggcc tgaaacagtc tgtttgtgac cagtcctttt acaacggacc cgtctcgaaa 5040

ttctgggctt acgcctttgt gctctccaaa gctcccgaac ttggcgatac catcttcatc 5100

attctgcgaa agcagaaact catcttcctg cactggtatc accacatcac cgtcctcctg 5160

tactcttggt actcctacaa ggacatggtg gctggaggtg gctggttcat gactatgaac 5220

tacggtgtcc acgccgtgat gtactcctac tacgccctcc gagctgccgg tttccgagtc 5280

tctcgaaagt ttgccatgtt catcaccctg tcgcagatca ctcagatgct catgggctgt 5340

gtcattaact acctggtctt caactggatg cagcatgaca atgaccagtg ctactcccac 5400

tttcagaaca tcttctggtc ctctctcatg tacctctcct accttctgct cttctgccat 5460

ttcttctttg aggcctacat tggcaaagtg aagaaagcca ccaaggctga gtaagcggcc 5520

gcaagtgtgg atggggaagt gagtgcccgg ttctgtgtgc acaattggca atccaagatg 5580

gatggattca acacagggat atagcgagct acgtggtggt gcgaggatat agcaacggat 5640

atttatgttt gacacttgag aatgtacgat acaagcactg tccaagtaca atactaaaca 5700

tactgtacat actcatactc gtacccgggc aacggtttca cttgagtgca gtggctagtg 5760

ctcttactcg tacagtgtgc aatactgcgt atcatagtct ttgatgtata tcgtattcat 5820

tcatgttagt tgc 5833

<210> 160
<211> 12663
<212> DNA
<213> Artificial Sequence

<220>
<223> Plasmid pKO2UM25E

<400> 160
agcttccact ggtcggcgtg gtagtggggc agagtggggt cggtgtgctg caggtaggtg 60

atggccacga gccagtggtt gacccacagg taggggatca ggtagtagag ggtgacggaa 120

gccaggcccc atcggttgat ggagtatgcg atgacggaca tggtgatacc aataccgacg 180

ttagagatcc agatgttgaa ccagtccttc ttctcaaaca gcggggcgtt ggggttgaag 240

tggttgacag cccatttgtt gagcttgggg tacttctgtc cggtaacgta agacagcaga 300

tacagaggcc atccaaacac ctgctgggtg atgaggccgt agagggtcat gaggggagcg 360

tcctcagcaa gctcagacca gtcatgggcg cctcggttct ccataaactc ctttcggtcc 420

ttgggcacaa acaccatatc acgggtgagg tgaccagtgg acttgtggtg catggagtgg 480

gtcagcttcc aggcgtagta agggaccagc atggaggagt gcagaaccca tccggtgacg 540

ttgttgacgg tgttagagtc ggagaaagca gagtggccac actcgtgggc aagaacccac 600

agaccggtgc caaacagacc ctggacaatg gagtacatgg cccaggccac agctcggccg 660

gaagccgagg gaataagagg caggtacgcg taggccatgt aggcaaaaac ggcgataaag 720

aagcaggcgc gccagctgca ttaatgaatc ggccaacgcg cggggagagg cggtttgcgt 780

attgggcgct cttccgcttc ctcgctcact gactcgctgc gctcggtcgt tcggctgcgg 840

cgagcggtat cagctcactc aaaggcggta atacggttat ccacagaatc aggggataac 900

gcaggaaaga acatgtgagc aaaaggccag caaaaggcca ggaaccgtaa aaaggccgcg 960

ttgctggcgt ttttccatag gctccgcccc cctgacgagc atcacaaaaa tcgacgctca 1020

agtcagaggt ggcgaaaccc gacaggacta taaagatacc aggcgtttcc ccctggaagc 1080

tccctcgtgc gctctcctgt tccgaccctg ccgcttaccg gatacctgtc cgcctttctc 1140

ccttcgggaa gcgtggcgct ttctcatagc tcacgctgta ggtatctcag ttcggtgtag 1200

gtcgttcgct ccaagctggg ctgtgtgcac gaaccccccg ttcagcccga ccgctgcgcc 1260

ttatccggta actatcgtct tgagtccaac ccggtaagac acgacttatc gccactggca 1320

gcagccactg gtaacaggat tagcagagcg aggtatgtag gcggtgctac agagttcttg 1380

aagtggtggc ctaactacgg ctacactaga agaacagtat ttggtatctg cgctctgctg 1440

aagccagtta ccttcggaaa aagagttggt agctcttgat ccggcaaaca aaccaccgct 1500

ggtagcggtg gttttttgt ttgcaagcag cagattacgc gcagaaaaaa aggatctcaa 1560

gaagatcctt tgatcttttc tacggggtct gacgctcagt ggaacgaaaa ctcacgttaa 1620

gggattttgg tcatgagatt atcaaaaagg atcttcacct agatcctttt aaattaaaaa 1680

tgaagtttta aatcaatcta aagtatatat gagtaaactt ggtctgacag ttaccaatgc 1740

ttaatcagtg aggcacctat ctcagcgatc tgtctatttc gttcatccat agttgcctga 1800

ctccccgtcg tgtagataac tacgatacgg gagggcttac catctggccc cagtgctgca 1860

atgataccgc gagacccacg ctcaccggct ccagatttat cagcaataaa ccagccagcc 1920

ggaagggccg agcgcagaag tggtcctgca actttatccg cctccatcca gtctattaat 1980

tgttgccggg aagctagagt aagtagttcg ccagttaata gtttgcgcaa cgttgttgcc 2040

attgctacag gcatcgtggt gtcacgctcg tcgtttggta tggcttcatt cagctccggt 2100

tcccaacgat caaggcgagt tacatgatcc cccatgttgt gcaaaaaagc ggttagctcc 2160

ttcggtcctc cgatcgttgt cagaagtaag ttggccgcag tgttatcact catggttatg 2220

gcagcactgc ataattctct tactgtcatg ccatccgtaa gatgcttttc tgtgactggt 2280

gagtactcaa ccaagtcatt ctgagaatag tgtatgcggc gaccgagttg ctcttgcccg 2340

gcgtcaatac gggataatac cgcgccacat agcagaactt taaaagtgct catcattgga 2400

aaacgttctt cggggcgaaa actctcaagg atcttaccgc tgttgagatc cagttcgatg 2460

taacccactc gtgcacccaa ctgatcttca gcatctttta ctttcaccag cgtttctggg 2520

tgagcaaaaa caggaaggca aaatgccgca aaaaagggaa taagggcgac acggaaatgt 2580

tgaatactca tactcttcct ttttcaatat tattgaagca tttatcaggg ttattgtctc 2640

atgagcggat acatatttga atgtatttag aaaaataaac aaataggggt tccgcgcaca 2700

tttccccgaa aagtgccacc tgatgcggtg tgaaataccg cacagatgcg taaggagaaa 2760

ataccgcatc aggaaattgt aagcgttaat attttgttaa aattcgcgtt aaatttttgt 2820

taaatcagct catttttaa ccaataggcc gaaatcggca aaatccctta taaatcaaaa 2880

gaatagaccg agatagggtt gagtgttgtt ccagtttgga acaagagtcc actattaaag 2940

aacgtggact ccaacgtcaa agggcgaaaa accgtctatc agggcgatgg cccactacgt 3000

gaaccatcac cctaatcaag ttttttgggg tcgaggtgcc gtaaagcact aaatcggaac 3060

cctaaaggga gcccccgatt tagagcttga cggggaaagc cggcgaacgt ggcgagaaag 3120

gaagggaaga aagcgaaagg agcgggcgct agggcgctgg caagtgtagc ggtcacgctg 3180

cgcgtaacca ccacacccgc cgcgcttaat gcgccgctac agggcgcgtc cattcgccat 3240

tcaggctgcg caactgttgg gaagggcgat cggtgcgggc ctcttcgcta ttacgccagc 3300

tggcgaaagg gggatgtgct gcaaggcgat taagttgggt aacgccaggg ttttcccagt 3360

cacgacgttg taaaacgacg gccagtgaat tgtaatacga ctcactatag ggcgaattgg 3420

gcccgacgtc gcatgcttga atctacaagt aggagggttg gagtgattaa gtgaaacttc 3480

tttaacggct ctatgccagt tctattgata tccgaaacat cagtatgaag gtctgataag 3540

ggtgacttct tcccacagat tcgtatcagt acgagtacga gaccggtact tgtaacagta 3600

ttgatactaa agggaaacta caacggttgt cagcgtaatg tgacttcgcc catgaacgca 3660

gacacgcagt gccgagtgcg gtgatatcgc ctactcgtta cgtccatgga ctacacaacc 3720

cctcggcttc gcttggctta gcctcgggct cggtgctgtt cagttaaaac acaatcaaat 3780

aacatttcta ctttttagaa ggcaggccgt caggagcaac tccgactcca ttgacgtttc 3840

taaacatctg aatgccttcc ttaccttcaa caaactggca ggttcgggcg acagtgtaaa 3900

gagacttgat gaagttggtg tcgtcgtgtc ggtagtgctt gcccatgacc ttcttgatct 3960

tctcagtggc gattcgggcg ttgtagaagg gaattcagga gagaccgggt tggcggcgta 4020

tttgtgtccc aaaaaacagc cccaattgcc ccaattgacc ccaaattgac ccagtagcgg 4080

gcccaacccc ggcgagagcc cccttcaccc cacatatcaa acctcccccg gttcccacac 4140

ttgccgttaa gggcgtaggg tactgcagtc tggaatctac gcttgttcag actttgtact 4200

agtttctttg tctggccatc cgggtaaccc atgccggacg caaaatagac tactgaaaat 4260

ttttttgctt tgtggttggg actttagcca agggtataaa agaccaccgt ccccgaatta 4320

cctttcctct tcttttctct ctctccttgt caactcacac ccgaaatcgt taagcatttc 4380

cttctgagta taagaatcat tcaccatggt cgctggcaag tccggagctg cagcccacgt 4440

gacccactct tccactctcc ctcgagagta ccacggtgct accaacgact cccgatctga 4500

ggctgccgat gtcaccgtct cgtctatcga cgccgagaag gagatgatta tcaacggacg 4560

agtgtacgac gtctcctcgt tcgtgaagcg acaccctggt ggctccgtca tcaagttcca 4620

gctcggagca gatgcttctg acgcttacaa caacttccac gtccgatcga agaaggctga 4680

caagatgctg tactctcttc cctcccgacc tgccgaggct ggctatgcac aggacgacat 4740

ctctagagac tttgagaagc tgcgactgga actgaaagag gaaggttact tcgagcccaa 4800

tctggtgcac gtgtcctacc gatgtgtcga ggtgcttgcc atgtactggg ctggcgtcca 4860

gctgatctgg tccggatact ggttcctcgg tgccatcgtt gctggaattg ctcaaggtcg 4920

atgcggatgg ctccagcatg aaggcggaca ctactcgctc actggcaaca tcaagattga 4980

ccgacatctc cagatggcca tctatggact gggctgtggt atgtctggct gctactggag 5040

aaaccagcac aacaaacatc acgccactcc tcagaagctc ggaaccgatc ccgacctgca 5100

gaccatgcct ctcgttgcct tccacaagat tgtcggagcc aaggcacgag gcaagggtaa 5160

agcctggctt gcttggcaag ctcccctctt ctttggaggc atcatttgct ccctggtctc 5220

tttcggctgg cagttcgttc tccaccccaa tcatgcactg cgagtgcaca accatctcga 5280

actggcctac atgggtctcc gatacgttct ctggcacctt gcctttggcc atctgggact 5340

cctgtcctct cttcgactgt atgccttcta cgtggctgtc ggtggcacct acatcttcac 5400

caacttcgcc gtctcccata ctcacaagga tgtcgttcct cccaccaagc acatttcgtg 5460

ggctctgtac tctgccaacc acactaccaa ctgttccgac tctccctttg tcaactggtg 5520

gatggcctac ctcaacttcc agatcgagca ccatctgttc ccctccatgc ctcagtacaa 5580

ccaccccaag attgctcctc gagtgcgagc actcttcgag aagcacggag tcgagtacga 5640

cgtccgaccc tatctggaat gctttcgagt gacctacgtc aacctccttg ctgttggcaa 5700

ccctgagcac tcctaccacg agcatactca ctaagcggcc gcaagtgtgg atggggaagt 5760

gagtgcccgg ttctgtgtgc acaattggca atccaagatg gatggattca acacagggat 5820

atagcgagct acgtggtggt gcgaggatat agcaacggat atttatgttt gacacttgag 5880

aatgtacgat acaagcactg tccaagtaca atactaaaca tactgtacat actcatactc 5940

gtacccgggc aacggtttca cttgagtgca gtggctagtg ctcttactcg tacagtgtgc 6000

aatactgcgt atcatagtct ttgatgtata tcgtattcat tcatgtcgac gagtatctgt 6060

ctgactcgtc attgccgcct ttggagtacg actccaacta tgagtgtgct tggatcactt 6120

tgacgataca ttcttcgttg gaggctgtgg gtctgacagc tgcgttttcg gcgcggttgg 6180

ccgacaacaa tatcagctgc aacgtcattg ctggctttca tcatgatcac atttttgtcg 6240

gcaaaggcga cgcccagaga gccattgacg ttctttctaa tttggaccga tagccgtata 6300

gtccagtcta tctataagtt caactaactc gtaactatta ccataacata tacttcactg 6360

ccccagataa ggttccgata aaaagttctg cagactaaat ttatttcagt ctcctcttca 6420

ccaccaaaat gccctcctac gaagctcgag ctaacgtcca caagtccgcc tttgccgctc 6480

gagtgctcaa gctcgtggca gccaagaaaa ccaacctgtg tgcttctctg gatgttacca 6540

ccaccaagga gctcattgag cttgccgata aggtcggacc ttatgtgtgc atgatcaaaa 6600

cccatatcga catcattgac gacttcacct acgccggcac tgtgctcccc ctcaaggaac 6660

ttgctcttaa gcacggtttc ttcctgttcg aggacagaaa gttcgcagat attggcaaca 6720

ctgtcaagca ccagtaccgg tgtcaccgaa tcgccgagtg gtccgatatc accaacgccc 6780

acggtgtacc cggaaccgga atcattgctg gcctgcgagc tggtgccgag gaaactgtct 6840

ctgaacagaa gaaggaggac gtctctgact acgagaactc ccagtacaag gagttcctag 6900

tcccctctcc caacgagaag ctggccagag gtctgctcat gctggccgag ctgtcttgca 6960

agggctctct ggccactggc gagtactcca agcagaccat tgagcttgcc cgatccgacc 7020

ccgagtttgt ggttggcttc attgcccaga accgacctaa gggcgactct gaggactggc 7080

ttattctgac ccccggggtg ggtcttgacg acaagggaga cgctctcgga cagcagtacc 7140

gaactgttga ggatgtcatg tctaccggaa cggatatcat aattgtcggc cgaggtctgt 7200

acggccagaa ccgagatcct attgaggagg ccaagcgata ccagaaggct ggctgggagg 7260

cttaccagaa gattaactgt tagaggttag actatggata tgtaatttaa ctgtgtatat 7320

agagagcgtg caagtatgga gcgcttgttc agcttgtatg atggtcagac gacctgtctg 7380

atcgagtatg tatgatactg cacaacctgt gtatccgcat gatctgtcca atggggcatg 7440

ttgttgtgtt tctcgatacg gagatgctgg gtacagtgct aatacgttga actacttata 7500

cttatatgag gctcgaagaa agctgacttg tgtatgactt aattaatttg aatcgaatcg 7560

atgagcctaa aatgaacccg agtatatctc ataaaattct cggtgagagg tctgtgactg 7620

tcagtacaag gtgccttcat tatgccctca accttaccat acctcactga atgtagtgta 7680

cctctaaaaa tgaaatacag tgccaaaagc caaggcactg agctcgtcta acggacttga 7740

tatacaacca attaaaacaa atgaaaagaa atacagttct ttgtatcatt tgtaacaatt 7800

accctgtaca aactaaggta ttgaaatccc acaatattcc caaagtccac ccctttccaa 7860

attgtcatgc ctacaactca tataccaagc actaacctac cgtttaaaca gtgtacgcag 7920

atctactata gaggaacatt taaattgccc cggagaagac ggccaggccg cctagatgac 7980

aaattcaaca actcacagct gactttctgc cattgccact agggggggggc ctttttatat 8040

ggccaagcca agctctccac gtcggttggg ctgcacccaa caataaatgg gtagggttgc 8100

accaacaaag ggatgggatg gggggtagaa gatacgagga taacggggct caatggcaca 8160

aataagaacg aatactgcca ttaagactcg tgatccagcg actgacacca ttgcatcatc 8220

taagggcctc aaaactacct cggaactgct gcgctgatct ggacaccaca gaggttccga 8280

gcactttagg ttgcaccaaa tgtcccacca ggtgcaggca gaaaacgctg gaacagcgtg 8340

tacagtttgt cttaacaaaa agtgagggcg ctgaggtcga gcagggtggt gtgacttgtt 8400

atagccttta gagctgcgaa agcgcgtatg gatttggctc atcaggccag attgagggtc 8460

tgtggacaca tgtcatgtta gtgtacttca atcgcccccc ggatatagcc ccgacaatag 8520

gccgtggcct catttttttg ccttccgcac atttccattg ctcggtaccc acaccttgct 8580

tctcctgcac ttgccaacct taatactggt ttacattgac caacatctta caagcggggg 8640

gcttgtctag ggtatatata aacagtggct ctcccaatcg gttgccagtc tctttttccc 8700

tttctttccc cacagattcg aaatctaaac tacacatcac acaatgcctg ttactgacgt 8760

ccttaagcga aagtccggtg tcatcgtcgg cgacgatgtc cgagccgtga gtatccacga 8820

caagatcagt gtcgagacga cgcgttttgt gtaatgacac aatccgaaag tcgctagcaa 8880

cacacactct ctacacaaac taacccagct ctccatggca cctcccaaca ctatcgacgc 8940

tggcttgacc cagcgtcata tcaccaccac ggccgcccca acctcggcca agcccgcttt 9000

cgagcgcaac taccagctcc ccgagttcac catcaaggag atccgagagt gcatccctgc 9060

ccactgcttt gagcgctccg gtcttcgtgg tctctgccac gttgccattg atctgacctg 9120

ggcctcgctc ttgttcctgg ctgcaaccca gatcgacaag ttcgagaacc ccttgatccg 9180

ctatctggcc tggcctgcgt actggatcat gcagggcatt gtctgcaccg gcatatgggt 9240

gctggcccac gagtgcggtc accagtcctt ctcgacctcc aagactctca acaacaccgt 9300

cggctggatc ctgcactcga tgctcttggt cccctaccac tcctggagaa tctcgcactc 9360

gaagcaccac aaggccactg gccacatgac caaggaccag gtctttgttc ccaagacccg 9420

ctcccaggtt ggtttgcctc ccaaggagag cgctgctgct gccgttcaag aggaggacat 9480

gtccgtgcac ctggatgagg aggcccctat tgtgactttg ttctggatgg tgatccagtt 9540

cctgttcgga tggcctgcat acctgatcat gaacgcctct ggtcaggact atggccgctg 9600

gacctcgcac ttccacactt actcgcccat ctttgagccc cgcaacttct tcgacattat 9660

catctcggat ctcggtgtgt tggctgccct cggtgccctg atctacgctt ccatgcagct 9720

gtcgctcttg accgtgacca agtactacat catcccgtac ctgtttgtca acttttggtt 9780

ggtcctgatt actttcttgc agcacaccga ccccaagctg ccccattacc gtgagggtgc 9840

ctggaacttc cagcgtggag ccctctgcac cgttgaccgc tcgtttggca agttcttgga 9900

ccatatgttc cacggcatcg tccataccca tgtggcccat cacttgttct cgcagatgcc 9960

gttctaccat gctgaagaag ctacctacca tctcaagaaa ctgctgggag agtactacgt 10020

ttacgaccca tccccgatcg tcgttgcggt ctggaggtcg ttccgcgagt gccgattcgt 10080

ggaggatcag ggagacgtgg tctttttcaa gaagtaagcg gccgcaagtg tggatgggga 10140

agtgagtgcc cggttctgtg tgcacaattg gcaatccaag atggatggat tcaacacagg 10200

gatatagcga gctacgtggt ggtgcgagga tatagcaacg gatatttatg tttgacactt 10260

gagaatgtac gatacaagca ctgtccaagt acaatactaa acatactgta catactcata 10320

ctcgtacccg ggcaacggtt tcacttgagt gcagtggcta gtgctcttac tcgtacagtg 10380

tgcaatactg cgtatcatag tctttgatgt atatcgtatt cattcatgtt agttgcgtac 10440

gagatcgtca agggtttgtg gccaactggt atttaaatgt agctaacggt agcaggcgaa 10500

ctactggtac atacctcccc cggaatatgt acaggcataa tgcgtatctg tgggacatgt 10560

ggtcgttgcg ccattatgta agcagcgtgt actcctctga ctgtccatat ggtttgctcc 10620

atctcaccct catcgttttc attgttcaca ggcggccaca aaaaaactgt cttctctcct 10680

tctctcttcg ccttagtcta ctcggaccag ttttagttta gcttggcgcc actggataaa 10740

tgagacctca ggccttgtga tgaggaggtc acttatgaag catgttagga ggtgcttgta 10800

tggatagaga agcacccaaa ataataagaa taataataaa acaggggggcg ttgtcatttc 10860

atatcgtgtt ttcaccatca atacacctcc aaacaatgcc cttcatgtgg ccagccccaa 10920

tattgtcctg tagttcaact ctatgcagct cgtatcttat tgagcaagta aaactctgtc 10980

agccgatatt gcccgacccg cgacaagggt caacaaggtg gtgtaaggcc ttcgcagaag 11040

tcaaaactgt gccaaacaaa catctagagt ctctttggtg tttctcgcat atatttwatc 11100

ggctgtctta cgtatttgcg cctcggtacc ggactaattt cggatcatcc ccaatacgct 11160

ttttcttcgc agctgtcaac agtgtccatg atctatccac ctaaatgggt catatgaggc 11220

gtataatttc gtggtgctga taataattcc catatatttg acacaaaact tccccccccta 11280

gacatacatc tcacaatctc acttcttgtg cttctgtcac acatctcctc cagctgactt 11340

caactcacac ctctgcccca gttggtctac agcggtataa ggtttctccg catagaggtg 11400

caccactcct cccgatactt gtttgtgtga cttgtgggtc acgacatata tatctacaca 11460

cattgcgcca ccctttggtt cttccagcac aacaaaaaca cgacacgcta accatggagt 11520

ccattgctcc cttcctgccc tccaagatgc ctcaggacct gttcatggac ctcgccagcg 11580

ctatcggtgt ccgagctgct ccctacgtcg atcccctgga ggctgccctg gttgcccagg 11640

ccgagaagta cattcccacc attgtccatc acactcgagg cttcctggtt gccgtggagt 11700

ctcccctggc tcgagagctg cctctgatga acccccttcca cgtgctcctg atcgtgctcg 11760

cctacctggt caccgtgttt gtgggtatgc agatcatgaa gaactttgaa cgattcgagg 11820

tcaagacctt ctccctcctg cacaacttct gtctggtctc catctccgcc tacatgtgcg 11880

gtggcatcct gtacgaggct tatcaggcca actatggact gtttgagaac gctgccgatc 11940

acaccttcaa gggtctccct atggctaaga tgatctggct cttctacttc tccaagatca 12000

tggagtttgt cgacaccatg atcatggtcc tcaagaagaa caaccgacag atttcctttc 12060

tgcacgtgta ccaccactct tccatcttca ccatctggtg gctggtcacc ttcgttgctc 12120

ccaacggtga agcctacttc tctgctgccc tgaactcctt catccacgtc atcatgtacg 12180

gctactactt tctgtctgcc ctgggcttca agcaggtgtc gttcatcaag ttctacatca 12240

ctcgatccca gatgacccag ttctgcatga tgtctgtcca gtcttcctgg gacatgtacg 12300

ccatgaaggt ccttggccga cctggatacc ccttcttcat caccgctctg ctctggttct 12360

acatgtggac catgctcggt ctcttctaca acttttaccg aaagaacgcc aagctcgcca 12420

agcaggccaa ggctgacgct gccaaggaga aggccagaaa gctccagtaa gcggccgcca 12480

ccgcggcccg agattccggc ctcttcggcc gccaagcgac ccgggtggac gtctagaggt 12540

acctagcaat taacagatag tttgccggtg ataattctct taacctccca cactcctttg 12600

acataacgat ttatgtaacg aaactgaaat ttgaccagat attgtgtccg cggtggagct 12660

cca 12663

<210> 161
<211> 15996
<212> DNA
<213> Artificial Sequence

<220>
<223> Plasmid pDMW303

<220>
<221> misc_feature
<222> (5344)..(5344)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (5349)..(5352)
<223> n is a, c, g, or t

<400> 161
gtacgatagt tagtagacaa caatcagaac atctccctcc ttatataatc acacaggcca 60

gaacgcgcta aactaaagcg ctttggacac tatgttacat tggcattgat tgaactgaaa 120

ccacagtctc cctcgcctga atcgagcaat ggatgttgtc ggaagtcaac ttcactagaa 180

gagcggttct atgccttgtc aagatcatat cataaactca ctctgtatta ccccatctat 240

agaacacttg ttatgaatgg gcggaaacat tccgcttata tgcacctttc cacactaatg 300

caaagatgtg catcttcaac gggtagtaag actggttccg acttccgttg catggagagc 360

aatgacctcg ataatgcgaa catcccccac atatacactc ttacacaggc caatataatc 420

tgtgcattta ctaaatattt aagtctatgc acctgcttga tgaaaagcgg cagcggatgg 480

tatcatctag tttccgccaa tccaagaacc aactgtgttg gcagtggtgt agcccatggc 540

acacagacca aagatgaaaa tacagacatc ggcggttcga gccgtggtgc ctcgagcaac 600

acccttgtaa tgcaaaagag gagggtaaat gtacaccaga ggcacacatg caaacgatcc 660

ggtgagagcg acgaaccgat cgagatcgtc ggcacctccc catgcaacaa aggcggtgac 720

aaacacaagg aagaaccgga aaatgttctt ctgccacttg atggtagagt tgtacttgcc 780

tgatcgggtg aagagaccat tctcgatgat tcggatggcg cgccagctgc attaatgaat 840

cggccaacgc gcggggagag gcggtttgcg tattgggcgc tcttccgctt cctcgctcac 900

tgactcgctg cgctcggtcg ttcggctgcg gcgagcggta tcagctcact caaaggcggt 960

aatacggtta tccacagaat caggggataa cgcaggaaag aacatgtgag caaaaggcca 1020

gcaaaaggcc aggaaccgta aaaaggccgc gttgctggcg tttttccata ggctccgccc 1080

ccctgacgag catcacaaaa atcgacgctc aagtcagagg tggcgaaacc cgacaggact 1140

ataaagatac caggcgtttc cccctggaag ctccctcgtg cgctctcctg ttccgaccct 1200

gccgcttacc ggatacctgt ccgcctttct cccttcggga agcgtggcgc tttctcatag 1260

ctcacgctgt aggtatctca gttcggtgta ggtcgttcgc tccaagctgg gctgtgtgca 1320

cgaacccccc gttcagcccg accgctgcgc cttatccggt aactatcgtc ttgagtccaa 1380

cccggtaaga cacgacttat cgccactggc agcagccact ggtaacagga ttagcagagc 1440

gaggtatgta ggcggtgcta cagagttctt gaagtggtgg cctaactacg gctacactag 1500

aagaacagta tttggtatct gcgctctgct gaagccagtt accttcggaa aaagagttgg 1560

tagctcttga tccggcaaac aaaccaccgc tggtagcggt ggtttttttg tttgcaagca 1620

gcagattacg cgcagaaaaa aaggatctca agaagatcct ttgatctttt ctacggggtc 1680

tgacgctcag tggaacgaaa actcacgtta agggattttg gtcatgagat tatcaaaaag 1740

gatcttcacc tagatccttt taaattaaaa atgaagtttt aaatcaatct aaagtatata 1800

tgagtaaact tggtctgaca gttaccaatg cttaatcagt gaggcaccta tctcagcgat 1860

ctgtctattt cgttcatcca tagttgcctg actccccgtc gtgtagataa ctacgatacg 1920

ggagggctta ccatctggcc ccagtgctgc aatgataccg cgagacccac gctcaccggc 1980

tccagattta tcagcaataa accagccagc cggaagggcc gagcgcagaa gtggtcctgc 2040

aactttatcc gcctccatcc agtctattaa ttgttgccgg gaagctagag taagtagttc 2100

gccagttaat agtttgcgca acgttgttgc cattgctaca ggcatcgtgg tgtcacgctc 2160

gtcgtttggt atggcttcat tcagctccgg ttcccaacga tcaaggcgag ttacatgatc 2220

ccccatgttg tgcaaaaaag cggttagctc cttcggtcct ccgatcgttg tcagaagtaa 2280

gttggccgca gtgttatcac tcatggttat ggcagcactg cataattctc ttactgtcat 2340

gccatccgta agatgctttt ctgtgactgg tgagtactca accaagtcat tctgagaata 2400

gtgtatgcgg cgaccgagtt gctcttgccc ggcgtcaata cgggataata ccgcgccaca 2460

tagcagaact ttaaaagtgc tcatcattgg aaaacgttct tcggggcgaa aactctcaag 2520

gatcttaccg ctgttgagat ccagttcgat gtaacccact cgtgcaccca actgatcttc 2580

agcatctttt actttcacca gcgtttctgg gtgagcaaaa acaggaaggc aaaatgccgc 2640

aaaaaaggga ataagggcga cacggaaatg ttgaatactc atactcttcc tttttcaata 2700

ttattgaagc atttatcagg gttattgtct catgagcgga tacatatttg aatgtattta 2760

gaaaaataaa caaatagggg ttccgcgcac atttccccga aaagtgccac ctgatgcggt 2820

gtgaaatacc gcacagatgc gtaaggagaa aataccgcat caggaaattg taagcgttaa 2880

tattttgtta aaattcgcgt taaatttttg ttaaatcagc tcatttttta accaataggc 2940

cgaaatcggc aaaatcccttataaatcaaa agaatagacc gagatagggt tgagtgttgt 3000

tccagtttgg aacaagagtc cactattaaa gaacgtggac tccaacgtca aagggcgaaa 3060

aaccgtctat cagggcgatg gcccactacg tgaaccatca ccctaatcaa gtttttttggg 3120

gtcgaggtgc cgtaaagcac taaatcggaa ccctaaaggg agcccccgat ttagagcttg 3180

acggggaaag ccggcgaacg tggcgagaaa ggaagggaag aaagcgaaag gagcgggcgc 3240

tagggcgctg gcaagtgtag cggtcacgct gcgcgtaacc accacacccg ccgcgcttaa 3300

tgcgccgcta cagggcgcgt ccattcgcca ttcaggctgc gcaactgttg ggaagggcga 3360

tcggtgcggg cctcttcgct attacgccag ctggcgaaag ggggatgtgc tgcaaggcga 3420

ttaagttggg taacgccagg gttttcccag tcacgacgtt gtaaaacgac ggccagtgaa 3480

ttgtaatacg actcactata gggcgaattg ggcccgacgt cgcatgctta gaagtgagga 3540

ttacaagaag cctctggata tcaatgatga acgtactcag cggctggtca agcatttcga 3600

ccgtcgaatc gacgaggtgt tcacctttga caagcgaggg ttcccaattg atcacgttct 3660

cgagttgttc aaatcttctc tcaacatctc tctgcatgaa ctatctctgt tgacgaacgt 3720

gtcacccact gttcctcgaa cgcccttctc cgagtttggt ctgaacatct tcgatctcaa 3780

actgaccccc gcagtgatca atagtgccat ggccactgcc gatgcggtgc gaacatccct 3840

ggagggattc tcggagctct acacaatgca gattctgtcg tcgagtactc tctaccttgc 3900

tcggatgact tattgtgcta ctactgcact catgcttcga tcatgtgccc tactgcaccc 3960

caaatttggt gatctgattg agacagagta ccctcttcag ctgattcaga agatcatcag 4020

caacatgaat gatgtggttg accaggcagg ctgttgtagt cacgtccttc acttcaagtt 4080

cattcttcat ctgcttctgt tttactttga caggcaaatg aagacatggt acgacttgat 4140

ggaggccaag aacgccattt caccccgaga caccgaagtg cctgaaatcc tggctgcccc 4200

cattgataac atcggaaact acggtattcc ggaaagtgta tatagaacct ttccccagct 4260

tgtgtctgtg gatatggatg gtgtaatccc cttaattaaa aggcgttgaa acagaatgag 4320

ccagacagca aggacaaggt ggccaacagc aaggagtcca aaaagccctc tattgacgag 4380

atccacgatg ttattgctca tgaggtttcc gagctcgatg ctgggaagaa gaagtgattt 4440

gtatataaga aataaatgag atatagtaaa ggagtgcaag agaatggcaa ggtggtcaaa 4500

ttctatatta cttgcagtca ctggttcctc gttgacatga atgaagttac cgttggcata 4560

gctgatttaa tatataactg tccaactaac tctcacctag atataaccca tgtgtgtgtt 4620

tccaatcatc aatgcggccg cgcctactta agcaacgggc ttgataacag cgggggggt 4680

gcccacgttg ttgcggttgc ggaagaacag aacaccctta ccagcaccct cggcaccagc 4740

gctgggctca acccactggc acatacgcgc actgcggtac atggcgcgga tgaagccacg 4800

aggaccatcc tggacatcag cccggtagtg cttgcccatg atgggcttaa tggcctcggt 4860

ggcctcgtcc gcgttgtaga aggggatgct gctgacgtag tggtggagga catgagtctc 4920

gatgatgccg tggagaaggt ggcggccgat gaagcccatc tcacggtcaa tggtagcagc 4980

ggcaccacgg acgaagttcc actcgtcgtt ggtgtagtgg ggaagggtag ggtcggtgtg 5040

ctggaggaag gtgatggcaa cgagccagtg gttaacccag aggtagggaa caaagtacca 5100

gatggccatg ttgtagaaac cgaacttctg aacgaggaag tacagagcag tggccatcag 5160

accgatacca atatcgctga ggacgatgag cttagcgtca ctgttctcgt acagagggct 5220

gcggggatcg aagtggttaa caccaccgcc gaggccgtta tgcttgccct tgccgcgacc 5280

ctcacgctgg cgctcgtggt agttgtggcc ggtaacattg gtgatgaggt agttgggcca 5340

gccnacgann nnctcagtaa gatgagcgag ctcgtgggtc atctttccga gacgagtagc 5400

ctgctgctcg cgggttcggg gaacgaagac catgtcacgc tccatgttgc cagtggcctt 5460

gtggtgcttt cggtgggaga tttgccagct gaagtagggg acaaggaggg aagagtgaag 5520

aacccagcca gtaatgtcgt tgatgatgcg agaatcggag aaagcaccgt gaccgcactc 5580

atgggcaata acccagagac cagtaccgaa aagaccctga agaacggtgt acacggccca 5640

cagaccagcg cgggcggggg tggaggggat atattcgggg gtcacaaagt tgtaccagat 5700

gctgaaagtg gtagtcagga ggacaatgtc gcggaggata taaccgtatc ccttgagagc 5760

ggagcgcttg aagcagtgct tagggatggc attgtagatg tccttgatgg taaagtcggg 5820

aacctcgaac tggttgccgt aggtgtcgag catgacacca tactcggact tgggcttggc 5880

gatatcaacc tcggacatgg acgagagcga tgtggaagag gccgagtggc ggggagagtc 5940

tgaaggagag acggcggcag actcagaatc cgtcacagta gttgaggtga cggtgcgtct 6000

aagcgcaggg ttctgcttgg gcagagccga agtggacgcc atggtgaatg attcttatac 6060

tcagaaggaa atgcttaacg atttcgggtg tgagttgaca aggagagaga gaaaagaaga 6120

ggaaaggtaa ttcggggacg gtggtctttt ataccccttgg ctaaagtccc aaccacaaag 6180

caaaaaaatt ttcagtagtc tattttgcgt ccggcatggg ttacccggat ggccagacaa 6240

agaaactagt acaaagtctg aacaagcgta gattccagac tgcagtaccc tacgccctta 6300

acggcaagtg tgggaaccgg gggaggtttg atatgtgggg tgaaggggggc tctcgccggg 6360

gttgggcccg ctactgggtc aatttggggt caattggggc aattggggct gtttttttggg 6420

acacaaatac gccgccaacc cggtctctcc tgaattctct ctcttgagct tttccataac 6480

aagttcttct gcctccagga agtccatggg tggtttgatc atggttttgg tgtagtggta 6540

gtgcagtggt ggtattgtga ctggggatgt agttgagaat aagtcataca caagtcagct 6600

ttcttcgagc ctcatataag tataagtagt tcaacgtatt agcactgtac ccagcatctc 6660

cgtatcgaga aacacaacaa catgccccat tggacagatc atgcggatac acaggttgtg 6720

cagtatcata catactcgat cagacaggtc gtctgaccat catacaagct gaacaagcgc 6780

tccatacttg cacgctctct atatacacag ttaaattaca tatccatagt ctaacctcta 6840

acagttaatc ttctggtaag cctcccagcc agccttctgg tatcgcttgg cctcctcaat 6900

aggatctcgg ttctggccgt acagacctcg gccgacaatt atgatatccg ttccggtaga 6960

catgacatcc tcaacagttc ggtactgctg tccgagagcg tctcccttgt cgtcaagacc 7020

caccccgggg gtcagaataa gccagtcctc agagtcgccc ttaggtcggt tctgggcaat 7080

gaagccaacc acaaactcgg ggtcggatcg ggcaagctca atggtctgct tggagtactc 7140

gccagtggcc agagagccct tgcaagacag ctcggccagc atgagcagac ctctggccag 7200

cttctcgttg ggagagggga ctaggaactc cttgtactgg gagttctcgt agtcagagac 7260

gtcctccttc ttctgttcag agacagtttc ctcggcacca gctcgcaggc cagcaatgat 7320

tccggttccg ggtacaccgt gggcgttggt gatatcggac cactcggcga ttcggtgaca 7380

ccggtactgg tgcttgacag tgttgccaat atctgcgaac tttctgtcct cgaacaggaa 7440

gaaaccgtgc ttaagagcaa gttccttgag ggggagcaca gtgccggcgt aggtgaagtc 7500

gtcaatgatg tcgatatggg ttttgatcat gcacacataa ggtccgacct tatcggcaag 7560

ctcaatgagc tccttggtgg tggtaacatc cagagaagca cacaggttgg ttttcttggc 7620

tgccacgagc ttgagcactc gagcggcaaa ggcggacttg tggacgttag ctcgagcttc 7680

gtaggagggc attttggtgg tgaagaggag actgaaataa atttagtctg cagaactttt 7740

tatcggaacc ttatctgggg cagtgaagta tatgttatgg taatagttac gagttagttg 7800

aacttataga tagactggac tatacggcta tcggtccaaa ttagaaagaa cgtcaatggc 7860

tctctgggcg tcgcctttgc cgacaaaaat gtgatcatga tgaaagccag caatgacgtt 7920

gcagctgata ttgttgtcgg ccaaccgcgc cgaaaacgca gctgtcagac ccacagcctc 7980

caacgaagaa tgtatcgtca aagtgatcca agcacactca tagttggagt cgtactccaa 8040

aggcggcaat gacgagtcag acagatactc gtcgactcat cgatcgagga agaggacaag 8100

cggctgcttc ttaagtttgt gacatcagta tccaaggcac cattgcaagg attcaaggct 8160

ttgaacccgt catttgccat tcgtaacgct ggtagacagg ttgatcggtt ccctacggcc 8220

tccacctgtg tcaatcttct caagctgcct gactatcagg acattgatca acttcggaag 8280

aaacttttgt atgccattcg atcacatgct ggtttcgatt tgtcttagag gaacgcatat 8340

acagtaatca tagagaataa acgatattca tttattaaag tagatagttg aggtagaagt 8400

tgtaaagagt gataaatagc ggccgcttac tggagctttc tggccttctc cttggcagcg 8460

tcagccttgg cctgcttggc gagcttggcg ttctttcggt aaaagttgta gaagagaccg 8520

agcatggtcc acatgtagaa ccagagcaga gcggtgatga agaaggggta tccaggtcgg 8580

ccaaggacct tcatggcgta catgtcccag gaagactgga cagacatcat gcagaactgg 8640

gtcatctggg atcgagtgat gtagaacttg atgaacgaca cctgcttgaa gcccagggca 8700

gacagaaagt agtagccgta catgatgacg tggatgaagg agttcagggc agcagagaag 8760

taggcttcac cgttgggagc aacgaaggtg accagccacc agatggtgaa gatggaagag 8820

tggtggtaca cgtgcagaaa ggaaatctgt cggttgttct tcttgaggac catgatcatg 8880

gtgtcgacaa actccatgat cttggagaag tagaagagcc agatcatctt agccataggg 8940

agacccttga aggtgtgatc ggcagcgttc tcaaacagtc catagttggc ctgataagcc 9000

tcgtacagga tgccaccgca catgtaggcg gagatggaga ccagacagaa gttgtgcagg 9060

agggagaagg tcttgacctc gaatcgttca aagttcttca tgatctgcat acccacaaac 9120

acggtgacca ggtaggcgag cacgatcagg agcacgtgga aggggttcat cagaggcagc 9180

tctcgagcca ggggagactc cacggcaacc aggaagcctc gagtgtgatg gacaatggtg 9240

ggaatgtact tctcggcctg ggcaaccagg gcagcctcca ggggatcgac gtagggagca 9300

gctcggacac cgatagcgct ggcgaggtcc atgaacaggt cctgaggcat cttggagggc 9360

aggaagggag caatggactc catgggcagg acctgtgtta gtacattgtc ggggagtcat 9420

caattggttc gacaggttgt cgactgttag tatgagctca attgggctct ggtgggtcga 9480

tgacacttgt catctgtttc tgttgggtca tgtttccatc accttctatg gtactcacaa 9540

ttcgtccgat tcgcccgaat ccgttaatac cgactttgat ggccatgttg atgtgtgttt 9600

aattcaagaa tgaatataga gaagagaaga agaaaaaaga ttcaattgag ccggcgatgc 9660

agacccttat ataaatgttg ccttggacag acggagcaag cccgcccaaa cctacgttcg 9720

gtataatatg ttaagctttt taacacaaag gtttggcttg gggtaacctg atgtggtgca 9780

aaagaccggg cgttggcgag ccattgcgcg ggcgaatggg gccgtgactc gtctcaaatt 9840

cgagggcgtg cctcaattcg tgcccccgtg gctttttccc gccgtttccg ccccgtttgc 9900

accactgcag ccgcttcttt ggttcggaca ccttgctgcg agctaggtgc cttgtgctac 9960

ttaaaaagtg gcctcccaac accaacatga catgagtgcg tgggccaaga cacgttggcg 10020

gggtcgcagt cggctcaatg gcccggaaaa aacgctgctg gagctggttc ggacgcagtc 10080

cgccgcggcg tatggatatc cgcaaggttc catagcgcca ttgccctccg tcggcgtcta 10140

tcccgcaacc tctaaataga gcgggaatat aacccaagct tctttttttt cctttaacac 10200

gcacaccccc aactatcatg ttgctgctgc tgtttgactc tactctgtgg aggggtgctc 10260

ccacccaacc caacctacag gtggatccgg cgctgtgatt ggctgataag tctcctatcc 10320

ggactaattc tgaccaatgg gacatgcgcg caggacccaa atgccgcaat tacgtaaccc 10380

caacgaaatg cctacccctc tttggagccc agcggcccca aatcccccca agcagcccgg 10440

ttctaccggc ttccatctcc aagcacaagc agcccggttc taccggcttc catctccaag 10500

cacccctttc tccacacccc acaaaaagac ccgtgcagga catcctactg cgtgtttaaa 10560

caccactaaa accccacaaa atatatctta ccgaatatac agatctacta tagaggaaca 10620

attgccccgg agaagacggc caggccgcct agatgacaaa ttcaacaact cacagctgac 10680

tttctgccat tgccactagg gggggggcctt tttatatggc caagccaagc tctccacgtc 10740

ggttgggctg cacccaacaa taaatgggta gggttgcacc aacaaaggga tgggatgggg 10800

ggtagaagat acgaggataa cggggctcaa tggcacaaat aagaacgaat actgccatta 10860

agactcgtga tccagcgact gacaccattg catcatctaa gggcctcaaa actacctcgg 10920

aactgctgcg ctgatctgga caccacagag gttccgagca ctttaggttg caccaaatgt 10980

cccaccaggt gcaggcagaa aacgctggaa cagcgtgtac agtttgtctt aacaaaaagt 11040

gagggcgctg aggtcgagca gggtggtgtg acttgttata gcctttagag ctgcgaaagc 11100

gcgtatggat ttggctcatc aggccagatt gagggtctgt ggacacatgt catgttagtg 11160

tacttcaatc gccccctgga tatagccccg acaataggcc gtggcctcat ttttttgcct 11220

tccgcacatt tccattgctc ggtacccaca ccttgcttct cctgcacttg ccaaccttaa 11280

tactggttta cattgaccaa catcttacaa gcgggggggct tgtctagggt atatataaac 11340

agtggctctc ccaatcggtt gccagtctct tttttccttt ctttccccac agattcgaaa 11400

tctaaactac acatcacaca atgcctgtta ctgacgtcct taagcgaaag tccggtgtca 11460

tcgtcggcga cgatgtccga gccgtgagta tccacgacaa gatcagtgtc gagacgacgc 11520

gttttgtgta atgacacaat ccgaaagtcg ctagcaacac acactctcta cacaaactaa 11580

cccagctctc catggctgcc gctccctctg tgcgaacctt tacccgagcc gaggttctga 11640

acgctgaggc tctgaacgag ggcaagaagg acgctgaggc tcccttcctg atgatcatcg 11700

acaacaaggt gtacgacgtc cgagagttcg tccctgacca tcctggaggc tccgtgattc 11760

tcacccacgt tggcaaggac ggcaccgacg tctttgacac ctttcatccc gaggctgctt 11820

gggagactct cgccaacttc tacgttggag acattgacga gtccgaccga gacatcaaga 11880

acgatgactt tgccgctgag gtccgaaagc tgcgaaccct gttccagtct ctcggctact 11940

acgactcctc taaggcctac tacgccttca aggtctcctt caacctctgc atctggggac 12000

tgtccaccgt cattgtggcc aagtggggtc agacctccac cctcgccaac gtgctctctg 12060

ctgccctgct cggcctgttc tggcagcagt gcggatggct ggctcacgac tttctgcacc 12120

accaggtctt ccaggaccga ttctggggtg atctcttcgg agccttcctg ggaggtgtct 12180

gccagggctt ctcctcttcc tggtggaagg acaagcacaa cactcaccat gccgctccca 12240

acgtgcatgg cgaggatcct gacattgaca cccaccctct cctgacctgg tccgagcacg 12300

ctctggagat gttctccgac gtccccgatg aggagctgac ccgaatgtgg tctcgattca 12360

tggtcctgaa ccagacctgg ttctacttcc ccattctctc cttcgctcga ctgtcttggt 12420

gcctccagtc cattctcttt gtgctgccca acggtcaggc tcacaagccc tccggagctc 12480

gagtgcccat ctccctggtc gagcagctgt ccctcgccat gcactggacc tggtacctcg 12540

ctaccatgtt cctgttcatc aaggatcctg tcaacatgct cgtgtacttc ctggtgtctc 12600

aggctgtgtg cggaaacctg ctcgccatcg tgttctccct caaccacaac ggtatgcctg 12660

tgatctccaa ggaggaggct gtcgacatgg atttctttac caagcagatc atcactggtc 12720

gagatgtcca tcctggactg ttcgccaact ggttcaccgg tggcctgaac taccagatcg 12780

agcatcacct gttcccttcc atgcctcgac acaacttctc caagatccag cctgccgtcg 12840

agaccctgtg caagaagtac aacgtccgat accacaccac tggtatgatc gagggaactg 12900

ccgaggtctt ctcccgactg aacgaggtct ccaaggccac ctccaagatg ggcaaggctc 12960

agtaagcggc cgcatgagaa gataaatata taaatacatt gagatattaa atgcgctaga 13020

ttagagagcc tcatactgct cggagagaag ccaagacgag tactcaaagg ggattacacc 13080

atccatatcc acagacacaa gctggggaaa ggttctatat acactttccg gaataccgta 13140

gtttccgatg ttatcaatgg gggcagccag gatttcaggc acttcggtgt ctcggggtga 13200

aatggcgttc ttggcctcca tcaagtcgta ccatgtcttc atttgcctgt caaagtaaaa 13260

cagaagcaga tgaagaatga acttgaagtg aaggaattta aatgtagcta acggtagcag 13320

gcgaactact ggtacatacc tcccccggaa tatgtacagg cataatgcgt atctgtggga 13380

catgtggtcg ttgcgccatt atgtaagcag cgtgtactcc tctgactgtc catatggttt 13440

gctccatctc accctcatcg ttttcattgt tcacaggcgg ccacaaaaaa actgtcttct 13500

ctccttctct cttcgcctta gtctactcgg accagtttta gtttagcttg gcgccactgg 13560

ataaatgaga cctcaggcct tgtgatgagg aggtcactta tgaagcatgt taggaggtgc 13620

ttgtatggat agagaagcac ccaaaataat aagaataata ataaaacagg gggcgttgtc 13680

atttcatatc gtgttttcac catcaataca cctccaaaca atgcccttca tgtggccagc 13740

cccaatattg tcctgtagtt caactctatg cagctcgtat cttattgagc aagtaaaact 13800

ctgtcagccg atattgcccg acccgcgaca agggtcaaca aggtggtgta aggccttcgc 13860

agaagtcaaa actgtgccaa acaaacatct agagtctctt tggtgtttct cgcatatatt 13920

twatcggctg tcttacgtat ttgcgcctcg gtaccggact aatttcggat catccccaat 13980

acgctttttc ttcgcagctg tcaacagtgt ccatgatcta tccacctaaa tgggtcatat 14040

gaggcgtata atttcgtggt gctgataata attcccatat atttgacaca aaacttcccc 14100

ccctagacat acatctcaca atctcacttc ttgtgcttct gtcacacatc tcctccagct 14160

gacttcaact cacacctctg ccccagttgg tctacagcgg tataaggttt ctccgcatag 14220

aggtgcacca ctcctcccga tacttgtttg tgtgacttgt gggtcacgac atatatatct 14280

acacacattg cgccaccctt tggttcttcc agcacaacaa aaacacgaca cgctaaccat 14340

ggatggctcc cgaccctgtc gctgccgaga ccgctgccca gggtcccact ccccgatact 14400

tcacctggga cgaggtcgcc cagcgatccg gttgcgagga acgatggctg gtcatcgacc 14460

gaaaggtgta caacatctct gagttcaccc gacgacatcc cggtggctcc cgagtgatct 14520

cgcactacgc tggacaggac gccactgacc ccttcgttgc ctttcacatt aacaagggcc 14580

tggttaagaa gtacatgaac tccctgctca ttggagagct gtctcccgaa cagccttcgt 14640

ttgagcctac caagaacaag gagctgaccg acgagtttcg agagctccga gccaccgttg 14700

agcgaatggg actgatgaag gccaaccatg tcttctttct gctctacctg ctccacattc 14760

ttctccttga cggagctgcc tggcttaccc tgtgggtctt cggcacttcc tttctgccct 14820

ttcttctctg cgccgtcctg ctctctgccg tgcaggctca ggctggttgg cttcagcatg 14880

actttggtca cctttccgtg ttctctacct ccaagtggaa ccacctgctc catcacttcg 14940

tgatcggcca cctcaagggt gctcctgcct cgtggtggaa ccacatgcat ttccagcacc 15000

atgccaagcc caactgtttt cgaaaggatc ccgacatcaa catgcacccc ttctttttcg 15060

ctcttggcaa gatcctgtcc gtcgagctcg gaaagcagaa gaagaagtac atgccctaca 15120

accaccagca caagtacttc ttcctgattg gacctcccgc tctcctgcct ctttactttc 15180

agtggtacat cttttacttt gttattcagc gaaagaagtg ggttgatctt gcctggatga 15240

tcaccttcta cgtccgattc ttcctgacct acgtccctct ccttggactg aaggcctttc 15300

tcggtctgtt ctttatcgtc cgattcctgg agtccaactg gttcgtgtgg gtgacccaga 15360

tgaaccacat tcccatgcac attgaccatg atcgaaacat ggactgggtg tcgactcagc 15420

tgcaggccac ctgcaacgtt cacaagtctg ctttcaacga ctggttttcc ggtcacctca 15480

actttcagat tgagcaccat ctgtttccca ccatgcctcg acacaactac cacaaggttg 15540

ctcccctggt ccagtcgctc tgtgccaagc atggcatcga gtaccagtcc aagcccctgc 15600

tctctgcctt cgctgacatc attcactcgc tgaaggaatc tggccagctc tggctcgatg 15660

cctacctgca ccagtaagcg gccgcaagtg tggatgggga agtgagtgcc cggttctgtg 15720

tgcacaattg gcaatccaag atggatggat tcaacacagg gatatagcga gctacgtggt 15780

ggtgcgagga tatagcaacg gatatttatg tttgacactt gagaatgtac gatacaagca 15840

ctgtccaagt acaatactaa acatactgta catactcata ctcgtacccg ggcaacggtt 15900

tcacttgagt gcagtggcta gtgctcttac tcgtacagtg tgcaatactg cgtatcatag 15960

tctttgatgt atatcgtatt cattcatgtt agttgc 15996

<210> 162
<211> 8165
<212> DNA
<213> Artificial Sequence

<220>
<223> Plasmid pZUF17

<400> 162
gtacgagccg gaagcataaa gtgtaaagcc tggggtgcct aatgagtgag ctaactcaca 60

ttaattgcgt tgcgctcact gcccgctttc agtcgggaa acctgtcgtg ccagctgcat 120

taatgaatcg gccaacgcgc ggggagaggc ggtttgcgta ttgggcgctc ttccgcttcc 180

tcgctcactg actcgctgcg ctcggtcgtt cggctgcggc gagcggtatc agctcactca 240

aaggcggtaa tacggttatc cacagaatca ggggataacg caggaaagaa catgtgagca 300

aaaggccagc aaaaggccag gaaccgtaaa aaggccgcgt tgctggcgtt tttccatagg 360

ctccgccccc ctgacgagca tcacaaaaat cgacgctcaa gtcagaggtg gcgaaacccg 420

acaggactat aaagatacca ggcgtttccc cctggaagct ccctcgtgcg ctctcctgtt 480

ccgaccctgc cgcttaccgg atacctgtcc gcctttctcc cttcgggaag cgtggcgctt 540

tctcatagct cacgctgtag gtatctcagt tcggtgtagg tcgttcgctc caagctgggc 600

tgtgtgcacg aaccccccgt tcagcccgac cgctgcgcct tatccggtaa ctatcgtctt 660

gagtccaacc cggtaagaca cgacttatcg ccactggcag cagccactgg taacaggatt 720

agcagagcga ggtatgtagg cggtgctaca gagttcttga agtggtggcc taactacggc 780

tacactagaa ggacagtatt tggtatctgc gctctgctga agccagttac cttcggaaaa 840

agagttggta gctcttgatc cggcaaacaa accaccgctg gtagcggtgg ttttttgtt 900

tgcaagcagc agattacgcg cagaaaaaaa ggatctcaag aagatccttt gatcttttct 960

acggggtctg acgctcagtg gaacgaaaac tcacgttaag ggattttggt catgagatta 1020

tcaaaaagga tcttcaccta gatcctttta aattaaaaat gaagttttaa atcaatctaa 1080

agtatatatg agtaaacttg gtctgacagt taccaatgct taatcagtga ggcacctatc 1140

tcagcgatct gtctatttcg ttcatccata gttgcctgac tccccgtcgt gtagataact 1200

acgatacggg agggcttacc atctggcccc agtgctgcaa tgataccgcg agacccacgc 1260

tcaccggctc cagatttatc agcaataaac cagccagccg gaagggccga gcgcagaagt 1320

ggtcctgcaa ctttatccgc ctccatccag tctattaatt gttgccggga agctagagta 1380

agtagttcgc cagttaatag tttgcgcaac gttgttgcca ttgctacagg catcgtggtg 1440

tcacgctcgt cgtttggtat ggcttcattc agctccggtt cccaacgatc aaggcgagtt 1500

acatgatccc ccatgttgtg caaaaaagcg gttagctcct tcggtcctcc gatcgttgtc 1560

agaagtaagt tggccgcagt gttatcactc atggttatgg cagcactgca taattctctt 1620

actgtcatgc catccgtaag atgcttttct gtgactggtg agtactcaac caagtcattc 1680

tgagaatagt gtatgcggcg accgagttgc tcttgcccgg cgtcaatacg ggataatacc 1740

gcgccacata gcagaacttt aaaagtgctc atcattggaa aacgttcttc ggggcgaaaa 1800

ctctcaagga tcttaccgct gttgagatcc agttcgatgt aacccactcg tgcacccaac 1860

tgatcttcag catcttttac tttcaccagc gtttctgggt gagcaaaaac aggaaggcaa 1920

aatgccgcaa aaaagggaat aagggcgaca cggaaatgtt gaatactcat actcttcctt 1980

tttcaatatt attgaagcat ttatcagggt tattgtctca tgagcggata catatttgaa 2040

tgtatttaga aaaataaaca aatagggggtt ccgcgcacat ttccccgaaa agtgccacct 2100

gacgcgccct gtagcggcgc attaagcgcg gcgggtgtgg tggttacgcg cagcgtgacc 2160

gctacacttg ccagcgccct agcgcccgct cctttcgctt tcttcccttc ctttctcgcc 2220

acgttcgccg gctttccccg tcaagctcta aatcgggggc tccctttagg gttccgattt 2280

agtgctttac ggcacctcga ccccaaaaaa cttgattagg gtgatggttc acgtagtggg 2340

ccatcgccct gatagacggt ttttcgccct ttgacgttgg agtccacgtt ctttaatagt 2400

ggactcttgt tccaaactgg aacaacactc aaccctatct cggtctattc ttttgattta 2460

taagggattt tgccgatttc ggcctattgg ttaaaaaatg agctgattta acaaaaattt 2520

aacgcgaatt ttaacaaaat attaacgctt acaatttcca ttcgccattc aggctgcgca 2580

actgttggga agggcgatcg gtgcgggcct cttcgctatt acgccagctg gcgaaagggg 2640

gatgtgctgc aaggcgatta agttgggtaa cgccagggtt ttcccagtca cgacgttgta 2700

aaacgacggc cagtgaattg taatacgact cactataggg cgaattgggt accgggcccc 2760

ccctcgaggt cgatggtgtc gataagcttg atatcgaatt catgtcacac aaaccgatct 2820

tcgcctcaag gaaacctaat tctacatccg agagactgcc gagatccagt ctacactgat 2880

taattttcgg gccaataatt taaaaaaatc gtgttatata atattatatg tattatatat 2940

atacatcatg atgatactga cagtcatgtc ccattgctaa atagacagac tccatctgcc 3000

gcctccaact gatgttctca atatttaagg ggtcatctcg cattgtttaa taataaacag 3060

actccatcta ccgcctccaa atgatgttct caaaatatat tgtatgaact tatttttatt 3120

acttagtatt attagacaac ttacttgctt tatgaaaaac acttcctatt taggaaacaa 3180

tttataatgg cagttcgttc atttaacaat ttatgtagaa taaatgttat aaatgcgtat 3240

gggaaatctt aaatatggat agcataaatg atatctgcat tgcctaattc gaaatcaaca 3300

gcaacgaaaa aaatcccttg tacaacataa atagtcatcg agaaatatca actatcaaag 3360

aacagctatt cacacgttac tattgagatt attattggac gagaatcaca cactcaactg 3420

tctttctctc ttctagaaat acaggtacaa gtatgtacta ttctcattgt tcatacttct 3480

agtcatttca tcccacatat tccttggatt tctctccaat gaatgacatt ctatcttgca 3540

aattcaacaa ttataataag atataccaaa gtagcggtat agtggcaatc aaaaagcttc 3600

tctggtgtgc ttctcgtatt tatttttatt ctaatgatcc attaaaggta tatatttatt 3660

tcttgttata taatcctttt gtttattaca tgggctggat acataaaggt attttgattt 3720

aatttttgc ttaaattcaa tccccccctcg ttcagtgtca actgtaatgg taggaaatta 3780

ccatactttt gaagaagcaa aaaaaatgaa agaaaaaaaa aatcgtattt ccaggttaga 3840

cgttccgcag aatctagaat gcggtatgcg gtacattgtt cttcgaacgt aaaagttgcg 3900

ctccctgaga tattgtacat ttttgctttt acaagtacaa gtacatcgta caactatgta 3960

ctactgttga tgcatccaca acagtttgtt ttgtttttt ttgtttttt tttttctaat 4020

gattcattac cgctatgtat acctacttgt acttgtagta agccgggtta ttggcgttca 4080

attaatcata gacttatgaa tctgcacggt gtgcgctgcg agttactttt agcttatgca 4140

tgctacttgg gtgtaatatt gggatctgtt cggaaatcaa cggatgctca atcgatttcg 4200

acagtaatta attaagtcat acacaagtca gctttcttcg agcctcatat aagtataagt 4260

agttcaacgt attagcactg tacccagcat ctccgtatcg agaaacacaa caacatgccc 4320

cattggacag atcatgcgga tacacaggtt gtgcagtatc atacatactc gatcagacag 4380

gtcgtctgac catcatacaa gctgaacaag cgctccatac ttgcacgctc tctatataca 4440

cagttaaatt acatatccat agtctaacct ctaacagtta atcttctggt aagcctccca 4500

gccagccttc tggtatcgct tggcctcctc aataggatct cggttctggc cgtacagacc 4560

tcggccgaca attatgatat ccgttccggt agacatgaca tcctcaacag ttcggtactg 4620

ctgtccgaga gcgtctccct tgtcgtcaag acccaccccg ggggtcagaa taagccagtc 4680

ctcagagtcg cccttaggtc ggttctgggc aatgaagcca accacaaact cggggtcgga 4740

tcgggcaagc tcaatggtct gcttggagta ctcgccagtg gccagagagc ccttgcaaga 4800

cagctcggcc agcatgagca gacctctggc cagcttctcg ttgggagagg ggactaggaa 4860

ctccttgtac tgggagttct cgtagtcaga gacgtcctcc ttcttctgtt cagagacagt 4920

ttcctcggca ccagctcgca ggccagcaat gattccggtt ccgggtacac cgtgggcgtt 4980

ggtgatatcg gaccactcgg cgattcggtg acaccggtac tggtgcttga cagtgttgcc 5040

aatatctgcg aactttctgt cctcgaacag gaagaaaccg tgcttaagag caagttcctt 5100

gaggggggagc acagtgccgg cgtaggtgaa gtcgtcaatg atgtcgatat gggttttgat 5160

catgcacaca taaggtccga ccttatcggc aagctcaatg agctccttgg tggtggtaac 5220

atccagagaa gcacacaggt tggttttctt ggctgccacg agcttgagca ctcgagcggc 5280

aaaggcggac ttgtggacgt tagctcgagc ttcgtaggag ggcattttgg tggtgaagag 5340

gagactgaaa taaatttagt ctgcagaact ttttatcgga accttatctg gggcagtgaa 5400

gtatatgtta tggtaatagt tacgagttag ttgaacttat agatagactg gactatacgg 5460

ctatcggtcc aaattagaaa gaacgtcaat ggctctctgg gcgtcgcctt tgccgacaaa 5520

aatgtgatca tgatgaaagc cagcaatgac gttgcagctg atattgttgt cggccaaccg 5580

cgccgaaaac gcagctgtca gacccacagc ctccaacgaa gaatgtatcg tcaaagtgat 5640

ccaagcacac tcatagttgg agtcgtactc caaaggcggc aatgacgagt cagacagata 5700

ctcgtcgact caggcgacga cggaattcct gcagcccatc tgcagaattc aggagagacc 5760

gggttggcgg cgtatttgtg tcccaaaaaa cagccccaat tgccccggag aagacggcca 5820

ggccgcctag atgacaaatt caacaactca cagctgactt tctgccattg ccactagggg 5880

ggggccttt tatatggcca agccaagctc tccacgtcgg ttgggctgca cccaacaata 5940

aatgggtagg gttgcaccaa caaagggatg ggatgggggg tagaagatac gaggataacg 6000

gggctcaatg gcacaaataa gaacgaatac tgccattaag actcgtgatc cagcgactga 6060

caccattgca tcatctaagg gcctcaaaac tacctcggaa ctgctgcgct gatctggaca 6120

ccacagaggt tccgagcact ttaggttgca ccaaatgtcc caccaggtgc aggcagaaaa 6180

cgctggaaca gcgtgtacag tttgtcttaa caaaaagtga gggcgctgag gtcgagcagg 6240

gtggtgtgac ttgttatagc ctttagagct gcgaaagcgc gtatggattt ggctcatcag 6300

gccagattga gggtctgtgg acacatgtca tgttagtgta cttcaatcgc cccctggata 6360

tagccccgac aataggccgt ggcctcattt ttttgccttc cgcacatttc cattgctcgg 6420

tacccacacc ttgcttctcc tgcacttgcc aaccttaata ctggtttaca ttgaccaaca 6480

tcttacaagc ggggggcttg tctagggtat atataaacag tggctctccc aatcggttgc 6540

cagtctcttt tttcctttct ttccccacag attcgaaatc taaactacac atcacacaat 6600

gcctgttact gacgtcctta agcgaaagtc cggtgtcatc gtcggcgacg atgtccgagc 6660

cgtgagtatc cacgacaaga tcagtgtcga gacgacgcgt tttgtgtaat gacacaatcc 6720

gaaagtcgct agcaacacac actctctaca caaactaacc cagctctcca tggctgagga 6780

taagaccaag gtcgagttcc ctaccctgac tgagctgaag cactctatcc ctaacgcttg 6840

ctttgagtcc aacctcggac tctcgctcta ctacactgcc cgagcgatct tcaacgcatc 6900

tgcctctgct gctctgctct acgctgcccg atctactccc ttcattgccg ataacgttct 6960

gctccacgct ctggtttgcg ccacctacat ctacgtgcag ggtgtcatct tctggggttt 7020

ctttaccgtc ggtcacgact gtggtcactc tgccttctcc cgataccact ccgtcaactt 7080

catcattggc tgcatcatgc actctgccat tctgactccc ttcgagtcct ggcgagtgac 7140

ccaccgacac catcacaaga acactggcaa cattgataag gacgagatct tctaccctca 7200

tcggtccgtc aaggacctcc aggacgtgcg acaatgggtc tacaccctcg gaggtgcttg 7260

gtttgtctac ctgaaggtcg gatatgctcc tcgaaccatg tcccactttg acccctggga 7320

ccctctcctg cttcgacgag cctccgctgt catcgtgtcc ctcggagtct gggctgcctt 7380

cttcgctgcc tacgcctacc tcacatactc gctcggcttt gccgtcatgg gcctctacta 7440

ctatgctcct ctctttgtct ttgcttcgtt cctcgtcatt actaccttct tgcatcacaa 7500

cgacgaagct actccctggt acggtgactc ggagtggacc tacgtcaagg gcaacctgag 7560

ctccgtcgac cgatcgtacg gagctttcgt ggacaacctg tctcaccaca ttggcaccca 7620

ccaggtccat cacttgttcc ctatcattcc ccactacaag ctcaacgaag ccaccaagca 7680

ctttgctgcc gcttaccctc acctcgtgag acgtaacgac gagcccatca ttactgcctt 7740

cttcaagacc gctcacctct ttgtcaacta cggagctgtg cccgagactg ctcagatttt 7800

caccctcaaa gagtctgccg ctgcagccaa ggccaagagc gactaagcgg ccgcaagtgt 7860

ggatggggaa gtgagtgccc ggttctgtgt gcacaattgg caatccaaga tggatggatt 7920

caacacaggg atatagcgag ctacgtggtg gtgcgaggat atagcaacgg atatttatgt 7980

ttgacacttg agaatgtacg atacaagcac tgtccaagta caatactaaa catactgtac 8040

atactcatac tcgtacccgg gcaacggttt cacttgagtg cagtggctag tgctcttact 8100

cgtacagtgt gcaatactgc gtatcatagt ctttgatgta tatcgtattc attcatgtta 8160

gttgc 8165

<210> 163
<211> 8633
<212> DNA
<213> Artificial Sequence

<220>
<223> Plasmid pZUF4

<400> 163
catggtcggc ttcgacgaga ccgtcactat ggacaccgtc cgaaaccaca acatgcccga 60

cgatgcctgg tgcgctatcc acggcaccgt gtacgacatc accaagttct ccaaggtgca 120

tcctggtgga gacattatca tgctggccgc tggcaaggag gccaccatcc tgttcgagac 180

ctaccacatc aagggtgttc ccgacgctgt cctgcgaaag tacaaggtcg gcaagctccc 240

tcagggcaag aagggtgaga cctcccacat gcccaccgga ctcgactctg cctcctacta 300

ctcctgggac tctgagttct accgagtcct ccgagagcga gtcgccaaga agctggccga 360

gcccggtctc atgcagcgag ctcgaatgga gctctgggcc aaggctatct ttctcctggc 420

tggtttctgg ggatccctgt acgccatgtg cgtgctcgac cctcacggag gtgctatggt 480

cgctgccgtc accctcggtg tgtttgctgc cttcgtcgga acctgcatcc agcacgacgg 540

ctcccacgga gccttctcca agtctcgatt catgaacaag gctgccggtt ggaccctcga 600

catgattggt gcctccgcta tgacctggga gatgcagcac gtcctcggtc accatcccta 660

caccaacctg atcgagatgg agaacggtct ggccaaggtc aagggtgccg acgtggaccc 720

caagaaggtc gaccaggagt ccgaccctga cgtcttctct acctacccta tgctgcgact 780

gcatccctgg caccgacagc gattctacca caagcttcag cacctgtacg ctcccctgat 840

cttcggcttc atgaccatca acaaggtgat ctcccaggac gtcggtgtcg tgctgcgaaa 900

gcgactgttt cagatcgacg ccaactgccg atatggctct ccctggaacg tggctcgatt 960

ctggatcatg aagctcctga ccactctcta catggtggct ctgcccatgt acatgcaggg 1020

acccgctcag ggtctgaagc tgttctttat ggctcacttc acctgcggag aggtcctcgc 1080

caccatgttc attgttaacc acattatcga gggtgtctcc tacgcctcca aggacgctgt 1140

caagggcgtc atggctcctc cccgaaccgt gcacggtgtc actcccatgc aggtgaccca 1200

gaaggctctc tctgctgccg agtccaccaa gtctgacgct gacaagacca ctatgatccc 1260

tctcaacgac tgggctgccg tccagtgcca gacctccgtc aactgggctg tcggatcctg 1320

gttctggaac cacttctctg gaggtctcaa ccaccagatt gagcaccatt gctttcctca 1380

gaacccccac accgtcaacg tctacatctc cggcatcgtc aaggagacct gcgaggagta 1440

cggcgtgccc taccaggctg agatctccct cttctctgcc tatttcaaga tgctgtccca 1500

cctccgaacc ctcggcaacg aggacctcac cgcctggtcc acctaagcgg ccgcaagtgt 1560

ggatggggaa gtgagtgccc ggttctgtgt gcacaattgg caatccaaga tggatggatt 1620

caacacaggg atatagcgag ctacgtggtg gtgcgaggat atagcaacgg atatttatgt 1680

ttgacacttg agaatgtacg atacaagcac tgtccaagta caatactaaa catactgtac 1740

atactcatac tcgtacccgg gcaacggttt cacttgagtg cagtggctag tgctcttact 1800

cgtacagtgt gcaatactgc gtatcatagt ctttgatgta tatcgtattc attcatgtta 1860

gttgcgtacg agccggaagc ataaagtgta aagcctgggg tgcctaatga gtgagctaac 1920

tcacattaat tgcgttgcgc tcactgcccg ctttccagtc gggaaacctg tcgtgccagc 1980

tgcattaatg aatcggccaa cgcgcgggga gaggcggttt gcgtattggg cgctcttccg 2040

cttcctcgct cactgactcg ctgcgctcgg tcgttcggct gcggcgagcg gtatcagctc 2100

actcaaaggc ggtaatacgg ttatccacag aatcagggga taacgcagga aagaacatgt 2160

gagcaaaagg ccagcaaaag gccaggaacc gtaaaaaggc cgcgttgctg gcgtttttcc 2220

ataggctccg ccccccctgac gagcatcaca aaaatcgacg ctcaagtcag aggtggcgaa 2280

acccgacagg actataaaga taccaggcgt ttcccccctgg aagctccctc gtgcgctctc 2340

ctgttccgac cctgccgctt accggatacc tgtccgcctt tctcccttcg ggaagcgtgg 2400

cgctttctca tagctcacgc tgtaggtatc tcagttcggt gtaggtcgtt cgctccaagc 2460

tgggctgtgt gcacgaaccc cccgttcagc ccgaccgctg cgccttatcc ggtaactatc 2520

gtcttgagtc caacccggta agacacgact tatcgccact ggcagcagcc actggtaaca 2580

ggattagcag agcgaggtat gtaggcggtg ctacagagtt cttgaagtgg tggcctaact 2640

acggctacac tagaaggaca gtatttggta tctgcgctct gctgaagcca gttaccttcg 2700

gaaaaagagt tggtagctct tgatccggca aacaaaccac cgctggtagc ggtggttttt 2760

ttgtttgcaa gcagcagatt acgcgcagaa aaaaaggatc tcaagaagat cctttgatct 2820

tttctacggg gtctgacgct cagtggaacg aaaactcacg ttaagggatt ttggtcatga 2880

gattatcaaa aaggatcttc acctagatcc ttttaaatta aaaatgaagt tttaaatcaa 2940

tctaaagtat atatgagtaa acttggtctg acagttacca atgcttaatc agtgaggcac 3000

ctatctcagc gatctgtcta tttcgttcat ccatagttgc ctgactcccc gtcgtgtaga 3060

taactacgat acgggagggc ttaccatctg gccccagtgc tgcaatgata ccgcgagacc 3120

cacgctcacc ggctccagat ttatcagcaa taaaccagcc agccggaagg gccgagcgca 3180

gaagtggtcc tgcaacttta tccgcctcca tccagtctat taattgttgc cgggaagcta 3240

gagtaagtag ttcgccagtt aatagtttgc gcaacgttgt tgccattgct acaggcatcg 3300

tggtgtcacg ctcgtcgttt ggtatggctt cattcagctc cggttcccaa cgatcaaggc 3360

gagttacatg atcccccatg ttgtgcaaaa aagcggttag ctccttcggt cctccgatcg 3420

ttgtcagaag taagttggcc gcagtgttat cactcatggt tatggcagca ctgcataatt 3480

ctcttactgt catgccatcc gtaagatgct tttctgtgac tggtgagtac tcaaccaagt 3540

cattctgaga atagtgtatg cggcgaccga gttgctcttg cccggcgtca atacgggata 3600

ataccgcgcc acatagcaga actttaaaag tgctcatcat tggaaaacgt tcttcggggc 3660

gaaaactctc aaggatctta ccgctgttga gatccagttc gatgtaaccc actcgtgcac 3720

ccaactgatc ttcagcatct tttactttca ccagcgtttc tgggtgagca aaaacaggaa 3780

ggcaaaatgc cgcaaaaaag ggaataaggg cgacacggaa atgttgaata ctcatactct 3840

tccttttca atattattga agcatttatc agggttattg tctcatgagc ggatacatat 3900

ttgaatgtat ttagaaaaat aaacaaatag gggttccgcg cacatttccc cgaaaagtgc 3960

cacctgacgc gccctgtagc ggcgcattaa gcgcggcggg tgtggtggtt acgcgcagcg 4020

tgaccgctac acttgccagc gccctagcgc ccgctccttt cgctttcttc ccttcctttc 4080

tcgccacgtt cgccggcttt ccccgtcaag ctctaaatcg ggggctccct ttagggttcc 4140

gatttagtgc tttacggcac ctcgacccca aaaaacttga ttagggtgat ggttcacgta 4200

gtgggccatc gccctgatag acggtttttc gccctttgac gttggagtcc acgttcttta 4260

atagtggact cttgttccaa actggaacaa cactcaaccc tatctcggtc tattcttttg 4320

atttataagg gattttgccg atttcggcct attggttaaa aaatgagctg atttaacaaa 4380

aatttaacgc gaattttaac aaaatattaa cgcttacaat ttccattcgc cattcaggct 4440

gcgcaactgt tgggaagggc gatcggtgcg ggcctcttcg ctattacgcc agctggcgaa 4500

aggggatgt gctgcaaggc gattaagttg ggtaacgcca gggttttccc agtcacgacg 4560

ttgtaaaacg acggccagtg aattgtaata cgactcacta tagggcgaat tgggtaccgg 4620

gccccccctc gaggtcgatg gtgtcgataa gcttgatatc gaattcatgt cacacaaacc 4680

gatcttcgcc tcaaggaaac ctaattctac atccgagaga ctgccgagat ccagtctaca 4740

ctgattaatt ttcgggccaa taatttaaaa aaatcgtgtt atataatatt atatgtatta 4800

tatatataca tcatgatgat actgacagtc atgtcccatt gctaaataga cagactccat 4860

ctgccgcctc caactgatgt tctcaatatt taaggggtca tctcgcattg tttaataata 4920

aacagactcc atctaccgcc tccaaatgat gttctcaaaa tatattgtat gaacttattt 4980

ttattactta gtattattag acaacttact tgctttatga aaaacacttc ctatttagga 5040

aacaatttat aatggcagtt cgttcattta acaatttatg tagaataaat gttataaatg 5100

cgtatgggaa atcttaaata tggatagcat aaatgatatc tgcattgcct aattcgaaat 5160

caacagcaac gaaaaaaatc ccttgtacaa cataaatagt catcgagaaa tatcaactat 5220

caaagaacag ctattcacac gttactattg agattattat tggacgagaa tcacacactc 5280

aactgtcttt ctctcttcta gaaatacagg tacaagtatg tactattctc attgttcata 5340

cttctagtca tttcatccca catattcctt ggatttctct ccaatgaatg acattctatc 5400

ttgcaaattc aacaattata ataagatata ccaaagtagc ggtatagtgg caatcaaaaa 5460

gcttctctgg tgtgcttctc gtatttattt ttattctaat gatccattaa aggtatatat 5520

ttatttcttg ttatataatc cttttgttta ttacatgggc tggatacata aaggtatttt 5580

gatttaattt tttgcttaaa ttcaatcccc cctcgttcag tgtcaactgt aatggtagga 5640

aattaccata cttttgaaga agcaaaaaaa atgaaagaaa aaaaaaatcg tatttccagg 5700

ttagacgttc cgcagaatct agaatgcggt atgcggtaca ttgttcttcg aacgtaaaag 5760

ttgcgctccc tgagatattg tacatttttg cttttacaag tacaagtaca tcgtacaact 5820

atgtactact gttgatgcat ccacaacagt ttgttttgtt ttttttttgtt ttttttttttt 5880

ctaatgattc attaccgcta tgtataccta cttgtacttg tagtaagccg ggttattggc 5940

gttcaattaa tcatagactt atgaatctgc acggtgtgcg ctgcgagtta cttttagctt 6000

atgcatgcta cttgggtgta atattgggat ctgttcggaa atcaacggat gctcaatcga 6060

tttcgacagt aattaattaa gtcatacaca agtcagcttt cttcgagcct catataagta 6120

taagtagttc aacgtattag cactgtaccc agcatctccg tatcgagaaa cacaacaaca 6180

tgccccattg gacagatcat gcggatacac aggttgtgca gtatcataca tactcgatca 6240

gacaggtcgt ctgaccatca tacaagctga acaagcgctc catacttgca cgctctctat 6300

atacacagtt aaattacata tccatagtct aacctctaac agttaatctt ctggtaagcc 6360

tcccagccag ccttctggta tcgcttggcc tcctcaatag gatctcggtt ctggccgtac 6420

agacctcggc cgacaattat gatatccgtt ccggtagaca tgacatcctc aacagttcgg 6480

tactgctgtc cgagagcgtc tcccttgtcg tcaagaccca ccccgggggt cagaataagc 6540

cagtcctcag agtcgccctt aggtcggttc tgggcaatga agccaaccac aaactcgggg 6600

tcggatcggg caagctcaat ggtctgcttg gagtactcgc cagtggccag agagcccttg 6660

caagacagct cggccagcat gagcagacct ctggccagct tctcgttggg agaggggact 6720

aggaactcct tgtactggga gttctcgtag tcagagacgt cctccttctt ctgttcagag 6780

acagtttcct cggcaccagc tcgcaggcca gcaatgattc cggttccggg tacaccgtgg 6840

gcgttggtga tatcggacca ctcggcgatt cggtgacacc ggtactggtg cttgacagtg 6900

ttgccaatat ctgcgaactt tctgtcctcg aacaggaaga aaccgtgctt aagagcaagt 6960

tccttgaggg ggagcacagt gccggcgtag gtgaagtcgt caatgatgtc gatatgggtt 7020

ttgatcatgc acacataagg tccgacctta tcggcaagct caatgagctc cttggtggtg 7080

gtaacatcca gagaagcaca caggttggtt ttcttggctg ccacgagctt gagcactcga 7140

gcggcaaagg cggacttgtg gacgttagct cgagcttcgt aggagggcat tttggtggtg 7200

aagaggagac tgaaataaat ttagtctgca gaacttttta tcggaacctt atctggggca 7260

gtgaagtata tgttatggta atagttacga gttagttgaa cttatagata gactggacta 7320

tacggctatc ggtccaaatt agaaagaacg tcaatggctc tctgggcgtc gcctttgccg 7380

acaaaaatgt gatcatgatg aaagccagca atgacgttgc agctgatatt gttgtcggcc 7440

aaccgcgccg aaaacgcagc tgtcagaccc acagcctcca acgaagaatg tatcgtcaaa 7500

gtgatccaag cacactcata gttggagtcg tactccaaag gcggcaatga cgagtcagac 7560

agatactcgt cgactcaggc gacgacggaa ttcctgcagc ccatctgcag aattcaggag 7620

agaccgggtt ggcggcgtat ttgtgtccca aaaaacagcc ccaattgccc cggagaagac 7680

ggccaggccg cctagatgac aaattcaaca actcacagct gactttctgc cattgccact 7740

agggggggc ctttttatat ggccaagcca agctctccac gtcggttggg ctgcacccaa 7800

caataaatgg gtagggttgc accaacaaag ggatgggatg gggggtagaa gatacgagga 7860

taacggggct caatggcaca aataagaacg aatactgcca ttaagactcg tgatccagcg 7920

actgacacca ttgcatcatc taagggcctc aaaactacct cggaactgct gcgctgatct 7980

ggacaccaca gaggttccga gcactttagg ttgcaccaaa tgtcccacca ggtgcaggca 8040

gaaaacgctg gaacagcgtg tacagtttgt cttaacaaaa agtgagggcg ctgaggtcga 8100

gcagggtggt gtgacttgtt atagccttta gagctgcgaa agcgcgtatg gatttggctc 8160

atcaggccag attgagggtc tgtggacaca tgtcatgtta gtgtacttca atcgcccccct 8220

ggatatagcc ccgacaatag gccgtggcct catttttttg ccttccgcac atttccattg 8280

ctcggtaccc acaccttgct tctcctgcac ttgccaacct taatactggt ttacattgac 8340

caacatctta caagcggggg gcttgtctag ggtatatata aacagtggct ctcccaatcg 8400

gttgccagtc tctttttttcc tttctttccc cacagattcg aaatctaaac tacacatcac 8460

acaatgcctg ttactgacgt ccttaagcga aagtccggtg tcatcgtcgg cgacgatgtc 8520

cgagccgtga gtatccacga caagatcagt gtcgagacga cgcgttttgt gtaatgacac 8580

aatccgaaag tcgctagcaa cacacactct ctacacaaac taacccagct ctc 8633

<210> 164
<211> 8633
<212> DNA
<213> Artificial Sequence

<220>
<223> Plasmid pZUF4S

<400> 164
catggtcggc ttcgacgaga ccgtcactat ggacaccgtc cgaaaccaca acatgcccga 60

cgatgcctgg tgcgctatcc acggcaccgt gtacgacatc accaagttct ccaaggtgca 120

tcctggtgga gacattatca tgctggccgc tggcaaggag gccaccatcc tgttcgagac 180

ctaccacatc aagggtgttc ccgacgctgt cctgcgaaag tacaaggtcg gcaagctccc 240

tcagggcaag aagggtgaga cctcccacat gcccaccgga ctcgactctg cctcctacta 300

ctcctgggac tctgagttct accgagtcct ccgagagcga gtcgccaaga agctggccga 360

gcccggtctc atgcagcgag ctcgaatgga gctctgggcc aaggctatct ttctcctggc 420

tggtttctgg ggatccctgt acgccatgtg cgtgctcgac cctcacggag gtgctatggt 480

cgctgccgtc accctcggtg tgtttgctgc cttcgtcgga acctgcatcc agcacgacgg 540

ctcccacgga gccttctcca agtctcgatt catgaacaag gctgccggtt ggaccctcga 600

catgattggt gcctccgcta tgacctggga gatgcagcac gtcctcggtc accatcccta 660

caccaacctg atcgagatgg agaacggtct ggccaaggtc aagggtgccg acgtggaccc 720

caagaaggtc gaccaggagt ccgaccctga cgtcttctct acctacccta tgctgcgact 780

gcatccctgg caccgacagc gattctacca caagtttcag cacctgtacg ctcccctgat 840

cttcggcttc atgaccatca acaaggtgat ctcccaggac gtcggtgtcg tgctgcgaaa 900

gcgactgttt cagatcgacg ccaactgccg atatggctct ccctggaacg tggctcgatt 960

ctggatcatg aagctcctga ccactctcta catggtggct ctgcccatgt acatgcaggg 1020

acccgctcag ggtctgaagc tgttctttat ggctcacttc acctgcggag aggtcctcgc 1080

caccatgttc attgttaacc acattatcga gggtgtctcc tacgcctcca aggacgctgt 1140

caagggcgtc atggctcctc cccgaaccgt gcacggtgtc actcccatgc aggtgaccca 1200

gaaggctctc tctgctgccg agtccaccaa gtctgacgct gacaagacca ctatgatccc 1260

tctcaacgac tgggctgccg tccagtgcca gacctccgtc aactgggctg tcggatcctg 1320

gttctggaac cacttctctg gaggtctcaa ccaccagatt gagcaccatt gctttcctca 1380

gaaccccccac accgtcaacg tctacatctc cggcatcgtc aaggagacct gcgaggagta 1440

cggcgtgccc taccaggctg agatctccct cttctctgcc tatttcaaga tgctgtccca 1500

cctccgaacc ctcggcaacg aggacctcac cgcctggtcc acctaagcgg ccgcaagtgt 1560

ggatggggaa gtgagtgccc ggttctgtgt gcacaattgg caatccaaga tggatggatt 1620

caacacaggg atatagcgag ctacgtggtg gtgcgaggat atagcaacgg atatttatgt 1680

ttgacacttg agaatgtacg atacaagcac tgtccaagta caatactaaa catactgtac 1740

atactcatac tcgtacccgg gcaacggttt cacttgagtg cagtggctag tgctcttact 1800

cgtacagtgt gcaatactgc gtatcatagt ctttgatgta tatcgtattc attcatgtta 1860

gttgcgtacg agccggaagc ataaagtgta aagcctgggg tgcctaatga gtgagctaac 1920

tcacattaat tgcgttgcgc tcactgcccg ctttccagtc gggaaacctg tcgtgccagc 1980

tgcattaatg aatcggccaa cgcgcgggga gaggcggttt gcgtattggg cgctcttccg 2040

cttcctcgct cactgactcg ctgcgctcgg tcgttcggct gcggcgagcg gtatcagctc 2100

actcaaaggc ggtaatacgg ttatccacag aatcagggga taacgcagga aagaacatgt 2160

gagcaaaagg ccagcaaaag gccaggaacc gtaaaaaggc cgcgttgctg gcgtttttcc 2220

ataggctccg ccccccctgac gagcatcaca aaaatcgacg ctcaagtcag aggtggcgaa 2280

acccgacagg actataaaga taccaggcgt ttcccccctgg aagctccctc gtgcgctctc 2340

ctgttccgac cctgccgctt accggatacc tgtccgcctt tctcccttcg ggaagcgtgg 2400

cgctttctca tagctcacgc tgtaggtatc tcagttcggt gtaggtcgtt cgctccaagc 2460

tgggctgtgt gcacgaaccc cccgttcagc ccgaccgctg cgccttatcc ggtaactatc 2520

gtcttgagtc caacccggta agacacgact tatcgccact ggcagcagcc actggtaaca 2580

ggattagcag agcgaggtat gtaggcggtg ctacagagtt cttgaagtgg tggcctaact 2640

acggctacac tagaaggaca gtatttggta tctgcgctct gctgaagcca gttaccttcg 2700

gaaaaagagt tggtagctct tgatccggca aacaaaccac cgctggtagc ggtggttttt 2760

ttgtttgcaa gcagcagatt acgcgcagaa aaaaaggatc tcaagaagat cctttgatct 2820

tttctacggg gtctgacgct cagtggaacg aaaactcacg ttaagggatt ttggtcatga 2880

gattatcaaa aaggatcttc acctagatcc ttttaaatta aaaatgaagt tttaaatcaa 2940

tctaaagtat atatgagtaa acttggtctg acagttacca atgcttaatc agtgaggcac 3000

ctatctcagc gatctgtcta tttcgttcat ccatagttgc ctgactcccc gtcgtgtaga 3060

taactacgat acgggagggc ttaccatctg gccccagtgc tgcaatgata ccgcgagacc 3120

cacgctcacc ggctccagat ttatcagcaa taaaccagcc agccggaagg gccgagcgca 3180

gaagtggtcc tgcaacttta tccgcctcca tccagtctat taattgttgc cgggaagcta 3240

gagtaagtag ttcgccagtt aatagtttgc gcaacgttgt tgccattgct acaggcatcg 3300

tggtgtcacg ctcgtcgttt ggtatggctt cattcagctc cggttcccaa cgatcaaggc 3360

gagttacatg atcccccatg ttgtgcaaaa aagcggttag ctccttcggt cctccgatcg 3420

ttgtcagaag taagttggcc gcagtgttat cactcatggt tatggcagca ctgcataatt 3480

ctcttactgt catgccatcc gtaagatgct tttctgtgac tggtgagtac tcaaccaagt 3540

cattctgaga atagtgtatg cggcgaccga gttgctcttg cccggcgtca atacgggata 3600

ataccgcgcc acatagcaga actttaaaag tgctcatcat tggaaaacgt tcttcggggc 3660

gaaaactctc aaggatctta ccgctgttga gatccagttc gatgtaaccc actcgtgcac 3720

ccaactgatc ttcagcatct tttactttca ccagcgtttc tgggtgagca aaaacaggaa 3780

ggcaaaatgc cgcaaaaaag ggaataaggg cgacacggaa atgttgaata ctcatactct 3840

tccttttca atattattga agcatttatc agggttattg tctcatgagc ggatacatat 3900

ttgaatgtat ttagaaaaat aaacaaatag gggttccgcg cacatttccc cgaaaagtgc 3960

cacctgacgc gccctgtagc ggcgcattaa gcgcggcggg tgtggtggtt acgcgcagcg 4020

tgaccgctac acttgccagc gccctagcgc ccgctccttt cgctttcttc ccttcctttc 4080

tcgccacgtt cgccggcttt ccccgtcaag ctctaaatcg ggggctccct ttagggttcc 4140

gatttagtgc tttacggcac ctcgacccca aaaaacttga ttagggtgat ggttcacgta 4200

gtgggccatc gccctgatag acggtttttc gccctttgac gttggagtcc acgttcttta 4260

atagtggact cttgttccaa actggaacaa cactcaaccc tatctcggtc tattcttttg 4320

atttataagg gattttgccg atttcggcct attggttaaa aaatgagctg atttaacaaa 4380

aatttaacgc gaattttaac aaaatattaa cgcttacaat ttccattcgc cattcaggct 4440

gcgcaactgt tgggaagggc gatcggtgcg ggcctcttcg ctattacgcc agctggcgaa 4500

aggggggatgt gctgcaaggc gattaagttg ggtaacgcca gggttttccc agtcacgacg 4560

ttgtaaaacg acggccagtg aattgtaata cgactcacta tagggcgaat tgggtaccgg 4620

gccccccctc gaggtcgatg gtgtcgataa gcttgatatc gaattcatgt cacacaaacc 4680

gatcttcgcc tcaaggaaac ctaattctac atccgagaga ctgccgagat ccagtctaca 4740

ctgattaatt ttcgggccaa taatttaaaa aaatcgtgtt atataatatt atatgtatta 4800

tatatataca tcatgatgat actgacagtc atgtcccatt gctaaataga cagactccat 4860

ctgccgcctc caactgatgt tctcaatatt taaggggtca tctcgcattg tttaataata 4920

aacagactcc atctaccgcc tccaaatgat gttctcaaaa tatattgtat gaacttattt 4980

ttattactta gtattattag acaacttact tgctttatga aaaacacttc ctatttagga 5040

aacaatttat aatggcagtt cgttcattta acaatttatg tagaataaat gttataaatg 5100

cgtatgggaa atcttaaata tggatagcat aaatgatatc tgcattgcct aattcgaaat 5160

caacagcaac gaaaaaaatc ccttgtacaa cataaatagt catcgagaaa tatcaactat 5220

caaagaacag ctattcacac gttactattg agattattat tggacgagaa tcacacactc 5280

aactgtcttt ctctcttcta gaaatacagg tacaagtatg tactattctc attgttcata 5340

cttctagtca tttcatccca catattcctt ggatttctct ccaatgaatg acattctatc 5400

ttgcaaattc aacaattata ataagatata ccaaagtagc ggtatagtgg caatcaaaaa 5460

gcttctctgg tgtgcttctc gtatttattt ttattctaat gatccattaa aggtatatat 5520

ttatttcttg ttatataatc cttttgttta ttacatgggc tggatacata aaggtatttt 5580

gatttaattt tttgcttaaa ttcaatcccc cctcgttcag tgtcaactgt aatggtagga 5640

aattaccata cttttgaaga agcaaaaaaa atgaaagaaa aaaaaaatcg tatttccagg 5700

ttagacgttc cgcagaatct agaatgcggt atgcggtaca ttgttcttcg aacgtaaaag 5760

ttgcgctccc tgagatattg tacatttttg cttttacaag tacaagtaca tcgtacaact 5820

atgtactact gttgatgcat ccacaacagt ttgttttgtt ttttttttgtt ttttttttt 5880

ctaatgattc attaccgcta tgtataccta cttgtacttg tagtaagccg ggttattggc 5940

gttcaattaa tcatagactt atgaatctgc acggtgtgcg ctgcgagtta cttttagctt 6000

atgcatgcta cttgggtgta atattgggat ctgttcggaa atcaacggat gctcaatcga 6060

tttcgacagt aattaattaa gtcatacaca agtcagcttt cttcgagcct catataagta 6120

taagtagttc aacgtattag cactgtaccc agcatctccg tatcgagaaa cacaacaaca 6180

tgccccattg gacagatcat gcggatacac aggttgtgca gtatcataca tactcgatca 6240

gacaggtcgt ctgaccatca tacaagctga acaagcgctc catacttgca cgctctctat 6300

atacacagtt aaattacata tccatagtct aacctctaac agttaatctt ctggtaagcc 6360

tcccagccag ccttctggta tcgcttggcc tcctcaatag gatctcggtt ctggccgtac 6420

agacctcggc cgacaattat gatatccgtt ccggtagaca tgacatcctc aacagttcgg 6480

tactgctgtc cgagagcgtc tcccttgtcg tcaagaccca ccccgggggt cagaataagc 6540

cagtcctcag agtcgccctt aggtcggttc tgggcaatga agccaaccac aaactcgggg 6600

tcggatcggg caagctcaat ggtctgcttg gagtactcgc cagtggccag agagcccttg 6660

caagacagct cggccagcat gagcagacct ctggccagct tctcgttggg agaggggact 6720

aggaactcct tgtactggga gttctcgtag tcagagacgt cctccttctt ctgttcagag 6780

acagtttcct cggcaccagc tcgcaggcca gcaatgattc cggttccggg tacaccgtgg 6840

gcgttggtga tatcggacca ctcggcgatt cggtgacacc ggtactggtg cttgacagtg 6900

ttgccaatat ctgcgaactt tctgtcctcg aacaggaaga aaccgtgctt aagagcaagt 6960

tccttgaggg ggagcacagt gccggcgtag gtgaagtcgt caatgatgtc gatatgggtt 7020

ttgatcatgc acacataagg tccgacctta tcggcaagct caatgagctc cttggtggtg 7080

gtaacatcca gagaagcaca caggttggtt ttcttggctg ccacgagctt gagcactcga 7140

gcggcaaagg cggacttgtg gacgttagct cgagcttcgt aggagggcat tttggtggtg 7200

aagaggagac tgaaataaat ttagtctgca gaacttttta tcggaacctt atctggggca 7260

gtgaagtata tgttatggta atagttacga gttagttgaa cttatagata gactggacta 7320

tacggctatc ggtccaaatt agaaagaacg tcaatggctc tctgggcgtc gcctttgccg 7380

acaaaaatgt gatcatgatg aaagccagca atgacgttgc agctgatatt gttgtcggcc 7440

aaccgcgccg aaaacgcagc tgtcagaccc acagcctcca acgaagaatg tatcgtcaaa 7500

gtgatccaag cacactcata gttggagtcg tactccaaag gcggcaatga cgagtcagac 7560

agatactcgt cgactcaggc gacgacggaa ttcctgcagc ccatctgcag aattcaggag 7620

agaccgggtt ggcggcgtat ttgtgtccca aaaaacagcc ccaattgccc cggagaagac 7680

ggccaggccg cctagatgac aaattcaaca actcacagct gactttctgc cattgccact 7740

agggggggggc ctttttatat ggccaagcca agctctccac gtcggttggg ctgcacccaa 7800

caataaatgg gtagggttgc accaacaaag ggatgggatg gggggtagaa gatacgagga 7860

taacggggct caatggcaca aataagaacg aatactgcca ttaagactcg tgatccagcg 7920

actgacacca ttgcatcatc taagggcctc aaaactacct cggaactgct gcgctgatct 7980

ggacaccaca gaggttccga gcactttagg ttgcaccaaa tgtcccacca ggtgcaggca 8040

gaaaacgctg gaacagcgtg tacagtttgt cttaacaaaa agtgagggcg ctgaggtcga 8100

gcagggtggt gtgacttgtt atagccttta gagctgcgaa agcgcgtatg gatttggctc 8160

atcaggccag attgagggtc tgtggacaca tgtcatgtta gtgtacttca atcgcccccct 8220

ggatatagcc ccgacaatag gccgtggcct catttttttg ccttccgcac atttccattg 8280

ctcggtaccc acaccttgct tctcctgcac ttgccaacct taatactggt ttacattgac 8340

caacatctta caagcggggg gcttgtctag ggtatatata aacagtggct ctcccaatcg 8400

gttgccagtc tctttttttcc tttctttccc cacagattcg aaatctaaac tacacatcac 8460

acaatgcctg ttactgacgt ccttaagcga aagtccggtg tcatcgtcgg cgacgatgtc 8520

cgagccgtga gtatccacga caagatcagt gtcgagacga cgcgttttgt gtaatgacac 8580

aatccgaaag tcgctagcaa cacacactct ctacacaaac taacccagct ctc 8633

<210> 165
<211> 10660
<212> DNA

<213> Artificial Sequence

<220>
<223> Plasmid pZP2FOEN4

<400> 165
aaatagttgg agcaagggag aaatgtagag tgtgaaagac tcactatggt ccgggcttat 60

ctcgaccaat agccaaagtc tggagtttct gagagaaaaa ggcaagatac gtatgtaaca 120

aagcgacgca tggtacaata ataccggagg catgtatcat agagagttag tggttcgatg 180

atggcactgg tgcctggtat gactttatac ggctgactac atatttgtcc tcagacatac 240

aattacagtc aagcacttac ccttggacat ctgtaggtac cccccggcca agacgatctc 300

agcgtgtcgt atgtcggatt ggcgtagctc cctcgctcgt caattggctc ccatctactt 360

tcttctgctt ggctacaccc agcatgtctg ctatggctcg ttttcgtgcc ttatctatcc 420

tcccagtatt accaactcta aatgacatga tgtgattggg tctacacttt catatcagag 480

ataaggagta gcacagttgc ataaaaagcc caactctaat cagcttcttc ctttcttgta 540

attagtacaa aggtgattag cgaaatctgg aagcttagtt ggccctaaaa aaatcaaaaa 600

aagcaaaaaa cgaaaaacga aaaaccacag ttttgagaac agggaggtaa cgaaggatcg 660

tatatatata tatatatata tatacccacg gatcccgaga ccggcctttg attcttccct 720

acaaccaacc attctcacca ccctaattca caaccatggt cggcttcgac gagaccgtca 780

ctatggacac cgtccgaaac cacaacatgc ccgacgatgc ctggtgcgct atccacggca 840

ccgtgtacga catcaccaag ttctccaagg tgcatcctgg tggagacatt atcatgctgg 900

ccgctggcaa ggaggccacc atcctgttcg agacctacca catcaagggt gttcccgacg 960

ctgtcctgcg aaagtacaag gtcggcaagc tcgctcaggg caagaagggt gagacctccc 1020

acatgcccac cggactcgac tctgcctcct actactcctg ggactctgag ttctaccgag 1080

tcctccgaga gcgagtcgcc aagaagctgg ccgagcccgg tctcatgcag cgagctcgaa 1140

tggagctctg ggccaaggct atctttctcc tggctggttt ctggggatcc ctgtacgcca 1200

tgtgcgtgct cgaccctcac ggaggtgcta tggtcgctgc cgtcaccctc ggtgtgtttg 1260

ctgccttcgt cggaacctgc atccagcacg acggctccca cggagccttc tccaagtctc 1320

gattcatgaa caaggctgcc ggttggaccc tcgacatgat tggtgcctcc gctatgacct 1380

gggagatgca gcacgtcctc ggtcaccatc cctacaccaa cctgatcgag atggagaacg 1440

gtctggccaa ggtcaagggt gccgacgtgg accccaagaa ggtcgaccag gagtccgacc 1500

ctgacgtctt ctctacctac cctatgctgc gactgcatcc ctggcaccga cagcgattct 1560

accacaagtt ccagcacctg tacgctcccc tgatcttcgg cttcatgacc atcaacaagg 1620

tgatctccca ggacgtcggt gtcgtgctgc gaaagcgact gtttcagatc gacgccaact 1680

gccgatatgg ctctccctgg aacgtggctc gattctggat catgaagctc ctgaccactc 1740

tctacatggt ggctctgccc atgtacatgc agggacccgc tcagggtctg aagctgttct 1800

ttatggctca cttcacctgc ggagaggtcc tcgccaccat gttcattgtt aaccacatta 1860

tcgagggtgt ctcctacgcc tccaaggacg ctgtcaaggg cgtcatggct cctccccgaa 1920

ccgtgcacgg tgtcactccc atgcaggtga cccagaaggc tctctctgct gccgagtcca 1980

ccaagtctga cgctgacaag accactatga tccctctcaa cgactgggct gccgtccagt 2040

gccagacctc cgtcaactgg gctgtcggat cctggttctg gaaccacttc tctggaggtc 2100

tcaaccacca gattgagcac cattgctttc ctcagaaccc ccacaccgtc aacgtctaca 2160

tctccggcat cgtcaaggag acctgcgagg agtacggcgt gccctaccag gctgagatct 2220

ccctcttctc tgcctatttc aagatgctgt cccacctccg aaccctcggc aacgaggacc 2280

tcaccgcctg gtccacctaa gcggccgcat tgatgattgg aaacacacac atgggttata 2340

tctaggtgag agttagttgg acagttatat attaaatcag ctatgccaac ggtaacttca 2400

ttcatgtcaa cgaggaacca gtgactgcaa gtaatataga atttgaccac cttgccattc 2460

tcttgcactc ctttactata tctcatttat ttcttatata caaatcactt cttcttccca 2520

gcatcgagct cggaaacctc atgagcaata acatcgtgga tctcgtcaat agagggcttt 2580

ttggactcct tgctgttggc caccttgtcc ttgctgtctg gctcattctg tttcaacgcc 2640

ttttaattaa gacgcccaag aacttccttg gtgattatgg aatacgtggt ctgctatatc 2700

tcaattttgc tgtaatgaat cattagaatt aaaaaaaaaa ccccatttt gtgtgattgt 2760

cggccaagag atggaacagg aagaatacgt gaacaagcga gcacgaatgc catatgctct 2820

tctgaacaac cgagtccgaa tccgatttgt gggtatcaca tgtctcaagt agctgaaatg 2880

tatttcgcta gaataaaata aatgagatta agaattaaaa atattggaat atattttcct 2940

agaatagaaa ctttggattt tttttcggct attacagtct gaactggaca aacggctgac 3000

tatatataaa tattattggg tctgttttct tgtttatgtc gaaattatct gggttttact 3060

actgtgtcgt cgagtataga gtggcctgac tggagaaaat gcagtagtat ggacagtagg 3120

tactgccagc cagagaagtt tttggaattg atacttgagt cattttttcca ttccccattc 3180

cccattccaa cacaatcaac tgtttctgaa cattttccaa aacgcggaga tgtatgtcac 3240

ttgggcactg caagtttggg attcaaaatg catctctttc agaccaaagt gtcatcagcg 3300

catgcgacgt cgggcccaat tcgccctata gtgagtcgta ttacaattca ctggccgtcg 3360

ttttacaacg tcgtgactgg gaaaaccctg gcgttaccca acttaatcgc cttgcagcac 3420

atccccctttt cgccagctgg cgtaatagcg aagaggcccg caccgatcgc ccttcccaac 3480

agttgcgcag cctgaatggc gaatggacgc gccctgtagc ggcgcattaa gcgcggcggg 3540

tgtggtggtt acgcgcagcg tgaccgctac acttgccagc gccctagcgc ccgctccttt 3600

cgctttcttc ccttcctttc tcgccacgtt cgccggcttt ccccgtcaag ctctaaatcg 3660

ggggctccct ttagggttcc gatttagtgc tttacggcac ctcgacccca aaaaacttga 3720

ttagggtgat ggttcacgta gtgggccatc gccctgatag acggtttttc gccctttgac 3780

gttggagtcc acgttcttta atagtggact cttgttccaa actggaacaa cactcaaccc 3840

tatctcggtc tattcttttg atttataagg gattttgccg atttcggcct attggttaaa 3900

aaatgagctg atttaacaaa aatttaacgc gaattttaac aaaatattaa cgcttacaat 3960

ttcctgatgc ggtattttct ccttacgcat ctgtgcggta tttcacaccg catcaggtgg 4020

cacttttcgg ggaaatgtgc gcggaacccc tatttgttta ttttttctaaa tacattcaaa 4080

tatgtatccg ctcatgagac aataaccctg ataaatgctt caataatatt gaaaaaggaa 4140

gagtatgagt attcaacatt tccgtgtcgc ccttattccc ttttttgcgg cattttgcct 4200

tcctgttttt gctcacccag aaacgctggt gaaagtaaaa gatgctgaag atcagttggg 4260

tgcacgagtg ggttacatcg aactggatct caacagcggt aagatccttg agagttttcg 4320

ccccgaagaa cgttttccaa tgatgagcac ttttaaagtt ctgctatgtg gcgcggtatt 4380

atcccgtatt gacgccgggc aagagcaact cggtcgccgc atacactatt ctcagaatga 4440

cttggttgag tactcaccag tcacagaaaa gcatcttacg gatggcatga cagtaagaga 4500

attatgcagt gctgccataa ccatgagtga taacactgcg gccaacttac ttctgacaac 4560

gatcggagga ccgaaggagc taaccgcttt tttgcacaac atgggggatc atgtaactcg 4620

ccttgatcgt tgggaaccgg agctgaatga agccatacca aacgacgagc gtgacaccac 4680

gatgcctgta gcaatggcaa caacgttgcg caaactatta actggcgaac tacttactct 4740

agcttcccgg caacaattaa tagactggat ggaggcggat aaagttgcag gaccacttct 4800

gcgctcggcc cttccggctg gctggtttat tgctgataaa tctggagccg gtgagcgtgg 4860

gtctcgcggt atcattgcag cactggggcc agatggtaag ccctcccgta tcgtagttat 4920

ctacacgacg gggagtcagg caactatgga tgaacgaaat agacagatcg ctgagatagg 4980

tgcctcactg attaagcatt ggtaactgtc agaccaagtt tactcatata tactttagat 5040

tgatttaaaa cttcattttt aatttaaaag gatctaggtg aagatccttt ttgataatct 5100

catgaccaaa atcccttaac gtgagttttc gttccactga gcgtcagacc ccgtagaaaa 5160

gatcaaagga tcttcttgag atccttttttt tctgcgcgta atctgctgct tgcaaacaaa 5220

aaaaccaccg ctaccagcgg tggtttgttt gccggatcaa gagctaccaa ctcttttttcc 5280

gaaggtaact ggcttcagca gagcgcagat accaaatact gttcttctag tgtagccgta 5340

gttaggccac cacttcaaga actctgtagc accgcctaca tacctcgctc tgctaatcct 5400

gttaccagtg gctgctgcca gtggcgataa gtcgtgtctt accgggttgg actcaagacg 5460

atagttaccg gataaggcgc agcggtcggg ctgaacgggg ggttcgtgca cacagcccag 5520

cttggagcga acgacctaca ccgaactgag atacctacag cgtgagctat gagaaagcgc 5580

cacgcttccc gaagggagaa aggcggacag gtatccggta agcggcaggg tcggaacagg 5640

agagcgcacg agggagcttc cagggggaaa cgcctggtat ctttatagtc ctgtcgggtt 5700

tcgccacctc tgacttgagc gtcgattttt gtgatgctcg tcaggggggc ggagcctatg 5760

gaaaaacgcc agcaacgcgg cctttttacg gttcctggcc ttttgctggc cttttgctca 5820

catgttcttt cctgcgttat cccctgattc tgtggataac cgtattaccg cctttgagtg 5880

agctgatacc gctcgccgca gccgaacgac cgagcgcagc gagtcagtga gcgaggaagc 5940

ggaagagcgc ccaatacgca aaccgcctct ccccgcgcgt tggccgattc attaatgcag 6000

ctggcgcgcc ggcaatacgg gggagaagcg atcgcccgtc tgttcaatcg ccacaagtcc 6060

gagtaatgct cgagtatcga agtcttgtac ctccctgtca atcatggcac cactggtctt 6120

gacttgtcta ttcatactgg acaagcgcca gagttaagct tgtagcgaat ttcgccctcg 6180

gacatcaccc catacgacgg acacacatgc ccgacaaaca gcctctctta ttgtagctga 6240

aagtatattg aatgtgaacg tgtacaatat caggtaccag cgggaggtta cggccaaggt 6300

gataccggaa taaccctggc ttggagatgg tcggtccatt gtactgaagt gtccgtgtcg 6360

tttccgtcac tgccccaatt ggacatgttt gtttttccga tctttcgggc gccctctcct 6420

tgtctccttg tctgtctcct ggactgttgc taccccattt ctttggcctc cattggttcc 6480

tccccgtctt tcacgtcgtc tatggttgca tggtttccct tatacttttc cccacagtca 6540

catgttatgg aggggtctag agtcatacac aagtcagctt tcttcgagcc tcatataagt 6600

ataagtagtt caacgtatta gcactgtacc cagcatctcc gtatcgagaa acacaacaac 6660

atgccccatt ggacagatca tgcggataca caggttgtgc agtatcatac atactcgatc 6720

agacaggtcg tctgaccatc atacaagctg aacaagcgct ccatacttgc acgctctcta 6780

tatacacagt taaattacat atccatagtc taacctctaa cagttaatct tctggtaagc 6840

ctcccagcca gccttctggt atcgcttggc ctcctcaata ggatctcggt tctggccgta 6900

cagacctcgg ccgacaatta tgatatccgt tccggtagac atgacatcct caacagttcg 6960

gtactgctgt ccgagagcgt ctcccttgtc gtcaagaccc accccggggg tcagaataag 7020

ccagtcctca gagtcgccct taggtcggtt ctgggcaatg aagccaacca caaactcggg 7080

gtcggatcgg gcaagctcaa tggtctgctt ggagtactcg ccagtggcca gagagccctt 7140

gcaagacagc tcggccagca tgagcagacc tctggccagc ttctcgttgg gagaggggac 7200

taggaactcc ttgtactggg agttctcgta gtcagagacg tcctccttct tctgttcaga 7260

gacagtttcc tcggcaccag ctcgcaggcc agcaatgatt ccggttccgg gtacaccgtg 7320

ggcgttggtg atatcggacc actcggcgat tcggtgacac cggtactggt gcttgacagt 7380

gttgccaata tctgcgaact ttctgtcctc gaacaggaag aaaccgtgct taagagcaag 7440

ttccttgagg gggagcacag tgccggcgta ggtgaagtcg tcaatgatgt cgatatgggt 7500

tttgatcatg cacacataag gtccgacctt atcggcaagc tcaatgagct ccttggtggt 7560

ggtaacatcc agagaagcac acaggttggt tttcttggct gccacgagct tgagcactcg 7620

agcggcaaag gcggacttgt ggacgttagc tcgagcttcg taggagggca ttttggtggt 7680

gaagaggaga ctgaaataaa tttagtctgc agaacttttt atcggaacct tatctggggc 7740

agtgaagtat atgttatggt aatagttacg agttagttga acttatagat agactggact 7800

atacggctat cggtccaaat tagaaagaac gtcaatggct ctctgggcgt cgcctttgcc 7860

gacaaaaatg tgatcatgat gaaagccagc aatgacgttg cagctgatat tgttgtcggc 7920

caaccgcgcc gaaaacgcag ctgtcagacc cacagcctcc aacgaagaat gtatcgtcaa 7980

agtgatccaa gcacactcat agttggagtc gtactccaaa ggcggcaatg acgagtcaga 8040

cagatactca ctagtgcttc caaacactat acagtggttc tgggtcgttt atggcaagga 8100

atactcactc agcagttgcc ggctgggcag caggggcttt tgtgttggaa acatatcgca 8160

gaccaagaga atacttggat ctggagattc tggccaggtt agtggcccgt cgggcgagtc 8220

gcgattgcat gttagtatac agagacaaat caagacgctt gcaatcttat aaaccacccg 8280

gtgagtaatt gcggagacaa ggacgtgaaa aattgccaac cttaatccgg ctgcagggta 8340

agcgataaac accactaaaa ccccacaaaa tatatcttac cgaatataca gatcccgtca 8400

acagttttat atatcgtagt tacaaccatc aacacttttt ggtaagtgta ccattctata 8460

ctccaactgg tctgcaactg tacaagtaga catgttaatg gtagttaata acatctacag 8520

cagaacctat ggtaaagaca ttgcattttt acaggaagta tcgtcctaca cgttgataaa 8580

tccaaagatg cggaacttct tccactttta tcatcatccc ctactcgtac actcgtactc 8640

tttgttcgat cgcgattcat ttctataaat aatcttgtat gtacatgcgg ccgcttagtc 8700

gatctttcga gatcgtgttc ttcgaacgga gggtgctcgg gtagtctcgg caggtttgac 8760

ggagctggca ccgtcgcctc gagacttgtt ggagtaggct ttgaggtaga agttgccaaa 8820

cagcacgagc atgttggtca tgacgaacat ctgagcccaa ggaagtgtaa cgggacagtg 8880

cttctgtctc aagacaaaca cggcatgagc gaagacaatg acgaattgca gcatctgtgc 8940

ctgagtgatg tatcgcttcc agggacatcg aatgccaaga gcagacatga gatagtacga 9000

gtacatgacg atgtgaatga aggaattgca ggcagcacca aagtaggcgt cgatgcaatc 9060

gttggtagcc atgagatgac acaccagcca ccaggcccaa atgagcagtg cgtgatggta 9120

aacgtggaga aacgacagct gcttggtctt cttccgagca accatgaaca cggtgtcaag 9180

cagctccaga tacttgttgt tgtagtgcag ccacacaccg agcagaatct tgaaggactt 9240

tcgatcgctc caaggcatgg tggaacccca gacgggctgt cccagtccag agatttctcg 9300

agcaaacatt ccgagcacga caacgttgaa agcagtctgg taggcgttgt atgccagcat 9360

gaagcccttg ggatcaaagg cctctcgctt ggccatgagt cgaggaccga ccaggcagaa 9420

gagcaggtac atggtagtgg caatagcggg aagtcgcaag gacagggctc cgatccactc 9480

tcgtgcgtcc acgttgtcga tgccgttggc gtgcgaccac tcgaaggcgt aggcgtaggt 9540

agcgtaggcg taggcagcgg aggcaatggc aggaagcaga gcaccagagg cagccatgga 9600

gagctgggtt agtttgtgta gagagtgtgt gttgctagcg actttcggat tgtgtcatta 9660

cacaaaacgc gtcgtctcga cactgatctt gtcgtggata ctcacggctc ggacatcgtc 9720

gccgacgatg acaccggact ttcgcttaag gacgtcagta acaggcattg tgtgatgtgt 9780

agtttagatt tcgaatctgt ggggaaagaa aggaaaaaag agactggcaa ccgattggga 9840

gagccactgt ttatatatac cctagacaag ccccccgctt gtaagatgtt ggtcaatgta 9900

aaccagtatt aaggttggca agtgcaggag aagcaaggtg tgggtaccga gcaatggaaa 9960

tgtgcggaag gcaaaaaaat gaggccacgg cctattgtcg gggctatatc cagggggcga 10020

ttgaagtaca ctaacatgac atgtgtccac agaccctcaa tctggcctga tgagccaaat 10080

ccatacgcgc tttcgcagct ctaaaggcta taacaagtca caccaccctg ctcgacctca 10140

gcgccctcac tttttgttaa gacaaactgt acacgctgtt ccagcgtttt ctgcctgcac 10200

ctggtgggac atttggtgca acctaaagtg ctcggaacct ctgtggtgtc cagatcagcg 10260

cagcagttcc gaggtagttt tgaggccctt agatgatgca atggtgtcag tcgctggatc 10320

acgagtctta atggcagtat tcgttcttat ttgtgccatt gagccccgtt atcctcgtat 10380

cttctacccc ccatcccatc cctttgttgg tgcaacccta cccatttatt gttgggtgca 10440

gcccaaccga cgtggagagc ttggcttggc catataaaaa ggccccccccc tagtggcaat 10500

ggcagaaagt cagctgtgag ttgttgaatt tgtcatctag gcggcctggc cgtcttctcc 10560

ggggcaattg gggctgtttt ttgggacaca aatacgccgc caacccggtc tctcctgaat 10620

tctgcagatg ggctgcagga attccgtcgt cgcctgagtc 10660

<210> 166

<211> 6457
<212> DNA
<213> Artificial Sequence

<220>
<223> Plasmid pKUNT2

<400> 166
tttgaatcga atcgatgagc ctaaaatgaa cccgagtata tctcataaaa ttctcggtga 60

gaggtctgtg actgtcagta caaggtgcct tcattatgcc ctcaacctta ccatacctca 120

ctgaatgtag tgtacctcta aaaatgaaat acagtgccaa aagccaaggc actgagctcg 180

tctaacggac ttgatataca accaattaaa acaaatgaaa agaaatacag ttctttgtat 240

catttgtaac aattaccctg tacaaactaa ggtattgaaa tcccacaata ttcccaaagt 300

ccaccccttt ccaaattgtc atgcctacaa ctcatatacc aagcactaac ctaccgttta 360

aacaccacta aaaccccaca aaatatatct taccgaatat acagatctgc gacgacggaa 420

ttcctgcagc ccatctgcag aattcaggag agaccgggtt ggcggcgtat ttgtgtccca 480

aaaaacagcc ccaattgccc caattgaccc caaattgacc cagtagcggg cccaaccccg 540

gcgagagccc ccttcacccc acatatcaaa cctcccccgg ttcccacact tgccgttaag 600

ggcgtagggt actgcagtct ggaatctacg cttgttcaga ctttgtacta gtttctttgt 660

ctggccatcc gggtaaccca tgccggacgc aaaatagact actgaaaatt tttttgcttt 720

gtggttggga ctttagccaa gggtataaaa gaccaccgtc cccgaattac ctttcctctt 780

cttttctctc tctccttgtc aactcacacc cgaaatcgtt aagcatttcc ttctgagtat 840

aagaatcatt caccatggat tcgaccacgc agaccaacac cggcaccggc aaggtggccg 900

tgcagccccc cacggccttc attaagccca ttgagaaggt gtccgagccc gtctacgaca 960

cctttggcaa cgagttcact cctccagact actctatcaa ggatattctg gatgccattc 1020

cccaggagtg ctacaagcgg tcctacgtta agtcctactc gtacgtggcc cgagactgct 1080

tctttatcgc cgtttttgcc tacatggcct acgcgtacct gcctcttatt ccctcggctt 1140

ccggccgagc tgtggcctgg gccatgtact ccattgtcca gggtctgttt ggcaccggtc 1200

tgtgggttct tgcccacgag tgtggccact ctgctttctc cgactctaac accgtcaaca 1260

acgtcaccgg atgggttctg cactcctcca tgctggtccc ttactacgcc tggaagctga 1320

cccactccat gcaccacaag tccactggtc acctcacccg tgatatggtg tttgtgccca 1380

aggaccgaaa ggagtttatg gagaaccgag gcgcccatga ctggtctgag cttgctgagg 1440

acgctcccct catgaccctc tacggcctca tcacccagca ggtgtttgga tggcctctgt 1500

atctgctgtc ttacgttacc ggacagaagt accccaagct caacaaatgg gctgtcaacc 1560

acttcaaccc caacgccccg ctgtttgaga agaaggactg gttcaacatc tggatctcta 1620

acgtcggtat tggtatcacc atgtccgtca tcgcatactc catcaaccga tggggcctgg 1680

cttccgtcac cctctactac ctgatcccct acctgtgggt caaccactgg ctcgtggcca 1740

tcacctacct gcagcacacc gaccccactc tgccccacta ccacgccgac cagtggaact 1800

tcacccgagg agccgccgcc accatcgacc gagagtttgg cttcatcggc tccttctgct 1860

tccatgacat catcgagacc cacgttctgc accactacgt gtctcgaatt cccttctaca 1920

acgcccgaat cgccactgag aagatcaaga aggtcatggg caagcactac cgacacgacg 1980

acaccaactt catcaagtct ctttacactg tcgcccgaac ctgccagttt gttgaaggta 2040

aggaaggcat tcagatgttt agaaacgtca atggagtcgg agttgctcct gacggcctgc 2100

cttctaaaaa gtaggcggcc gcaagtgtgg atggggaagt gagtgcccgg ttctgtgtgc 2160

acaattggca atccaagatg gatggattca acacagggat atagcgagct acgtggtggt 2220

gcgaggatat agcaacggat atttatgttt gacacttgag aatgtacgat acaagcactg 2280

tccaagtaca atactaaaca tactgtacat actcatactc gtacccgggc aacggtttca 2340

cttgagtgca gtggctagtg ctcttactcg tacagtgtgc aatactgcgt atcatagtct 2400

ttgatgtata tcgtattcat tcatgttagt tgcgtacgaa gtcgtcaatg atgtcgatat 2460

gggttttgat catgcacaca taaggtccga ccttatcggc aagctcaatg agctccttgg 2520

tggtggtaac atccagagaa gcacacaggt tggttttctt ggctgccacg agcttgagca 2580

ctcgagcggc aaaggcggac ttgtggacgt tagctcgagc ttcgtaggag ggcattttgg 2640

tggtgaagag gagactgaaa taaatttagt ctgcagaact ttttatcgga accttatctg 2700

gggcagtgaa gtatatgtta tggtaatagt tacgagttag ttgaacttat agatagactg 2760

gactatacgg ctatcggtcc aaattagaaa gaacgtcaat ggctctctgg gcgtcgcctt 2820

tgccgacaaa aatgtgatca tgatgaaagc cagcaatgac gttgcagctg atattgttgt 2880

cggccaaccg cgccgaaaac gcagctgtca gacccacagc ctccaacgaa gaatgtatcg 2940

tcaaagtgat ccaagcacac tcatagttgg agtcgtactc caaaggcggc aatgacgagt 3000

cagacagata ctcgtcgacc ttttccttgg gaaccaccac cgtcagccct tctgactcac 3060

gtattgtagc caccgacaca ggcaacagtc cgtggatagc agaatatgtc ttgtcggtcc 3120

atttctcacc aactttaggc gtcaagtgaa tgttgcagaa gaagtatgtg ccttcattga 3180

gaatcggtgt tgctgatttc aataaagtct tgagatcagt ttggcgcgcc agctgcatta 3240

atgaatcggc caacgcgcgg ggagaggcgg tttgcgtatt gggcgctctt ccgcttcctc 3300

gctcactgac tcgctgcgct cggtcgttcg gctgcggcga gcggtatcag ctcactcaaa 3360

ggcggtaata cggttatcca cagaatcagg ggataacgca ggaaagaaca tgtgagcaaa 3420

aggccagcaa aaggccagga accgtaaaaa ggccgcgttg ctggcgtttt tccataggct 3480

ccgccccccct gacgagcatc acaaaaatcg acgctcaagt cagaggtggc gaaacccgac 3540

aggactataa agataccagg cgtttccccc tggaagctcc ctcgtgcgct ctcctgttcc 3600

gaccctgccg cttaccggat acctgtccgc ctttctccct tcgggaagcg tggcgctttc 3660

tcatagctca cgctgtaggt atctcagttc ggtgtaggtc gttcgctcca agctgggctg 3720

tgtgcacgaa ccccccgttc agcccgaccg ctgcgcctta tccggtaact atcgtcttga 3780

gtccaacccg gtaagacacg acttatcgcc actggcagca gccactggta acaggattag 3840

cagagcgagg tatgtaggcg gtgctacaga gttcttgaag tggtggccta actacggcta 3900

cactagaaga acagtatttg gtatctgcgc tctgctgaag ccagttacct tcggaaaaag 3960

agttggtagc tcttgatccg gcaaacaaac caccgctggt agcggtggtt tttttgtttg 4020

caagcagcag attacgcgca gaaaaaaagg atctcaagaa gatcctttga tcttttctac 4080

ggggtctgac gctcagtgga acgaaaactc acgttaaggg attttggtca tgagattatc 4140

aaaaaggatc ttcacctaga tccttttaaa ttaaaaatga agttttaaat caatctaaag 4200

tatatatgag taaacttggt ctgacagtta ccaatgctta atcagtgagg cacctatctc 4260

agcgatctgt ctatttcgtt catccatagt tgcctgactc cccgtcgtgt agataactac 4320

gatacgggag ggcttaccat ctggccccag tgctgcaatg ataccgcgag acccacgctc 4380

accggctcca gatttatcag caataaacca gccagccgga agggccgagc gcagaagtgg 4440

tcctgcaact ttatccgcct ccatccagtc tattaattgt tgccgggaag ctagagtaag 4500

tagttcgcca gttaatagtt tgcgcaacgt tgttgccatt gctacaggca tcgtggtgtc 4560

acgctcgtcg tttggtatgg cttcattcag ctccggttcc caacgatcaa ggcgagttac 4620

atgatccccc atgttgtgca aaaaagcggt tagctccttc ggtcctccga tcgttgtcag 4680

aagtaagttg gccgcagtgt tatcactcat ggttatggca gcactgcata attctcttac 4740

tgtcatgcca tccgtaagat gcttttctgt gactggtgag tactcaacca agtcattctg 4800

agaatagtgt atgcggcgac cgagttgctc ttgcccggcg tcaatacggg ataataccgc 4860

gccacatagc agaactttaa aagtgctcat cattggaaaa cgttcttcgg ggcgaaaact 4920

ctcaaggatc ttaccgctgt tgagatccag ttcgatgtaa cccactcgtg cacccaactg 4980

atcttcagca tcttttactt tcaccagcgt ttctgggtga gcaaaaacag gaaggcaaaa 5040

tgccgcaaaa aagggaataa gggcgacacg gaaatgttga atactcatac tcttcctttt 5100

tcaatattat tgaagcattt atcagggtta ttgtctcatg agcggataca tatttgaatg 5160

tatttagaaa aataaacaaa taggggttcc gcgcacattt ccccgaaaag tgccacctga 5220

tgcggtgtga aataccgcac agatgcgtaa ggagaaaata ccgcatcagg aaattgtaag 5280

cgttaatatt ttgttaaaat tcgcgttaaa tttttgttaa atcagctcat tttttaacca 5340

ataggccgaa atcggcaaaa tcccttataa atcaaaagaa tagaccgaga tagggttgag 5400

tgttgttcca gtttggaaca agagtccact attaaagaac gtggactcca acgtcaaagg 5460

gcgaaaaacc gtctatcagg gcgatggccc actacgtgaa ccatcaccct aatcaagttt 5520

tttggggtcg aggtgccgta aagcactaaa tcggaaccct aaagggagcc cccgatttag 5580

agcttgacgg ggaaagccgg cgaacgtggc gagaaaggaa gggaagaaag cgaaaggagc 5640

gggcgctagg gcgctggcaa gtgtagcggt cacgctgcgc gtaaccacca cacccgccgc 5700

gcttaatgcg ccgctacagg gcgcgtccat tcgccattca ggctgcgcaa ctgttgggaa 5760

gggcgatcgg tgcgggcctc ttcgctatta cgccagctgg cgaaaggggg atgtgctgca 5820

aggcgattaa gttgggtaac gccagggttt tcccagtcac gacgttgtaa aacgacggcc 5880

agtgaattgt aatacgactc actatagggc gaattgggcc cgacgtcgca tgcagtggtg 5940

gtattgtgac tggggatgta gttgagaata agtcatacac aagtcagctt tcttcgagcc 6000

tcatataagt ataagtagtt caacgtatta gcactgtacc cagcatctcc gtatcgagaa 6060

acacaacaac atgccccatt ggacagatca tgcggataca caggttgtgc agtatcatac 6120

atactcgatc agacaggtcg tctgaccatc atacaagctg aacaagcgct ccatacttgc 6180

acgctctcta tatacacagt taaattacat atccatagtc taacctctaa cagttaatct 6240

tctggtaagc ctcccagcca gccttctggt atcgcttggc ctcctcaata ggatctcggt 6300

tctggccgta cagacctcgg ccgacaatta tgatatccgt tccggtagac atgacatcct 6360

caacagttcg gtactgctgt ccgagagcgt ctcccttgtc gtcaagaccc accccggggg 6420

tcagaataag ccagtcctca gagtcgccct taattaa 6457

<210> 167
<211> 7879
<212> DNA
<213> Artificial Sequence

<220>
<223> Plasmid pDMW237

<400> 167
ggccgcaagt gtggatgggg aagtgagtgc ccggttctgt gtgcacaatt ggcaatccaa 60

gatggatgga ttcaacacag ggatatagcg agctacgtgg tggtgcgagg atatagcaac 120

ggatatttat gtttgacact tgagaatgta cgatacaagc actgtccaag tacaatacta 180

aacatactgt acatactcat actcgtaccc gggcaacggt ttcacttgag tgcagtggct 240

agtgctctta ctcgtacagt gtgcaatact gcgtatcata gtctttgatg tatatcgtat 300

tcattcatgt tagttgcgta cgagccggaa gcataaagtg taaagcctgg ggtgcctaat 360

gagtgagcta actcacatta attgcgttgc gctcactgcc cgctttccag tcgggaaacc 420

tgtcgtgcca gctgcattaa tgaatcggcc aacgcgcggg gagaggcggt ttgcgtattg 480

ggcgctcttc cgcttcctcg ctcactgact cgctgcgctc ggtcgttcgg ctgcggcgag 540

cggtatcagc tcactcaaag gcggtaatac ggttatccac agaatcaggg gataacgcag 600

gaaagaacat gtgagcaaaa ggccagcaaa aggccaggaa ccgtaaaaag gccgcgttgc 660

tggcgttttt ccataggctc cgccccccctg acgagcatca caaaaatcga cgctcaagtc 720

agaggtggcg aaacccgaca ggactataaa gataccaggc gtttcccccct ggaagctccc 780

tcgtgcgctc tcctgttccg accctgccgc ttaccggata cctgtccgcc tttctccctt 840

cgggaagcgt ggcgctttct catagctcac gctgtaggta tctcagttcg gtgtaggtcg 900

ttcgctccaa gctgggctgt gtgcacgaac cccccgttca gcccgaccgc tgcgccttat 960

ccggtaacta tcgtcttgag tccaacccgg taagacacga cttatcgcca ctggcagcag 1020

ccactggtaa caggattagc agagcgaggt atgtaggcgg tgctacagag ttcttgaagt 1080

ggtggcctaa ctacggctac actagaagga cagtatttgg tatctgcgct ctgctgaagc 1140

cagttacctt cggaaaaaga gttggtagct cttgatccgg caaacaaacc accgctggta 1200

gcggtggttt ttttgtttgc aagcagcaga ttacgcgcag aaaaaaagga tctcaagaag 1260

atcctttgat cttttctacg gggtctgacg ctcagtggaa 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 cccacgctca ccggctccag atttatcagc aataaaccag ccagccggaa 1620

gggccgagcg cagaagtggt cctgcaactt tatccgcctc catccagtct attaattgtt 1680

gccgggaagc tagagtaagt agttcgccag ttaatagttt gcgcaacgtt gttgccattg 1740

ctacaggcat cgtggtgtca cgctcgtcgt ttggtatggc ttcattcagc tccggttccc 1800

aacgatcaag gcgagttaca tgatccccca tgttgtgcaa aaaagcggtt agctccttcg 1860

gtcctccgat cgttgtcaga agtaagttgg ccgcagtgtt atcactcatg gttatggcag 1920

cactgcataa ttctcttact gtcatgccat ccgtaagatg cttttctgtg actggtgagt 1980

actcaaccaa gtcattctga gaatagtgta tgcggcgacc gagttgctct tgcccggcgt 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 caatattatt gaagcattta tcagggttat tgtctcatga 2340

gcggatacat atttgaatgt atttagaaaa ataaacaaat aggggttccg cgcacatttc 2400

cccgaaaagt gccacctgac gcgccctgta gcggcgcatt aagcgcggcg ggtgtggtgg 2460

ttacgcgcag cgtgaccgct acacttgcca gcgccctagc gcccgctcct ttcgctttct 2520

tcccttcctt tctcgccacg ttcgccggct ttccccgtca agctctaaat cgggggctcc 2580

ctttagggtt ccgatttagt gctttacggc acctcgaccc caaaaaactt gattagggtg 2640

atggttcacg tagtgggcca tcgccctgat agacggtttt tcgccctttg acgttggagt 2700

ccacgttctt taatagtgga ctcttgttcc 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 aaaggggatg tgctgcaag gcgattaagt tgggtaacgc cagggttttc 3000

ccagtcacga cgttgtaaaa cgacggccag tgaattgtaa tacgactcac tatagggcga 3060

attgggtacc gggcccccccc tcgaggtcga tggtgtcgat 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 gttctcaata tttaaggggt 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 tcccttgtac aacataaata gtcatcgaga 3660

aatatcaact atcaaagaac agctattcac acgttactat tgagattatt attggacgag 3720

aatcacacac tcaactgtct ttctctcttc tagaaataca ggtacaagta tgtactattc 3780

tcattgttca tacttctagt catttcatcc cacatattcc 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 cccctcgttc agtgtcaact 4080

gtaatggtag gaaattacca tacttttgaa gaagcaaaaa aaatgaaaga aaaaaaaaat 4140

cgtatttcca ggttagacgt tccgcagaat ctagaatgcg gtatgcggta cattgttctt 4200

cgaacgtaaa agttgcgctc cctgagatat tgtacatttt tgcttttaca agtacaagta 4260

catcgtacaa ctatgtacta ctgttgatgc atccacaaca gtttgttttg ttttttttg 4320

ttttttttt ttctaatgat 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 gtctgaccat 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 caccccgggg 4980

gtcagaataa gccagtcctc agagtcgccc ttaggtcggt tctgggcaat gaagccaacc 5040

acaaactcgg ggtcggatcg ggcaagctca atggtctgct tggagtactc gccagtggcc 5100

agagagccct tgcaagacag ctcggccagc atgagcagac ctctggccag cttctcgttg 5160

ggagagggga ctaggaactc cttgtactgg gagttctcgt agtcagagac gtcctccttc 5220

ttctgttcag agacagtttc ctcggcacca gctcgcaggc cagcaatgat tccggttccg 5280

ggtacaccgt gggcgttggt gatatcggac cactcggcga ttcggtgaca ccggtactgg 5340

tgcttgacag tgttgccaat 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

attttggtgg tgaagaggag actgaaataa atttagtctg cagaactttt tatcggaacc 5700

ttatctgggg cagtgaagta tatgttatgg taatagttac gagttagttg 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 cgtactccaa aggcggcaat 6000

gacgagtcag acagatactc gtcgactcag gcgacgacgg aattcctgca gcccatctgc 6060

agaattcagg agagaccggg ttggcggcgt atttgtgtcc caaaaaacag ccccaattgc 6120

cccggagaag acggccaggc cgcctagatg acaaattcaa caactcacag ctgactttct 6180

gccattgcca ctaggggggg gccttttttat atggccaagc caagctctcc acgtcggttg 6240

ggctgcaccc aacaataaat gggtagggtt gcaccaacaa agggatggga tgggggggtag 6300

aagatacgag gataacgggg ctcaatggca caaataagaa cgaatactgc cattaagact 6360

cgtgatccag cgactgacac cattgcatca tctaagggcc tcaaaactac ctcggaactg 6420

ctgcgctgat ctggacacca cagaggttcc gagcacttta ggttgcacca aatgtcccac 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 aggccgtggc ctcatttttt tgccttccgc 6720

acatttccat tgctcggtac ccacaccttg cttctcctgc acttgccaac cttaatactg 6780

gtttacattg accaacatct tacaagcggg gggcttgtct agggtatata taaacagtgg 6840

ctctcccaat cggttgccag tctcttttttt ccttttctttc cccacagatt cgaaatctaa 6900

actacacatc acacaatgcc tgttactgac gtccttaagc gaaagtccgg tgtcatcgtc 6960

ggcgacgatg tccgagccgt gagtatccac gacaagatca gtgtcgagac gacgcgtttt 7020

gtgtaatgac acaatccgaa agtcgctagc aacacacact ctctacacaa actaacccag 7080

ctctccatgg ctctggccaa cgacgctggc gagcgaatct gggctgccgt caccgatccc 7140

gaaatcctca ttggcacctt ctcctacctg ctcctgaagc ctctcctgcg aaactctggt 7200

ctcgtggacg agaagaaagg agcctaccga acctccatga tctggtacaa cgtcctcctg 7260

gctctcttct ctgccctgtc cttctacgtg actgccaccg ctctcggctg ggactacggt 7320

actggagcct ggctgcgaag acagaccggt gatactcccc agcctctctt tcagtgtccc 7380

tctcctgtct gggactccaa gctgttcacc tggactgcca aggccttcta ctattctaag 7440

tacgtggagt acctcgacac cgcttggctg gtcctcaagg gcaagcgagt gtcctttctg 7500

caggccttcc atcactttgg agctccctgg gacgtctacc tcggcattcg actgcacaac 7560

gagggtgtgt ggatcttcat gttctttaac tcgttcattc acaccatcat gtacacctac 7620

tatggactga ctgccgctgg ctacaagttc aaggccaagc ctctgatcac tgccatgcag 7680

atttgccagt tcgtcggtgg ctttctcctg gtctgggact acatcaacgt tccctgcttc 7740

aactctgaca agggcaagct gttctcctgg gctttcaact acgcctacgt cggatctgtc 7800

tttctcctgt tctgtcactt cttttaccag gacaacctgg ccaccaagaa atccgctaag 7860

gctggtaagc agctttagc 7879

<210> 168
<211> 8502
<212> DNA
<213> Artificial Sequence

<220>
<223> Plasmid pY54PC

<400> 168
ggccgccacc gcggcccgag attccggcct cttcggccgc caagcgaccc gggtggacgt 60

ctagaggtac ctagcaatta acagatagtt tgccggtgat aattctctta acctcccaca 120

ctcctttgac ataacgattt atgtaacgaa actgaaattt gaccagatat tgtgtccgcg 180

gtggagctcc agcttttgtt ccctttagtg agggttaatt aatcgagctt ggcgtaatca 240

tggtcatagc tgtttcctgt gtgaaattgt tatccgctca caattccaca caacatacga 300

gccggaagca taaagtgtaa agcctggggt gcctaatgag tgagctaact cacattaatt 360

gcgttgcgct cactgcccgc tttccagtcg ggaaacctgt cgtgccagct gcattaatga 420

atcggccaac gcgcggggag aggcggtttg cgtattgggc gctcttccgc ttcctcgctc 480

actgactcgc tgcgctcggt cgttcggctg cggcgagcgg tatcagctca ctcaaaggcg 540

gtaatacggt tatccacaga atcaggggat aacgcaggaa agaacatgtg agcaaaaggc 600

cagcaaaagg ccaggaaccg taaaaaggcc gcgttgctgg cgtttttcca taggctccgc 660

cccccctgacg agcatcacaa aaatcgacgc tcaagtcaga ggtggcgaaa cccgacagga 720

ctataaagat accaggcgtt tccccctgga agctccctcg tgcgctctcc tgttccgacc 780

ctgccgctta ccggatacct gtccgccttt ctcccttcgg gaagcgtggc gctttctcat 840

agctcacgct gtaggtatct cagttcggtg taggtcgttc gctccaagct gggctgtgtg 900

cacgaacccc ccgttcagcc cgaccgctgc gccttatccg gtaactatcg tcttgagtcc 960

aacccggtaa gacacgactt atcgccactg gcagcagcca ctggtaacag gattagcaga 1020

gcgaggtatg taggcggtgc tacagagttc ttgaagtggt ggcctaacta cggctacact 1080

agaaggacag tatttggtat ctgcgctctg ctgaagccag ttaccttcgg aaaaagagtt 1140

ggtagctctt gatccggcaa acaaaccacc gctggtagcg gtggtttttt tgtttgcaag 1200

cagcagatta cgcgcagaaa aaaaggatct caagaagatc ctttgatctt ttctacgggg 1260

tctgacgctc agtggaacga aaactcacgt taagggattt tggtcatgag attatcaaaa 1320

aggatcttca cctagatcct tttaaattaa aaatgaagtt ttaaatcaat ctaaagtata 1380

tatgagtaaa cttggtctga cagttaccaa tgcttaatca gtgaggcacc tatctcagcg 1440

atctgtctat ttcgttcatc catagttgcc tgactccccg tcgtgtagat aactacgata 1500

cgggagggct taccatctgg ccccagtgct gcaatgatac cgcgagaccc acgctcaccg 1560

gctccagatt tatcagcaat aaaccagcca gccggaaggg ccgagcgcag aagtggtcct 1620

gcaactttat ccgcctccat ccagtctatt aattgttgcc gggaagctag agtaagtagt 1680

tcgccagtta atagtttgcg caacgttgtt gccattgcta caggcatcgt ggtgtcacgc 1740

tcgtcgtttg gtatggcttc attcagctcc ggttcccaac gatcaaggcg agttacatga 1800

tcccccatgt tgtgcaaaaa agcggttagc tccttcggtc ctccgatcgt tgtcagaagt 1860

aagttggccg cagtgttatc actcatggtt atggcagcac tgcataattc tcttactgtc 1920

atgccatccg taagatgctt ttctgtgact ggtgagtact caaccaagtc attctgagaa 1980

tagtgtatgc ggcgaccgag ttgctcttgc ccggcgtcaa tacgggataa taccgcgcca 2040

catagcagaa ctttaaaagt gctcatcatt ggaaaacgtt cttcggggcg aaaactctca 2100

aggatcttac cgctgttgag atccagttcg atgtaaccca ctcgtgcacc caactgatct 2160

tcagcatctt ttactttcac cagcgtttct gggtgagcaa aaacaggaag gcaaaatgcc 2220

gcaaaaaagg gaataagggc gacacggaaa tgttgaatac tcatactctt cctttttcaa 2280

tattattgaa gcatttatca gggttattgt ctcatgagcg gatacatatt tgaatgtatt 2340

tagaaaaata aacaaatagg ggttccgcgc acatttcccc gaaaagtgcc acctgacgcg 2400

ccctgtagcg gcgcattaag cgcggcgggt gtggtggtta cgcgcagcgt gaccgctaca 2460

cttgccagcg ccctagcgcc cgctcctttc gctttcttcc cttcctttct cgccacgttc 2520

gccggctttc cccgtcaagc tctaaatcgg gggctccctt tagggttccg atttagtgct 2580

ttacggcacc tcgaccccaa aaaacttgat tagggtgatg gttcacgtag tgggccatcg 2640

ccctgataga cggtttttcg ccctttgacg ttggagtcca cgttctttaa tagtggactc 2700

ttgttccaaa ctggaacaac actcaaccct atctcggtct attcttttga tttataaggg 2760

attttgccga tttcggccta ttggttaaaa aatgagctga tttaacaaaa atttaacgcg 2820

aattttaaca aaatattaac gcttacaatt tccattcgcc attcaggctg cgcaactgtt 2880

gggaagggcg atcggtgcgg gcctcttcgc tattacgcca gctggcgaaa gggggatgtg 2940

ctgcaaggcg attaagttgg gtaacgccag ggttttccca gtcacgacgt tgtaaaacga 3000

cggccagtga attgtaatac gactcactat agggcgaatt gggtaccggg ccccccctcg 3060

aggtcgacgg tatcgataag cttgatatcg aattcatgtc acacaaaccg atcttcgcct 3120

caaggaaacc taattctaca tccgagagac tgccgagatc cagtctacac tgattaattt 3180

tcgggccaat aatttaaaaa aatcgtgtta tataatatta tatgtattat atatatacat 3240

catgatgata ctgacagtca tgtcccattg ctaaatagac agactccatc tgccgcctcc 3300

aactgatgtt ctcaatattt aaggggtcat ctcgcattgt ttaataataa acagactcca 3360

tctaccgcct ccaaatgatg ttctcaaaat atattgtatg aacttatttt tattacttag 3420

tattattaga caacttactt gctttatgaa aaacacttcc tatttaggaa acaatttata 3480

atggcagttc gttcatttaa caatttatgt agaataaatg ttataaatgc gtatgggaaa 3540

tcttaaatat ggatagcata aatgatatct gcattgccta attcgaaatc aacagcaacg 3600

aaaaaaatcc cttgtacaac ataaatagtc atcgagaaat atcaactatc aaagaacagc 3660

tattcacacg ttactattga gattattatt ggacgagaat cacacactca actgtctttc 3720

tctcttctag aaatacaggt acaagtatgt actattctca ttgttcatac ttctagtcat 3780

ttcatcccac atattccttg gatttctctc caatgaatga cattctatct tgcaaattca 3840

acaattataa taagatatac caaagtagcg gtatagtggc aatcaaaaag cttctctggt 3900

gtgcttctcg tatttatttt tattctaatg atccattaaa ggtatatatt tatttcttgt 3960

tatataatcc ttttgtttat tacatgggct ggatacataa aggtattttg atttaatttt 4020

ttgcttaaat tcaatccccc ctcgttcagt gtcaactgta atggtaggaa attaccatac 4080

ttttgaagaa gcaaaaaaaa tgaaagaaaa aaaaaatcgt atttccaggt tagacgttcc 4140

gcagaatcta gaatgcggta tgcggtacat tgttcttcga acgtaaaagt tgcgctccct 4200

gagatattgt acatttttgc ttttacaagt acaagtacat cgtacaacta tgtactactg 4260

ttgatgcatc cacaacagtt tgttttgttt ttttttgttt ttttttttttc taatgattca 4320

ttaccgctat gtatacctac ttgtacttgt agtaagccgg gttattggcg ttcaattaat 4380

catagactta tgaatctgca cggtgtgcgc tgcgagttac ttttagctta tgcatgctac 4440

ttgggtgtaa tattgggatc tgttcggaaa tcaacggatg ctcaaccgat ttcgacagta 4500

ataatttgaa tcgaatcgga gcctaaaatg aacccgagta tatctcataa aattctcggt 4560

gagaggtctg tgactgtcag tacaaggtgc cttcattatg ccctcaacct taccatacct 4620

cactgaatgt agtgtacctc taaaaatgaa atacagtgcc aaaagccaag gcactgagct 4680

cgtctaacgg acttgatata caaccaatta aaacaaatga aaagaaatac agttctttgt 4740

atcatttgta acaattaccc tgtacaaact aaggtattga aatcccacaa tattcccaaa 4800

gtccacccct ttccaaattg tcatgcctac aactcatata ccaagcacta acctaccaaa 4860

caccactaaa accccacaaa atatatctta ccgaatatac agtaacaagc taccaccaca 4920

ctcgttgggt gcagtcgcca gcttaaagat atctatccac atcagccaca actcccttcc 4980

tttaataaac cgactacacc cttggctatt gaggttatga gtgaatatac tgtagacaag 5040

acactttcaa gaagactgtt tccaaaacgt accactgtcc tccactacaa acacacccaa 5100

tctgcttctt ctagtcaagg ttgctacacc ggtaaattat aaatcatcat ttcattagca 5160

gggcagggcc ctttttatag agtcttatac actagcggac cctgccggta gaccaacccg 5220

caggcgcgtc agtttgctcc ttccatcaat gcgtcgtaga aacgacttac tccttcttga 5280

gcagctcctt gaccttgttg gcaacaagtc tccgacctcg gaggtggagg aagagcctcc 5340

gatatcggcg gtagtgatac cagcctcgac ggactccttg acggcagcct caacagcgtc 5400

accggcgggc ttcatgttaa gagagaactt gagcatcatg gcggcagaca gaatggtggc 5460

aatggggttg accttctgct tgccgagatc gggggcagat ccgtgacagg gctcgtacag 5520

accgaacgcc tcgttggtgt cgggcagaga agccagagag gcggagggca gcagacccag 5580

agaaccgggg atgacggagg cctcgtcgga gatgatatcg ccaaacatgt tggtggtgat 5640

gatgatacca ttcatcttgg agggctgctt gatgaggatc atggcggccg agtcgatcag 5700

ctggtggttg agctcgagct gggggaattc gtccttgagg actcgagtga cagtctttcg 5760

ccaaagtcga gaggaggcca gcacgttggc cttgtcaaga gaccacacgg gaagaggggg 5820

gttgtgctga agggccagga aggcggccat tcgggcaatt cgctcaacct caggaacgga 5880

gtaggtctcg gtgtcggaag cgacgccaga tccgtcatcc tcctttcgct ctccaaagta 5940

gatacctccg acgagctctc ggacaatgat gaagtcggtg ccctcaacgt ttcggatggg 6000

ggagagatcg gcgagcttgg gcgacagcag ctggcagggt cgcaggttgg cgtacaggtt 6060

caggtccttt cgcagcttga ggagaccctg ctcgggtcgc acgtcggttc gtccgtcggg 6120

agtggtccat acggtgttgg cagcgcctcc gacagcaccg agcataatag agtcagcctt 6180

tcggcagatg tcgagagtag cgtcggtgat gggctcgccc tccttctcaa tggcagctcc 6240

tccaatgagt cggtcctcaa acacaaactc ggtgccggag gcctcagcaa cagacttgag 6300

caccttgacg gcctcggcaa tcacctcggg gccacagaag tcgccgccga gaagaacaat 6360

cttcttggag tcagtcttgg tcttcttagt ttcgggttcc attgtggatg tgtgtggttg 6420

tatgtgtgat gtggtgtgtg gagtgaaaat ctgtggctgg caaacgctct tgtatatata 6480

cgcacttttg cccgtgctat gtggaagact aaacctccga agattgtgac tcaggtagtg 6540

cggtatcggc tagggaccca aaccttgtcg atgccgatag cgctatcgaa cgtaccccag 6600

ccggccggga gtatgtcgga ggggacatac gagatcgtca agggtttgtg gccaactggt 6660

aaataaatga tgactcaggc gacgacggaa ttcctgcagc ccatcgatgc agaattcagg 6720

agagaccggg ttggcggcgt atttgtgtcc caaaaaacag ccccaattgc cccaattgac 6780

cccaaattga cccagtagcg ggcccaaccc cggcgagagc ccccttcacc ccacatatca 6840

aacctccccc ggttcccaca cttgccgtta agggcgtagg gtactgcagt ctggaatcta 6900

cgcttgttca gactttgtac tagtttcttt gtctggccat ccgggtaacc catgccggac 6960

gcaaaataga ctactgaaaa tttttttgct ttgtggttgg gactttagcc aagggtataa 7020

aagaccaccg tccccgaatt acctttcctc ttcttttctc tctctccttg tcaactcaca 7080

cccgaaatcg ttaagcattt ccttctgagt ataagaatca ttcaccatgg ctgctgctcc 7140

cagtgtgagg acgtttactc gggccgaggt tttgaatgcc gaggctctga atgagggcaa 7200

gaaggatgcc gaggcaccct tcttgatgat catcgacaac aaggtgtacg atgtccgcga 7260

gttcgtccct gatcatcccg gtggaagtgt gattctcacg cacgttggca aggacggcac 7320

tgacgtcttt gacacttttc accccgaggc tgcttgggag actcttgcca acttttacgt 7380

tggtgatatt gacgagagcg accgcgatat caagaatgat gactttgcgg ccgaggtccg 7440

caagctgcgt accttgttcc agtctcttgg ttactacgat tcttccaagg catactacgc 7500

cttcaaggtc tcgttcaacc tctgcatctg gggtttgtcg acggtcattg tggccaagtg 7560

gggccagacc tcgaccctcg ccaacgtgct ctcggctgcg cttttgggtc tgttctggca 7620

gcagtgcgga tggttggctc acgacttttt gcatcaccag gtcttccagg accgtttctg 7680

gggtgatctt ttcggcgcct tcttgggagg tgtctgccag ggcttctcgt cctcgtggtg 7740

gaaggacaag cacaacactc accacgccgc ccccaacgtc cacggcgagg atcccgacat 7800

tgacacccac cctctgttga cctggagtga gcatgcgttg gagatgttct cggatgtccc 7860

agatgaggag ctgacccgca tgtggtcgcg tttcatggtc ctgaaccaga cctggtttta 7920

cttccccatt ctctcgtttg cccgtctctc ctggtgcctc cagtccattc tctttgtgct 7980

gcctaacggt caggcccaca agccctcggg cgcgcgtgtg cccatctcgt tggtcgagca 8040

gctgtcgctt gcgatgcact ggacctggta cctcgccacc atgttcctgt tcatcaagga 8100

tcccgtcaac atgctggtgt acttttttggt gtcgcaggcg gtgtgcggaa acttgttggc 8160

gatcgtgttc tcgctcaacc acaacggtat gcctgtgatc tcgaaggagg aggcggtcga 8220

tatggatttc ttcacgaagc agatcatcac gggtcgtgat gtccacccgg gtctatttgc 8280

caactggttc acgggtggat tgaactatca gatcgagcac cacttgttcc cttcgatgcc 8340

tcgccacaac ttttcaaaga tccagcctgc tgtcgagacc ctgtgcaaaa agtacaatgt 8400

ccgataccac accaccggta tgatcgaggg aactgcagag gtctttagcc gtctgaacga 8460

ggtctccaag gctacctcca agatgggtaa ggcgcagtaa gc 8502

<210> 169
<211> 7145
<212> DNA
<213> Artificial Sequence

<220>
<223> Plasmid pKUNFmkF2

<220>
<221> misc_feature
<222> (692)..(695)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (700)..(700)
<223> n is a, c, g, or t

<400> 169
catggcgtcc acttcggctc tgcccaagca gaaccctgcg cttagacgca ccgtcacctc 60

aactactgtg acggattctg agtctgccgc cgtctctcct tcagactctc cccgccactc 120

ggcctcttcc acatcgctct cgtccatgtc cgaggttgat atcgccaagc ccaagtccga 180

gtatggtgtc atgctcgaca cctacggcaa ccagttcgag gttcccgact ttaccatcaa 240

ggacatctac aatgccatcc ctaagcactg cttcaagcgc tccgctctca agggatacgg 300

ttatatcctc cgcgacattg tcctcctgac taccactttc agcatctggt acaactttgt 360

gacccccgaa tatatcccct ccacccccgc ccgcgctggt ctgtgggccg tgtacaccgt 420

tcttcagggt cttttcggta ctggtctctg ggttattgcc catgagtgcg gtcacggtgc 480

tttctccgat tctcgcatca tcaacgacat tactggctgg gttcttcact cttccctcct 540

tgtcccctac ttcagctggc aaatctccca ccgaaagcac cacaaggcca ctggcaacat 600

ggagcgtgac atggtcttcg ttccccgaac ccgcgagcag caggctactc gtctcggaaa 660

gatgacccac gagctcgctc atcttactga gnnnntcgtn ggctggccca actacctcat 720

caccaatgtt accggccaca actaccacga gcgccagcgt gagggtcgcg gcaagggcaa 780

gcataacggc ctcggcggtg gtgttaacca cttcgatccc cgcagccctc tgtacgagaa 840

cagtgacgct aagctcatcg tcctcagcga tattggtatc ggtctgatgg ccactgctct 900

gtacttcctc gttcagaagt tcggtttcta caacatggcc atctggtact ttgttcccta 960

cctctgggtt aaccactggc tcgttgccat caccttcctc cagcacaccg accctaccct 1020

tccccactac accaacgacg agtggaactt cgtccgtggt gccgctgcta ccattgaccg 1080

tgagatgggc ttcatcggcc gccaccttct ccacggcatc atcgagactc atgtcctcca 1140

ccactacgtc agcagcatcc ccttctacaa cgcggacgag gccaccgagg ccattaagcc 1200

catcatgggc aagcactacc gggctgatgt ccaggatggt cctcgtggct tcatccgcgc 1260

catgtaccgc agtgcgcgta tgtgccagtg ggttgagccc agcgctggtg ccgagggtgc 1320

tggtaagggt gttctgttct tccgcaaccg caacaacgtg ggcacccccc ccgctgttat 1380

caagcccgtt gcttaagtag gcgcggccgc tatttatcac tctttacaac ttctacctca 1440

actatctact ttaataaatg aatatcgttt attctctatg attactgtat atgcgttcct 1500

ctaagacaaa tcgaaaccag catgtgatcg aatggcatac aaaagtttct tccgaagttg 1560

atcaatgtcc tgatagtcag gcagcttgag aagattgaca caggtggagg ccgtagggaa 1620

ccgatcaacc tgtctaccag cgttacgaat ggcaaatgac gggttcaaag ccttgaatcc 1680

ttgcaatggt gccttggata ctgatgtcac aaacttaaga agcagccgct tgtcctcttc 1740

ctcgatcgat ggtcatagct gtttcctgtg tgaaattgtt atccgctcac aattccacac 1800

aacgtacgaa gtcgtcaatg atgtcgatat gggttttgat catgcacaca taaggtccga 1860

ccttatcggc aagctcaatg agctccttgg tggtggtaac atccagagaa gcacacaggt 1920

tggttttctt ggctgccacg agcttgagca ctcgagcggc aaaggcggac ttgtggacgt 1980

tagctcgagc ttcgtaggag ggcattttgg tggtgaagag gagactgaaa taaatttagt 2040

ctgcagaact ttttatcgga accttatctg gggcagtgaa gtatatgtta tggtaatagt 2100

tacgagttag ttgaacttat agatagactg gactatacgg ctatcggtcc aaattagaaa 2160

gaacgtcaat ggctctctgg gcgtcgcctt tgccgacaaa aatgtgatca tgatgaaagc 2220

cagcaatgac gttgcagctg atattgttgt cggccaaccg cgccgaaaac gcagctgtca 2280

gacccacagc ctccaacgaa gaatgtatcg tcaaagtgat ccaagcacac tcatagttgg 2340

agtcgtactc caaaggcggc aatgacgagt cagacagata ctcgtcgacc ttttccttgg 2400

gaaccaccac cgtcagccct tctgactcac gtattgtagc caccgacaca ggcaacagtc 2460

cgtggatagc agaatatgtc ttgtcggtcc atttctcacc aactttaggc gtcaagtgaa 2520

tgttgcagaa gaagtatgtg ccttcattga gaatcggtgt tgctgatttc aataaagtct 2580

tgagatcagt ttggcgcgcc agctgcatta atgaatcggc caacgcgcgg ggagaggcgg 2640

tttgcgtatt gggcgctctt ccgcttcctc gctcactgac tcgctgcgct cggtcgttcg 2700

gctgcggcga gcggtatcag ctcactcaaa ggcggtaata cggttatcca cagaatcagg 2760

ggataacgca ggaaagaaca tgtgagcaaa aggccagcaa aaggccagga accgtaaaaa 2820

ggccgcgttg ctggcgtttt tccataggct ccgcccccct gacgagcatc acaaaaatcg 2880

acgctcaagt cagaggtggc gaaacccgac aggactataa agataccagg cgtttccccc 2940

tggaagctcc ctcgtgcgct ctcctgttcc gaccctgccg cttaccggat acctgtccgc 3000

ctttctccct tcgggaagcg tggcgctttc tcatagctca cgctgtaggt atctcagttc 3060

ggtgtaggtc gttcgctcca agctgggctg tgtgcacgaa ccccccgttc agcccgaccg 3120

ctgcgcctta tccggtaact atcgtcttga gtccaacccg gtaagacacg acttatcgcc 3180

actggcagca gccactggta acaggattag cagagcgagg tatgtaggcg gtgctacaga 3240

gttcttgaag tggtggccta actacggcta cactagaaga acagtatttg gtatctgcgc 3300

tctgctgaag ccagttacct tcggaaaaag agttggtagc tcttgatccg gcaaacaaac 3360

caccgctggt agcggtggtt tttttgtttg caagcagcag attacgcgca gaaaaaaagg 3420

atctcaagaa gatcctttga tcttttctac ggggtctgac gctcagtgga acgaaaactc 3480

acgttaaggg attttggtca tgagattatc aaaaaggatc ttcacctaga tccttttaaa 3540

ttaaaaatga agttttaaat caatctaaag tatatatgag taaacttggt ctgacagtta 3600

ccaatgctta atcagtgagg cacctatctc agcgatctgt ctatttcgtt catccatagt 3660

tgcctgactc cccgtcgtgt agataactac gatacgggag ggcttaccat ctggccccag 3720

tgctgcaatg ataccgcgag acccacgctc accggctcca gatttatcag caataaacca 3780

gccagccgga agggccgagc gcagaagtgg tcctgcaact ttatccgcct ccatccagtc 3840

tattaattgt tgccgggaag ctagagtaag tagttcgcca gttaatagtt tgcgcaacgt 3900

tgttgccatt gctacaggca tcgtggtgtc acgctcgtcg tttggtatgg cttcattcag 3960

ctccggttcc caacgatcaa ggcgagttac atgatccccc atgttgtgca aaaaagcggt 4020

tagctccttc ggtcctccga tcgttgtcag aagtaagttg gccgcagtgt tatcactcat 4080

ggttatggca gcactgcata attctcttac tgtcatgcca tccgtaagat gcttttctgt 4140

gactggtgag tactcaacca agtcattctg agaatagtgt atgcggcgac cgagttgctc 4200

ttgcccggcg tcaatacggg ataataccgc gccacatagc agaactttaa aagtgctcat 4260

cattggaaaa cgttcttcgg ggcgaaaact ctcaaggatc ttaccgctgt tgagatccag 4320

ttcgatgtaa cccactcgtg cacccaactg atcttcagca tcttttactt tcaccagcgt 4380

ttctgggtga gcaaaaacag gaaggcaaaa tgccgcaaaa aagggaataa gggcgacacg 4440

gaaatgttga atactcatac tcttcctttt tcaatattat tgaagcattt atcagggtta 4500

ttgtctcatg agcggataca tatttgaatg tatttagaaa aataaacaaa taggggttcc 4560

gcgcacattt ccccgaaaag tgccacctga tgcggtgtga aataccgcac agatgcgtaa 4620

ggagaaaata ccgcatcagg aaattgtaag cgttaatatt ttgttaaaat tcgcgttaaa 4680

tttttgttaa atcagctcat tttttaacca ataggccgaa atcggcaaaa tcccttataa 4740

atcaaaagaa tagaccgaga tagggttgag tgttgttcca gtttggaaca agagtccact 4800

attaaagaac gtggactcca acgtcaaagg gcgaaaaacc gtctatcagg gcgatggccc 4860

actacgtgaa ccatcaccct aatcaagttt tttggggtcg aggtgccgta aagcactaaa 4920

tcggaaccct aaagggagcc cccgatttag agcttgacgg ggaaagccgg cgaacgtggc 4980

gagaaaggaa gggaagaaag cgaaaggagc gggcgctagg gcgctggcaa gtgtagcggt 5040

cacgctgcgc gtaaccacca cacccgccgc gcttaatgcg ccgctacagg gcgcgtccat 5100

tcgccattca ggctgcgcaa ctgttgggaa gggcgatcgg tgcgggcctc ttcgctatta 5160

cgccagctgg cgaaaggggg atgtgctgca aggcgattaa gttgggtaac gccagggttt 5220

tcccagtcac gacgttgtaa aacgacggcc agtgaattgt aatacgactc actatagggc 5280

gaattgggcc cgacgtcgca tgcagtggtg gtattgtgac tggggatgta gttgagaata 5340

agtcatacac aagtcagctt tcttcgagcc tcatataagt ataagtagtt caacgtatta 5400

gcactgtacc cagcatctcc gtatcgagaa acacaacaac atgccccatt ggacagatca 5460

tgcggataca caggttgtgc agtatcatac atactcgatc agacaggtcg tctgaccatc 5520

atacaagctg aacaagcgct ccatacttgc acgctctcta tatacacagt taaattacat 5580

atccatagtc taacctctaa cagttaatct tctggtaagc ctcccagcca gccttctggt 5640

atcgcttggc ctcctcaata ggatctcggt tctggccgta cagacctcgg ccgacaatta 5700

tgatatccgt tccggtagac atgacatcct caacagttcg gtactgctgt ccgagagcgt 5760

ctcccttgtc gtcaagaccc accccggggg tcagaataag ccagtcctca gagtcgccct 5820

taattaattt gaatcgaatc gatgagccta aaatgaaccc gagtatatct cataaaattc 5880

tcggtgagag gtctgtgact gtcagtacaa ggtgccttca ttatgccctc aaccttacca 5940

tacctcactg aatgtagtgt acctctaaaa atgaaataca gtgccaaaag ccaaggcact 6000

gagctcgtct aacggacttg atatacaacc aattaaaaca aatgaaaaga aatacagttc 6060

tttgtatcat ttgtaacaat taccctgtac aaactaaggt attgaaatcc cacaatattc 6120

ccaaagtcca ccccttccca aattgtcatg cctacaactc atataccaag cactaaccta 6180

ccgtttaaac agtgtacgca gatctactat agaggaacat ttaaattgcc ccggagaaga 6240

cggccaggcc gcctagatga caaattcaac aactcacagc tgactttctg ccattgccac 6300

taggggggggg ccttttttata tggccaagcc aagctctcca cgtcggttgg gctgcaccca 6360

acaataaatg ggtagggttg caccaacaaa gggatgggat ggggggtaga agatacgagg 6420

ataacggggc tcaatggcac aaataagaac gaatactgcc attaagactc gtgatccagc 6480

gactgacacc attgcatcat ctaagggcct caaaactacc tcggaactgc tgcgctgatc 6540

tggacaccac agaggttccg agcactttag gttgcaccaa atgtcccacc aggtgcaggc 6600

agaaaacgct ggaacagcgt gtacagtttg tcttaacaaa aagtgagggc gctgaggtcg 6660

agcagggtgg tgtgacttgt tatagccttt agagctgcga aagcgcgtat ggatttggct 6720

catcaggcca gattgagggt ctgtggacac atgtcatgtt agtgtacttc aatcgccccc 6780

tggatatagc cccgacaata ggccgtggcc tcatttttttt gccttccgca catttccatt 6840

gctcgatacc cacaccttgc ttctcctgca cttgccaacc ttaatactgg tttacattga 6900

ccaacatctt acaagcgggg ggcttgtcta gggtatatat aaacagtggc tctcccaatc 6960

ggttgccagt ctctttttttc ctttctttcc ccacagattc gaaatctaaa ctacacatca 7020

cagaattccg agccgtgagt atccacgaca agatcagtgt cgagacgacg cgttttgtgt 7080

aatgacacaa tccgaaagtc gctagcaaca cacactctct acacaaacta acccagctct 7140

ggtac 7145

<210> 170
<211> 5553
<212> DNA
<213> Artificial Sequence

<220>
<223> Plasmid pZF5T-PPC

<400> 170
ggccgcattg atgattggaa acacacacat gggttatatc taggtgagag ttagttggac 60

agttatatat taaatcagct atgccaacgg taacttcatt catgtcaacg aggaaccagt 120

gactgcaagt aatatagaat ttgaccacct tgccattctc ttgcactcct ttactatatc 180

tcatttattt cttatataca aatcacttct tcttcccagc atcgagctcg gaaacctcat 240

gagcaataac atcgtggatc tcgtcaatag agggcttttt ggactccttg ctgttggcca 300

ccttgtcctt gctgtctggc tcattctgtt tcaacgcctt ttaattaatc gagcttggcg 360

taatcatggt catagctgtt tcctgtgtga aattgttatc cgctcacaat tccacacaac 420

atacgagccg gaagcataaa gtgtaaagcc tggggtgcct aatgagtgag ctaactcaca 480

ttaattgcgt tgcgctcact gcccgctttc cagtcgggaa acctgtcgtg ccagctgcat 540

taatgaatcg gccaacgcgc ggggagaggc ggtttgcgta ttgggcgctc ttccgcttcc 600

tcgctcactg actcgctgcg ctcggtcgtt cggctgcggc gagcggtatc agctcactca 660

aaggcggtaa tacggttatc cacagaatca ggggataacg caggaaagaa catgtgagca 720

aaaggccagc aaaaggccag gaaccgtaaa aaggccgcgt tgctggcgtt tttccatagg 780

ctccgccccc ctgacgagca tcacaaaaat cgacgctcaa gtcagaggtg gcgaaacccg 840

acaggactat aaagatacca ggcgtttccc cctggaagct ccctcgtgcg ctctcctgtt 900

ccgaccctgc cgcttaccgg atacctgtcc gcctttctcc cttcgggaag cgtggcgctt 960

tctcatagct cacgctgtag gtatctcagt tcggtgtagg tcgttcgctc caagctgggc 1020

tgtgtgcacg aacccccccgt tcagcccgac cgctgcgcct tatccggtaa ctatcgtctt 1080

gagtccaacc cggtaagaca cgacttatcg ccactggcag cagccactgg taacaggatt 1140

agcagagcga ggtatgtagg cggtgctaca gagttcttga agtggtggcc taactacggc 1200

tacactagaa ggacagtatt tggtatctgc gctctgctga agccagttac cttcggaaaa 1260

agagttggta gctcttgatc cggcaaacaa accaccgctg gtagcggtgg ttttttttgtt 1320

tgcaagcagc agattacgcg cagaaaaaaa ggatctcaag aagatccttt gatcttttct 1380

acggggtctg acgctcagtg gaacgaaaac tcacgttaag ggattttggt catgagatta 1440

tcaaaaagga tcttcaccta gatcctttta aattaaaaat gaagttttaa atcaatctaa 1500

agtatatatg agtaaacttg gtctgacagt taccaatgct taatcagtga ggcacctatc 1560

tcagcgatct gtctatttcg ttcatccata gttgcctgac tccccgtcgt gtagataact 1620

acgatacggg agggcttacc atctggcccc agtgctgcaa tgataccgcg agacccacgc 1680

tcaccggctc cagatttatc agcaataaac cagccagccg gaagggccga gcgcagaagt 1740

ggtcctgcaa ctttatccgc ctccatccag tctattaatt gttgccggga agctagagta 1800

agtagttcgc cagttaatag tttgcgcaac gttgttgcca ttgctacagg catcgtggtg 1860

tcacgctcgt cgtttggtat ggcttcattc agctccggtt cccaacgatc aaggcgagtt 1920

acatgatccc ccatgttgtg caaaaaagcg gttagctcct tcggtcctcc gatcgttgtc 1980

agaagtaagt tggccgcagt gttatcactc atggttatgg cagcactgca taattctctt 2040

actgtcatgc catccgtaag atgcttttct gtgactggtg agtactcaac caagtcattc 2100

tgagaatagt gtatgcggcg accgagttgc tcttgcccgg cgtcaatacg ggataatacc 2160

gcgccacata gcagaacttt aaaagtgctc atcattggaa aacgttcttc ggggcgaaaa 2220

ctctcaagga tcttaccgct gttgagatcc agttcgatgt aacccactcg tgcacccaac 2280

tgatcttcag catcttttac tttcaccagc gtttctgggt gagcaaaaac aggaaggcaa 2340

aatgccgcaa aaaagggaat aagggcgaca cggaaatgtt gaatactcat actcttcctt 2400

tttcaatatt attgaagcat ttatcagggt tattgtctca tgagcggata catatttgaa 2460

tgtatttaga aaaataaaca aataggggtt ccgcgcacat ttccccgaaa agtgccacct 2520

gacgcgccct gtagcggcgc attaagcgcg gcgggtgtgg tggttacgcg cagcgtgacc 2580

gctacacttg ccagcgccct agcgcccgct cctttcgctt tcttcccttc ctttctcgcc 2640

acgttcgccg gctttccccg tcaagctcta aatcgggggc tccctttagg gttccgattt 2700

agtgctttac ggcacctcga ccccaaaaaa cttgattagg gtgatggttc acgtagtggg 2760

ccatcgccct gatagacggt ttttcgccct ttgacgttgg agtccacgtt ctttaatagt 2820

ggactcttgt tccaaactgg aacaacactc aaccctatct cggtctattc ttttgattta 2880

taagggattt tgccgatttc ggcctattgg ttaaaaaatg agctgattta acaaaaattt 2940

aacgcgaatt ttaacaaaat attaacgctt acaatttcca ttcgccattc aggctgcgca 3000

actgttggga agggcgatcg gtgcgggcct cttcgctatt acgccagctg gcgaaagggg 3060

gatgtgctgc aaggcgatta agttgggtaa cgccagggtt ttcccagtca cgacgttgta 3120

aaacgacggc cagtgaattg taatacgact cactataggg cgaattgggt accgggcccc 3180

ccctcgaggt cgacgtttaa acagtgtacg cagtactata gaggaacatc gattgccccg 3240

gagaagacgg ccaggccgcc tagatgacaa attcaacaac tcacagctga ctttctgcca 3300

ttgccactag gggggggcct ttttatatgg ccaagccaag ctctccacgt cggttgggct 3360

gcacccaaca ataaatgggt agggttgcac caacaaaggg atgggatggg gggtagaaga 3420

tacgaggata acggggctca atggcacaaa taagaacgaa tactgccatt aagactcgtg 3480

atccagcgac tgacaccatt gcatcatcta agggcctcaa aactacctcg gaactgctgc 3540

gctgatctgg acaccacaga ggttccgagc actttaggtt gcaccaaatg tcccaccagg 3600

tgcaggcaga aaacgctgga acagcgtgta cagtttgtct taacaaaaag tgagggcgct 3660

gaggtcgagc agggtggtgt gacttgttat agcctttaga gctgcgaaag cgcgtatgga 3720

tttggctcat caggccagat tgagggtctg tggacacatg tcatgttagt gtacttcaat 3780

cgccccctgg atatagcccc gacaataggc cgtggcctca ttttttttgcc ttccgcacat 3840

ttccattgct cggtacccac accttgcttc tcctgcactt gccaacctta atactggttt 3900

acattgacca acatcttaca agcgggggggc ttgtctaggg tatatataaa cagtggctct 3960

cccaatcggt tgccagtctc ttttttcctt tctttcccca cagattcgaa atctaaacta 4020

cacatcacac aatgcctgtt actgacgtcc ttaagcgaaa gtccggtgtc atcgtcggcg 4080

acgatgtccg agccgtgagt atccacgaca agatcagtgt cgagacgacg cgttttgtgt 4140

aatgacacaa tccgaaagtc gctagcaaca cacactctct acacaaacta acccagctct 4200

ccatgggaac ggaccaagga aaaaccttca cctgggaaga gctggcggcc cataacacca 4260

aggacgacct actcttggcc atccgcggca gggtgtacga tgtcacaaag ttcttgagcc 4320

gccatcctgg tggagtggac actctcctgc tcggagctgg ccgagatgtt actccggtct 4380

ttgagatgta tcacgcgttt ggggctgcag atgccattat gaagaagtac tatgtcggta 4440

cactggtctc gaatgagctg cccatcttcc cggagccaac ggtgttccac aaaaccatca 4500

agacgagagt cgagggctac tttacggatc ggaacattga tcccaagaat agaccagaga 4560

tctggggacg atacgctctt atctttggat ccttgatcgc ttcctactac gcgcagctct 4620

ttgtgccttt cgttgtcgaa cgcacatggc ttcaggtggt gtttgcaatc atcatgggat 4680

ttgcgtgcgc acaagtcgga ctcaaccctc ttcatgatgc gtctcacttt tcagtgaccc 4740

acaaccccac tgtctggaag attctgggag ccacgcacga ctttttcaac ggagcatcgt 4800

acctggtgtg gatgtaccaa catatgctcg gccatcaccc ctacaccaac attgctggag 4860

cagatcccga cgtgtcgacg tctgagcccg atgttcgtcg tatcaagccc aaccaaaagt 4920

ggtttgtcaa ccacatcaac cagcacatgt ttgttccttt cctgtacgga ctgctggcgt 4980

tcaaggtgcg cattcaggac atcaacattt tgtactttgt caagaccaat gacgctattc 5040

gtgtcaatcc catctcgaca tggcacactg tgatgttctg gggcggcaag gctttctttg 5100

tctggtatcg cctgattgtt cccctgcagt atctgcccct gggcaaggtg ctgctcttgt 5160

tcacggtcgc ggacatggtg tcgtcttact ggctggcgct gaccttccag gcgaaccacg 5220

ttgttgagga agttcagtgg ccgttgcctg acgagaacgg gatcatccaa aaggactggg 5280

cagctatgca ggtcgagact acgcaggatt acgcacacga ttcgcacctc tggaccagca 5340

tcactggcag cttgaactac caggctgtgc accatctgtt ccccaacgtg tcgcagcacc 5400

attatcccga tattctggcc atcatcaaga acacctgcag cgagtacaag gttccatacc 5460

ttgtcaagga tacgttttgg caagcatttg cttcacattt ggagcacttg cgtgttcttg 5520

gactccgtcc caaggaagag taggcagcta agc 5553

<210> 171
<211> 10448
<212> DNA
<213> Artificial Sequence

<220>
<223> Plasmid pDMW297

<400> 171
cgattgcccc ggagaagacg gccaggccgc ctagatgaca aattcaacaa ctcacagctg 60

actttctgcc attgccacta gggggggccc tttttatatg gccaagccaa gctctccacg 120

tcggttgggc tgcacccaac aataaatggg tagggttgca ccaacaaagg gatgggatgg 180

ggggtagaag atacgaggat aacggggctc aatggcacaa ataagaacga atactgccat 240

taagactcgt gatccagcga ctgacaccat tgcatcatct aagggcctca aaactacctc 300

ggaactgctg cgctgatctg gacaccacag aggttccgag cactttaggt tgcaccaaat 360

gtcccaccag gtgcaggcag aaaacgctgg aacagcgtgt acagtttgtc ttaacaaaaa 420

gtgagggcgc tgaggtcgag cagggtggtg tgacttgtta tagcctttag agctgcgaaa 480

gcgcgtatgg atttggctca tcaggccaga ttgagggtct gtggacacat gtcatgttag 540

tgtacttcaa tcgccccctg gatatagccc cgacaatagg ccgtggcctc atttttttgc 600

cttccgcaca tttccattgc tcggtaccca caccttgctt ctcctgcact tgccaacctt 660

aatactggtt tacattgacc aacatcttac aagcgggggg cttgtctagg gtatatataa 720

acagtggctc tcccaatcgg ttgccagtct ctttttttcct ttctttcccc acagattcga 780

aatctaaact acacatcaca caatgcctgt tactgacgtc cttaagcgaa agtccggtgt 840

catcgtcggc gacgatgtcc gagccgtgag tatccacgac aagatcagtg tcgagacgac 900

gcgttttgtg taatgacaca atccgaaagt cgctagcaac acacactctc tacacaaact 960

aacccagctc tccatggtga agtccaagcg acaggctctg cccctcacca tcgacggaac 1020

tacctacgac gtctccgctt gggtgaactt ccaccctggt ggagctgaaa tcattgagaa 1080

ctaccaggga cgagatgcta ctgacgcctt catggttatg cactctcagg aagccttcga 1140

caagctcaag cgaatgccca agatcaaccc ctcctccgag ctgcctcccc aggctgccgt 1200

caacgaagct caggaggatt tccgaaagct ccgagaagag ctgatcgcca ctggcatgtt 1260

tgacgcctct cccctctggt actcgtacaa gatctccacc accctgggtc ttggcgtgct 1320

tggatacttc ctgatggtcc agtaccagat gtacttcatt ggtgctgtgc tgctcggtat 1380

gcactaccag caaatgggat ggctgtctca tgacatctgc caccaccaga ccttcaagaa 1440

ccgaaactgg aataacctcg tgggtctggt ctttggcaac ggactccagg gcttctccgt 1500

gacctggtgg aaggacagac acaacgccca tcattctgct accaacgttc agggtcacga 1560

tcccgacatt gataacctgc ctctgctcgc ctggtccgag gacgatgtca ctcgagcttc 1620

tcccatctcc cgaaagctca ttcagttcca acagtactat ttcctggtca tctgtattct 1680

cctgcgattc atctggtgtt tccagtctgt gctgaccgtt cgatccctca aggaccgaga 1740

caaccagttc taccgatctc agtacaagaa agaggccatt ggactcgctc tgcactggac 1800

tctcaagacc ctgttccacc tcttctttat gccctccatc ctgacctcgc tcctggtgtt 1860

ctttgtttcc gagctcgtcg gtggcttcgg aattgccatc gtggtcttca tgaaccacta 1920

ccctctggag aagatcggtg attccgtctg ggacggacat ggcttctctg tgggtcagat 1980

ccatgagacc atgaacattc gacgaggcat cattactgac tggttctttg gaggcctgaa 2040

ctaccagatc gagcaccatc tctggcccac cctgcctcga cacaacctca ctgccgtttc 2100

ctaccaggtg gaacagctgt gccagaagca caacctcccc taccgaaacc ctctgcccca 2160

tgaaggtctc gtcatcctgc tccgatacct ggccgtgttc gctcgaatgg ccgagaagca 2220

gcccgctggc aaggctctct aagcggccgc attgatgatt ggaaacacac acatgggtta 2280

tatctaggtg agagttagtt ggacagttat atattaaatc agctatgcca acggtaactt 2340

cattcatgtc aacgaggaac cagtgactgc aagtaatata gaatttgacc accttgccat 2400

tctcttgcac tcctttacta tatctcattt atttcttata tacaaatcac ttcttcttcc 2460

cagcatcgag ctcggaaacc tcatgagcaa taacatcgtg gatctcgtca atagagggct 2520

ttttggactc cttgctgttg gccaccttgt ccttgctgtc tggctcattc tgtttcaacg 2580

ccttttaatt aagtcataca caagtcagct ttcttcgagc ctcatataag tataagtagt 2640

tcaacgtatt agcactgtac ccagcatctc cgtatcgaga aacacaacaa catgccccat 2700

tggacagatc atgcggatac acaggttgtg cagtatcata catactcgat cagacaggtc 2760

gtctgaccat catacaagct gaacaagcgc tccatacttg cacgctctct atatacacag 2820

ttaaattaca tatccatagt ctaacctcta acagttaatc ttctggtaag cctcccagcc 2880

agccttctgg tatcgcttgg cctcctcaat aggatctcgg ttctggccgt acagacctcg 2940

gccgacaatt atgatatccg ttccggtaga catgacatcc tcaacagttc ggtactgctg 3000

tccgagagcg tctcccttgt cgtcaagacc cacccgggg gtcagaataa gccagtcctc 3060

agagtcgccc ttaggtcggt tctgggcaat gaagccaacc acaaactcgg ggtcggatcg 3120

ggcaagctca atggtctgct tggagtactc gccagtggcc agagagccct tgcaagacag 3180

ctcggccagc atgagcagac ctctggccag cttctcgttg ggagagggga ctaggaactc 3240

cttgtactgg gagttctcgt agtcagagac gtcctccttc ttctgttcag agacagtttc 3300

ctcggcacca gctcgcaggc cagcaatgat tccggttccg ggtacaccgt gggcgttggt 3360

gatatcggac cactcggcga ttcggtgaca ccggtactgg tgcttgacag tgttgccaat 3420

atctgcgaac tttctgtcct cgaacaggaa gaaaccgtgc ttaagagcaa gttccttgag 3480

ggggagcaca gtgccggcgt aggtgaagtc gtcaatgatg tcgatatggg ttttgatcat 3540

gcacacataa ggtccgacct tatcggcaag ctcaatgagc tccttggtgg tggtaacatc 3600

cagagaagca cacaggttgg ttttcttggc tgccacgagc ttgagcactc gagcggcaaa 3660

ggcggacttg tggacgttag ctcgagcttc gtaggagggc attttggtgg tgaagaggag 3720

actgaaataa atttagtctg cagaactttt tatcggaacc ttatctgggg cagtgaagta 3780

tatgttatgg taatagttac gagttagttg aacttataga tagactggac tatacggcta 3840

tcggtccaaa ttagaaagaa cgtcaatggc tctctgggcg tcgcctttgc cgacaaaaat 3900

gtgatcatga tgaaagccag caatgacgtt gcagctgata ttgttgtcgg ccaaccgcgc 3960

cgaaaacgca gctgtcagac ccacagcctc caacgaagaa tgtatcgtca aagtgatcca 4020

agcacactca tagttggagt cgtactccaa aggcggcaat gacgagtcag acagatactc 4080

gtcgactcag gcgacgacgg aattcctgca gcccatctgc agaattcagg agagaccggg 4140

ttggcggcgt atttgtgtcc caaaaaacag ccccaattgc cccggagaag acggccaggc 4200

cgcctagatg acaaattcaa caactcacag ctgactttct gccattgcca ctaggggggg 4260

gccttttttat atggccaagc caagctctcc acgtcggttg ggctgcaccc aacaataaat 4320

gggtagggtt gcaccaacaa agggatggga tgggggggtag aagatacgag gataacgggg 4380

ctcaatggca caaataagaa cgaatactgc cattaagact cgtgatccag cgactgacac 4440

cattgcatca tctaagggcc tcaaaactac ctcggaactg ctgcgctgat ctggacacca 4500

cagaggttcc gagcacttta ggttgcacca aatgtcccac caggtgcagg cagaaaacgc 4560

tggaacagcg tgtacagttt gtcttaacaa aaagtgaggg cgctgaggtc gagcagggtg 4620

gtgtgacttg ttatagcctt tagagctgcg aaagcgcgta tggatttggc tcatcaggcc 4680

agattgaggg tctgtggaca catgtcatgt tagtgtactt caatcgcccc ctggatatag 4740

ccccgacaat aggccgtggc ctcatttttt tgccttccgc acatttccat tgctcggtac 4800

ccacaccttg cttctcctgc acttgccaac cttaatactg gtttacattg accaacatct 4860

tacaagcggg gggcttgtct agggtatata taaacagtgg ctctcccaat cggttgccag 4920

tctcttttt cctttctttc cccacagatt cgaaatctaa actacacatc acacaatgcc 4980

tgttactgac gtccttaagc gaaagtccgg tgtcatcgtc ggcgacgatg tccgagccgt 5040

gagtatccac gacaagatca gtgtcgagac gacgcgtttt gtgtaatgac acaatccgaa 5100

agtcgctagc aacacacact ctctacacaa actaacccag ctctccatgg ctctggccaa 5160

cgacgctggc gagcgaatct gggctgccgt caccgatccc gaaatcctca ttggcacctt 5220

ctcctacctg ctcctgaagc ctctcctgcg aaactctggt ctcgtggacg agaagaaagg 5280

agcctaccga acctccatga tctggtacaa cgtcctcctg gctctcttct ctgccctgtc 5340

cttctacgtg actgccaccg ctctcggctg ggactacggt actggagcct ggctgcgaag 5400

acagaccggt gatactcccc agcctctctt tcagtgtccc tctcctgtct gggactccaa 5460

gctgttcacc tggactgcca aggccttcta ctattctaag tacgtggagt acctcgacac 5520

cgcttggctg gtcctcaagg gcaagcgagt gtcctttctg caggccttcc atcactttgg 5580

agctccctgg gacgtctacc tcggcattcg actgcacaac gagggtgtgt ggatcttcat 5640

gttctttaac tcgttcattc acaccatcat gtacacctac tatggactga ctgccgctgg 5700

ctacaagttc aaggccaagc ctctgatcac tgccatgcag atttgccagt tcgtcggtgg 5760

ctttctcctg gtctgggact acatcaacgt tccctgcttc aactctgaca agggcaagct 5820

gttctcctgg gctttcaact acgcctacgt cggatctgtc tttctcctgt tctgtcactt 5880

cttttaccag gacaacctgg ccaccaagaa atccgctaag gctggtaagc agctttagcg 5940

gccgcaagtg tggatgggga agtgagtgcc cggttctgtg tgcacaattg gcaatccaag 6000

atggatggat tcaacacagg gatatagcga gctacgtggt ggtgcgagga tatagcaacg 6060

gatatttatg tttgacactt gagaatgtac gatacaagca ctgtccaagt acaatactaa 6120

acatactgta catactcata ctcgtacccg ggcaacggtt tcacttgagt gcagtggcta 6180

gtgctcttac tcgtacagtg tgcaatactg cgtatcatag tctttgatgt atatcgtatt 6240

cattcatgtt agttgcgtac gagccggaag cataaagtgt aaagcctggg gtgcctaatg 6300

agtgagctaa ctcacattaa ttgcgttgcg ctcactgccc gctttccagt cgggaaacct 6360

gtcgtgccag ctgcattaat gaatcggcca acgcgcgggg agaggcggtt tgcgtattgg 6420

gcgctcttcc gcttcctcgc tcactgactc gctgcgctcg gtcgttcggc tgcggcgagc 6480

ggtatcagct cactcaaagg cggtaatacg gttatccaca gaatcagggg ataacgcagg 6540

aaagaacatg tgagcaaaag gccagcaaaa ggccaggaac cgtaaaaagg ccgcgttgct 6600

ggcgtttttc cataggctcc gccccccctga cgagcatcac aaaaatcgac gctcaagtca 6660

gaggtggcga aacccgacag gactataaag ataccaggcg tttccccctg gaagctccct 6720

cgtgcgctct cctgttccga ccctgccgct taccggatac ctgtccgcct ttctcccttc 6780

gggaagcgtg gcgctttctc atagctcacg ctgtaggtat ctcagttcgg tgtaggtcgt 6840

tcgctccaag ctgggctgtg tgcacgaacc ccccgttcag cccgaccgct gcgccttatc 6900

cggtaactat cgtcttgagt ccaacccggt aagacacgac ttatcgccac tggcagcagc 6960

cactggtaac aggattagca gagcgaggta tgtaggcggt gctacagagt tcttgaagtg 7020

gtggcctaac tacggctaca ctagaaggac agtatttggt atctgcgctc tgctgaagcc 7080

agttaccttc ggaaaaagag ttggtagctc ttgatccggc aaacaaacca ccgctggtag 7140

cggtggtttt tttgtttgca agcagcagat tacgcgcaga aaaaaaggat ctcaagaaga 7200

tcctttgatc ttttctacgg ggtctgacgc tcagtggaac gaaaactcac gttaagggat 7260

tttggtcatg agattatcaa aaaggatctt cacctagatc cttttaaatt aaaaatgaag 7320

ttttaaatca atctaaagta tatatgagta aacttggtct gacagttacc aatgcttaat 7380

cagtgaggca cctatctcag cgatctgtct atttcgttca tccatagttg cctgactccc 7440

cgtcgtgtag ataactacga tacgggaggg cttaccatct ggccccagtg ctgcaatgat 7500

accgcgagac ccacgctcac cggctccaga tttatcagca ataaaccagc cagccggaag 7560

ggccgagcgc agaagtggtc ctgcaacttt atccgcctcc atccagtcta ttaattgttg 7620

ccgggaagct agagtaagta gttcgccagt taatagtttg cgcaacgttg ttgccattgc 7680

tacaggcatc gtggtgtcac gctcgtcgtt tggtatggct tcattcagct ccggttccca 7740

acgatcaagg cgagttacat gatcccccat gttgtgcaaa aaagcggtta gctccttcgg 7800

tcctccgatc gttgtcagaa gtaagttggc cgcagtgtta tcactcatgg ttatggcagc 7860

actgcataat tctcttactg tcatgccatc cgtaagatgc ttttctgtga ctggtgagta 7920

ctcaaccaag tcattctgag aatagtgtat gcggcgaccg agttgctctt gcccggcgtc 7980

aatacgggat aataccgcgc cacatagcag aactttaaaa gtgctcatca ttggaaaacg 8040

ttcttcgggg cgaaaactct caaggatctt accgctgttg agatccagtt cgatgtaacc 8100

cactcgtgca cccaactgat cttcagcatc ttttactttc accagcgttt ctgggtgagc 8160

aaaaacagga aggcaaaatg ccgcaaaaaa gggaataagg gcgacacgga aatgttgaat 8220

actcatactc ttccttttc aatattattg aagcatttat cagggttatt gtctcatgag 8280

cggatacata tttgaatgta tttagaaaaa taaacaaata ggggttccgc gcacatttcc 8340

ccgaaaagtg ccacctgacg cgccctgtag cggcgcatta agcgcggcgg gtgtggtggt 8400

tacgcgcagc gtgaccgcta cacttgccag cgccctagcg cccgctcctt tcgctttctt 8460

cccttccttt ctcgccacgt tcgccggctt tccccgtcaa gctctaaatc gggggctccc 8520

tttagggttc cgatttagtg ctttacggca cctcgacccc aaaaaacttg attagggtga 8580

tggttcacgt agtgggccat cgccctgata gacggttttt cgccctttga cgttggagtc 8640

cacgttcttt aatagtggac tcttgttcca aactggaaca acactcaacc ctatctcggt 8700

ctattctttt gatttataag ggattttgcc gatttcggcc tattggttaa aaaatgagct 8760

gatttaacaa aaatttaacg cgaattttaa caaaatatta acgcttacaa tttccattcg 8820

ccattcaggc tgcgcaactg ttgggaaggg cgatcggtgc gggcctcttc gctattacgc 8880

cagctggcga aaggggggatg tgctgcaagg cgattaagtt gggtaacgcc agggttttcc 8940

cagtcacgac gttgtaaaac gacggccagt gaattgtaat acgactcact atagggcgaa 9000

ttgggtaccg ggcccccccct cgaggtcgat ggtgtcgata agcttgatat cgaattcatg 9060

tcacacaaac cgatcttcgc ctcaaggaaa cctaattcta catccgagag actgccgaga 9120

tccagtctac actgattaat tttcgggcca ataatttaaa aaaatcgtgt tatataatat 9180

tatatgtatt atatatatac atcatgatga tactgacagt catgtcccat tgctaaatag 9240

acagactcca tctgccgcct ccaactgatg ttctcaatat ttaaggggtc atctcgcatt 9300

gtttaataat aaacagactc catctaccgc ctccaaatga tgttctcaaa atatattgta 9360

tgaacttatt tttattactt agtattatta gacaacttac ttgctttatg aaaaacactt 9420

cctatttagg aaacaattta taatggcagt tcgttcattt aacaatttat gtagaataaa 9480

tgttataaat gcgtatggga aatcttaaat atggatagca taaatgatat ctgcattgcc 9540

taattcgaaa tcaacagcaa cgaaaaaaat cccttgtaca acataaatag tcatcgagaa 9600

atatcaacta tcaaagaaca gctattcaca cgttactatt gagattatta ttggacgaga 9660

atcacacact caactgtctt tctctcttct agaaatacag gtacaagtat gtactattct 9720

cattgttcat acttctagtc atttcatccc acatattcct tggatttctc tccaatgaat 9780

gacattctat cttgcaaatt caacaattat aataagatat accaaagtag cggtatagtg 9840

gcaatcaaaa agcttctctg gtgtgcttct cgtatttatt tttattctaa tgatccatta 9900

aaggtatata tttatttctt gttatataat ccttttgttt attacatggg ctggatacat 9960

aaaggtattt tgatttaatt ttttgcttaa attcaatccc ccctcgttca gtgtcaactg 10020

taatggtagg aaattaccat acttttgaag aagcaaaaaa aatgaaagaa aaaaaaaatc 10080

gtatttccag gttagacgtt ccgcagaatc tagaatgcgg tatgcggtac attgttcttc 10140

gaacgtaaaa gttgcgctcc ctgagatatt gtacattttt gcttttacaa gtacaagtac 10200

atcgtacaac tatgtactac tgttgatgca tccacaacag tttgttttgt ttttttttgt 10260

ttttttttt tctaatgatt cattaccgct atgtatacct acttgtactt gtagtaagcc 10320

gggttattgg cgttcaatta atcatagact tatgaatctg cacggtgtgc gctgcgagtt 10380

acttttagct tatgcatgct acttgggtgt aatattggga tctgttcgga aatcaacgga 10440

tgctcaat 10448

<210> 172
<211> 15543
<212> DNA

<213> Artificial Sequence

<220>
<223> Plasmid pZP2C16M899

<400> 172
catggttagc gtgtcgtgtt tttgttgtgc tggaagaacc aaagggtggc gcaatgtgtg 60

tagatatata tgtcgtgacc cacaagtcac acaaacaagt atcgggagga gtggtgcacc 120

tctatgcgga gaaaccttat accgctgtag accaactggg gcagaggtgt gagttgaagt 180

cagctggagg agatgtgtga cagaagcaca agaagtgaga ttgtgagatg tatgtctagg 240

gggggaagtt ttgtgtcaaa tatatgggaa ttattatcag caccacgaaa ttatacgcct 300

catatgaccc atttaggtgg atagatcatg gacactgttg acagctgcga agaaaaagcg 360

tattggggat gatccgaaat tagtccggta ccgaggcgca aatacgtaag acagccgatw 420

aaatatatgc gagaaacacc aaagagactc tagatgtttg tttggcacag ttttgacttc 480

tgcgaaggcc ttacaccacc ttgttgaccc ttgtcgcggg tcgggcaata tcggctgaca 540

gagttttact tgctcaataa gatacgagct gcatagagtt gaactacagg acaatattgg 600

ggctggccac atgaagggca ttgtttggag gtgtattgat ggtgaaaaca cgatatgaaa 660

tgacaacgcc ccctgtttta ttattattct tattattttg ggtgcttctc tatccataca 720

agcacctcct aacatgcttc ataagtgacc tcctcatcac aaggcctgag gtctcattta 780

tccagtggcg ccaagctaaa ctaaaactgg tccgagtaga ctaaggcgaa gagagaagga 840

gagaagacag ttttttttgtg gccgcctgtg aacaatgaaa acgatgaggg tgagatggag 900

caaaccatat ggacagtcag aggagtacgt acgctgctta cataatggcg caacgaccac 960

atgtcccaca gatacgcatc gattccggtt ccgggtacac cgtgggcgtt ggtgatatca 1020

ccaacgccca cggtgtaccc ggaaccggaa tcgataacgt ctcgtaccaa ccacagatta 1080

cgacccattc gcagtcacag ttcactaggg tttgggttgc atccgttgag agcggtttgt 1140

ttttaacctt ctccatgtgc tcactcaggt tttgggttca gatcaaatca aggcgtgaac 1200

cactttgttt gaggacaaat gtgacacaac caaccagtgt caggggcaag tccgtgacaa 1260

aggggaagat acaatgcaat tactgacagt tacagactgc ctcgatgccc taaccttgcc 1320

ccaaaataag acaactgtcc tcgtttaagc gcaaccctat tcagcgtcac gtcataatag 1380

cgtttggata gcactagtct atgaggagcg ttttatgttg cggtgagggc gattggtgct 1440

catatgggtt caattgaggt ggcggaacga gcttagtctt caattgaggt gcgagcgaca 1500

caattgggtg tcacgtggcc taattgacct cgggtcgtgg agtccccagt tatacagcaa 1560

ccacgaggtg catgggtagg agacgtcacc agacaatagg gttttttttg gactggagag 1620

ggttgggcaa aagcgctcaa cgggctgttt ggggagctgt gggggaggaa ttggcgatat 1680

ttgtgaggtt aacggctccg atttgcgtgt tttgtcgctc ctgcatctcc ccatacccat 1740

atcttccctc cccacctctt tccacgataa ttttacggat cagcaataag gttccttctc 1800

ctagtttcca cgtccatata tatctatgct gcgtcgtcct tttcgtgaca tcaccaaaac 1860

acatacaaca atggctgtta ctgacgtcct taagcgaaag tccggtgtca tcgtcggcga 1920

cgatgtccga gccgtgagta tccacgacaa gatcagtgtc gagacgacgc gttttgtgta 1980

atgacacaat ccgaaagtcg ctagcaacac acactctcta cacaaactaa cccagctctc 2040

catggacatg tccgtcctga ctctccaaga gtacgagttc gagaagcagt tcaacgagaa 2100

tgaagccatc caatggatgc aggaaaactg gaagaaatcc ttcctgtttt ctgccctcta 2160

cgctgccttt atctttggtg gacgacatct gatgaacaag cgagccaagt ttgagctgcg 2220

aaaacctctc gtgctctggt ccctgaccct cgctgtcttc tctatcttcg gtgctctgcg 2280

aactggagcc tacatgctct acatcctgat gaccaaaggc ctgaaacagt ctgtttgtga 2340

ccagtccttt tacaacggac ccgtctcgaa attctgggct tacgcctttg tgctctccaa 2400

agctcccgaa cttggcgata ccatcttcat cattctgcga aagcagaaac tcatcttcct 2460

gcactggtat caccacatca ccgtcctcct gtactcttgg tactcctaca aggacatggt 2520

ggctggaggt ggctggttca tgactatgaa ctacggtgtc cacgccgtga tgtactccta 2580

ctacgccctc cgagctgccg gtttccgagt ctctcgaaag tttgccatgt tcatcaccct 2640

gtcgcagatc actcagatgc tcatgggctg tgtcattaac tacctggtct tcaactggat 2700

gcagcatgac aatgaccagt gctactccca ctttcagaac atcttctggt cctctctcat 2760

gtacctctcc taccttctgc tcttctgcca tttcttcttt gaggcctaca ttggcaaagt 2820

gaagaaagcc accaaggctg agtaagcggc cgcatgtaca tacaagatta tttatagaaa 2880

tgaatcgcga tcgaacaaag agtacgagtg tacgagtagg ggatgatgat aaaagtggaa 2940

gaagttccgc atctttggat ttatcaacgt gtaggacgat acttcctgta aaaatgcaat 3000

gtctttacca taggttctgc tgtagatgtt attaactacc attaacatgt ctacttgtac 3060

agttgcagac cagttggagt atagaatggt acacttacca aaaagtgttg atggttgtaa 3120

ctacgatata taaaactgtt gacgggatct gtatattcgg taagatatat tttgtggggt 3180

tttagtggtg tttatcgctt accctgcagc cggattaagg ttggcaattt ttcacgtcct 3240

tgtctccgca attactcacc gggtggttta taagattgca agcgtcttga tttgtctctg 3300

tatactaaca tgcaatcgcg actcgcccga cgggccacta acctggccag aatctccaga 3360

tccaagtatt ctcttggtct gcgatatgtt tccaacacaa aagcccctgc tgcccagccg 3420

gcaactgctg agtgagtatt ccttgccata aacgacccag aaccactgta tagtgtttgg 3480

aagcactagt cagaagacca gcgaaaacag gtggaaaaaa ctgagacgaa aagcaacgac 3540

cagaaatgta atgtgtggaa aagcgacaca cacagagcag ataaagaggt gacaaataac 3600

gacaaatgaa atatcagtat cttcccacaa tcactacctc tcagctgtct gaaggtgcgg 3660

ctgatatatc catcccacgt ctaacgtatg gagtgtgata gaatatgacg acacaagcat 3720

gagaactcgc tctctatcca accaccgaaa cactgtcact acagccgttc ttgttgctcc 3780

attcgctttt gtgattccat gccttctctg gtgactgaca acattccttc cttttctcca 3840

gccctgttgt tatctgctca tgacctacgg ccactctcta tcgcatacta acatagacga 3900

tcccagcccg ctccccactt ccagggcacc gttggcaagc ctcctatcct caagaaggct 3960

gaggctgcca acgctgacat ggacgagtcc ttcatcggaa tgtctggagg agagatcttc 4020

cacgagatga tgctgcgaca caacgtcgac actgtcttcg gttaccccgg tggagccatt 4080

ctccccgtct ttgacgccat tcacaactct gagtacttca actttgtgct ccctcgacac 4140

gagcagggtg ccggccacat ggccgagggc tacgctcgag cctctggtaa gcccggtgtc 4200

gttctcgtca cctctggccc cggtgccacc aacgtcatca cccccatgca ggacgctctt 4260

tccgatggta cccccatggt tgtcttcacc ggtcaggtcc tgacctccgt tatcggcact 4320

gacgccttcc aggaggccga tgttgtcggc atctcccgat cttgcaccaa gtggaacgtc 4380

atggtcaaga acgttgctga gctccccga cgaatcaacg aggcctttga gattgctact 4440

tccggccgac ccggtcccgt tctcgtcgat ctgcccaagg atgttactgc tgccatcctg 4500

cgagagccca tccccaccaa gtccaccatt ccctcgcatt ctctgaccaa cctcacctct 4560

gccgccgcca ccgagttcca gaagcaggct atccagcgag ccgccaacct catcaaccag 4620

tccaagaagc ccgtccttta cgtcggacag ggtatccttg gctccgagga gggtcctaag 4680

ctgcttaagg agctggctga gaaggccgag attcccgtca ccactactct gcagggtctt 4740

ggtgcctttg acgagcgaga ccccaagtct ctgcacatgc tcggtatgca cggttccggc 4800

tacgccaaca tggccatgca gaacgctgac tgtatcattg ctctcggcgc ccgatttgat 4860

gaccgagtta ccggctccat ccccaagttt gcccccgagg ctcgagccgc tgcccttgag 4920

ggtcgaggtg gtattgttca ctttgagatc caggccaaga acatcaacaa ggttgttcag 4980

gccaccgaag ccgttgaggg agacgttacc gagtctgtcc gacagctcat ccccctcatc 5040

aacaaggtct ctgccgctga gcgagctccc tggactgaga ctatccagtc ctggaagcag 5100

cagttccccc tcctcttcga ggctgaaggt gaggatggtg ttatcaagcc ccagtccgtc 5160

attgctctgc tctctgacct gacagagaac aacaaggaca agaccatcat caccaccggt 5220

gttggtcagc atcagatgtg gactgcccag catttccgat ggcgacaccc tcgaaccatg 5280

atcacttctg gtggtcttgg aactatgggt tacggcctgc ccgccgctat cggcgccaag 5340

gttgcccgac ctgactgcga cgtcattgac atcgatggtg acgcttcttt caacatgact 5400

ctgaccgagc tgtccaccgc cgttcagttc aacattggcg tcaaggctat tgtcctcaac 5460

aacgaggaac agggtatggt cacccagctg cagtctctct tctacgagaa ccgatactgc 5520

cacactcatc agaagaaccc cgacttcatg aagctggccg agtccatggg catgaagggt 5580

atccgaatca ctcacattga ccagctggag gccggtctca aggagatgct cgcatacaag 5640

ggccctgtgc tcgttgaggt tgttgtcgac aagaagatcc ccgttcttcc catggttccc 5700

gctggtaagg ctttgcatga gttccttgtc tacgacgctg acgccgaggc tgcttctcga 5760

cccgatcgac tgaagaatgc ccccgccccc cacgtccacc agaccacctt tgagaactaa 5820

gtggaaagga acacaagcaa tccgaaccaa aaataattgg ggtcccgtgc ccacagagtc 5880

tagtgcagac ctaaaatgac cacagtaaat tatagctgtt attaaacatg agattttgac 5940

caacaagagc gtaggaatgt tattagctac tacttgtaca tacacagcat ttgttttaaa 6000

taatgttgcc tccaggggca gtgagatcag gacccagatc cgtggccagc tctctgactt 6060

cagaccgctt gtacttaagc agctcgcaac actgttgtcg aggattgaac ttgccatatt 6120

cgattttgtg gtcatgaatc cagcacacct cgtacgggcg tcgttgcttg tgtgattttt 6180

gaggacccat ccctttggta tataagtata ctctggggtt aaggttgccc gtgtagtcta 6240

ggttatagtt ttcatgtgaa ataccgagag ccgagggaga ataaacgggg gtatttggac 6300

ttgttttttt cgcggaaaag cgtcgaatca accctgcggg ccttgcacca tgtccacgac 6360

gtgtttctcg ccccaattcg ccccttgcac gtcaaaatta ggcctccatc tagacccctc 6420

cataacatgt gactgtgggg aaaagtataa gggaaaccat gcaaccatag acgacgtgaa 6480

agacggggag gaaccaatgg aggccaaaga aatggggtag caacagtcca ggagacagac 6540

aaggagacaa ggagagggcg cccgaaagat cggaaaaaca aacatgtcca attggggcag 6600

tgacggaaac gacacggaca cttcagtaca atggaccgac catctccaag ccagggttat 6660

tccggtatca ccttggccgt aacctcccgc tggtacctga tattgtacac gttcacattc 6720

aatatacttt cagctacaat aagagaggct gtttgtcggg catgtgtgtc cgtcgtatgg 6780

ggtgatgtcc gagggcgaaa ttcgctacaa gcttaactct ggcgcttgtc cagtatgaat 6840

agacaagtca agaccagtgg tgccatgatt gacagggagg tacaagactt cgatactcga 6900

gcattactcg gacttgtggc gattgaacag acgggcgatc gcttctcccc cgtattgccg 6960

gcgcgccagc tgcattaatg aatcggccaa cgcgcgggga gaggcggttt gcgtattggg 7020

cgctcttccg cttcctcgct cactgactcg ctgcgctcgg tcgttcggct gcggcgagcg 7080

gtatcagctc actcaaaggc ggtaatacgg ttatccacag aatcagggga taacgcagga 7140

aagaacatgt gagcaaaagg ccagcaaaag gccaggaacc gtaaaaaggc cgcgttgctg 7200

gcgtttttcc ataggctccg ccccccctgac gagcatcaca aaaatcgacg ctcaagtcag 7260

aggtggcgaa acccgacagg actataaaga taccaggcgt ttcccccctgg aagctccctc 7320

gtgcgctctc ctgttccgac cctgccgctt accggatacc tgtccgcctt tctcccttcg 7380

ggaagcgtgg cgctttctca tagctcacgc tgtaggtatc tcagttcggt gtaggtcgtt 7440

cgctccaagc tgggctgtgt gcacgaaccc cccgttcagc ccgaccgctg cgccttatcc 7500

ggtaactatc gtcttgagtc caacccggta agacacgact tatcgccact ggcagcagcc 7560

actggtaaca ggattagcag agcgaggtat gtaggcggtg ctacagagtt cttgaagtgg 7620

tggcctaact acggctacac tagaagaaca gtatttggta tctgcgctct gctgaagcca 7680

gttaccttcg gaaaaagagt tggtagctct tgatccggca aacaaaccac cgctggtagc 7740

ggtggttttt ttgtttgcaa gcagcagatt acgcgcagaa aaaaaggatc tcaagaagat 7800

cctttgatct tttctacggg gtctgacgct cagtggaacg aaaactcacg ttaagggatt 7860

ttggtcatga gattatcaaa aaggatcttc acctagatcc ttttaaatta aaaatgaagt 7920

tttaaatcaa tctaaagtat atatgagtaa acttggtctg acagttacca atgcttaatc 7980

agtgaggcac ctatctcagc gatctgtcta tttcgttcat ccatagttgc ctgactcccc 8040

gtcgtgtaga taactacgat acgggagggc ttaccatctg gccccagtgc tgcaatgata 8100

ccgcgagacc cacgctcacc ggctccagat ttatcagcaa taaaccagcc agccggaagg 8160

gccgagcgca gaagtggtcc tgcaacttta tccgcctcca tccagtctat taattgttgc 8220

cgggaagcta gagtaagtag ttcgccagtt aatagtttgc gcaacgttgt tgccattgct 8280

acaggcatcg tggtgtcacg ctcgtcgttt ggtatggctt cattcagctc cggttcccaa 8340

cgatcaaggc gagttacatg atcccccatg ttgtgcaaaa aagcggttag ctccttcggt 8400

cctccgatcg ttgtcagaag taagttggcc gcagtgttat cactcatggt tatggcagca 8460

ctgcataatt ctcttactgt catgccatcc gtaagatgct tttctgtgac tggtgagtac 8520

tcaaccaagt cattctgaga atagtgtatg cggcgaccga gttgctcttg cccggcgtca 8580

atacgggata ataccgcgcc acatagcaga actttaaaag tgctcatcat tggaaaacgt 8640

tcttcggggc gaaaactctc aaggatctta ccgctgttga gatccagttc gatgtaaccc 8700

actcgtgcac ccaactgatc ttcagcatct tttactttca ccagcgtttc tgggtgagca 8760

aaaacaggaa ggcaaaatgc cgcaaaaaag ggaataaggg cgacacggaa atgttgaata 8820

ctcatactct tccttttca atattattga agcatttatc agggttattg tctcatgagc 8880

ggatacatat ttgaatgtat ttagaaaaat aaacaaatag gggttccgcg cacatttccc 8940

cgaaaagtgc cacctgatgc ggtgtgaaat accgcacaga tgcgtaagga gaaaataccg 9000

catcaggaaa ttgtaagcgt taatattttg ttaaaattcg cgttaaattt ttgttaaatc 9060

agctcatttt ttaaccaata ggccgaaatc ggcaaaatcc cttataaatc aaaagaatag 9120

accgagatag ggttgagtgt tgttccagtt tggaacaaga gtccactatt aaagaacgtg 9180

gactccaacg tcaaagggcg aaaaaccgtc tatcagggcg atggcccact acgtgaacca 9240

tcaccctaat caagtttttt ggggtcgagg tgccgtaaag cactaaatcg gaaccctaaa 9300

gggagccccc gatttagagc ttgacgggga aagccggcga acgtggcgag aaaggaaggg 9360

aagaaagcga aaggagcggg cgctagggcg ctggcaagtg tagcggtcac gctgcgcgta 9420

accaccacac ccgccgcgct taatgcgccg ctacagggcg cgtccattcg ccattcaggc 9480

tgcgcaactg ttgggaaggg cgatcggtgc gggcctcttc gctattacgc cagctggcga 9540

aaggggggatg tgctgcaagg cgattaagtt gggtaacgcc agggttttcc cagtcacgac 9600

gttgtaaaac gacggccagt gaattgtaat acgactcact atagggcgaa ttgggcccga 9660

cgtcgcatgc gctgatgaca ctttggtctg aaagagatgc attttgaatc ccaaacttgc 9720

agtgcccaag tgacatacat ctccgcgttt tggaaaatgt tcagaaacag ttgattgtgt 9780

tggaatgggg aatggggaat ggaaaaatga ctcaagtatc aattccaaaa acttctctgg 9840

ctggcagtac ctactgtcca tactactgca ttttctccag tcaggccact ctatactcga 9900

cgacacagta gtaaaaccca gataatttcg acataaacaa gaaaacagac ccaataatat 9960

ttatatatag tcagccgttt gtccagttca gactgtaata gccgaaaaaa aatccaaagt 10020

ttctattcta ggaaaatata ttccaatatt tttaattctt aatctcattt attttattct 10080

agcgaaatac atttcagcta cttgagacat gtgataccca caaatcggat tcggactcgg 10140

ttgttcagaa gagcatatgg cattcgtgct cgcttgttca cgtattcttc ctgttccatc 10200

tcttggccga caatcacaca aaaatggggt ttttttttta attctaatga ttcattacag 10260

caaaattgag atatagcaga ccacgtattc cataatcacc aaggaagttc ttgggcgtct 10320

taattaaaag gcgttgaaac agaatgagcc agacagcaag gacaaggtgg ccaacagcaa 10380

ggagtccaaa aagccctcta ttgacgagat ccacgatgtt attgctcatg aggtttccga 10440

gctcgatgct gggaagaaga agtgatttgt atataagaaa taaatgagat atagtaaagg 10500

agtgcaagag aatggcaagg tggtcaaatt ctatattact tgcagtcact ggttcctcgt 10560

tgacatgaat gaagttaccg ttggcatagc tgatttaata tataactgtc caactaactc 10620

tcacctagat ataacccatg tgtgtgtttc caatcatcaa tgcggccgct tagagagcct 10680

tgccagcggg ctgcttctcg gccattcgag cgaacacggc caggtatcgg agcaggatga 10740

cgagaccttc atggggcaga gggtttcggt aggggaggtt gtgcttctgg cacagctgtt 10800

ccacctggta ggaaacggca gtgaggttgt gtcgaggcag ggtgggccag agatggtgct 10860

cgatctggta gttcaggcct ccaaagaacc agtcagtaat gatgcctcgt cgaatgttca 10920

tggtctcatg gatctgaccc acagagaagc catgtccgtc ccagacggaa tcaccgatct 10980

tctccagagg gtagtggttc atgaagacca cgatggcaat tccgaagcca ccgacgagct 11040

cggaaacaaa gaacaccagg agcgaggtca ggatggaggg cataaagaag aggtggaaca 11100

gggtcttgag agtccagtgc agagcgagtc caatggcctc tttcttgtac tgagatcggt 11160

agaactggtt gtctcggtcc ttgagggatc gaacggtcag cacagactgg aaacaccaga 11220

tgaatcgcag gagaatacag atgaccagga aatagtactg ttggaactga atgagctttc 11280

gggagatggg agaagctcga gtgacatcgt cctcggacca ggcgagcaga ggcaggttat 11340

caatgtcggg atcgtgaccc tgaacgttgg tagcagaatg atgggcgttg tgtctgtcct 11400

tccaccaggt cacggagaag ccctggagtc cgttgccaaa gaccagaccc acgaggttat 11460

tccagtttcg gttcttgaag gtctggtggt ggcagatgtc atgagacagc catcccattt 11520

gctggtagtg cataccgagc agcacagcac caatgaagta catctggtac tggaccatca 11580

ggaagtatcc aagcacgcca agacccaggg tggtggagat cttgtacgag taccagaggg 11640

gagaggcgtc aaacatgcca gtggcgatca gctcttctcg gagctttcgg aaatcctcct 11700

gagcttcgtt gacggcagcc tggggaggca gctcggagga ggggttgatc ttgggcattc 11760

gcttgagctt gtcgaaggct tcctgagagt gcataaccat gaaggcgtca gtagcatctc 11820

gtccctggta gttctcaatg atttcagctc caccagggtg gaagttcacc caagcggaga 11880

cgtcgtaggt agttccgtcg atggtgaggg gcagagcctg tcgcttggac ttcaccatgg 11940

agagctgggt tagtttgtgt agagagtgtg tgttgctagc gactttcgga ttgtgtcatt 12000

acacaaaacg cgtcgtctcg acactgatct tgtcgtggat actcacggct cggacatcgt 12060

cgccgacgat gacaccggac tttcgcttaa ggacgtcagt aacaggcatt gtgtgatgtg 12120

tagtttagat ttcgaatctg tggggaaaga aaggaaaaaa gagactggca accgattggg 12180

agagccactg tttatatata ccctagacaa gccccccgct tgtaagatgt tggtcaatgt 12240

aaaccagtat taaggttggc aagtgcagga gaagcaaggt gtgggtaccg agcaatggaa 12300

atgtgcggaa ggcaaaaaaa tgaggccacg gcctattgtc ggggctatat ccaggggcg 12360

attgaagtac actaacatga catgtgtcca cagaccctca atctggcctg atgagccaaa 12420

tccatacgcg ctttcgcagc tctaaaggct ataacaagtc acaccaccct gctcgacctc 12480

agcgccctca ctttttgtta agacaaactg tacacgctgt tccagcgttt tctgcctgca 12540

cctggtggga catttggtgc aacctaaagt gctcggaacc tctgtggtgt ccagatcagc 12600

gcagcagttc cgaggtagtt ttgaggccct tagatgatgc aatggtgtca gtcgctggat 12660

cacgagtctt aatggcagta ttcgttctta tttgtgccat tgagccccgt tatcctcgta 12720

tcttctaccc cccatcccat ccctttgttg gtgcaaccct acccatttat tgttgggtgc 12780

agcccaaccg acgtggagag cttggcttgg ccatataaaa aggccccccc ctagtggcaa 12840

tggcagaaag tcagctgtga gttgttgaat ttgtcatcta ggcggcctgg ccgtcttctc 12900

cggggcaatt taaattcctt cacttcaagt tcattcttca tctgcttctg ttttactttg 12960

acaggcaaat gaagacatgg tacgacttga tggaggccaa gaacgccatt tcaccccgag 13020

acaccgaagt gcctgaaatc ctggctgccc ccattgataa catcggaaac tacggtattc 13080

cggaaagtgt atatagaacc tttccccagc ttgtgtctgt ggatatggat ggtgtaatcc 13140

cctttgagta ctcgtcttgg cttctctccg agcagtatga ggctctctaa tctagcgcat 13200

ttaatatctc aatgtattta tatatttatc ttctcatgcg gccgctaaag ctgcttacca 13260

gccttagcgg atttcttggt ggccaggttg tcctggtaaa agaagtgaca gaacaggaga 13320

aagacagatc cgacgtaggc gtagttgaaa gcccaggaga acagcttgcc cttgtcagag 13380

ttgaagcagg gaacgttgat gtagtcccag accaggagaa agccaccgac gaactggcaa 13440

atctgcatgg cagtgatcag aggcttggcc ttgaacttgt agccagcggc agtcagtcca 13500

tagtaggtgt acatgatggt gtgaatgaac gagttaaaga acatgaagat ccacacaccc 13560

tcgttgtgca gtcgaatgcc gaggtagacg tcccagggag ctccaaagtg atggaaggcc 13620

tgcagaaagg acactcgctt gcccttgagg accagccaag cggtgtcgag gtactccacg 13680

tacttagaat agtagaaggc cttggcagtc caggtgaaca gcttggagtc ccagacagga 13740

gagggacact gaaagagagg ctggggagta tcaccggtct gtcttcgcag ccaggctcca 13800

gtaccgtagt cccagccgag agcggtggca gtcacgtaga aggacagggc agagaagaga 13860

gccaggagga cgttgtacca gatcatggag gttcggtagg ctcctttctt ctcgtccacg 13920

agaccagagt ttcgcaggag aggcttcagg agcaggtagg agaaggtgcc aatgaggatt 13980

tcgggatcgg tgacggcagc ccagattcgc tcgccagcgt cgttggccag agccatggtg 14040

aatgattctt atactcagaa ggaaatgctt aacgatttcg ggtgtgagtt gacaaggaga 14100

gagagaaaag aagaggaaag gtaattcggg gacggtggtc ttttataccc ttggctaaag 14160

tcccaaccac aaagcaaaaa aattttcagt agtctatttt gcgtccggca tgggttaccc 14220

ggatggccag acaaagaaac tagtacaaag tctgaacaag cgtagattcc agactgcagt 14280

accctacgcc cttaacggca agtgtgggaa ccgggggagg tttgatatgt ggggtgaagg 14340

gggctctcgc cggggttggg cccgctactg ggtcaatttg gggtcaattg gggcaattgg 14400

ggctgttttt tgggacacaa atacgccgcc aacccggtct ctcctgaatt cgaatacgat 14460

atacatcaaa gactatgata cgcagtattg cacactgtac gagtaagagc actagccact 14520

gcactcaagt gaaaccgttg cccgggtacg agtatgagta tgtacagtat gtttagtatt 14580

gtacttggac agtgcttgta tcgtacattc tcaagtgtca aacataaata tccgttgcta 14640

tatcctcgca ccaccacgta gctcgctata tccctgtgtt gaatccatcc atcttggatt 14700

gccaattgtg cacacagaac cgggcactca cttccccatc cacacttgcg gccgctaaag 14760

ctgcttacca gccttagcgg atttcttggt ggccaggttg tcctggtaaa agaagtgaca 14820

gaacaggaga aagacagatc cgacgtaggc gtagttgaaa gcccaggaga acagcttgcc 14880

cttgtcagag ttgaagcagg gaacgttgat gtagtcccag accaggagaa agccaccgac 14940

gaactggcaa atctgcatgg cagtgatcag aggcttggcc ttgaacttgt agccagcggc 15000

agtcagtcca tagtaggtgt acatgatggt gtgaatgaac gagttaaaga acatgaagat 15060

ccacacaccc tcgttgtgca gtcgaatgcc gaggtagacg tcccagggag ctccaaagtg 15120

atggaaggcc tgcagaaagg acactcgctt gcccttgagg accagccaag cggtgtcgag 15180

gtactccacg tacttagaat agtagaaggc cttggcagtc caggtgaaca gcttggagtc 15240

ccagacagga gagggacact gaaagagagg ctggggagta tcaccggtct gtcttcgcag 15300

ccaggctcca gtaccgtagt cccagccgag agcggtggca gtcacgtaga aggacagggc 15360

agagaagaga gccaggagga cgttgtacca gatcatggag gttcggtagg ctcctttctt 15420

ctcgtccacg agaccagagt ttcgcaggag aggcttcagg agcaggtagg agaaggtgcc 15480

aatgaggatt tcgggatcgg tgacggcagc ccagattcgc tcgccagcgt cgttggccag 15540

agc 15543

<210> 173
<211> 13295
<212> DNA
<213> Artificial Sequence

<220>
<223> Plasmid pDMW314

<220>
<221> misc_feature
<222> (10915)..(10918)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (10923)..(10923)
<223> n is a, c, g, or t

<400> 173
aacccggtaa gacacgactt atcgccactg gcagcagcca ctggtaacag gattagcaga 60

gcgaggtatg taggcggtgc tacagagttc ttgaagtggt ggcctaacta cggctacact 120

agaagaacag tatttggtat ctgcgctctg ctgaagccag ttaccttcgg aaaaagagtt 180

ggtagctctt gatccggcaa acaaaccacc gctggtagcg gtggtttttt tgtttgcaag 240

cagcagatta cgcgcagaaa aaaaggatct caagaagatc ctttgatctt ttctacgggg 300

tctgacgctc agtggaacga aaactcacgt taagggattt tggtcatgag attatcaaaa 360

aggatcttca cctagatcct tttaaattaa aaatgaagtt ttaaatcaat ctaaagtata 420

tatgagtaaa cttggtctga cagttaccaa tgcttaatca gtgaggcacc tatctcagcg 480

atctgtctat ttcgttcatc catagttgcc tgactccccg tcgtgtagat aactacgata 540

cgggagggct taccatctgg ccccagtgct gcaatgatac cgcgagaccc acgctcaccg 600

gctccagatt tatcagcaat aaaccagcca gccggaaggg ccgagcgcag aagtggtcct 660

gcaactttat ccgcctccat ccagtctatt aattgttgcc gggaagctag agtaagtagt 720

tcgccagtta atagtttgcg caacgttgtt gccattgcta caggcatcgt ggtgtcacgc 780

tcgtcgtttg gtatggcttc attcagctcc ggttcccaac gatcaaggcg agttacatga 840

tcccccatgt tgtgcaaaaa agcggttagc tccttcggtc ctccgatcgt tgtcagaagt 900

aagttggccg cagtgttatc actcatggtt atggcagcac tgcataattc tcttactgtc 960

atgccatccg taagatgctt ttctgtgact ggtgagtact caaccaagtc attctgagaa 1020

tagtgtatgc ggcgaccgag ttgctcttgc ccggcgtcaa tacgggataa taccgcgcca 1080

catagcagaa ctttaaaagt gctcatcatt ggaaaacgtt cttcggggcg aaaactctca 1140

aggatcttac cgctgttgag atccagttcg atgtaaccca ctcgtgcacc caactgatct 1200

tcagcatctt ttactttcac cagcgtttct gggtgagcaa aaacaggaag gcaaaatgcc 1260

gcaaaaaagg gaataagggc gacacggaaa tgttgaatac tcatactctt cctttttcaa 1320

tattattgaa gcatttatca gggttattgt ctcatgagcg gatacatatt tgaatgtatt 1380

tagaaaaata aacaaatagg ggttccgcgc acatttcccc gaaaagtgcc acctgatgcg 1440

gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcaggaaat tgtaagcgtt 1500

aatattttgt taaaattcgc gttaaatttt tgttaaatca gctcattttt taaccaatag 1560

gccgaaatcg gcaaaatccc ttataaatca aaagaataga ccgagatagg gttgagtgtt 1620

gttccagttt ggaacaagag tccactatta aagaacgtgg actccaacgt caaagggcga 1680

aaaaccgtct atcagggcga tggcccacta cgtgaaccat caccctaatc aagtttttg 1740

gggtcgaggt gccgtaaagc actaaatcgg aaccctaaag ggagcccccg atttagagct 1800

tgacggggaa agccggcgaa cgtggcgaga aaggaaggga agaaagcgaa aggagcgggc 1860

gctagggcgc tggcaagtgt agcggtcacg ctgcgcgtaa ccaccacacc cgccgcgctt 1920

aatgcgccgc tacagggcgc gtccattcgc cattcaggct gcgcaactgt tgggaagggc 1980

gatcggtgcg ggcctcttcg ctattacgcc agctggcgaa aggggatgt gctgcaaggc 2040

gattaagttg ggtaacgcca gggttttccc agtcacgacg ttgtaaaacg acggccagtg 2100

aattgtaata cgactcacta tagggcgaat tgggcccgac gtcgcatgca gtggtggtat 2160

tgtgactggg gatgtagttg agaataagtc atacacaagt cagctttctt cgagcctcat 2220

ataagtataa gtagttcaac gtattagcac tgtacccagc atctccgtat cgagaaacac 2280

aacaacatgc cccattggac agatcatgcg gatacacagg ttgtgcagta tcatacatac 2340

tcgatcagac aggtcgtctg accatcatac aagctgaaca agcgctccat acttgcacgc 2400

tctctatata cacagttaaa ttacatatcc atagtctaac ctctaacagt taatcttctg 2460

gtaagcctcc cagccagcct tctggtatcg cttggcctcc tcaataggat ctcggttctg 2520

gccgtacaga cctcggccga caattatgat atccgttccg gtagacatga catcctcaac 2580

agttcggtac tgctgtccga gagcgtctcc cttgtcgtca agacccaccc cgggggtcag 2640

aataagccag tcctcagagt cgcccttaat taaaaggcgt tgaaacagaa tgagccagac 2700

agcaaggaca aggtggccaa cagcaaggag tccaaaaagc cctctattga cgagatccac 2760

gatgttattg ctcatgaggt ttccgagctc gatgctggga agaagaagtg atttgtatat 2820

aagaaataaa tgagatatag taaaggagtg caagagaatg gcaaggtggt caaattctat 2880

attacttgca gtcactggtt cctcgttgac atgaatgaag ttaccgttgg catagctgat 2940

ttaatatata actgtccaac taactctcac ctagatataa cccatgtgtg tgtttccaat 3000

catcaatgcg gccgcttaga gagccttgcc agcgggctgc ttctcggcca ttcgagcgaa 3060

cacggccagg tatcggagca ggatgacgag accttcatgg ggcagagggt ttcggtaggg 3120

gaggttgtgc ttctggcaca gctgttccac ctggtaggaa acggcagtga ggttgtgtcg 3180

aggcagggtg ggccagagat ggtgctcgat ctggtagttc aggcctccaa agaaccagtc 3240

agtaatgatg cctcgtcgaa tgttcatggt ctcatggatc tgacccacag agaagccatg 3300

tccgtcccag acggaatcac cgatcttctc cagagggtag tggttcatga agaccacgat 3360

ggcaattccg aagccaccga cgagctcgga aacaaagaac accaggagcg aggtcaggat 3420

ggagggcata aagaagaggt ggaacagggt cttgagagtc cagtgcagag cgagtccaat 3480

ggcctctttc ttgtactgag atcggtagaa ctggttgtct cggtccttga gggatcgaac 3540

ggtcagcaca gactggaaac accagatgaa tcgcaggaga atacagatga ccaggaaata 3600

gtactgttgg aactgaatga gctttcggga gatgggagaa gctcgagtga catcgtcctc 3660

ggaccaggcg agcagaggca ggttatcaat gtcgggatcg tgaccctgaa cgttggtagc 3720

agaatgatgg gcgttgtgtc tgtccttcca ccaggtcacg gagaagccct ggagtccgtt 3780

gccaaagacc agacccacga ggttattcca gtttcggttc ttgaaggtct ggtggtggca 3840

gatgtcatga gacagccatc ccatttgctg gtagtgcata ccgagcagca cagcaccaat 3900

gaagtacatc tggtactgga ccatcaggaa gtatccaagc acgccaagac ccagggtggt 3960

ggagatcttg tacgagtacc agaggggaga ggcgtcaaac atgccagtgg cgatcagctc 4020

ttctcggagc tttcggaaat cctcctgagc ttcgttgacg gcagcctggg gaggcagctc 4080

ggaggagggg ttgatcttgg gcattcgctt gagcttgtcg aaggcttcct gagagtgcat 4140

aaccatgaag gcgtcagtag catctcgtcc ctggtagttc tcaatgattt cagctccacc 4200

agggtggaag ttcacccaag cggagacgtc gtaggtagtt ccgtcgatgg tgaggggcag 4260

agcctgtcgc ttggacttca ccatggagag ctgggttagt ttgtgtagag agtgtgtgtt 4320

gctagcgact ttcggattgt gtcattacac aaaacgcgtc gtctcgacac tgatcttgtc 4380

gtggatactc acggctcgga catcgtcgcc gacgatgaca ccggactttc gcttaaggac 4440

gtcagtaaca ggcattgtgt gatgtgtagt ttagatttcg aatctgtggg gaaagaaagg 4500

aaaaaagaga ctggcaaccg attgggagag ccactgttta tatataccct agacaagccc 4560

cccgcttgta agatgttggt caatgtaaac cagtattaag gttggcaagt gcaggagaag 4620

caaggtgtgg gtaccgagca atggaaatgt gcggaaggca aaaaaatgag gccacggcct 4680

attgtcgggg ctatatccag ggggcgattg aagtacacta acatgacatg tgtccacaga 4740

ccctcaatct ggcctgatga gccaaatcca tacgcgcttt cgcagctcta aaggctataa 4800

caagtcacac caccctgctc gacctcagcg ccctcacttt ttgttaagac aaactgtaca 4860

cgctgttcca gcgttttctg cctgcacctg gtgggacatt tggtgcaacc taaagtgctc 4920

ggaacctctg tggtgtccag atcagcgcag cagttccgag gtagttttga ggcccttaga 4980

tgatgcaatg gtgtcagtcg ctggatcacg agtcttaatg gcagtattcg ttcttatttg 5040

tgccattgag ccccgttatc ctcgtatctt ctacccccca tcccatccct ttgttggtgc 5100

aaccctaccc atttattgtt gggtgcagcc caaccgacgt ggagagcttg gcttggccat 5160

ataaaaaggc ccccccctag tggcaatggc agaaagtcag ctgtgagttg ttgaatttgt 5220

catctaggcg gcctggccgt cttctccggg gcaatttaaa ttccttcact tcaagttcat 5280

tcttcatctg cttctgtttt actttgacag gcaaatgaag acatggtacg acttgatgga 5340

ggccaagaac gccatttcac cccgagacac cgaagtgcct gaaatcctgg ctgcccccat 5400

tgataacatc ggaaactacg gtattccgga aagtgtatat agaacctttc cccagcttgt 5460

gtctgtggat atggatggtg taatcccctt tgagtactcg tcttggcttc tctccgagca 5520

gtatgaggct ctctaatcta gcgcatttaa tatctcaatg tatttatata tttatcttct 5580

catgcggccg ctaaagctgc ttaccagcct tagcggattt cttggtggcc aggttgtcct 5640

ggtaaaagaa gtgacagaac aggagaaaga cagatccgac gtaggcgtag ttgaaagccc 5700

aggagaacag cttgcccttg tcagagttga agcagggaac gttgatgtag tcccagacca 5760

ggagaaagcc accgacgaac tggcaaatct gcatggcagt gatcagaggc ttggccttga 5820

acttgtagcc agcggcagtc agtccatagt aggtgtacat gatggtgtga atgaacgagt 5880

taaagaacat gaagatccac acaccctcgt tgtgcagtcg aatgccgagg tagacgtccc 5940

agggagctcc aaagtgatgg aaggcctgca gaaaggacac tcgcttgccc ttgaggacca 6000

gccaagcggt gtcgaggtac tccacgtact tagaatagta gaaggccttg gcagtccagg 6060

tgaacagctt ggagtcccag acaggagagg gacactgaaa gagaggctgg ggagtatcac 6120

cggtctgtct tcgcagccag gctccagtac cgtagtccca gccgagagcg gtggcagtca 6180

cgtagaagga cagggcagag aagagagcca ggaggacgtt gtaccagatc atggaggttc 6240

ggtaggctcc tttcttctcg tccacgagac cagagtttcg caggagaggc ttcaggagca 6300

ggtaggagaa ggtgccaatg aggatttcgg gatcggtgac ggcagcccag attcgctcgc 6360

cagcgtcgtt ggccagagcc atggtgaatg attcttatac tcagaaggaa atgcttaacg 6420

atttcgggtg tgagttgaca aggagagaga gaaaagaaga ggaaaggtaa ttcggggacg 6480

gtggtctttt ataccccttgg ctaaagtccc aaccacaaag caaaaaaatt ttcagtagtc 6540

tattttgcgt ccggcatggg ttacccggat ggccagacaa agaaactagt acaaagtctg 6600

aacaagcgta gattccagac tgcagtaccc tacgccctta acggcaagtg tgggaaccgg 6660

gggaggtttg atatgtgggg tgaaggggggc tctcgccggg gttgggcccg ctactgggtc 6720

aatttggggt caattggggc aattggggct gttttttggg acacaaatac gccgccaacc 6780

cggtctctcc tgaattcgaa tacgatatac atcaaagact atgatacgca gtattgcaca 6840

ctgtacgagt aagagcacta gccactgcac tcaagtgaaa ccgttgcccg ggtacgagta 6900

tgagtatgta cagtatgttt agtattgtac ttggacagtg cttgtatcgt acattctcaa 6960

gtgtcaaaca taaatatccg ttgctatatc ctcgcaccac cacgtagctc gctatatccc 7020

tgtgttgaat ccatccatct tggattgcca attgtgcaca cagaaccggg cactcacttc 7080

cccatccaca cttgcggccg ctaaagctgc ttaccagcct tagcggattt cttggtggcc 7140

aggttgtcct ggtaaaagaa gtgacagaac aggagaaaga cagatccgac gtaggcgtag 7200

ttgaaagccc aggagaacag cttgcccttg tcagagttga agcagggaac gttgatgtag 7260

tcccagacca ggagaaagcc accgacgaac tggcaaatct gcatggcagt gatcagaggc 7320

ttggccttga acttgtagcc agcggcagtc agtccatagt aggtgtacat gatggtgtga 7380

atgaacgagt taaagaacat gaagatccac acaccctcgt tgtgcagtcg aatgccgagg 7440

tagacgtccc agggagctcc aaagtgatgg aaggcctgca gaaaggacac tcgcttgccc 7500

ttgaggacca gccaagcggt gtcgaggtac tccacgtact tagaatagta gaaggccttg 7560

gcagtccagg tgaacagctt ggagtcccag acaggagagg gacactgaaa gagaggctgg 7620

ggagtatcac cggtctgtct tcgcagccag gctccagtac cgtagtccca gccgagagcg 7680

gtggcagtca cgtagaagga cagggcagag aagagagcca ggaggacgtt gtaccagatc 7740

atggaggttc ggtaggctcc tttcttctcg tccacgagac cagagtttcg caggagaggc 7800

ttcaggagca ggtaggagaa ggtgccaatg aggatttcgg gatcggtgac ggcagcccag 7860

attcgctcgc cagcgtcgtt ggccagagcc atggttagcg tgtcgtgttt ttgttgtgct 7920

ggaagaacca aagggtggcg caatgtgtgt agatatatat gtcgtgaccc acaagtcaca 7980

caaacaagta tcgggaggag tggtgcacct ctatgcggag aaaccttata ccgctgtaga 8040

ccaactgggg cagaggtgtg agttgaagtc agctggagga gatgtgtgac agaagcacaa 8100

gaagtgagat tgtgagatgt atgtctaggg ggggaagttt tgtgtcaaat atatgggaat 8160

tattatcagc accacgaaat tatacgcctc atatgaccca tttaggtgga tagatcatgg 8220

acactgttga cagctgcgaa gaaaaagcgt attggggatg atccgaaatt agtccggtac 8280

cgaggcgcaa atacgtaaga cagccgatwa aatatatgcg agaaacacca aagagactct 8340

agatgtttgt ttggcacagt tttgacttct gcgaaggcct tacaccacct tgttgaccct 8400

tgtcgcgggt cgggcaatat cggctgacag agttttactt gctcaataag atacgagctg 8460

catagagttg aactacagga caatattggg gctggccaca tgaagggcat tgtttggagg 8520

tgtattgatg gtgaaaacac gatatgaaat gacaacgccc cctgttttat tattattctt 8580

attattttgg gtgcttctct atccatacaa gcacctccta acatgcttca taagtgacct 8640

cctcatcaca aggcctgagg tctcatttat ccagtggcgc caagctaaac taaaactggt 8700

ccgagtagac taaggcgaag agagaaggag agaagacagt ttttttgtgg ccgcctgtga 8760

acaatgaaaa cgatgagggt gagatggagc aaaccatatg gacagtcaga ggagtacgta 8820

cgctgcttac ataatggcgc aacgaccaca tgtcccacag atacgcatcg atgagcctaa 8880

aatgaacccg agtatatctc ataaaattct cggtgagagg tctgtgactg tcagtacaag 8940

gtgccttcat tatgccctca accttaccat acctcactga atgtagtgta cctctaaaaa 9000

tgaaatacag tgccaaaagc caaggcactg agctcgtcta acggacttga tatacaacca 9060

attaaaacaa atgaaaagaa atacagttct ttgtatcatt tgtaacaatt accctgtaca 9120

aactaaggta ttgaaatccc acaatattcc caaagtccac ccctttccaa attgtcatgc 9180

ctacaactca tataccaagc actaacctac cgtttaaaca gtgtacgcag atctactata 9240

gaggaacatt taaattgccc cggagaagac ggccaggccg cctagatgac aaattcaaca 9300

actcacagct gactttctgc cattgccact agggggggggc ctttttatat ggccaagcca 9360

agctctccac gtcggttggg ctgcacccaa caataaatgg gtagggttgc accaacaaag 9420

ggatgggatg gggggtagaa gatacgagga taacggggct caatggcaca aataagaacg 9480

aatactgcca ttaagactcg tgatccagcg actgacacca ttgcatcatc taagggcctc 9540

aaaactacct cggaactgct gcgctgatct ggacaccaca gaggttccga gcactttagg 9600

ttgcaccaaa tgtcccacca ggtgcaggca gaaaacgctg gaacagcgtg tacagtttgt 9660

cttaacaaaa agtgagggcg ctgaggtcga gcagggtggt gtgacttgtt atagccttta 9720

gagctgcgaa agcgcgtatg gatttggctc atcaggccag attgagggtc tgtggacaca 9780

tgtcatgtta gtgtacttca atcgcccccc ggatatagcc ccgacaatag gccgtggcct 9840

catttttttg ccttccgcac atttccattg ctcggtaccc acaccttgct tctcctgcac 9900

ttgccaacct taatactggt ttacattgac caacatctta caagcggggg gcttgtctag 9960

ggtatatata aacagtggct ctcccaatcg gttgccagtc tctttttttcc tttctttccc 10020

cacagattcg aaatctaaac tacacatcac acaatgcctg ttactgacgt ccttaagcga 10080

aagtccggtg tcatcgtcgg cgacgatgtc cgagccgtga gtatccacga caagatcagt 10140

gtcgagacga cgcgttttgt gtaatgacac aatccgaaag tcgctagcaa cacacactct 10200

ctacacaaac taacccagct ctccatggcg tccacttcgg ctctgcccaa gcagaaccct 10260

gcgcttagac gcaccgtcac ctcaactact gtgacggatt ctgagtctgc cgccgtctct 10320

ccttcagact ctccccgcca ctcggcctct tccacatcgc tctcgtccat gtccgaggtt 10380

gatatcgcca agcccaagtc cgagtatggt gtcatgctcg acacctacgg caaccagttc 10440

gaggttcccg actttaccat caaggacatc tacaatgcca tccctaagca ctgcttcaag 10500

cgctccgctc tcaagggata cggttatatc ctccgcgaca ttgtcctcct gactaccact 10560

ttcagcatct ggtacaactt tgtgaccccc gaatatatcc cctccacccc cgcccgcgct 10620

ggtctgtggg ccgtgtacac cgttcttcag ggtcttttcg gtactggtct ctgggttatt 10680

gcccatgagt gcggtcacgg tgctttctcc gattctcgca tcatcaacga cattactggc 10740

tgggttcttc actcttccct ccttgtcccc tacttcagct ggcaaatctc ccaccgaaag 10800

caccacaagg ccactggcaa catggagcgt gacatggtct tcgttccccg aacccgcgag 10860

cagcaggcta ctcgtctcgg aaagatgacc cacgagctcg ctcatcttac tgagnnnntc 10920

gtnggctggc ccaactacct catcaccaat gttaccggcc acaactacca cgagcgccag 10980

cgtgagggtc gcggcaaggg caagcataac ggcctcggcg gtggtgttaa ccacttcgat 11040

ccccgcagcc ctctgtacga gaacagtgac gctaagctca tcgtcctcag cgatattggt 11100

atcggtctga tggccactgc tctgtacttc ctcgttcaga agttcggttt ctacaacatg 11160

gccatctggt actttgttcc ctacctctgg gttaaccact ggctcgttgc catcaccttc 11220

ctccagcaca ccgaccctac ccttccccac tacaccaacg acgagtggaa cttcgtccgt 11280

ggtgccgctg ctaccattga ccgtgagatg ggcttcatcg gccgccacct tctccacggc 11340

atcatcgaga ctcatgtcct ccaccactac gtcagcagca tccccttcta caacgcggac 11400

gaggccaccg aggccattaa gcccatcatg ggcaagcact accgggctga tgtccaggat 11460

ggtcctcgtg gcttcatccg cgccatgtac cgcagtgcgc gtatgtgcca gtgggttgag 11520

cccagcgctg gtgccgaggg tgctggtaag ggtgttctgt tcttccgcaa ccgcaacaac 11580

gtgggcaccc cccccgctgt tatcaagccc gttgcttaag taggcgcggc cgcaagtgtg 11640

gatggggaag tgagtgcccg gttctgtgtg cacaattggc aatccaagat ggatggattc 11700

aacacaggga tatagcgagc tacgtggtgg tgcgaggata tagcaacgga tatttatgtt 11760

tgacacttga gaatgtacga tacaagcact gtccaagtac aatactaaac atactgtaca 11820

tactcatact cgtacccggg caacggtttc acttgagtgc agtggctagt gctcttactc 11880

gtacagtgtg caatactgcg tatcatagtc tttgatgtat atcgtattca ttcatgttag 11940

ttgcgtacga agtcgtcaat gatgtcgata tgggttttga tcatgcacac ataaggtccg 12000

accttatcgg caagctcaat gagctccttg gtggtggtaa catccagaga agcacacagg 12060

ttggttttct tggctgccac gagcttgagc actcgagcgg caaaggcgga cttgtggacg 12120

ttagctcgag cttcgtagga gggcattttg gtggtgaaga ggagactgaa ataaatttag 12180

tctgcagaac tttttatcgg aaccttatct ggggcagtga agtatatgtt atggtaatag 12240

ttacgagtta gttgaactta tagatagact ggactatacg gctatcggtc caaattagaa 12300

agaacgtcaa tggctctctg ggcgtcgcct ttgccgacaa aaatgtgatc atgatgaaag 12360

ccagcaatga cgttgcagct gatattgttg tcggccaacc gcgccgaaaa cgcagctgtc 12420

agacccacag cctccaacga agaatgtatc gtcaaagtga tccaagcaca ctcatagttg 12480

gagtcgtact ccaaaggcgg caatgacgag tcagacagat actcgtcgac cttttccttg 12540

ggaaccacca ccgtcagccc ttctgactca cgtattgtag ccaccgacac aggcaacagt 12600

ccgtggatag cagaatatgt cttgtcggtc catttctcac caactttagg cgtcaagtga 12660

atgttgcaga agaagtatgt gccttcattg agaatcggtg ttgctgattt caataaagtc 12720

ttgagatcag tttggcgcgc cagctgcatt aatgaatcgg ccaacgcgcg gggagaggcg 12780

gtttgcgtat tgggcgctct tccgcttcct cgctcactga ctcgctgcgc tcggtcgttc 12840

ggctgcggcg agcggtatca gctcactcaa aggcggtaat acggttatcc acagaatcag 12900

gggataacgc aggaaagaac atgtgagcaa aaggccagca aaaggccagg aaccgtaaaa 12960

aggccgcgtt gctggcgttt ttccataggc tccgcccccc tgacgagcat cacaaaaatc 13020

gacgctcaag tcagaggtgg cgaaacccga caggactata aagataccag gcgtttcccc 13080

ctggaagctc cctcgtgcgc tctcctgttc cgaccctgcc gcttaccgga tacctgtccg 13140

cctttctccc ttcgggaagc gtggcgcttt ctcatagctc acgctgtagg tatctcagtt 13200

cggtgtaggt cgttcgctcc aagctgggct gtgtgcacga acccccccgtt cagcccgacc 13260

gctgcgcctt atccggtaac tatcgtcttg agtcc 13295

<210> 174
<211> 15559
<212> DNA
<213> Artificial Sequence

<220>
<223> Plasmid pDMW325

<400> 174
catggctgag gataagacca aggtcgagtt ccctaccctg actgagctga agcactctat 60

ccctaacgct tgctttgagt ccaacctcgg actctcgctc tactacactg cccgagcgat 120

cttcaacgca tctgcctctg ctgctctgct ctacgctgcc cgatctactc ccttcattgc 180

cgataacgtt ctgctccacg ctctggtttg cgccacctac atctacgtgc agggtgtcat 240

cttctggggt ttctttaccg tcggtcacga ctgtggtcac tctgccttct cccgatacca 300

ctccgtcaac ttcatcattg gctgcatcat gcactctgcc attctgactc ccttcgagtc 360

ctggcgagtg acccaccgac accatcacaa gaacactggc aacattgata aggacgagat 420

cttctaccct catcggtccg tcaaggacct ccaggacgtg cgacaatggg tctacaccct 480

cggaggtgct tggtttgtct acctgaaggt cggatatgct cctcgaacca tgtcccactt 540

tgacccctgg gaccctctcc tgcttcgacg agcctccgct gtcatcgtgt ccctcggagt 600

ctgggctgcc ttcttcgctg cctacgccta cctcacatac tcgctcggct ttgccgtcat 660

gggcctctac tactatgctc ctctctttgt ctttgcttcg ttcctcgtca ttactacctt 720

cttgcatcac aacgacgaag ctactccctg gtacggtgac tcggagtgga cctacgtcaa 780

gggcaacctg agctccgtcg accgatcgta cggagctttc gtggacaacc tgtctcacca 840

cattggcacc caccaggtcc atcacttgtt ccctatcatt ccccactaca agctcaacga 900

agccaccaag cactttgctg ccgcttaccc tcacctcgtg agacgtaacg acgagcccat 960

cattactgcc ttcttcaaga ccgctcacct ctttgtcaac tacggagctg tgcccgagac 1020

tgctcagatt ttcaccctca aagagtctgc cgctgcagcc aaggccaaga gcgactaagc 1080

ggccgcattg atgattggaa acacacacat gggttatatc taggtgagag ttagttggac 1140

agttatatat taaatcagct atgccaacgg taacttcatt catgtcaacg aggaaccagt 1200

gactgcaagt aatatagaat ttgaccacct tgccattctc ttgcactcct ttactatatc 1260

tcatttattt cttatataca aatcacttct tcttcccagc atcgagctcg gaaacctcat 1320

gagcaataac atcgtggatc tcgtcaatag agggcttttt ggactccttg ctgttggcca 1380

ccttgtcctt gctgtctggc tcattctgtt tcaacgcctt ttaattaacg gagtaggtct 1440

cggtgtcgga agcgacgcca gatccgtcat cctcctttcg ctctccaaag tagatacctc 1500

cgacgagctc tcggacaatg atgaagtcgg tgccctcaac gtttcggatg ggggagagat 1560

cggcgagctt gggcgacagc agctggcagg gtcgcaggtt ggcgtacagg ttcaggtcct 1620

ttcgcagctt gaggagaccc tgctcgggtc gcacgtcggt tcgtccgtcg ggagtggtcc 1680

atacggtgtt ggcagcgcct ccgacagcac cgagcataat agagtcagcc tttcggcaga 1740

tgtcgagagt agcgtcggtg atgggctcgc cctccttctc aatggcagct cctccaatga 1800

gtcggtcctc aaacacaaac tcggtgccgg aggcctcagc aacagacttg agcaccttga 1860

cggcctcggc aatcacctcg gggccacaga agtcgccgcc gagaagaaca atcttcttgg 1920

agtcagtctt ggtcttctta gtttcgggtt ccattgtgga tgtgtgtggt tgtatgtgtg 1980

atgtggtgtg tggagtgaaa atctgtggct ggcaaacgct cttgtatata tacgcacttt 2040

tgcccgtgct atgtggaaga ctaaacctcc gaagattgtg actcaggtag tgcggtatcg 2100

gctagggacc caaaccttgt cgatgccgat agcatgcgac gtcgggccca attcgcccta 2160

tagtgagtcg tattacaatt cactggccgt cgttttacaa cgtcgtgact gggaaaaccc 2220

tggcgttacc caacttaatc gccttgcagc acatccccct ttcgccagct ggcgtaatag 2280

cgaagaggcc cgcaccgatc gcccttccca acagttgcgc agcctgaatg gcgaatggac 2340

gcgccctgta gcggcgcatt aagcgcggcg ggtgtggtgg ttacgcgcag cgtgaccgct 2400

acacttgcca gcgccctagc gcccgctcct ttcgctttct tcccttcctt tctcgccacg 2460

ttcgccggct ttccccgtca agctctaaat cgggggctcc ctttagggtt ccgatttagt 2520

gctttacggc acctcgaccc caaaaaactt gattagggtg atggttcacg tagtgggcca 2580

tcgccctgat agacggtttt tcgccctttg acgttggagt ccacgttctt taatagtgga 2640

ctcttgttcc aaactggaac aacactcaac cctatctcgg tctattcttt tgatttataa 2700

gggattttgc cgatttcggc ctattggtta aaaaatgagc tgatttaaca aaaatttaac 2760

gcgaatttta acaaaatatt aacgcttaca atttcctgat gcggtatttt ctccttacgc 2820

atctgtgcgg tatttcacac cgcatcaggt ggcacttttc ggggaaatgt gcgcggaacc 2880

cctatttgtt tatttttcta aatacattca aatatgtatc cgctcatgag acaataaccc 2940

tgataaatgc ttcaataata ttgaaaaagg aagagtatga gtattcaaca tttccgtgtc 3000

gcccttattc cctttttgc ggcattttgc cttcctgttt ttgctcaccc agaaacgctg 3060

gtgaaagtaa aagatgctga agatcagttg ggtgcacgag tgggttacat cgaactggat 3120

ctcaacagcg gtaagatcct tgagagtttt cgccccgaag aacgttttcc aatgatgagc 3180

acttttaaag ttctgctatg tggcgcggta ttatcccgta ttgacgccgg gcaagagcaa 3240

ctcggtcgcc gcatacacta ttctcagaat gacttggttg agtactcacc agtcacagaa 3300

aagcatctta cggatggcat gacagtaaga gaattatgca gtgctgccat aaccatgagt 3360

gataacactg cggccaactt acttctgaca acgatcggag gaccgaagga gctaaccgct 3420

tttttgcaca acatgggggа tcatgtaact cgccttgatc gttgggaacc ggagctgaat 3480

gaagccatac caaacgacga gcgtgacacc acgatgcctg tagcaatggc aacaacgttg 3540

cgcaaactat taactggcga actacttact ctagcttccc ggcaacaatt aatagactgg 3600

atggaggcgg ataaagttgc aggaccactt ctgcgctcgg cccttccggc tggctggttt 3660

attgctgata aatctggagc cggtgagcgt gggtctcgcg gtatcattgc agcactgggg 3720

ccagatggta agccctcccg tatcgtagtt atctacacga cggggagtca ggcaactatg 3780

gatgaacgaa atagacagat cgctgagata ggtgcctcac tgattaagca ttggtaactg 3840

tcagaccaag tttactcata tatactttag attgatttaa aacttcattt ttaatttaaa 3900

aggatctagg tgaagatcct ttttgataat ctcatgacca aaatccctta acgtgagttt 3960

tcgttccact gagcgtcaga ccccgtagaa aagatcaaag gatcttcttg agatcctttt 4020

tttctgcgcg taatctgctg cttgcaaaca aaaaaaccac cgctaccagc ggtggtttgt 4080

ttgccggatc aagagctacc aactcttttt ccgaaggtaa ctggcttcag cagagcgcag 4140

ataccaaata ctgttcttct agtgtagccg tagttaggcc accacttcaa gaactctgta 4200

gcaccgccta catacctcgc tctgctaatc ctgttaccag tggctgctgc cagtggcgat 4260

aagtcgtgtc ttaccgggtt ggactcaaga cgatagttac cggataaggc gcagcggtcg 4320

ggctgaacgg ggggttcgtg cacacagccc agcttggagc gaacgaccta caccgaactg 4380

agatacctac agcgtgagct atgagaaagc gccacgcttc ccgaagggag aaaggcggac 4440

aggtatccgg taagcggcag ggtcggaaca ggagagcgca cgagggagct tccagggga 4500

aacgcctggt atctttatag tcctgtcggg tttcgccacc tctgacttga gcgtcgattt 4560

ttgtgatgct cgtcaggggg gcggagccta tggaaaaacg ccagcaacgc ggccttttta 4620

cggttcctgg ccttttgctg gccttttgct cacatgttct ttcctgcgtt atcccctgat 4680

tctgtggata accgtattac cgcctttgag tgagctgata ccgctcgccg cagccgaacg 4740

accgagcgca gcgagtcagt gagcgaggaa gcggaagagc gcccaatacg caaaccgcct 4800

ctccccgcgc gttggccgat tcattaatgc agctggcgcg cccactgagc tcgtctaacg 4860

gacttgatat acaaccaatt aaaacaaatg aaaagaaata cagttctttg tatcatttgt 4920

aacaattacc ctgtacaaac taaggtattg aaatcccaca atattcccaa agtccacccc 4980

tttccaaatt gtcatgccta caactcatat accaagcact aacctaccaa acaccactaa 5040

aaccccacaa aatatatctt accgaatata cagtaacaag ctaccaccac actcgttggg 5100

tgcagtcgcc agcttaaaga tatctatcca catcagccac aactcccttc ctttaataaa 5160

ccgactacac ccttggctat tgaggttatg agtgaatata ctgtagacaa gacactttca 5220

agaagactgt ttccaaaacg taccactgtc ctccactaca aacacaccca atctgcttct 5280

tctagtcaag gttgctacac cggtaaatta taaatcatca tttcattagc agggcagggc 5340

ccttttttata gagtcttata cactagcgga ccctgccggt agaccaaccc gcaggcgcgt 5400

cagtttgctc cttccatcaa tgcgtcgtag aaacgactta ctccttcttg agcagctcct 5460

tgaccttgtt ggcaacaagt ctccgacctc ggaggtggag gaagagcctc cgatatcggc 5520

ggtagtgata ccagcctcga cggactcctt gacggcagcc tcaacagcgt caccggcggg 5580

cttcatgtta agagagaact tgagcatcat ggcggcagac agaatggtgg cgtacgcaac 5640

taacatgaat gaatacgata tacatcaaag actatgatac gcagtattgc acactgtacg 5700

agtaagagca ctagccactg cactcaagtg aaaccgttgc ccgggtacga gtatgagtat 5760

gtacagtatg tttagtattg tacttggaca gtgcttgtat cgtacattct caagtgtcaa 5820

acataaatat ccgttgctat atcctcgcac caccacgtag ctcgctatat ccctgtgttg 5880

aatccatcca tcttggattg ccaattgtgc acacagaacc gggcactcac ttccccatcc 5940

acacttgcgg ccgcttagct gcctactctt ccttgggacg gagtccaaga acacgcaagt 6000

gctccaaatg tgaagcaaat gcttgccaaa acgtatcctt gacaaggtat ggaaccttgt 6060

actcgctgca ggtgttcttg atgatggcca gaatatcggg ataatggtgc tgcgacacgt 6120

tggggaacag atggtgcaca gcctggtagt tcaagctgcc agtgatgctg gtccagaggt 6180

gcgaatcgtg tgcgtaatcc tgcgtagtct cgacctgcat agctgcccag tccttttgga 6240

tgatcccgtt ctcgtcaggc aacggccact gaacttcctc aacaacgtgg ttcgcctgga 6300

aggtcagcgc cagccagtaa gacgacacca tgtccgcgac cgtgaacaag agcagcacct 6360

tgcccagggg cagatactgc aggggaacaa tcaggcgata ccagacaaag aaagccttgc 6420

cgccccagaa catcacagtg tgccatgtcg agatgggatt gacacgaata gcgtcattgg 6480

tcttgacaaa gtacaaaatg ttgatgtcct gaatgcgcac cttgaacgcc agcagtccgt 6540

acaggaaagg aacaaacatg tgctggttga tgtggttgac aaaccacttt tggttgggct 6600

tgatacgacg aacatcgggc tcagacgtcg acacgtcggg atctgctcca gcaatgttgg 6660

tgtaggggtg atggccgagc atatgttggt acatccacac caggtacgat gctccgttga 6720

aaaagtcgtg cgtggctccc agaatcttcc agacagtggg gttgtgggtc actgaaaagt 6780

gagacgcatc atgaagaggg ttgagtccga cttgtgcgca cgcaaatccc atgatgattg 6840

caaacaccac ctgaagccat gtgcgttcga caacgaaagg cacaaagagc tgcgcgtagt 6900

aggaagcgat caaggatcca aagataagag cgtatcgtcc ccagatctct ggtctattct 6960

tgggatcaat gttccgatcc gtaaagtagc cctcgactct cgtcttgatg gttttgtgga 7020

acaccgttgg ctccgggaag atgggcagct cattcgagac cagtgtaccg acatagtact 7080

tcttcataat ggcatctgca gccccaaacg cgtgatacat ctcaaagacc ggagtaacat 7140

ctcggccagc tccgagcagg agagtgtcca ctccaccagg atggcggctc aagaactttg 7200

tgacatcgta caccctgccg cggatggcca agagtaggtc gtccttggtg ttatgggccg 7260

ccagctcttc ccaggtgaag gtttttcctt ggtccgttcc catggagagc tgggttagtt 7320

tgtgtagaga gtgtgtgttg ctagcgactt tcggattgtg tcattacaca aaacgcgtcg 7380

tctcgacact gatcttgtcg tggatactca cggctcggac atcgtcgccg acgatgacac 7440

cggactttcg cttaaggacg tcagtaacag gcattgtgtg atgtgtagtt tagatttcga 7500

atctgtgggg aaagaaagga aaaaagagac tggcaaccga ttgggagagc cactgtttat 7560

atatacccta gacaagcccc ccgcttgtaa gatgttggtc aatgtaaacc agtattaagg 7620

ttggcaagtg caggagaagc aaggtgtggg taccgagcaa tggaaatgtg cggaaggcaa 7680

aaaaatgagg ccacggccta ttgtcggggc tatatccagg gggcgattga agtacactaa 7740

catgacatgt gtccacagac cctcaatctg gcctgatgag ccaaatccat acgcgctttc 7800

gcagctctaa aggctataac aagtcacacc accctgctcg acctcagcgc cctcactttt 7860

tgttaagaca aactgtacac gctgttccag cgttttctgc ctgcacctgg tgggacattt 7920

ggtgcaacct aaagtgctcg gaacctctgt ggtgtccaga tcagcgcagc agttccgagg 7980

tagttttgag gcccttagat gatgcaatgg tgtcagtcgc tggatcacga gtcttaatgg 8040

cagtattcgt tcttatttgt gccattgagc cccgttatcc tcgtatcttc tacccccccat 8100

cccatccctt tgttggtgca accctaccca tttattgttg ggtgcagccc aaccgacgtg 8160

gagagcttgg cttggccata taaaaaggcc ccccccctagt ggcaatggca gaaagtcagc 8220

tgtgagttgt tgaatttgtc atctaggcgg cctggccgtc ttctccgggg caatttaaac 8280

agtgtacgca gatcccgtca acagttttat atatcgtagt tacaaccatc aacacttttt 8340

ggtaagtgta ccattctata ctccaactgg tctgcaactg tacaagtaga catgttaatg 8400

gtagttaata acatctacag cagaacctat ggtaaagaca ttgcattttt acaggaagta 8460

tcgtcctaca cgttgataaa tccaaagatg cggaacttct tccactttta tcatcatccc 8520

ctactcgtac actcgtactc tttgttcgat cgcgattcat ttctataaat aatcttgtat 8580

gtacatgcgg ccgcttagtg agtatgctcg tggtaggagt gctcagggtt gccaacagca 8640

aggaggttga cgtaggtcac tcgaaagcat tccagatagg gtcggacgtc gtactcgact 8700

ccgtgcttct cgaagagtgc tcgcactcga ggagcaatct tggggtggtt gtactgaggc 8760

atggagggga acagatggtg ctcgatctgg aagttgaggt aggccatcca ccagttgaca 8820

aagggagagt cggaacagtt ggtagtgtgg ttggcagagt acagagccca cgaaatgtgc 8880

ttggtgggag gaacgacatc cttgtgagta tgggagacgg cgaagttggt gaagatgtag 8940

gtgccaccga cagccacgta gaaggcatac agtcgaagag aggacaggag tcccagatgg 9000

ccaaaggcaa ggtgccagag aacgtatcgg agacccatgt aggccagttc gagatggttg 9060

tgcactcgca gtgcatgatt ggggtggaga acgaactgcc agccgaaaga gaccagggag 9120

caaatgatgc ctccaaagaa gaggggagct tgccaagcaa gccaggcttt acccttgcct 9180

cgtgccttgg ctccgacaat cttgtggaag gcaacgagag gcatggtctg caggtcggga 9240

tcggttccga gcttctgagg agtggcgtga tgtttgttgt gctggtttct ccagtagcag 9300

ccagacatac cacagcccag tccatagatg gccatctgga gatgtcggtc aatcttgatg 9360

ttgccagtga gcgagtagtg tccgccttca tgctggagcc atccgcatcg accttgagca 9420

attccagcaa cgatggcacc gaggaaccag tatccggacc agatcagctg gacgccagcc 9480

cagtacatgg caagcacctc gacacatcgg taggacacgt gcaccagatt gggctcgaag 9540

taaccttcct ctttcagttc cagtcgcagc ttctcaaagt ctctagagat gtcgtcctgt 9600

gcatagccag cctcggcagg tcgggaggga agagagtaca gcatcttgtc agccttcttc 9660

gatcggacgt ggaagttgtt gtaagcgtca gaagcatctg ctccgagctg gaacttgatg 9720

acggagccac cagggtgtcg cttcacgaac gaggagacgt cgtacactcg tccgttgata 9780

atcatctcct tctcggcgtc gatagacgag acggtgacat cggcagcctc agatcgggag 9840

tcgttggtag caccgtggta ctctcgaggg agagtggaag agtgggtcac gtgggctgca 9900

gctccggact tgccagcgac catggagagc tgggttagtt tgtgtagaga gtgtgtgttg 9960

ctagcgactt tcggattgtg tcattacaca aaacgcgtcg tctcgacact gatcttgtcg 10020

tggatactca cggctcggac atcgtcgccg acgatgacac cggactttcg cttaaggacg 10080

tcagtaacag ccattgttgt atgtgttttg gtgatgtcac gaaaaggacg acgcagcata 10140

gatatatatg gacgtggaaa ctaggagaag gaaccttatt gctgatccgt aaaattatcg 10200

tggaaagagg tggggaggga agatatgggt atggggagat gcaggagcga caaaacacgc 10260

aaatcggagc cgttaacctc acaaatatcg ccaattcctc ccccacagct ccccaaacag 10320

cccgttgagc gcttttgccc aaccctctcc agtccaaaaa aaaccctatt gtctggtgac 10380

gtctcctacc catgcacctc gtggttgctg tataactggg gactccacga cccgaggtca 10440

attaggccac gtgacaccca attgtgtcgc tcgcacctca attgaagact aagctcgttc 10500

cgccacctca attgaaccca tatgagcacc aatcgccctc accgcaacat aaaacgctcc 10560

tcatagacta gtgctatcca aacgctatta tgacgtgacg ctgaataggg ttgcgcttaa 10620

acgaggacag ttgtcttatt ttggggcaag gttagggcat cgaggcagtc tgtaactgtc 10680

agtaattgca ttgtatcttc ccctttgtca cggacttgcc cctgacactg gttggttgtg 10740

tcacatttgt cctcaaacaa agtggttcac gccttgattt gatctgaacc caaaacctga 10800

gtgagcacat ggagaaggtt aaaaacaaac cgctctcaac ggatgcaacc caaaccctag 10860

tgaactgtga ctgcgaattc tacaatacgt gagtcagaag ggctgacggt ggtggttccc 10920

aaggaaaagg tcgacgagta tctgtctgac tcgtcattgc cgcctttgga gtacgactcc 10980

aactatgagt gtgcttggat cactttgacg atacattctt cgttggaggc tgtgggtctg 11040

acagctgcgt tttcggcgcg gttggccgac aacaatatca gctgcaacgt cattgctggc 11100

tttcatcatg atcacatttt tgtcggcaaa ggcgacgccc agagagccat tgacgttctt 11160

tctaatttgg accgatagcc gtatagtcca gtctatctat aagttcaact aactcgtaac 11220

tattaccata acatatactt cactgcccca gataaggttc cgataaaaag ttctgcagac 11280

taaatttatt tcagtctcct cttcaccacc aaaatgccct cctacgaagc tcgagctaac 11340

gtccacaagt ccgcctttgc cgctcgagtg ctcaagctcg tggcagccaa gaaaaccaac 11400

ctgtgtgctt ctctggatgt taccaccacc aaggagctca ttgagcttgc cgataaggtc 11460

ggaccttatg tgtgcatgat caaaacccat atcgacatca ttgacgactt cacctacgcc 11520

ggcactgtgc tccccctcaa ggaacttgct cttaagcacg gtttcttcct gttcgaggac 11580

agaaagttcg cagatattgg caacactgtc aagcaccagt accggtgtca ccgaatcgcc 11640

gagtggtccg atatcaccaa cgcccacggt gtacccggaa ccggaatcat tgctggcctg 11700

cgagctggtg ccgaggaaac tgtctctgaa cagaagaagg aggacgtctc tgactacgag 11760

aactcccagt acaaggagtt cctagtcccc tctcccaacg agaagctggc cagaggtctg 11820

ctcatgctgg ccgagctgtc ttgcaagggc tctctggcca ctggcgagta ctccaagcag 11880

accattgagc ttgcccgatc cgaccccgag tttgtggttg gcttcattgc ccagaaccga 11940

cctaagggcg actctgagga ctggcttatt ctgacccccg gggtgggtct tgacgacaag 12000

ggagacgctc tcggacagca gtaccgaact gttgaggatg tcatgtctac cggaacggat 12060

atcataattg tcggccgagg tctgtacggc cagaaccgag atcctattga ggaggccaag 12120

cgataccaga aggctggctg ggaggcttac cagaagatta actgttagag gttagactat 12180

ggatatgtaa tttaactgtg tatatagaga gcgtgcaagt atggagcgct tgttcagctt 12240

gtatgatggt cagacgacct gtctgatcga gtatgtatga tactgcacaa cctgtgtatc 12300

cgcatgatct gtccaatggg gcatgttgtt gtgtttctcg atacggagat gctgggtaca 12360

gtgctaatac gttgaactac ttatacttat atgaggctcg aagaaagctg acttgtgtat 12420

gacttattct caactacatc cccagtcaca ataccaccac tgcactacca ctacaccaga 12480

tctgcgtaca ctgtttaaac agagtgtgaa agactcacta tggtccgggc ttatctcgac 12540

caatagccaa agtctggagt ttctgagaga aaaaggcaag atacgtatgt aacaaagcga 12600

cgcatggtac aataataccg gaggcatgta tcatagagag ttagtggttc gatgatggca 12660

ctggtgcctg gtatgacttt atacggctga ctacatattt gtcctcagac atacaattac 12720

agtcaagcac ttacccttgg acatctgtag gtacccccg gccaagacga tctcagcgtg 12780

tcgtatgtcg gattggcgta gctccctcgc tcgtcaattg gctcccatct actttcttct 12840

gcttggctac acccagcatg tctgctatgg ctcgttttcg tgccttatct atcctcccag 12900

tattaccaac tctaaatgac atgatgtgat tgggtctaca ctttcatatc agagataagg 12960

agtagcacag ttgcataaaa agcccaactc taatcagctt cttcctttct tgtaattagt 13020

acaaaggtga ttagcgaaat ctggaagctt agttggccct aaaaaaatca aaaaaagcaa 13080

aaaacgaaaa acgaaaaacc acagttttga gaacagggag gtaacgaagg atcgtatata 13140

tatatatata tatatatacc cacggatccc gagaccggcc tttgattctt ccctacaacc 13200

aaccattctc accaccctaa ttcacaacca tggctgagga taagaccaag gtcgagttcc 13260

ctaccctgac tgagctgaag cactctatcc ctaacgcttg ctttgagtcc aacctcggac 13320

tctcgctcta ctacactgcc cgagcgatct tcaacgcatc tgcctctgct gctctgctct 13380

acgctgcccg atctactccc ttcattgccg ataacgttct gctccacgct ctggtttgcg 13440

ccacctacat ctacgtgcag ggtgtcatct tctggggttt ctttaccgtc ggtcacgact 13500

gtggtcactc tgccttctcc cgataccact ccgtcaactt catcattggc tgcatcatgc 13560

actctgccat tctgactccc ttcgagtcct ggcgagtgac ccaccgacac catcacaaga 13620

acactggcaa cattgataag gacgagatct tctaccctca tcggtccgtc aaggacctcc 13680

aggacgtgcg acaatgggtc tacaccctcg gaggtgcttg gtttgtctac ctgaaggtcg 13740

gatatgctcc tcgaaccatg tcccactttg acccctggga ccctctcctg cttcgacgag 13800

cctccgctgt catcgtgtcc ctcggagtct gggctgcctt cttcgctgcc tacgcctacc 13860

tcacatactc gctcggcttt gccgtcatgg gcctctacta ctatgctcct ctctttgtct 13920

ttgcttcgtt cctcgtcatt actaccttct tgcatcacaa cgacgaagct actccctggt 13980

acggtgactc ggagtggacc tacgtcaagg gcaacctgag ctccgtcgac cgatcgtacg 14040

gagctttcgt ggacaacctg tctcaccaca ttggcaccca ccaggtccat cacttgttcc 14100

ctatcattcc ccactacaag ctcaacgaag ccaccaagca ctttgctgcc gcttaccctc 14160

acctcgtgag acgtaacgac gagcccatca ttactgcctt cttcaagacc gctcacctct 14220

ttgtcaacta cggagctgtg cccgagactg ctcagatttt caccctcaaa gagtctgccg 14280

ctgcagccaa ggccaagagc gactaagcgg ccgctattta tcactcttta caacttctac 14340

ctcaactatc tactttaata aatgaatatc gtttattctc tatgattact gtatatgcgt 14400

tcctctaaga caaatcgaaa ccagcatgtg atcgaatggc atacaaaagt ttcttccgaa 14460

gttgatcaat gtcctgatag tcaggcagct tgagaagatt gacacaggtg gaggccgtag 14520

ggaaccgatc aacctgtcta ccagcgttac gaatggcaaa tgacgggttc aaagccttga 14580

atccttgcaa tggtgccttg gatactgatg tcacaaactt aagaagcagc cgcttgtcct 14640

cttcctcgat cgatggaagc cggtagaacc gggctgcttg tgcttggaga tggaagccgg 14700

tagaaccggg ctgcttgggg ggatttgggg ccgctgggct ccaaagaggg gtaggcattt 14760

cgttggggtt acgtaattgc ggcatttggg tcctgcgcgc atgtcccatt ggtcagaatt 14820

agtccggata ggagacttat cagccaatca cagcgccgga tccacctgta ggttgggttg 14880

ggtgggagca cccctccaca gagtagagtc aaacagcagc agcaacatga tagttggggg 14940

tgtgcgtgtt aaaggaaaaa aaagaagctt gggttatatt cccgctctat ttagaggttg 15000

cgggatagac gccgacggag ggcaatggcg ctatggaacc ttgcggatat ccatacgccg 15060

cggcggactg cgtccgaacc agctccagca gcgtttttttc cgggccattg agccgactgc 15120

gaccccgcca acgtgtcttg gcccacgcac tcatgtcatg ttggtgttgg gaggccactt 15180

tttaagtagc acaaggcacc tagctcgcag caaggtgtcc gaaccaaaga agcggctgca 15240

gtggtgcaaa cggggcggaa acggcgggaa aaagccacgg gggcacgaat tgaggcacgc 15300

cctcgaattt gagacgagtc acggccccat tcgcccgcgc aatggctcgc caacgcccgg 15360

tcttttgcac cacatcaggt taccccaagc caaacctttg tgttaaaaag cttaacatat 15420

tataccgaac gtaggtttgg gcgggcttgc tccgtctgtc caaggcaaca tttatataag 15480

ggtctgcatc gccggctcaa ttgaatcttt tttcttcttc tcttctctat attcattctt 15540

gaattaaaca cacatcaac 15559

<210> 175
<211> 16325
<212> DNA
<213> Artificial Sequence

<220>
<223> Plasmid pZKSL5598

<400> 175
cgatgagcct aaaatgaacc cgagtatatc tcataaaatt ctcggtgaga ggtctgtgac 60

tgtcagtaca aggtgccttc attatgccct caaccttacc atacctcact gaatgtagtg 120

tacctctaaa aatgaaatac agtgccaaaa gccaaggcac tgagctcgtc taacggactt 180

gatatacaac caattaaaac aaatgaaaag aaatacagtt ctttgtatca tttgtaacaa 240

ttaccctgta caaactaagg tattgaaatc ccacaatatt cccaaagtcc acccctttcc 300

aaattgtcat gcctacaact catataccaa gcactaacct accgtttaaa cagtgtacgc 360

agatcccgtc aacagtttta tatatcgtag ttacaaccat caacactttt tggtaagtgt 420

accattctat actccaactg gtctgcaact gtacaagtag acatgttaat ggtagttaat 480

aacatctaca gcagaaccta tggtaaagac attgcatttt tacaggaagt atcgtcctac 540

acgttgataa atccaaagat gcggaacttc ttccactttt atcatcatcc cctactcgta 600

cactcgtact ctttgttcga tcgcgattca tttctataaa taatcttgta tgtacatgcg 660

gccgctaaag ctgcttacca gccttagcgg atttcttggt ggccaggttg tcctggtaaa 720

agaagtgaca gaacaggaga aagacagatc cgacgtaggc gtagttgaaa gcccaggaga 780

acagcttgcc cttgtcagag ttgaagcagg gaacgttgat gtagtcccag accaggagaa 840

agccaccgac gaactggcaa atctgcatgg cagtgatcag aggcttggcc ttgaacttgt 900

agccagcggc agtcagtcca tagtaggtgt acatgatggt gtgaatgaac gagttaaaga 960

acatgaagat ccacacaccc tcgttgtgca gtcgaatgcc gaggtagacg tcccagggag 1020

ctccaaagtg atggaaggcc tgcagaaagg acactcgctt gcccttgagg accagccaag 1080

cggtgtcgag gtactccacg tacttagaat agtagaaggc cttggcagtc caggtgaaca 1140

gcttggagtc ccagacagga gagggacact gaaagagagg ctggggagta tcaccggtct 1200

gtcttcgcag ccaggctcca gtaccgtagt cccagccgag agcggtggca gtcacgtaga 1260

aggacagggc agagaagaga gccaggagga cgttgtacca gatcatggag gttcggtagg 1320

ctcctttctt ctcgtccacg agaccagagt ttcgcaggag aggcttcagg agcaggtagg 1380

agaaggtgcc aatgaggatt tcgggatcgg tgacggcagc ccagattcgc tcgccagcgt 1440

cgttggccag agccatggta ccagagctgg gttagtttgt gtagagagtg tgtgttgcta 1500

gcgactttcg gattgtgtca ttacacaaaa cgcgtcgtct cgacactgat cttgtcgtgg 1560

atactcacgg ctcggaattc tgtgatgtgt agtttagatt tcgaatctgt ggggaaagaa 1620

aggaaaaaag agactggcaa ccgattggga gagccactgt ttatatatac cctagacaag 1680

ccccccgctt gtaagatgtt ggtcaatgta aaccagtatt aaggttggca agtgcaggag 1740

aagcaaggtg tgggtaccga gcaatggaaa tgtgcggaag gcaaaaaaat gaggccacgg 1800

cctattgtcg gggctatatc caggggggcga ttgaagtaca ctaacatgac atgtgtccac 1860

agaccctcaa tctggcctga tgagccaaat ccatacgcgc tttcgcagct ctaaaggcta 1920

taacaagtca caccaccctg ctcgacctca gcgccctcac tttttgttaa gacaaactgt 1980

acacgctgtt ccagcgtttt ctgcctgcac ctggtgggac atttggtgca acctaaagtg 2040

ctcggaacct ctgtggtgtc cagatcagcg cagcagttcc gaggtagttt tgaggccctt 2100

agatgatgca atggtgtcag tcgctggatc acgagtctta atggcagtat tcgttcttat 2160

ttgtgccatt gagccccgtt atcctcgtat cttctacccc ccatcccatc cctttgttgg 2220

tgcaacccta cccatttatt gttgggtgca gcccaaccga cgtggagagc ttggcttggc 2280

catataaaaa ggcccccccc tagtggcaat ggcagaaagt cagctgtgag ttgttgaatt 2340

tgtcatctag gcggcctggc cgtcttctcc ggggcaattt aaatgttcct ctatagtaga 2400

tctgcgtaca ctgtttaaac aagctaccac cacactcgtt gggtgcagtc gccagcttaa 2460

agatatctat ccacatcagc cacaactccc ttcctttaat aaaccgacta cacccttggc 2520

tattgaggtt atgagtgaat atactgtaga caagacactt tcaagaagac tgtttccaaa 2580

acgtaccact gtcctccact acaaacacac ccaatctgct tcttctagtc aaggttgcta 2640

caccggtaaa ttataaatca tcatttcatt agcagggcag ggccctttt atagagtctt 2700

atacactagc ggaccctgcc ggtagaccaa cccgcaggcg cgtcagtttg ctccttccat 2760

caatgcgtcg tagaaacgac ttactccttc ttgagcagct ccttgacctt gttggcaaca 2820

agtctccgac ctcggaggtg gaggaagagc ctccgatatc ggcggtagtg ataccagcct 2880

cgacggactc cttgacggca gcctcaacag cgtcaccggc gggcttcatg ttaagagaga 2940

acttgagcat catggcggca gacagaatgg tggcaatggg gttgaccttc tgcttgccga 3000

gatcgggggc agatccgtga cagggctcgt acagaccgaa cgcctcgttg gtgtcgggca 3060

gagaagccag agaggcggag ggcagcagac ccagagaacc ggggatgacg gaggcctcgt 3120

cggagatgat atcgccaaac atgttggtgg tgatgatgat accattcatc ttggagggct 3180

gcttgatgag gatcatggcg gccgagtcga tcagctggtg gttgagctcg agctggggga 3240

attcgtcctt gaggactcga gtgacagtct ttcgccaaag tcgagaggag gccagcacgt 3300

tggccttgtc aagagaccac acgggaagag gggggttgtg ctgaagggcc aggaaggcgg 3360

ccattcgggc aattcgctca acctcaggaa cggagtaggt ctcggtgtcg gaagcgacgc 3420

cagatccgtc atcctccttt cgctctccaa agtagatacc tccgacgagc tctcggacaa 3480

tgatgaagtc ggtgccctca acgtttcgga tgggggagag atcggcgagc ttgggcgaca 3540

gcagctggca gggtcgcagg ttggcgtaca ggttcaggtc ctttcgcagc ttgaggagac 3600

cctgctcggg tcgcacgtcg gttcgtccgt cgggagtggt ccatacggtg ttggcagcgc 3660

ctccgacagc accgagcata atagagtcag cctttcggca gatgtcgaga gtagcgtcgg 3720

tgatgggctc gccctccttc tcaatggcag ctcctccaat gagtcggtcc tcaaacacaa 3780

actcggtgcc ggaggcctca gcaacagact tgagcacctt gacggcctcg gcaatcacct 3840

cggggccaca gaagtcgccg ccgagaagaa caatcttctt ggagtcagtc ttggtcttct 3900

tagtttcggg ttccattgtg gatgtgtgtg gttgtatgtg tgatgtggtg tgtggagtga 3960

aaatctgtgg ctggcaaacg ctcttgtata tatacgcact tttgcccgtg ctatgtggaa 4020

gactaaacct ccgaagattg tgactcaggt agtgcggtat cggctaggga cccaaacctt 4080

gtcgatgccg atagcgctat cgaacgtacc ccagccggcc gggagtatgt cggaggggac 4140

atacgagatc gtcaagggtt tgtggccaac tggtaaataa atgatgactc aggcgacgac 4200

ggaattcgac agcaactact cctttcacca accatgtgca ttttagctcg aataacattc 4260

acaggcttgg tgatctacat ccatggtgtc tggccgatta ccgtggtgtt ttggcagtaa 4320

cgagaatatt gagtgaactc ttcccatcac caataaagac tcatactaca atcacgagcg 4380

cttcagctgc cactatagtg ttggtgacac aataccccctc gatgctgggc attactgtag 4440

caagagatat tcatttcatg gcgcattttc cagtctacct gacttttag tgccgatttc 4500

ttctccacat tttacgctca gtgtgaaaag ttggagtcga cttgcgccat tatgtaagca 4560

gcgtgtactc ctctgactgt ccatatggtt tgctccatct caccctcatc gttttcattg 4620

ttcacaggcg gccacaaaaa aactgtcttc tctccttctc tcttcgcctt agtctactcg 4680

gaccagtttt agtttagctt ggcgccactg gataaatgag acctcaggcc ttgtgatgag 4740

gaggtcactt atgaagcatg ttaggaggtg cttgtatgga tagagaagca cccaaaataa 4800

taagaataat aataaaacag ggggcgttgt catttcatat cgtgttttca ccatcaatac 4860

acctccaaac aatgcccttc atgtggccag ccccaatatt gtcctgtagt tcaactctat 4920

gcagctcgta tcttattgag caagtaaaac tctgtcagcc gatattgccc gacccgcgac 4980

aagggtcaac aaggtggtgt aaggccttcg cagaagtcaa aactgtgcca aacaaacatc 5040

tagagtctct ttggtgtttc tcgcatatat ttwatcggct gtcttacgta tttgcgcctc 5100

ggtaccggac taatttcgga tcatccccaa tacgcttttt cttcgcagct gtcaacagtg 5160

tccatgatct atccacctaa atgggtcata tgaggcgtat aatttcgtgg tgctgataat 5220

aattcccata tatttgacac aaaacttccc cccctagaca tacatctcac aatctcactt 5280

cttgtgcttc tgtcacacat ctcctccagc tgacttcaac tcacacctct gccccagttg 5340

gtctacagcg gtataaggtt tctccgcata gaggtgcacc actcctcccg atacttgttt 5400

gtgtgacttg tgggtcacga catatatatc tacacacatt gcgccaccct ttggttcttc 5460

cagcacaaca aaaacacgac acgctaacca tgggaacgga ccaaggaaaa accttcacct 5520

gggaagagct ggcggcccat aacaccaagg acgacctact cttggccatc cgcggcaggg 5580

tgtacgatgt cacaaagttc ttgagccgcc atcctggtgg agtggacact ctcctgctcg 5640

gagctggccg agatgttact ccggtctttg agatgtatca cgcgtttggg gctgcagatg 5700

ccattatgaa gaagtactat gtcggtacac tggtctcgaa tgagctgccc atcttcccgg 5760

agccaacggt gttccacaaa accatcaaga cgagagtcga gggctacttt acggatcgga 5820

acattgatcc caagaataga ccagagatct ggggacgata cgctcttatc tttggatcct 5880

tgatcgcttc ctactacgcg cagctctttg tgcctttcgt tgtcgaacgc acatggcttc 5940

aggtggtgtt tgcaatcatc atgggatttg cgtgcgcaca agtcggactc aaccctcttc 6000

atgatgcgtc tcacttttca gtgacccaca accccactgt ctggaagatt ctgggagcca 6060

cgcacgactt tttcaacgga gcatcgtacc tggtgtggat gtaccaacat atgctcggcc 6120

atcaccccta caccaacatt gctggagcag atcccgacgt gtcgacgtct gagcccgatg 6180

ttcgtcgtat caagcccaac caaaagtggt ttgtcaacca catcaaccag cacatgtttg 6240

ttcctttcct gtacggactg ctggcgttca aggtgcgcat tcaggacatc aacattttgt 6300

actttgtcaa gaccaatgac gctattcgtg tcaatcccat ctcgacatgg cacactgtga 6360

tgttctgggg cggcaaggct ttctttgtct ggtatcgcct gattgttccc ctgcagtatc 6420

tgcccctggg caaggtgctg ctcttgttca cggtcgcgga catggtgtcg tcttactggc 6480

tggcgctgac cttccaggcg aaccacgttg ttgaggaagt tcagtggccg ttgcctgacg 6540

agaacgggat catccaaaag gactgggcag ctatgcaggt cgagactacg caggattacg 6600

cacacgattc gcacctctgg accagcatca ctggcagctt gaactaccag gctgtgcacc 6660

atctgttccc caacgtgtcg cagcaccatt atcccgatat tctggccatc atcaagaaca 6720

cctgcagcga gtacaaggtt ccataccttg tcaaggatac gttttggcaa gcatttgctt 6780

cacatttgga gcacttgcgt gttcttggac tccgtcccaa ggaagagtag gcggccgcaa 6840

gtgtggatgg ggaagtgagt gcccggttct gtgtgcacaa ttggcaatcc aagatggatg 6900

gattcaacac agggatatag cgagctacgt ggtggtgcga ggatatagca acggatattt 6960

atgtttgaca cttgagaatg tacgatacaa gcactgtcca agtacaatac taaacatact 7020

gtacatactc atactcgtac ccgggcaacg gtttcacttg agtgcagtgg ctagtgctct 7080

tactcgtaca gtgtgcaata ctgcgtatca tagtctttga tgtatatcgt attcattcat 7140

gttagttgcg tacgttatag ttggagcaag ggagaaatgt agagtgtgaa agactcacta 7200

tggtccgggc ttatctcgac caatagccaa agtctggagt ttctgagaga aaaaggcaag 7260

atacgtatgt aacaaagcga cgcatggtac aataataccg gaggcatgta tcatagagag 7320

ttagtggttc gatgatggca ctggtgcctg gtatgacttt atacggctga ctacatattt 7380

gtcctcagac atacaattac agtcaagcac ttacccttgg acatctgtag gtacccccg 7440

gccaagacga tctcagcgtg tcgtatgtcg gattggcgta gctccctcgc tcgtcaattg 7500

gctcccatct actttcttct gcttggctac acccagcatg tctgctatgg ctcgttttcg 7560

tgccttatct atcctcccag tattaccaac tctaaatgac atgatgtgat tgggtctaca 7620

ctttcatatc agagataagg agtagcacag ttgcataaaa agcccaactc taatcagctt 7680

cttcctttct tgtaattagt acaaaggtga ttagcgaaat ctggaagctt agttggccct 7740

aaaaaaatca aaaaaagcaa aaaacgaaaa acgaaaaacc acagttttga gaacagggag 7800

gtaacgaagg atcgtatata tatatatata tatatatacc cacggatccc gagaccggcc 7860

tttgattctt ccctacaacc aaccattctc accaccctaa ttcacaacca tggtcgctgg 7920

caagtccgga gctgcagccc acgtgaccca ctcttccact ctccctcgag agtaccacgg 7980

tgctaccaac gactcccgat ctgaggctgc cgatgtcacc gtctcgtcta tcgacgccga 8040

gaaggagatg attatcaacg gacgagtgta cgacgtctcc tcgttcgtga agcgacaccc 8100

tggtggctcc gtcatcaagt tccagctcgg agcagatgct tctgacgctt acaacaactt 8160

ccacgtccga tcgaagaagg ctgacaagat gctgtactct cttccctccc gacctgccga 8220

ggctggctat gcacaggacg acatctctag agactttgag aagctgcgac tggaactgaa 8280

agaggaaggt tacttcgagc ccaatctggt gcacgtgtcc taccgatgtg tcgaggtgct 8340

tgccatgtac tgggctggcg tccagctgat ctggtccgga tactggttcc tcggtgccat 8400

cgttgctgga attgctcaag gtcgatgcgg atggctccag catgaaggcg gacactactc 8460

gctcactggc aacatcaaga ttgaccgaca tctccagatg gccatctatg gactgggctg 8520

tggtatgtct ggctgctact ggagaaacca gcacaacaaa catcacgcca ctcctcagaa 8580

gctcggaacc gatcccgacc tgcagaccat gcctctcgtt gccttccaca agattgtcgg 8640

agccaaggca cgaggcaagg gtaaagcctg gcttgcttgg caagctcccc tcttctttgg 8700

aggcatcatt tgctccctgg tctctttcgg ctggcagttc gttctccacc ccaatcatgc 8760

actgcgagtg cacaaccatc tcgaactggc ctacatgggt ctccgatacg ttctctggca 8820

ccttgccttt ggccatctgg gactcctgtc ctctcttcga ctgtatgcct tctacgtggc 8880

tgtcggtggc acctacatct tcaccaactt cgccgtctcc catactcaca aggatgtcgt 8940

tcctcccacc aagcacattt cgtgggctct gtactctgcc aaccacacta ccaactgttc 9000

cgactctccc tttgtcaact ggtggatggc ctacctcaac ttccagatcg agcaccatct 9060

gttcccctcc atgcctcagt acaaccaccc caagattgct cctcgagtgc gagcactctt 9120

cgagaagcac ggagtcgagt acgacgtccg accctatctg gaatgctttc gagtgaccta 9180

cgtcaacctc cttgctgttg gcaaccctga gcactcctac cacgagcata ctcactaagc 9240

ggccgcatga gaagataaat atataaatac attgagatat taaatgcgct agattagaga 9300

gcctcatact gctcggagag aagccaagac gagtactcaa aggggattac accatccata 9360

tccacagaca caagctgggg aaaggttcta tatacacttt ccggaatacc gtagtttccg 9420

atgttatcaa tgggggcagc caggatttca ggcacttcgg tgtctcgggg tgaaatggcg 9480

ttcttggcct ccatcaagtc gtaccatgtc ttcatttgcc tgtcaaagta aaacagaagc 9540

agatgaagaa tgaacttgaa gtgaaggaat ttaaatgtaa cgaaactgaa atttgaccag 9600

atattgtgtc cgcggtggag ctccagcttt tgttcccttt agtgagggtt aatttcgagc 9660

ttggcgtaat catggtcata gctgtttcct gtgtgaaatt gttatccgct cacaagcttg 9720

acaatggatg gtgtgtcgta tcaggctggg gaacaattgt gcttaagtat gctgcagttg 9780

agtaagagtc atcgctccac caaaataaag tttgccatta gggttggaga gagagatggt 9840

ggctggaaga attaaatgac atcaagctga ggattgtggg tgtgcaataa cacatgttag 9900

gggtgacctg tggctcgaaa tctgataatt attttgtaac tttatgatta ttcttagatt 9960

ttttaatatt cctctatata acacataagt agctgtcgtc tagttgttca tagcctgact 10020

cctgcaatag attagtgcag agtgattttg tgcaattgag agccacggtt gagtcaagtg 10080

actttgtgtg tgaagtcatc ttacgtttca agtctcacag gttactcaat tggttggttg 10140

tctgcccttt acagatattt acagtacctg agcgtaaagt cgttcatcca cggaatgact 10200

gttcctgtca cgcagtcatg atcatggatg tggctggtca ggaaccattt tggataggag 10260

acttagggat tggactatta ttgaaaaaac tgagccgaat atgatatagt tctatttgaa 10320

tgcagaactt ctgatggtca attcacttat ttcaggcata tcggtcatgg tggcagctgc 10380

cacgatgtta tctcgttgga aacctcggcg cgccagctgc attaatgaat cggccaacgc 10440

gcggggagag gcggtttgcg tattgggcgc tcttccgctt cctcgctcac tgactcgctg 10500

cgctcggtcg ttcggctgcg gcgagcggta tcagctcact caaaggcggt aatacggtta 10560

tccacagaat caggggataa cgcaggaaag aacatgtgag caaaaggcca gcaaaaggcc 10620

aggaaccgta aaaaggccgc gttgctggcg tttttccata ggctccgccc ccctgacgag 10680

catcacaaaa atcgacgctc aagtcagagg tggcgaaacc cgacaggact ataaagatac 10740

caggcgtttc cccctggaag ctccctcgtg cgctctcctg ttccgaccct gccgcttacc 10800

ggatacctgt ccgcctttct cccttcggga agcgtggcgc tttctcatag ctcacgctgt 10860

aggtatctca gttcggtgta ggtcgttcgc tccaagctgg gctgtgtgca cgaacccccc 10920

gttcagcccg accgctgcgc cttatccggt aactatcgtc ttgagtccaa cccggtaaga 10980

cacgacttat cgccactggc agcagccact ggtaacagga ttagcagagc gaggtatgta 11040

ggcggtgcta cagagttctt gaagtggtgg cctaactacg gctacactag aagaacagta 11100

tttggtatct gcgctctgct gaagccagtt accttcggaa aaagagttgg tagctcttga 11160

tccggcaaac aaaccaccgc tggtagcggt ggtttttttg tttgcaagca gcagattacg 11220

cgcagaaaaa aaggatctca agaagatcct ttgatctttt ctacggggtc tgacgctcag 11280

tggaacgaaa actcacgtta agggattttg gtcatgagat tatcaaaaag gatcttcacc 11340

tagatccttt taaattaaaa atgaagtttt aaatcaatct aaagtatata tgagtaaact 11400

tggtctgaca gttaccaatg cttaatcagt gaggcaccta tctcagcgat ctgtctattt 11460

cgttcatcca tagttgcctg actccccgtc gtgtagataa ctacgatacg ggagggctta 11520

ccatctggcc ccagtgctgc aatgataccg cgagacccac gctcaccggc tccagattta 11580

tcagcaataa accagccagc cggaagggcc gagcgcagaa gtggtcctgc aactttatcc 11640

gcctccatcc agtctattaa ttgttgccgg gaagctagag taagtagttc gccagttaat 11700

agtttgcgca acgttgttgc cattgctaca ggcatcgtgg tgtcacgctc gtcgtttggt 11760

atggcttcat tcagctccgg ttcccaacga tcaaggcgag ttacatgatc ccccatgttg 11820

tgcaaaaaag cggttagctc cttcggtcct ccgatcgttg tcagaagtaa gttggccgca 11880

gtgttatcac tcatggttat ggcagcactg cataattctc ttactgtcat gccatccgta 11940

agatgctttt ctgtgactgg tgagtactca accaagtcat tctgagaata gtgtatgcgg 12000

cgaccgagtt gctcttgccc ggcgtcaata cgggataata ccgcgccaca tagcagaact 12060

ttaaaagtgc tcatcattgg aaaacgttct tcggggcgaa aactctcaag gatcttaccg 12120

ctgttgagat ccagttcgat gtaacccact cgtgcaccca actgatcttc agcatctttt 12180

actttcacca gcgtttctgg gtgagcaaaa acaggaaggc aaaatgccgc aaaaaaggga 12240

ataagggcga cacggaaatg ttgaatactc atactcttcc tttttcaata ttattgaagc 12300

atttatcagg gttattgtct catgagcgga tacatatttg aatgtattta gaaaaataaa 12360

caaatagggg ttccgcgcac atttccccga aaagtgccac ctgatgcggt gtgaaatacc 12420

gcacagatgc gtaaggagaa aataccgcat caggaaattg taagcgttaa tattttgtta 12480

aaattcgcgt taaatttttg ttaaatcagc tcatttttta accaataggc cgaaatcggc 12540

aaaatccctt ataaatcaaa agaatagacc gagatagggt tgagtgttgt tccagtttgg 12600

aacaagagtc cactattaaa gaacgtggac tccaacgtca aagggcgaaa aaccgtctat 12660

cagggcgatg gcccactacg tgaaccatca ccctaatcaa gttttttggg gtcgaggtgc 12720

cgtaaagcac taaatcggaa ccctaaaggg agcccccgat ttagagcttg acggggaaag 12780

ccggcgaacg tggcgagaaa ggaagggaag aaagcgaaag gagcgggcgc tagggcgctg 12840

gcaagtgtag cggtcacgct gcgcgtaacc accacacccg ccgcgcttaa tgcgccgcta 12900

cagggcgcgt ccattcgcca ttcaggctgc gcaactgttg ggaagggcga tcggtgcggg 12960

cctcttcgct attacgccag ctggcgaaag ggggatgtgc tgcaaggcga ttaagttggg 13020

taacgccagg gttttcccag tcacgacgtt gtaaaacgac ggccagtgaa ttgtaatacg 13080

actcactata gggcgaattg ggcccgacgt cgcatgcatt ccatagccac acctttgcct 13140

atggcttcac aaccgaaggc aattcgagag gtcgcgctta tggaatcgac tcgtataaag 13200

ctgaagggaa agggagacgt tccgagcgct cagatgcaat agtcgtccag ctaatgtgga 13260

ttcaaaaaca accccaacag taatcttgaa aatttgaacg gatcaatctg aacactcttg 13320

ctccaggtca ttcttctaac gcacatcccc agagtctaga gggagttgtg ttgtgaacat 13380

cctaataaac aatgcaatgg attcgggata tcttctgtct cgccccctac tcgatgtcga 13440

gtaaaccgat caccaactaa caatactcct ccgcgttctg ccattgactc tcaaacagac 13500

atcgctatca acggaacagc atattttagc ttcttaggac aataaatatt gataatgccg 13560

gctctccctc ggtatattaa gcaatccatt catacactca ttcatcaggt taattttata 13620

tatataattt gtctattcaa acaccgtaaa ttactggtac catcatctcc tccttttcaa 13680

atacacgtct atttgcatta atgaaattac tcgccaattc gcagaacgtg tttgtcgaac 13740

agagccttag ctcgggtcca gacaggagca gtgtctcgct gaggaagctg caggagagtt 13800

aattaaaagg cgttgaaaca gaatgagcca gacagcaagg acaaggtggc caacagcaag 13860

gagtccaaaa agccctctat tgacgagatc cacgatgtta ttgctcatga ggtttccgag 13920

ctcgatgctg ggaagaagaa gtgatttgta tataagaaat aaatgagata tagtaaagga 13980

gtgcaagaga atggcaaggt ggtcaaattc tatattactt gcagtcactg gttcctcgtt 14040

gacatgaatg aagttaccgt tggcatagct gatttaatat ataactgtcc aactaactct 14100

cacctagata taacccatgt gtgtgtttcc aatcatcaat gcggccgctt agagagcctt 14160

gccagcgggc tgcttctcgg ccattcgagc gaacacggcc aggtatcgga gcaggatgac 14220

gagaccttca tggggcagag ggtttcggta ggggaggttg tgcttctggc acagctgttc 14280

cacctggtag gaaacggcag tgaggttgtg tcgaggcagg gtgggccaga gatggtgctc 14340

gatctggtag ttcaggcctc caaagaacca gtcagtaatg atgcctcgtc gaatgttcat 14400

ggtctcatgg atctgaccca cagagaagcc atgtccgtcc cagacggaat caccgatctt 14460

ctccagaggg tagtggttca tgaagaccac gatggcaatt ccgaagccac cgacgagctc 14520

ggaaacaaag aacaccagga gcgaggtcag gatggagggc ataaagaaga ggtggaacag 14580

ggtcttgaga gtccagtgca gagcgagtcc aatggcctct ttcttgtact gagatcggta 14640

gaactggttg tctcggtcct tgagggatcg aacggtcagc acagactgga aacaccagat 14700

gaatcgcagg agaatacaga tgaccaggaa atagtactgt tggaactgaa tgagctttcg 14760

ggagatggga gaagctcgag tgacatcgtc ctcggaccag gcgagcagag gcaggttatc 14820

aatgtcggga tcgtgaccct gaacgttggt agcagaatga tgggcgttgt gtctgtcctt 14880

ccaccaggtc acggagaagc cctggagtcc gttgccaaag accagaccca cgaggttatt 14940

ccagtttcgg ttcttgaagg tctggtggtg gcagatgtca tgagacagcc atcccatttg 15000

ctggtagtgc ataccgagca gcacagcacc aatgaagtac atctggtact ggaccatcag 15060

gaagtatcca agcacgccaa gacccagggt ggtggagatc ttgtacgagt accagagggg 15120

agaggcgtca aacatgccag tggcgatcag ctcttctcgg agctttcgga aatcctcctg 15180

agcttcgttg acggcagcct ggggaggcag ctcggaggag gggttgatct tgggcattcg 15240

cttgagcttg tcgaaggctt cctgagagtg cataaccatg aaggcgtcag tagcatctcg 15300

tccctggtag ttctcaatga tttcagctcc accagggtgg aagttcaccc aagcggagac 15360

gtcgtaggta gttccgtcga tggtgagggg cagagcctgt cgcttggact tcaccatggt 15420

tgatgtgtgt ttaattcaag aatgaatata gagaagagaa gaagaaaaaa gattcaattg 15480

agccggcgat gcagaccctt atataaatgt tgccttggac agacggagca agcccgccca 15540

aacctacgtt cggtataata tgttaagctt tttaacacaa aggtttggct tggggtaacc 15600

tgatgtggtg caaaagaccg ggcgttggcg agccattgcg cgggcgaatg gggccgtgac 15660

tcgtctcaaa ttcgagggcg tgcctcaatt cgtgcccccg tggctttttc ccgccgtttc 15720

cgccccgttt gcaccactgc agccgcttct ttggttcgga caccttgctg cgagctaggt 15780

gccttgtgct acttaaaaag tggcctccca acaccaacat gacatgagtg cgtgggccaa 15840

gacacgttgg cggggtcgca gtcggctcaa tggcccggaa aaaacgctgc tggagctggt 15900

tcggacgcag tccgccgcgg cgtatggata tccgcaaggt tccatagcgc cattgccctc 15960

cgtcggcgtc tatcccgcaa cctctaaata gagcgggaat ataacccaag cttctttttt 16020

ttcctttaac acgcacaccc ccaactatca tgttgctgct gctgtttgac tctactctgt 16080

ggaggggtgc tcccacccaa cccaacctac aggtggatcc ggcgctgtga ttggctgata 16140

agtctcctat ccggactaat tctgaccaat gggacatgcg cgcaggaccc aaatgccgca 16200

attacgtaac cccaacgaaa tgcctacccc tctttggagc ccagcggccc caaatccccc 16260

caagcagccc ggttctaccg gcttccatct ccaagcacaa gcagcccggt tctaccggct 16320

tccat 16325

<210> 176
<211> 10189
<212> DNA
<213> Artificial Sequence

<220>

<223> Plasmid pY72

<400> 176

taactttggc cggccagtgt acgcagatct actatagagg aacatttaaa ttgccccgga 60

gaagacggcc aggccgccta gatgacaaat tcaacaactc acagctgact ttctgccatt 120

gccactaggg gggggccttt ttatatggcc aagccaagct ctccacgtcg gttgggctgc 180

acccaacaat aaatgggtag ggttgcacca acaaagggat gggatggggg gtagaagata 240

cgaggataac ggggctcaat ggcacaaata agaacgaata ctgccattaa gactcgtgat 300

ccagcgactg acaccattgc atcatctaag ggcctcaaaa ctacctcgga actgctgcgc 360

tgatctggac accacagagg ttccgagcac tttaggttgc accaaatgtc ccaccaggtg 420

caggcagaaa acgctggaac agcgtgtaca gtttgtctta acaaaaagtg agggcgctga 480

ggtcgagcag ggtggtgtga cttgttatag cctttagagc tgcgaaagcg cgtatggatt 540

tggctcatca ggccagattg agggtctgtg gacacatgtc atgttagtgt acttcaatcg 600

ccccctggat atagccccga caataggccg tggcctcatt tttttgcctt ccgcacattt 660

ccattgctcg gtacccacac cttgcttctc ctgcacttgc caaccttaat actggtttac 720

attgaccaac atcttacaag cgggggggctt gtctagggta tatataaaca gtggctctcc 780

caatcggttg ccagtctctt ttttcctttc tttccccaca gattcgaaat ctaaactaca 840

catcacagaa ttccgagccg tgagtatcca cgacaagatc agtgtcgaga cgacgcgttt 900

tgtgtaatga cacaatccga aagtcgctag caacacacac tctctacaca aactaaccca 960

gctctggtac catggcgact cgacagcgaa ctgccaccac tgttgtggtc gaggaccttc 1020

ccaaggtcac tcttgaggcc aagtctgaac ctgtgttccc cgatatcaag accatcaagg 1080

atgccattcc cgcgcactgc ttccagccct cgctcgtcac ctcattctac tacgtcttcc 1140

gcgattttgc tatggtctct gccctcgtct gggctgctct cacctacatc cccagcatcc 1200

ccgaccagac cctccgcgtc gcagcttgga tggtctacgg cttcgtccag ggtctgttct 1260

gcaccggtgt ctggattctc ggccatgagt gcggccacgg tgctttctct ctccacggaa 1320

aggtcaacaa tgtgaccggc tggttcctcc actcgttcct cctcgtcccc tacttcagct 1380

ggaagtactc tcaccaccgc caccaccgct tcaccggcca catggatctc gacatggctt 1440

tcgtccccaa gactgagccc aagccctcca agtcgctcat gattgctggc attgacgtcg 1500

ccgagcttgt tgaggacacc cccgctgctc agatggtcaa gctcatcttc caccagcttt 1560

tcggatggca ggcgtacctc ttcttcaacg ctagctctgg caagggcagc aagcagtggg 1620

agcccaagac tggcctctcc aagtggttcc gagtcagtca cttcgagcct accagcgctg 1680

tcttccgccc caacgaggcc atcttcatcc tcatctccga tatcggtctt gctctaatgg 1740

gaactgctct gtactttgct tccaagcaag ttggtgtttc gaccattctc ttcctctacc 1800

ttgttcccta cctgtgggtt caccactggc tcgttgccat tacctacctc caccaccacc 1860

acaccgagct ccctcactac accgctgagg gctggaccta cgtcaaggga gctctcgcca 1920

ctgtcgaccg tgagtttggc ttcatcggaa agcacctctt ccacggtatc attgagaagc 1980

acgttgttca ccatctcttc cctaagatcc ccttctacaa ggctgacgag gccaccgagg 2040

ccatcaagcc cgtcattggc gaccactact gccacgacga ccgaagcttc ctgggccagc 2100

tgtggaccat cttcggcacg ctcaagtacg tcgagcacga ccctgcccga cccggtgcca 2160

tgcgatggaa caaggactag gcggccgcta tttatcactc tttacaactt ctacctcaac 2220

tatctacttt aataaatgaa tatcgtttat tctctatgat tactgtatat gcgttcctct 2280

aagacaaatc gaaaccagca tgtgatcgaa tggcatacaa aagtttcttc cgaagttgat 2340

caatgtcctg atagtcaggc agcttgagaa gattgacaca ggtggaggcc gtagggaacc 2400

gatcaacctg tctaccagcg ttacgaatgg caaatgacgg gttcaaagcc ttgaatcctt 2460

gcaatggtgc cttggatact gatgtcacaa acttaagaag cagccgcttg tcctcttcct 2520

cgcctgcagg ataacttcgt ataatgtatg ctatacgaag ttatgaattc tctctcttga 2580

gcttttccat aacaagttct tctgcctcca ggaagtccat gggtggtttg atcatggttt 2640

tggtgtagtg gtagtgcagt ggtggtattg tgactgggga tgtagttgag aataagtcat 2700

acacaagtca gctttcttcg agcctcatat aagtataagt agttcaacgt attagcactg 2760

tacccagcat ctccgtatcg agaaacacaa caacatgccc cattggacag atcatgcgga 2820

tacacaggtt gtgcagtatc atacatactc gatcagacag gtcgtctgac catcatacaa 2880

gctgaacaag cgctccatac ttgcacgctc tctatataca cagttaaatt acatatccat 2940

agtctaacct ctaacagtta atcttctggt aagcctccca gccagccttc tggtatcgct 3000

tggcctcctc aataggatct cggttctggc cgtacagacc tcggccgaca attatgatat 3060

ccgttccggt agacatgaca tcctcaacag ttcggtactg ctgtccgaga gcgtctccct 3120

tgtcgtcaag acccaccccg ggggtcagaa taagccagtc ctcagagtcg cccttaggtc 3180

ggttctgggc aatgaagcca accacaaact cggggtcgga tcgggcaagc tcaatggtct 3240

gcttggagta ctcgccagtg gccagagagc ccttgcaaga cagctcggcc agcatgagca 3300

gacctctggc cagcttctcg ttgggagagg ggactaggaa ctccttgtac tgggagttct 3360

cgtagtcaga gacgtcctcc ttcttctgtt cagagacagt ttcctcggca ccagctcgca 3420

ggccagcaat gattccggtt ccgggtacac cgtgggcgtt ggtgatatcg gaccactcgg 3480

cgattcggtg acaccggtac tggtgcttga cagtgttgcc aatatctgcg aactttctgt 3540

cctcgaacag gaagaaaccg tgcttaagag caagttcctt gaggggggagc acagtgccgg 3600

cgtaggtgaa gtcgtcaatg atgtcgatat gggttttgat catgcacaca taaggtccga 3660

ccttatcggc aagctcaatg agctccttgg tggtggtaac atccagagaa gcacacaggt 3720

tggttttctt ggctgccacg agcttgagca ctcgagcggc aaaggcggac ttgtggacgt 3780

tagctcgagc ttcgtaggag ggcattttgg tggtgaagag gagactgaaa taaatttagt 3840

ctgcagaact ttttatcgga accttatctg gggcagtgaa gtatatgtta tggtaatagt 3900

tacgagttag ttgaacttat agatagactg gactatacgg ctatcggtcc aaattagaaa 3960

gaacgtcaat ggctctctgg gcgtcgcctt tgccgacaaa aatgtgatca tgatgaaagc 4020

cagcaatgac gttgcagctg atattgttgt cggccaaccg cgccgaaaac gcagctgtca 4080

gacccacagc ctccaacgaa gaatgtatcg tcaaagtgat ccaagcacac tcatagttgg 4140

agtcgtactc caaaggcggc aatgacgagt cagacagata ctcgtcgact catcgatcag 4200

gagagaccgg gttggcggcg tatttgtgtc ccaaaaaaca gccccaattg ccccaattga 4260

ccccaaattg acccagtagc gggcccaacc ccggcgagag ccccccttcac cccacatatc 4320

aaacctcccc cggttcccac acttgccgtt aagggcgtag ggtactgcag tctggaatct 4380

acgcttgttc agactttgta ctagtttctt tgtctggcca tccgggtaac ccatgccgga 4440

cgcaaaatag actactgaaa atttttttgc tttgtggttg ggactttagc caagggtata 4500

aaagaccacc gtccccgaat taccttttcct cttcttttct ctctctcctt gtcaactcac 4560

acccgaaatc gttaagcatt tccttctgag tataagaatc attcaaagga tccactagtt 4620

ctagagcggc cgcttaaacc atgaaaaagc ctgaactcac cgcgacgtct gtcgagaagt 4680

ttctgatcga aaagttcgac agcgtctccg acctgatgca gctctcggag ggcgaagaat 4740

ctcgtgcttt cagcttcgat gtaggagggc gtggatatgt cctgcgggta aatagctgcg 4800

ccgatggttt ctacaaagat cgttatgttt atcggcactt tgcatcggcc gcgctcccga 4860

ttccggaagt gcttgacatt ggggaattca gcgagagcct gacctattgc atctcccgcc 4920

gtgcacaggg tgtcacgttg caagacctgc ctgaaaccga actgcccgct gttctgcagc 4980

cggtcgcgga ggccatggat gcgatcgctg cggccgatct tagccagacg agcgggttcg 5040

gcccattcgg accgcaagga atcggtcaat acactacatg gcgtgatttc atatgcgcga 5100

ttgctgatcc ccatgtgtat cactggcaaa ctgtgatgga cgacaccgtc agtgcgtccg 5160

tcgcgcaggc tctcgatgag ctgatgcttt gggccgagga ctgccccgaa gtccggcacc 5220

tcgtgcacgc ggatttcggc tccaacaatg tcctgacgga caatggccgc ataacagcgg 5280

tcattgactg gagcgaggcg atgttcgggg attcccaata cgaggtcgcc aacatcttct 5340

tctggaggcc gtggttggct tgtatggagc agcagacgcg ctacttcgag cggaggcatc 5400

cggagcttgc aggatcgccg cggctccggg cgtatatgct ccgcattggt cttgaccaac 5460

tctatcagag cttggttgac ggcaatttcg atgatgcagc ttgggcgcag ggtcgatgcg 5520

acgcaatcgt ccgatccgga gccgggactg tcgggcgtac acaaatcgcc cgcagaagcg 5580

cggccgtctg gaccgatggc tgtgtagaag tactcgccga tagtggaaac cgacgcccca 5640

gcactcgtcc gagggcaaag gaatagtgag gtacctaaag cggccgccac cgcggcccga 5700

gattccggcc tcttcggccg ccaagcgacc cgggtggacg tctagaggta cctagcaatt 5760

aacagatagt ttgccggtga taattctctt aacctcccac actcctttga cataacgatt 5820

tatgtaacga aactgaaatt tgaccagata ttgtgtccgc gataacttcg tataatgtat 5880

gctatacgaa gttatcgtac gatagttagt agacaacaat cagaacatct ccctccttat 5940

ataatcacac aggccagaac gcgctaaact aaagcgcttt ggacactatg ttacattggc 6000

attgattgaa ctgaaaccac agtctccctc gcctgaatcg agcaatggat gttgtcggaa 6060

gtcaacttca ctagaagagc ggttctatgc cttgtcaaga tcatatcata aactcactct 6120

gtattacccc atctatagaa cacttgttat gaatgggcgg aaacattccg ctatatgcac 6180

ctttccacac taatgcaaag atgtgcatct tcaacgggta gtaagactgg ttccgacttc 6240

cgttgcatgg agagcaatga cctcgataat gcgaacatcc cccacatata cactcttaca 6300

caggccaata taatctgtgc atttactaaa tatttaagtc tatgcacctg cttgatgaaa 6360

agcggcacgg atggtatcat ctagtttccg ccaatccaag aaccaactgt gttggcagtg 6420

gtgtagccca tggcacacag accaaagatg aaaatacaga catcggcggt tcgagccgtg 6480

gtgcctcgag caacaccctt gtaatgcaaa agaggagggt aaatgtacac cagaggcaca 6540

catgcaaacg atccggtgag agcgacgaac cgatcgagat cgtcggcacc tccccatgca 6600

acaaaggcgg tgacaaacac aaggaagaac cggaaaatgt tcttctgcca cttgatggta 6660

gagttgtact tgcctgatcg ggtgaagaga ccattctcga tgattcggat ggcgcgccag 6720

ctgcattaat gaatcggcca acgcgcgggg agaggcggtt tgcgtattgg gcgctcttcc 6780

gcttcctcgc tcactgactc gctgcgctcg gtcgttcggc tgcggcgagc ggtatcagct 6840

cactcaaagg cggtaatacg gttatccaca gaatcagggg ataacgcagg aaagaacatg 6900

tgagcaaaag gccagcaaaa ggccaggaac cgtaaaaagg ccgcgttgct ggcgtttttc 6960

cataggctcc gccccccctga cgagcatcac aaaaatcgac gctcaagtca gaggtggcga 7020

aacccgacag gactataaag ataccaggcg tttccccctg gaagctccct cgtgcgctct 7080

cctgttccga ccctgccgct taccggatac ctgtccgcct ttctcccttc gggaagcgtg 7140

gcgctttctc atagctcacg ctgtaggtat ctcagttcgg tgtaggtcgt tcgctccaag 7200

ctgggctgtg tgcacgaacc ccccgttcag cccgaccgct gcgccttatc cggtaactat 7260

cgtcttgagt ccaacccggt aagacacgac ttatcgccac tggcagcagc cactggtaac 7320

aggattagca gagcgaggta tgtaggcggt gctacagagt tcttgaagtg gtggcctaac 7380

tacggctaca ctagaagaac agtatttggt atctgcgctc tgctgaagcc agttaccttc 7440

ggaaaaagag ttggtagctc ttgatccggc aaacaaacca ccgctggtag cggtggtttt 7500

tttgtttgca agcagcagat tacgcgcaga aaaaaaggat ctcaagaaga tcctttgatc 7560

ttttctacgg ggtctgacgc tcagtggaac gaaaactcac gttaagggat tttggtcatg 7620

agattatcaa aaaggatctt cacctagatc cttttaaatt aaaaatgaag ttttaaatca 7680

atctaaagta tatatgagta aacttggtct gacagttacc aatgcttaat cagtgaggca 7740

cctatctcag cgatctgtct atttcgttca tccatagttg cctgactccc cgtcgtgtag 7800

ataactacga tacgggaggg cttaccatct ggccccagtg ctgcaatgat accgcgagac 7860

ccacgctcac cggctccaga tttatcagca ataaaccagc cagccggaag ggccgagcgc 7920

agaagtggtc ctgcaacttt atccgcctcc atccagtcta ttaattgttg ccgggaagct 7980

agagtaagta gttcgccagt taatagtttg cgcaacgttg ttgccattgc tacaggcatc 8040

gtggtgtcac gctcgtcgtt tggtatggct tcattcagct ccggttccca acgatcaagg 8100

cgagttacat gatcccccat gttgtgcaaa aaagcggtta gctccttcgg tcctccgatc 8160

gttgtcagaa gtaagttggc cgcagtgtta tcactcatgg ttatggcagc actgcataat 8220

tctcttactg tcatgccatc cgtaagatgc ttttctgtga ctggtgagta ctcaaccaag 8280

tcattctgag aatagtgtat gcggcgaccg agttgctctt gcccggcgtc aatacgggat 8340

aataccgcgc cacatagcag aactttaaaa gtgctcatca ttggaaaacg ttcttcgggg 8400

cgaaaactct caaggatctt accgctgttg agatccagtt cgatgtaacc cactcgtgca 8460

cccaactgat cttcagcatc ttttactttc accagcgttt ctgggtgagc aaaaacagga 8520

aggcaaaatg ccgcaaaaaa gggaataagg gcgacacgga aatgttgaat actcatactc 8580

ttccttttc aatattattg aagcatttat cagggttatt gtctcatgag cggatacata 8640

tttgaatgta tttagaaaaa taaacaaata ggggttccgc gcacatttcc ccgaaaagtg 8700

ccacctgatg cggtgtgaaa taccgcacag atgcgtaagg agaaaatacc gcatcaggaa 8760

attgtaagcg ttaatatttt gttaaaattc gcgttaaatt tttgttaaat cagctcattt 8820

tttaaccaat aggccgaaat cggcaaaatc ccttataaat caaaagaata gaccgagata 8880

gggttgagtg ttgttccagt ttggaacaag agtccactat taaagaacgt ggactccaac 8940

gtcaaagggc gaaaaaccgt ctatcagggc gatggcccac tacgtgaacc atcaccctaa 9000

tcaagttttt tggggtcgag gtgccgtaaa gcactaaatc ggaaccctaa agggagcccc 9060

cgatttagag cttgacgggg aaagccggcg aacgtggcga gaaaggaagg gaagaaagcg 9120

aaaggagcgg gcgctagggc gctggcaagt gtagcggtca cgctgcgcgt aaccaccaca 9180

cccgccgcgc ttaatgcgcc gctacagggc gcgtccattc gccattcagg ctgcgcaact 9240

gttgggaagg gcgatcggtg cgggcctctt cgctattacg ccagctggcg aaaggggat 9300

gtgctgcaag gcgattaagt tgggtaacgc cagggttttc ccagtcacga cgttgtaaaa 9360

cgacggccag tgaattgtaa tacgactcac tatagggcga attgggcccg acgtcgcatg 9420

cttagaagtg aggattacaa gaagcctctg gatatcaatg atgaacgtac tcagcggctg 9480

gtcaagcatt tcgaccgtcg aatcgacgag gtgttcacct ttgacaagcg agggttccca 9540

attgatcacg ttctcgagtt gttcaaatct tctctcaaca tctctctgca tgaactatct 9600

ctgttgacga acgtgtcacc cactgttcct cgaacgccct tctccgagtt tggtctgaac 9660

atcttcgatc tcaaactgac ccccgcagtg atcaatagtg ccatgccact gccgatgcgg 9720

tgcgaacatc cctggaggga ttctcggagc tctacacaat gcagattctg tcgtcgagta 9780

ctctctacct tgctcgaatg acttattgtg ctactactgc actcatgctt cgatcatgtg 9840

ccctactgca ccccaaattt ggtgatctga ttgagacaga gtaccctctt cagctgattc 9900

agaagatcat cagcaacatg aatgatgtgg ttgaccaggc aggctgttgt agtcacgtcc 9960

ttcacttcaa gttcattctt catctgcttc tgttttactt tgacaggcaa atgaagacat 10020

ggtacgactt gatggaggcc aagaacgcca tttcaccccg agacaccgaa gtgcctgaaa 10080

tcctggctgc ccccattgat aacatcggaa actacggtat tccggaaagt gtatatagaa 10140

cctttcccca gcttgtgtct gtggatatgg atggtgtaat ccccttaat 10189

<210> 177
<211> 12558
<212> DNA
<213> Artificial Sequence

<220>
<223> Plasmid pY80

<400> 177
taacgcggac acaatatctg gtcaaatttc agtttcgtta cataaatcgt tatgtcaaag 60

gagtgtggga ggttaagaga attatcaccg gcaaactatc tgttaattgc taggtacctc 120

tagacgtcca cccgggtcgc ttggcggccg aagaggccgg aattcgggcc gcggtggcgg 180

ccgcctagtc cttgttccat cgcatggcac cgggtcgggc agggtcgtgc tcgacgtact 240

tgagcgtgcc gaagatggtc cacagctggc ccaggaagct tcggtcgtcg tggcagtagt 300

ggtcgccaat gacgggcttg atggcctcgg tggcctcgtc agccttgtag aaggggatct 360

tagggaagag atggtgaaca acgtgcttct caatgatacc gtggaagagg tgctttccga 420

tgaagccaaa ctcacggtcg acagtggcga gagctccctt gacgtaggtc cagccctcag 480

cggtgtagtg agggagctcg gtgtggtggt ggtggaggta ggtaatggca acgagccagt 540

ggtgaaccca caggtaggga acaaggtaga ggaagagaat ggtcgaaaca ccaacttgct 600

tggaagcaaa gtacagagca gttcccatta gagcaagacc gatatcggag atgaggatga 660

agatggcctc gttggggcgg aagacagcgc tggtaggctc gaagtgactg actcggaacc 720

acttggagag gccagtcttg ggctcccact gcttgctgcc cttgccagag ctagcgttga 780

agaagaggta cgcctgccat ccgaaaagct ggtggaagat gagcttgacc atctgagcag 840

cgggggtgtc ctcaacaagc tcggcgacgt caatgccagc aatcatgagc gacttggagg 900

gcttgggctc agtcttgggg acgaaagcca tgtcgagatc catgtggccg gtgaagcggt 960

ggtggcggtg gtgagagtac ttccagctga agtaggggac gaggaggaac gagtggagga 1020

accagccggt cacattgttg acctttccgt ggagagagaa agcaccgtgg ccgcactcat 1080

ggccgagaat ccagacaccg gtgcagaaca gaccctggac gaagccgtag accatccaag 1140

ctgcgacgcg gagggtctgg tcggggatgc tggggatgta ggtgagagca gcccagacga 1200

gggcagagac catggcaaaa tcgcggaaga cgtagtagaa tgaggtgacg agcgagggct 1260

ggaagcagtg cgcgggaatg gcatccttga tggtcttgat atcggggaac acaggttcag 1320

acttggcctc aagagtgacc ttgggaaggt cctcgaccac aacagtggtg gcagttcgct 1380

gtcgagtcgc cattgtgcgg ccgcggattg atgtgtgttt aattcaagaa tgaatataga 1440

gaagagaaga agaaaaaaga ttcaattgag ccggcgatgc agacccttat ataaatgttg 1500

ccttggacag acggagcaag cccgcccaaa cctacgttcg gtataatatg ttaagctttt 1560

taacacaaag gtttggcttg gggtaacctg atgtggtgca aaagaccggg cgttggcgag 1620

ccattgcgcg ggcgaatggg gccgtgactc gtctcaaatt cgagggcgtg cctcaattcg 1680

tgcccccgtg gctttttccc gccgtttccg ccccgtttgc accactgcag ccgcttcttt 1740

ggttcggaca ccttgctgcg agctaggtgc cttgtgctac ttaaaaagtg gcctcccaac 1800

accaacatga catgagtgcg tgggccaaga cacgttggcg gggtcgcagt cggctcaatg 1860

gcccggaaaa aacgctgctg gagctggttc ggacgcagtc cgccgcggcg tatcgatatc 1920

cgcaaggttc catggcgcca ttgccctccg tcggcgtcta tcccgcaacc tctaaataga 1980

gcgggaatat aacccaagct tctttttttt cctttaacac gcacaccccc aactatcatg 2040

ttgctgctgc tgtttgactc tactctgtgg aggggtgctc ccacccaacc caacctacag 2100

gtggatccgg cgctgtgatt ggctgataag tctcctatcc ggactaattc tgaccaatgg 2160

gacatgcgcg caggacccaa atgccgcaat tacgtaaccc caacgaaatg cctacccctc 2220

tttggagccc agcggcccca aatcccccca agcagcccgg ttctaccggc ttccatctcc 2280

aagcacaagc agcccggttc taccggcttc catctccaag cacccctttc tccacacccc 2340

acaaaaagac ccgtgcagga catcctactg cgtcgaggcc ggccagtgta cgcagatcta 2400

ctatagagga acatttaaat tgccccggag aagacggcca ggccgcctag atgacaaatt 2460

caacaactca cagctgactt tctgccattg ccactagggg ggggcctttt tatatggcca 2520

agccaagctc tccacgtcgg ttgggctgca cccaacaata aatgggtagg gttgcaccaa 2580

caaagggatg ggatgggggg tagaagatac gaggataacg gggctcaatg gcacaaataa 2640

gaacgaatac tgccattaag actcgtgatc cagcgactga caccattgca tcatctaagg 2700

gcctcaaaac tacctcggaa ctgctgcgct gatctggaca ccacagaggt tccgagcact 2760

ttaggttgca ccaaatgtcc caccaggtgc aggcagaaaa cgctggaaca gcgtgtacag 2820

tttgtcttaa caaaaagtga gggcgctgag gtcgagcagg gtggtgtgac ttgttatagc 2880

ctttagagct gcgaaagcgc gtatggattt ggctcatcag gccagattga gggtctgtgg 2940

acacatgtca tgttagtgta cttcaatcgc cccctggata tagccccgac aataggccgt 3000

ggcctcattt ttttgccttc cgcacatttc cattgctcgg tacccacacc ttgcttctcc 3060

tgcacttgcc aaccttaata ctggtttaca ttgaccaaca tcttacaagc ggggggcttg 3120

tctagggtat atataaacag tggctctccc aatcggttgc cagtctcttt tttcctttct 3180

ttccccacag attcgaaatc taaactacac atcacagaat tccgagccgt gagtatccac 3240

gacaagatca gtgtcgagac gacgcgtttt gtgtaatgac acaatccgaa agtcgctagc 3300

aacacacact ctctacacaa actaacccag ctctggtacc atggcgactc gacagcgaac 3360

tgccaccact gttgtggtcg aggaccttcc caaggtcact cttgaggcca agtctgaacc 3420

tgtgttcccc gatatcaaga ccatcaagga tgccattccc gcgcactgct tccagccctc 3480

gctcgtcacc tcattctact acgtcttccg cgattttgct atggtctctg ccctcgtctg 3540

ggctgctctc acctacatcc ccagcatccc cgaccagacc ctccgcgtcg cagcttggat 3600

ggtctacggc ttcgtccagg gtctgttctg caccggtgtc tggattctcg gccatgagtg 3660

cggccacggt gctttctctc tccacggaaa ggtcaacaat gtgaccggct ggttcctcca 3720

ctcgttcctc ctcgtcccct acttcagctg gaagtactct caccaccgcc accaccgctt 3780

caccggccac atggatctcg acatggcttt cgtccccaag actgagccca agccctccaa 3840

gtcgctcatg attgctggca ttgacgtcgc cgagcttgtt gaggacaccc ccgctgctca 3900

gatggtcaag ctcatcttcc accagctttt cggatggcag gcgtacctct tcttcaacgc 3960

tagctctggc aagggcagca agcagtggga gcccaagact ggcctctcca agtggttccg 4020

agtcagtcac ttcgagccta ccagcgctgt cttccgcccc aacgaggcca tcttcatcct 4080

catctccgat atcggtcttg ctctaatggg aactgctctg tactttgctt ccaagcaagt 4140

tggtgtttcg accattctct tcctctacct tgttccctac ctgtgggttc accactggct 4200

cgttgccatt acctacctcc accaccacca caccgagctc cctcactaca ccgctgaggg 4260

ctggacctac gtcaagggag ctctcgccac tgtcgaccgt gagtttggct tcatcggaaa 4320

gcacctcttc cacggtatca ttgagaagca cgttgttcac catctcttcc ctaagatccc 4380

cttctacaag gctgacgagg ccaccgaggc catcaagccc gtcattggcg accactactg 4440

ccacgacgac cgaagcttcc tgggccagct gtggaccatc ttcggcacgc tcaagtacgt 4500

cgagcacgac cctgcccgac ccggtgccat gcgatggaac aaggactagg cggccgctat 4560

ttatcactct ttacaacttc tacctcaact atctacttta ataaatgaat atcgtttatt 4620

ctctatgatt actgtatatg cgttcctcta agacaaatcg aaaccagcat gtgatcgaat 4680

ggcatacaaa agtttcttcc gaagttgatc aatgtcctga tagtcaggca gcttgagaag 4740

attgacacag gtggaggccg tagggaaccg atcaacctgt ctaccagcgt tacgaatggc 4800

aaatgacggg ttcaaagcct tgaatccttg caatggtgcc ttggatactg atgtcacaaa 4860

cttaagaagc agccgcttgt cctcttcctc gcctgcagga taacttcgta taatgtatgc 4920

tatacgaagt tatgaattct ctctcttgag cttttccata acaagttctt ctgcctccag 4980

gaagtccatg ggtggtttga tcatggtttt ggtgtagtgg tagtgcagtg gtggtattgt 5040

gactggggat gtagttgaga ataagtcata cacaagtcag ctttcttcga gcctcatata 5100

agtataagta gttcaacgta ttagcactgt acccagcatc tccgtatcga gaaacacaac 5160

aacatgcccc attggacaga tcatgcggat acacaggttg tgcagtatca tacatactcg 5220

atcagacagg tcgtctgacc atcatacaag ctgaacaagc gctccatact tgcacgctct 5280

ctatatacac agttaaatta catatccata gtctaacctc taacagttaa tcttctggta 5340

agcctcccag ccagccttct ggtatcgctt ggcctcctca ataggatctc ggttctggcc 5400

gtacagacct cggccgacaa ttatgatatc cgttccggta gacatgacat cctcaacagt 5460

tcggtactgc tgtccgagag cgtctccctt gtcgtcaaga cccaccccgg gggtcagaat 5520

aagccagtcc tcagagtcgc ccttaggtcg gttctgggca atgaagccaa ccacaaactc 5580

ggggtcggat cgggcaagct caatggtctg cttggagtac tcgccagtgg ccagagagcc 5640

cttgcaagac agctcggcca gcatgagcag acctctggcc agcttctcgt tgggagaggg 5700

gactaggaac tccttgtact gggagttctc gtagtcagag acgtcctcct tcttctgttc 5760

agagacagtt tcctcggcac cagctcgcag gccagcaatg attccggttc cgggtacacc 5820

gtgggcgttg gtgatatcgg accactcggc gattcggtga caccggtact ggtgcttgac 5880

agtgttgcca atatctgcga actttctgtc ctcgaacagg aagaaaccgt gcttaagagc 5940

aagttccttg agggggagca cagtgccggc gtaggtgaag tcgtcaatga tgtcgatatg 6000

ggttttgatc atgcacacat aaggtccgac cttatcggca agctcaatga gctccttggt 6060

ggtggtaaca tccagagaag cacacaggtt ggttttcttg gctgccacga gcttgagcac 6120

tcgagcggca aaggcggact tgtggacgtt agctcgagct tcgtaggagg gcattttggt 6180

ggtgaagagg agactgaaat aaatttagtc tgcagaactt tttatcggaa ccttatctgg 6240

ggcagtgaag tatatgttat ggtaatagtt acgagttagt tgaacttata gatagactgg 6300

actatacggc tatcggtcca aattagaaag aacgtcaatg gctctctggg cgtcgccttt 6360

gccgacaaaa atgtgatcat gatgaaagcc agcaatgacg ttgcagctga tattgttgtc 6420

ggccaaccgc gccgaaaacg cagctgtcag acccacagcc tccaacgaag aatgtatcgt 6480

caaagtgatc caagcacact catagttgga gtcgtactcc aaaggcggca atgacgagtc 6540

agacagatac tcgtcgactc atcgatcagg agagaccggg ttggcggcgt atttgtgtcc 6600

caaaaaacag ccccaattgc cccaattgac cccaaattga cccagtagcg ggcccaaccc 6660

cggcgagagc ccccttcacc ccacatatca aacctccccc ggttcccaca cttgccgtta 6720

agggcgtagg gtactgcagt ctggaatcta cgcttgttca gactttgtac tagtttcttt 6780

gtctggccat ccgggtaacc catgccggac gcaaaataga ctactgaaaa ttttttttgct 6840

ttgtggttgg gactttagcc aagggtataa aagaccaccg tccccgaatt acctttcctc 6900

ttcttttctc tctctccttg tcaactcaca cccgaaatcg ttaagcattt ccttctgagt 6960

ataagaatca ttcaaaggat ccactagttc tagagcggcc gcttaaacca tgaaaaagcc 7020

tgaactcacc gcgacgtctg tcgagaagtt tctgatcgaa aagttcgaca gcgtctccga 7080

cctgatgcag ctctcggagg gcgaagaatc tcgtgctttc agcttcgatg taggagggcg 7140

tggatatgtc ctgcgggtaa atagctgcgc cgatggtttc tacaaagatc gttatgttta 7200

tcggcacttt gcatcggccg cgctcccgat tccggaagtg cttgacattg gggaattcag 7260

cgagagcctg acctattgca tctcccgccg tgcacagggt gtcacgttgc aagacctgcc 7320

tgaaaccgaa ctgcccgctg ttctgcagcc ggtcgcggag gccatggatg cgatcgctgc 7380

ggccgatctt agccagacga gcgggttcgg cccattcgga ccgcaaggaa tcggtcaata 7440

cactacatgg cgtgatttca tatgcgcgat tgctgatccc catgtgtatc actggcaaac 7500

tgtgatggac gacaccgtca gtgcgtccgt cgcgcaggct ctcgatgagc tgatgctttg 7560

ggccgaggac tgccccgaag tccggcacct cgtgcacgcg gatttcggct ccaacaatgt 7620

cctgacggac aatggccgca taacagcggt cattgactgg agcgaggcga tgttcgggga 7680

ttcccaatac gaggtcgcca acatcttctt ctggaggccg tggttggctt gtatggagca 7740

gcagacgcgc tacttcgagc ggaggcatcc ggagcttgca ggatcgccgc ggctccgggc 7800

gtatatgctc cgcattggtc ttgaccaact ctatcagagc ttggttgacg gcaatttcga 7860

tgatgcagct tgggcgcagg gtcgatgcga cgcaatcgtc cgatccggag ccgggactgt 7920

cgggcgtaca caaatcgccc gcagaagcgc ggccgtctgg accgatggct gtgtagaagt 7980

actcgccgat agtggaaacc gacgccccag cactcgtccg agggcaaagg aatagtgagg 8040

tacctaaagc ggccgccacc gcggcccgag attccggcct cttcggccgc caagcgaccc 8100

gggtggacgt ctagaggtac ctagcaatta acagatagtt tgccggtgat aattctctta 8160

acctcccaca ctcctttgac ataacgattt atgtaacgaa actgaaattt gaccagatat 8220

tgtgtccgcg ataacttcgt ataatgtatg ctatacgaag ttatcgtacg atagttagta 8280

gacaacaatc agaacatctc cctccttata taatcacaca ggccagaacg cgctaaacta 8340

aagcgctttg gacactatgt tacattggca ttgattgaac tgaaaccaca gtctccctcg 8400

cctgaatcga gcaatggatg ttgtcggaag tcaacttcac tagaagagcg gttctatgcc 8460

ttgtcaagat catatcataa actcactctg tattacccca tctatagaac acttgttatg 8520

aatgggcgga aacattccgc tatatgcacc tttccacact aatgcaaaga tgtgcatctt 8580

caacgggtag taagactggt tccgacttcc gttgcatgga gagcaatgac ctcgataatg 8640

cgaacatccc ccacatatac actcttacac aggccaatat aatctgtgca tttactaaat 8700

atttaagtct atgcacctgc ttgatgaaaa gcggcacgga tggtatcatc tagtttccgc 8760

caatccaaga accaactgtg ttggcagtgg tgtagcccat ggcacacaga ccaaagatga 8820

aaatacagac atcggcggtt cgagccgtgg tgcctcgagc aacacccttg taatgcaaaa 8880

gaggagggta aatgtacacc agaggcacac atgcaaacga tccggtgaga gcgacgaacc 8940

gatcgagatc gtcggcacct ccccatgcaa caaaggcggt gacaaacaca aggaagaacc 9000

ggaaaatgtt cttctgccac ttgatggtag agttgtactt gcctgatcgg gtgaagagac 9060

cattctcgat gattcggatg gcgcgccagc tgcattaatg aatcggccaa cgcgcgggga 9120

gaggcggttt gcgtattggg cgctcttccg cttcctcgct cactgactcg ctgcgctcgg 9180

tcgttcggct gcggcgagcg gtatcagctc actcaaaggc ggtaatacgg ttatccacag 9240

aatcagggga taacgcagga aagaacatgt gagcaaaagg ccagcaaaag gccaggaacc 9300

gtaaaaaggc cgcgttgctg gcgtttttcc ataggctccg ccccccctgac gagcatcaca 9360

aaaatcgacg ctcaagtcag aggtggcgaa acccgacagg actataaaga taccaggcgt 9420

ttcccccctgg aagctccctc gtgcgctctc ctgttccgac cctgccgctt accggatacc 9480

tgtccgcctt tctcccttcg ggaagcgtgg cgctttctca tagctcacgc tgtaggtatc 9540

tcagttcggt gtaggtcgtt cgctccaagc tgggctgtgt gcacgaaccc cccgttcagc 9600

ccgaccgctg cgccttatcc ggtaactatc gtcttgagtc caacccggta agacacgact 9660

tatcgccact ggcagcagcc actggtaaca ggattagcag agcgaggtat gtaggcggtg 9720

ctacagagtt cttgaagtgg tggcctaact acggctacac tagaagaaca gtatttggta 9780

tctgcgctct gctgaagcca gttaccttcg gaaaaagagt tggtagctct tgatccggca 9840

aacaaaccac cgctggtagc ggtggttttt ttgtttgcaa gcagcagatt acgcgcagaa 9900

aaaaaggatc tcaagaagat cctttgatct tttctacggg gtctgacgct cagtggaacg 9960

aaaactcacg ttaagggatt ttggtcatga gattatcaaa aaggatcttc acctagatcc 10020

ttttaaatta aaaatgaagt tttaaatcaa tctaaagtat atatgagtaa acttggtctg 10080

acagttacca atgcttaatc agtgaggcac ctatctcagc gatctgtcta tttcgttcat 10140

ccatagttgc ctgactcccc gtcgtgtaga taactacgat acgggagggc ttaccatctg 10200

gccccagtgc tgcaatgata ccgcgagacc cacgctcacc ggctccagat ttatcagcaa 10260

taaaccagcc agccggaagg gccgagcgca gaagtggtcc tgcaacttta tccgcctcca 10320

tccagtctat taattgttgc cgggaagcta gagtaagtag ttcgccagtt aatagtttgc 10380

gcaacgttgt tgccattgct acaggcatcg tggtgtcacg ctcgtcgttt ggtatggctt 10440

cattcagctc cggttcccaa cgatcaaggc gagttacatg atcccccatg ttgtgcaaaa 10500

aagcggttag ctccttcggt cctccgatcg ttgtcagaag taagttggcc gcagtgttat 10560

cactcatggt tatggcagca ctgcataatt ctcttactgt catgccatcc gtaagatgct 10620

tttctgtgac tggtgagtac tcaaccaagt cattctgaga atagtgtatg cggcgaccga 10680

gttgctcttg cccggcgtca atacgggata ataccgcgcc acatagcaga actttaaaag 10740

tgctcatcat tggaaaacgt tcttcggggc gaaaactctc aaggatctta ccgctgttga 10800

gatccagttc gatgtaaccc actcgtgcac ccaactgatc ttcagcatct tttactttca 10860

ccagcgtttc tgggtgagca aaaacaggaa ggcaaaatgc cgcaaaaaag ggaataaggg 10920

cgacacggaa atgttgaata ctcatactct tccttttca atattattga agcatttatc 10980

agggttattg tctcatgagc ggatacatat ttgaatgtat ttagaaaaat aaacaaatag 11040

gggttccgcg cacatttccc cgaaaagtgc cacctgatgc ggtgtgaaat accgcacaga 11100

tgcgtaagga gaaaataccg catcaggaaa ttgtaagcgt taatattttg ttaaaattcg 11160

cgttaaattt ttgttaaatc agctcatttt ttaaccaata ggccgaaatc ggcaaaatcc 11220

cttataaatc aaaagaatag accgagatag ggttgagtgt tgttccagtt tggaacaaga 11280

gtccactatt aaagaacgtg gactccaacg tcaaagggcg aaaaaccgtc tatcagggcg 11340

atggcccact acgtgaacca tcaccctaat caagtttttt ggggtcgagg tgccgtaaag 11400

cactaaatcg gaaccctaaa gggagccccc gatttagagc ttgacgggga aagccggcga 11460

acgtggcgag aaaggaaggg aagaaagcga aaggagcggg cgctagggcg ctggcaagtg 11520

tagcggtcac gctgcgcgta accaccacac ccgccgcgct taatgcgccg ctacagggcg 11580

cgtccattcg ccattcaggc tgcgcaactg ttgggaaggg cgatcggtgc gggcctcttc 11640

gctattacgc cagctggcga aaggggggatg tgctgcaagg cgattaagtt gggtaacgcc 11700

agggttttcc cagtcacgac gttgtaaaac gacggccagt gaattgtaat acgactcact 11760

atagggcgaa ttgggcccga cgtcgcatgc ttagaagtga ggattacaag aagcctctgg 11820

atatcaatga tgaacgtact cagcggctgg tcaagcattt cgaccgtcga atcgacgagg 11880

tgttcacctt tgacaagcga gggttcccaa ttgatcacgt tctcgagttg ttcaaatctt 11940

ctctcaacat ctctctgcat gaactatctc tgttgacgaa cgtgtcaccc actgttcctc 12000

gaacgccctt ctccgagttt ggtctgaaca tcttcgatct caaactgacc cccgcagtga 12060

tcaatagtgc catgccactg ccgatgcggt gcgaacatcc ctggagggat tctcggagct 12120

ctacacaatg cagattctgt cgtcgagtac tctctacctt gctcgaatga cttattgtgc 12180

tactactgca ctcatgcttc gatcatgtgc cctactgcac cccaaatttg gtgatctgat 12240

tgagacagag taccctcttc agctgattca gaagatcatc agcaacatga atgatgtggt 12300

tgaccaggca ggctgttgta gtcacgtcct tcacttcaag ttcattcttc atctgcttct 12360

gttttacttt gacaggcaaa tgaagacatg gtacgacttg atggaggcca agaacgccat 12420

ttcaccccga gacaccgaag tgcctgaaat cctggctgcc cccattgata acatcggaaa 12480

ctacggtatt ccggaaagtg tatatagaac ctttccccag cttgtgtctg tggatatgga 12540

tggtgtaatc cccttaat 12558

<210> 178
<211> 8982
<212> DNA
<213> Artificial Sequence

<220>
<223> Plasmid pY79

<400> 178
ggtggagctc cagcttttgt tccctttagt gagggttaat ttcgagcttg gcgtaatcat 60

ggtcatagct gtttcctgtg tgaaattgtt atccgctcac aattccacac aacgtacgag 120

ccggaagcat aaagtgtaaa gcctggggtg cctaatgagt gagctaactc acattaattg 180

cgttgcgctc actgcccgct ttccagtcgg gaaacctgtc gtgccagctg cattaatgaa 240

tcggccaacg cgcggggaga ggcggtttgc gtattgggcg ctcttccgct tcctcgctca 300

ctgactcgct gcgctcggtc gttcggctgc ggcgagcggt atcagctcac tcaaaggcgg 360

taatacggtt atccacagaa tcaggggata acgcaggaaa gaacatgtga gcaaaaggcc 420

agcaaaaggc caggaaccgt aaaaaggccg cgttgctggc gtttttccat aggctccgcc 480

cccctgacga gcatcacaaa aatcgacgct caagtcagag gtggcgaaac ccgacaggac 540

tataaagata ccaggcgttt ccccctggaa gctccctcgt gcgctctcct gttccgaccc 600

tgccgcttac cggatacctg tccgcctttc tcccttcggg aagcgtggcg ctttctcata 660

gctcacgctg taggtatctc agttcggtgt aggtcgttcg ctccaagctg ggctgtgtgc 720

acgaaccccc cgttcagccc gaccgctgcg ccttatccgg taactatcgt cttgagtcca 780

acccggtaag acacgactta tcgccactgg cagcagccac tggtaacagg attagcagag 840

cgaggtatgt aggcggtgct acagagttct tgaagtggtg gcctaactac ggctacacta 900

gaaggacagt atttggtatc tgcgctctgc tgaagccagt taccttcgga aaaagagttg 960

gtagctcttg atccggcaaa caaaccaccg ctggtagcgg tggttttttt gtttgcaagc 1020

agcagattac gcgcagaaaa aaaggatctc aagaagatcc tttgatcttt tctacggggt 1080

ctgacgctca gtggaacgaa aactcacgtt aagggatttt ggtcatgaga ttatcaaaaa 1140

ggatcttcac ctagatcctt ttaaattaaa aatgaagttt taaatcaatc taaagtatat 1200

atgagtaaac ttggtctgac agttaccaat gcttaatcag tgaggcacct atctcagcga 1260

tctgtctatt tcgttcatcc atagttgcct gactccccgt cgtgtagata actacgatac 1320

gggagggctt accatctggc cccagtgctg caatgatacc gcgagaccca cgctcaccgg 1380

ctccagattt atcagcaata aaccagccag ccggaagggc cgagcgcaga agtggtcctg 1440

caactttatc cgcctccatc cagtctatta attgttgccg ggaagctaga gtaagtagtt 1500

cgccagttaa tagtttgcgc aacgttgttg ccattgctac aggcatcgtg gtgtcacgct 1560

cgtcgtttgg tatggcttca ttcagctccg gttcccaacg atcaaggcga gttacatgat 1620

cccccatgtt gtgcaaaaaa gcggttagct ccttcggtcc tccgatcgtt gtcagaagta 1680

agttggccgc agtgttatca ctcatggtta tggcagcact gcataattct cttactgtca 1740

tgccatccgt aagatgcttt tctgtgactg gtgagtactc aaccaagtca ttctgagaat 1800

agtgtatgcg gcgaccgagt tgctcttgcc cggcgtcaat acgggataat accgcgccac 1860

atagcagaac tttaaaagtg ctcatcattg gaaaacgttc ttcggggcga aaactctcaa 1920

ggatcttacc gctgttgaga tccagttcga tgtaacccac tcgtgcaccc aactgatctt 1980

cagcatcttt tactttcacc agcgtttctg ggtgagcaaa aacaggaagg caaaatgccg 2040

caaaaaaggg aataagggcg acacggaaat gttgaatact catactcttc ctttttcaat 2100

attattgaag catttatcag ggttattgtc tcatgagcgg atacatattt gaatgtattt 2160

agaaaaataa acaaataggg gttccgcgca catttccccg aaaagtgcca cctgacgcgc 2220

cctgtagcgg cgcattaagc gcggcgggtg tggtggttac gcgcagcgtg accgctacac 2280

ttgccagcgc cctagcgccc gctcctttcg ctttcttccc ttcctttctc gccacgttcg 2340

ccggctttcc ccgtcaagct ctaaatcggg ggctcccttt agggttccga tttagtgctt 2400

tacggcacct cgaccccaaa aaacttgatt agggtgatgg ttcacgtagt gggccatcgc 2460

cctgatagac ggtttttcgc cctttgacgt tggagtccac gttctttaat agtggactct 2520

tgttccaaac tggaacaaca ctcaacccta tctcggtcta ttcttttgat ttataaggga 2580

ttttgccgat ttcggcctat tggttaaaaa atgagctgat ttaacaaaaa tttaacgcga 2640

attttaacaa aatattaacg cttacaattt ccattcgcca ttcaggctgc gcaactgttg 2700

ggaagggcga tcggtgcggg cctcttcgct attacgccag ctggcgaaag ggggatgtgc 2760

tgcaaggcga ttaagttggg taacgccagg gttttcccag tcacgacgtt gtaaaacgac 2820

ggccagtgaa ttgtaatacg actcactata gggcgaattg ggtaccgggc ccccccctcga 2880

ggtcgatggt gtcgataagc ttgatatcga attcatgtca cacaaaccga tcttcgcctc 2940

aaggaaacct aattctacat ccgagagact gccgagatcc agtctacact gattaatttt 3000

cgggccaata atttaaaaaa atcgtgttat ataatattat atgtattata tatatacatc 3060

atgatgatac tgacagtcat gtcccattgc taaatagaca gactccatct gccgcctcca 3120

actgatgttc tcaatattta aggggtcatc tcgcattgtt taataataaa cagactccat 3180

ctaccgcctc caaatgatgt tctcaaaata tattgtatga acttattttt attacttagt 3240

attattagac aacttacttg ctttatgaaa aacacttcct atttaggaaa caatttataa 3300

tggcagttcg ttcatttaac aatttatgta gaataaatgt tataaatgcg tatgggaaat 3360

cttaaatatg gatagcataa atgatatctg cattgcctaa ttcgaaatca acagcaacga 3420

aaaaaatccc ttgtacaaca taaatagtca tcgagaaata tcaactatca aagaacagct 3480

attcacacgt tactattgag attattattg gacgagaatc acacactcaa ctgtctttct 3540

ctcttctaga aatacaggta caagtatgta ctattctcat tgttcatact tctagtcatt 3600

tcatcccaca tattccttgg atttctctcc aatgaatgac attctatctt gcaaattcaa 3660

caattataat aagatatacc aaagtagcgg tatagtggca atcaaaaagc ttctctggtg 3720

tgcttctcgt atttattttt attctaatga tccattaaag gtatatattt atttcttgtt 3780

atataatcct tttgtttatt acatgggctg gatacataaa ggtattttga tttaattttt 3840

tgcttaaatt caatcccccc tcgttcagtg tcaactgtaa tggtaggaaa ttaccatact 3900

tttgaagaag caaaaaaaat gaaagaaaaa aaaaatcgta tttccaggtt agacgttccg 3960

cagaatctag aatgcggtat gcggtacatt gttcttcgaa cgtaaaagtt gcgctccctg 4020

agatattgta catttttgct tttacaagta caagtacatc gtacaactat gtactactgt 4080

tgatgcatcc acaacagttt gttttgtttt tttttgtttt ttttttttct aatgattcat 4140

taccgctatg tatacctact tgtacttgta gtaagccggg ttattggcgt tcaattaatc 4200

atagacttat gaatctgcac ggtgtgcgct gcgagttact tttagcttat gcatgctact 4260

tgggtgtaat attgggatct gttcggaaat caacggatgc tcaaccgatt tcgacagtaa 4320

ttaattaatt ccctagtccc agtgtacacc cgccgatatc gcttaccctg cagccggatt 4380

aaggttggca atttttcacg tccttgtctc cgcaattact caccgggtgg tttataagat 4440

tgcaagcgtc ttgatttgtc tctgtatact aacatgcaat cgcgactcgc ccgacgggcc 4500

actaacctgg ccagaatctc cagatccaag tattctcttg gtctgcgata tgtttccaac 4560

acaaaagccc ctgctgccca gccggcaact gctgagtgag tattccttgc cataaacgac 4620

ccagaaccac tgtatagtgt ttggaagcac tagtcagaag accagcgaaa acaggtggaa 4680

aaaactgaga cgaaaagcaa cgaccagaaa tgtaatgtgt ggaaaagcga cacacacaga 4740

gcagataaag aggtgacaaa taacgacaaa tgaaatatca gtatcttccc acaatcacta 4800

cctctcagct gtctgaaggt gcggctgata tatccatccc acgtctaacg tatggagtgt 4860

gatagaatat gacgacacaa gcatgagaac tcgctctcta tccaaccacc gaaacactgt 4920

cactacagcc gttcttgttg ctccattcgc ttttgtgatt ccatgccttc tctggtgact 4980

gacaacattc cttccttttc tccagccctg ttgttatctg ctcatgacct acggccactc 5040

tctatcgcat actaacatag acgatcccag cccgctcccc acttccaggg caccgttggc 5100

aagcctccta tcctcaagaa ggctgaggct gccaacgctg acatggacga gtccttcatc 5160

ggaatgtctg gaggagagat cttccacgag atgatgctgc gacacaacgt cgacactgtc 5220

ttcggttacc ccggtggagc cattctcccc gtctttgacg ccattcacaa ctctgagtac 5280

ttcaactttg tgctccctcg acacgagcag ggtgccggcc acatggccga gggctacgct 5340

cgagcctctg gtaagcccgg tgtcgttctc gtcacctctg gccccggtgc caccaacgtc 5400

atcaccccca tgcaggacgc tctttccgat ggtaccccca tggttgtctt caccggtcag 5460

gtcctgacct ccgttatcgg cactgacgcc ttccaggagg ccgatgttgt cggcatctcc 5520

cgatcttgca ccaagtggaa cgtcatggtc aagaacgttg ctgagctccc ccgacgaatc 5580

aacgaggcct ttgagattgc tacttccggc cgacccggtc ccgttctcgt cgatctgccc 5640

aaggatgtta ctgctgccat cctgcgagag cccatcccca ccaagtccac cattccctcg 5700

cattctctga ccaacctcac ctctgccgcc gccaccgagt tccagaagca ggctatccag 5760

cgagccgcca acctcatcaa ccagtccaag aagcccgtcc tttacgtcgg acagggtatc 5820

cttggctccg aggagggtcc taagctgctt aaggagctgg ctgagaaggc cgagattccc 5880

gtcaccacta ctctgcaggg tcttggtgcc tttgacgagc gagaccccaa gtctctgcac 5940

atgctcggta tgcacggttc cggctacgcc aacatggcca tgcagaacgc tgactgtatc 6000

attgctctcg gcgcccgatt tgatgaccga gttaccggct ccatccccaa gtttgccccc 6060

gaggctcgag ccgctgccct tgagggtcga ggtggtattg ttcactttga gatccaggcc 6120

aagaacatca acaaggttgt tcaggccacc gaagccgttg agggagacgt taccgagtct 6180

gtccgacagc tcatccccct catcaacaag gtctctgccg ctgagcgagc tccctggact 6240

gagactatcc agtcctggaa gcagcagttc cccttcctct tcgaggctga aggtgaggat 6300

ggtgttatca agccccagtc cgtcattgct ctgctctctg acctgacaga gaacaacaag 6360

gacaagacca tcatcaccac cggtgttggt cagcatcaga tgtggactgc ccagcatttc 6420

cgatggcgac accctcgaac catgatcact tctggtggtc ttggaactat gggttacggc 6480

ctgcccgccg ctatcggcgc caaggttgcc cgacctgact gcgacgtcat tgacatcgat 6540

ggtgacgctt ctttcaacat gactctgacc gagctgtcca ccgccgttca gttcaacatt 6600

ggcgtcaagg ctattgtcct caacaacgag gaacagggta tggtcaccca gctgcagtct 6660

ctcttctacg agaaccgata ctgccacact catcagaaga accccgactt catgaagctg 6720

gccgagtcca tgggcatgaa gggtatccga atcactcaca ttgaccagct ggaggccggt 6780

ctcaaggaga tgctcgcata caagggccct gtgctcgttg aggttgttgt cgacaagaag 6840

atccccgttc ttcccatggt tcccgctggt aaggctttgc atgagttcct tgtctacgac 6900

gctgacgccg aggctgcttc tcgacccgat cgactgaaga atgcccccgc ccctcacgtc 6960

caccagacca cctttgagaa ctaagtggaa aggaacacaa gcaatccgaa ccaaaaataa 7020

ttggggtccc gtgcccacag agtctagtgc agacctaaaa tgaccacagt aaattatagc 7080

tgttattaaa catgagattt tgaccaacaa gagcgtagga atgttattag ctactacttg 7140

tacatacaca gcatttgttt taaataatgt tgcctccagg ggcagtgaga tcaggaccca 7200

gatccgtggc cagctctctg acttcagacc gcttgtactt aagcagctcg caacactgtt 7260

gtcgaggatt gaacttgcca tattcgattt tgtggtcatg aatccagcac acctcgtcga 7320

ctcaggcgac gacggaattc ctgcagccca tctgcagaat tcaggagaga ccgggttggc 7380

ggcgtatttg tgtcccaaaa aacagcccca attgccccaa ttgaccccaa attgacccag 7440

tagcgggccc aaccccggcg agagcccccc tcaccccaca tatcaaacct cccccggttc 7500

ccacacttgc cgttaagggc gtagggtact gcagtctgga atctacgctt gttcagactt 7560

tgtactagtt tctttgtctg gccatccggg taacccatgc cggacgcaaa atagactact 7620

gaaaattttt ttgctttgtg gttgggactt tagccaaggg tataaaagac caccgtcccc 7680

gaattacctt tcctcttctt ttctctctct ccttgtcaac tcacacccga aatcgttaag 7740

catttccttc tgagtataag aatcattcac catgtccaat ttactgaccg tacaccaaaa 7800

tttgcctgca ttaccggtcg atgcaacgag tgatgaggtt cgcaagaacc tgatggacat 7860

gttcagggat cgccaggcgt tttctgagca tacctggaaa atgcttctgt ccgtttgccg 7920

gtcgtgggcg gcatggtgca agttgaataa ccggaaatgg tttcccgcag aacctgaaga 7980

tgttcgcgat tatcttctat atcttcaggc gcgcggtctg gcagtaaaaa ctatccagca 8040

acatttgggc cagctaaaca tgcttcatcg tcggtccggg ctgccacgac caagtgacag 8100

caatgctgtt tcactggtta tgcggcggat ccgaaaagaa aacgttgatg ccggtgaacg 8160

tgcaaaacag gctctagcgt tcgaacgcac tgatttcgac caggttcgtt cactcatgga 8220

aaatagcgat cgctgccagg atatacgtaa tctggcattt ctggggattg cttataacac 8280

cctgttacgt atagccgaaa ttgccaggat cagggttaaa gatatctcac gtactgacgg 8340

tgggagaatg ttaatccata ttggcagaac gaaaacgctg gttagcaccg caggtgtaga 8400

gaaggcactt agcctggggg taactaaact ggtcgagcga tggatttccg tctctggtgt 8460

agctgatgat ccgaataact acctgttttg ccgggtcaga aaaaatggtg ttgccgcgcc 8520

atctgccacc agccagctat caactcgcgc cctggaaggg atttttgaag caactcatcg 8580

attgatttac ggcgctaagg atgactctgg tcagagatac ctggcctggt ctggacacag 8640

tgcccgtgtc ggagccgcgc gagatatggc ccgcgctgga gtttcaatac cggagatcat 8700

gcaagctggt ggctggacca atgtaaatat tgtcatgaac tatatccgta acctggatag 8760

tgaaacaggg gcaatggtgc gcctgctgga agatggcgat taaggccgcc accgcggccc 8820

gagattccgg cctcttcggc cgccaagcga cccgggtgga cgtctagagg tacctagcaa 8880

ttaacagata gtttgccggt gataattctc ttaacctccc acactccttt gacataacga 8940

tttatgtaac gaaactgaaa tttgaccaga tattgtgtcc gc 8982

<210> 179
<211> 10424
<212> DNA
<213> Artificial Sequence

<220>
<223> Plasmid pY86

<220>
<221> misc_feature
<222> (8734)..(8734)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (8739)..(8742)
<223> n is a, c, g, or t

<400> 179
acactgtggc cggcctttac ctgcaggata acttcgtata atgtatgcta tacgaagtta 60

tgaattctct ctcttgagct tttccataac aagttcttct gcctccagga agtccatggg 120

tggtttgatc atggttttgg tgtagtggta gtgcagtggt ggtattgtga ctggggatgt 180

agttgagaat aagtcataca caagtcagct ttcttcgagc ctcatataag tataagtagt 240

tcaacgtatt agcactgtac ccagcatctc cgtatcgaga aacacaacaa catgccccat 300

tggacagatc atgcggatac acaggttgtg cagtatcata catactcgat cagacaggtc 360

gtctgaccat catacaagct gaacaagcgc tccatacttg cacgctctct atatacacag 420

ttaaattaca tatccatagt ctaacctcta acagttaatc ttctggtaag cctcccagcc 480

agccttctgg tatcgcttgg cctcctcaat aggatctcgg ttctggccgt acagacctcg 540

gccgacaatt atgatatccg ttccggtaga catgacatcc tcaacagttc ggtactgctg 600

tccgagagcg tctcccttgt cgtcaagacc caccccgggg gtcagaataa gccagtcctc 660

agagtcgccc ttaggtcggt tctgggcaat gaagccaacc acaaactcgg ggtcggatcg 720

ggcaagctca atggtctgct tggagtactc gccagtggcc agagagccct tgcaagacag 780

ctcggccagc atgagcagac ctctggccag cttctcgttg ggagagggga ctaggaactc 840

cttgtactgg gagttctcgt agtcagagac gtcctccttc ttctgttcag agacagtttc 900

ctcggcacca gctcgcaggc cagcaatgat tccggttccg ggtacaccgt gggcgttggt 960

gatatcggac cactcggcga ttcggtgaca ccggtactgg tgcttgacag tgttgccaat 1020

atctgcgaac tttctgtcct cgaacaggaa gaaaccgtgc ttaagagcaa gttccttgag 1080

ggggagcaca gtgccggcgt aggtgaagtc gtcaatgatg tcgatatggg ttttgatcat 1140

gcacacataa ggtccgacct tatcggcaag ctcaatgagc tccttggtgg tggtaacatc 1200

cagagaagca cacaggttgg ttttcttggc tgccacgagc ttgagcactc gagcggcaaa 1260

ggcggacttg tggacgttag ctcgagcttc gtaggagggc attttggtgg tgaagaggag 1320

actgaaataa atttagtctg cagaactttt tatcggaacc ttatctgggg cagtgaagta 1380

tatgttatgg taatagttac gagttagttg aacttataga tagactggac tatacggcta 1440

tcggtccaaa ttagaaagaa cgtcaatggc tctctgggcg tcgcctttgc cgacaaaaat 1500

gtgatcatga tgaaagccag caatgacgtt gcagctgata ttgttgtcgg ccaaccgcgc 1560

cgaaaacgca gctgtcagac ccacagcctc caacgaagaa tgtatcgtca aagtgatcca 1620

agcacactca tagttggagt cgtactccaa aggcggcaat gacgagtcag acagatactc 1680

gtcgactcat cgatcaggag agaccgggtt ggcggcgtat ttgtgtccca aaaaacagcc 1740

ccaattgccc caattgaccc caaattgacc cagtagcggg cccaaccccg gcgagagccc 1800

ccttcacccc acatatcaaa cctcccccgg ttcccacact tgccgttaag ggcgtagggt 1860

actgcagtct ggaatctacg cttgttcaga ctttgtacta gtttctttgt ctggccatcc 1920

gggtaaccca tgccggacgc aaaatagact actgaaaatt tttttgcttt gtggttggga 1980

ctttagccaa gggtataaaa gaccaccgtc cccgaattac ctttcctctt cttttctctc 2040

tctccttgtc aactcacacc cgaaatcgtt aagcatttcc ttctgagtat aagaatcatt 2100

caaaggatcc actagttcta gagcggccgc ttaaaccatg aaaaagcctg aactcaccgc 2160

gacgtctgtc gagaagtttc tgatcgaaaa gttcgacagc gtctccgacc tgatgcagct 2220

ctcggagggc gaagaatctc gtgctttcag cttcgatgta ggagggcgtg gatatgtcct 2280

gcgggtaaat agctgcgccg atggtttcta caaagatcgt tatgtttatc ggcactttgc 2340

atcggccgcg ctcccgattc cggaagtgct tgacattggg gaattcagcg agagcctgac 2400

ctattgcatc tcccgccgtg cacagggtgt cacgttgcaa gacctgcctg aaaccgaact 2460

gcccgctgtt ctgcagccgg tcgcggaggc catggatgcg atcgctgcgg ccgatcttag 2520

ccagacgagc gggttcggcc cattcggacc gcaaggaatc ggtcaataca ctacatggcg 2580

tgatttcata tgcgcgattg ctgatcccca tgtgtatcac tggcaaactg tgatggacga 2640

caccgtcagt gcgtccgtcg cgcaggctct cgatgagctg atgctttggg ccgaggactg 2700

ccccgaagtc cggcacctcg tgcacgcgga tttcggctcc aacaatgtcc tgacggacaa 2760

tggccgcata acagcggtca ttgactggag cgaggcgatg ttcggggatt cccaatacga 2820

ggtcgccaac atcttcttct ggaggccgtg gttggcttgt atggagcagc agacgcgcta 2880

cttcgagcgg aggcatccgg agcttgcagg atcgccgcgg ctccgggcgt atatgctccg 2940

cattggtctt gaccaactct atcagagctt ggttgacggc aatttcgatg atgcagcttg 3000

ggcgcagggt cgatgcgacg caatcgtccg atccggagcc gggactgtcg ggcgtacaca 3060

aatcgcccgc agaagcgcgg ccgtctggac cgatggctgt gtagaagtac tcgccgatag 3120

tggaaaccga cgccccagca ctcgtccgag ggcaaaggaa tagtgaggta cctaaagcgg 3180

ccgccaccgc ggcccgagat tccggcctct tcggccgcca agcgacccgg gtggacgtct 3240

agaggtacct agcaattaac agatagtttg ccggtgataa ttctcttaac ctcccacact 3300

cctttgacat aacgatttat gtaacgaaac tgaaatttga ccagatattg tgtccgcgat 3360

aacttcgtat aatgtatgct atacgaagtt atcgtacgat agttagtaga caacaatcag 3420

aacatctccc tccttatata atcacacagg ccagaacgcg ctaaactaaa gcgctttgga 3480

cactatgtta cattggcatt gattgaactg aaaccacagt ctccctcgcc tgaatcgagc 3540

aatggatgtt gtcggaagtc aacttcacta gaagagcggt tctatgcctt gtcaagatca 3600

tatcataaac tcactctgta ttaccccatc tatagaacac ttgttatgaa tgggcggaaa 3660

cattccgcta tatgcacctt tccacactaa tgcaaagatg tgcatcttca acgggtagta 3720

agactggttc cgacttccgt tgcatggaga gcaatgacct cgataatgcg aacatccccc 3780

acatatacac tcttacacag gccaatataa tctgtgcatt tactaaatat ttaagtctat 3840

gcacctgctt gatgaaaagc ggcacggatg gtatcatcta gtttccgcca atccaagaac 3900

caactgtgtt ggcagtggtg tagcccatgg cacacagacc aaagatgaaa atacagacat 3960

cggcggttcg agccgtggtg cctcgagcaa cacccttgta atgcaaaaga ggagggtaaa 4020

tgtacaccag aggcacacat gcaaacgatc cggtgagagc gacgaaccga tcgagatcgt 4080

cggcacctcc ccatgcaaca aaggcggtga caaacacaag gaagaaccgg aaaatgttct 4140

tctgccactt gatggtagag ttgtacttgc ctgatcgggt gaagagacca ttctcgatga 4200

ttcggatggc gcgccagctg cattaatgaa tcggccaacg cgcggggaga ggcggtttgc 4260

gtattgggcg ctcttccgct tcctcgctca ctgactcgct gcgctcggtc gttcggctgc 4320

ggcgagcggt atcagctcac tcaaaggcgg taatacggtt atccacagaa tcaggggata 4380

acgcaggaaa gaacatgtga gcaaaaggcc agcaaaaggc caggaaccgt aaaaaggccg 4440

cgttgctggc gtttttccat aggctccgcc cccctgacga gcatcacaaa aatcgacgct 4500

caagtcagag gtggcgaaac ccgacaggac tataaagata ccaggcgttt cccccctggaa 4560

gctccctcgt gcgctctcct gttccgaccc tgccgcttac cggatacctg tccgcctttc 4620

tcccttcggg aagcgtggcg ctttctcata gctcacgctg taggtatctc agttcggtgt 4680

aggtcgttcg ctccaagctg ggctgtgtgc acgaaccccc cgttcagccc gaccgctgcg 4740

ccttatccgg taactatcgt cttgagtcca acccggtaag acacgactta tcgccactgg 4800

cagcagccac tggtaacagg attagcagag cgaggtatgt aggcggtgct acagagttct 4860

tgaagtggtg gcctaactac ggctacacta gaagaacagt atttggtatc tgcgctctgc 4920

tgaagccagt taccttcgga aaaagagttg gtagctcttg atccggcaaa caaaccaccg 4980

ctggtagcgg tggtttttt gtttgcaagc agcagattac gcgcagaaaa aaaggatctc 5040

aagaagatcc tttgatcttt tctacggggt ctgacgctca gtggaacgaa aactcacgtt 5100

aagggatttt ggtcatgaga ttatcaaaaa ggatcttcac ctagatcctt ttaaattaaa 5160

aatgaagttt taaatcaatc taaagtatat atgagtaaac ttggtctgac agttaccaat 5220

gcttaatcag tgaggcacct atctcagcga tctgtctatt tcgttcatcc atagttgcct 5280

gactccccgt cgtgtagata actacgatac gggagggctt accatctggc cccagtgctg 5340

caatgatacc gcgagaccca cgctcaccgg ctccagattt atcagcaata aaccagccag 5400

ccggaagggc cgagcgcaga agtggtcctg caactttatc cgcctccatc cagtctatta 5460

attgttgccg ggaagctaga gtaagtagtt cgccagttaa tagtttgcgc aacgttgttg 5520

ccattgctac aggcatcgtg gtgtcacgct cgtcgtttgg tatggcttca ttcagctccg 5580

gttcccaacg atcaaggcga gttacatgat cccccatgtt gtgcaaaaaa gcggttagct 5640

ccttcggtcc tccgatcgtt gtcagaagta agttggccgc agtgttatca ctcatggtta 5700

tggcagcact gcataattct cttactgtca tgccatccgt aagatgcttt tctgtgactg 5760

gtgagtactc aaccaagtca ttctgagaat agtgtatgcg gcgaccgagt tgctcttgcc 5820

cggcgtcaat acgggataat accgcgccac atagcagaac tttaaaagtg ctcatcattg 5880

gaaaacgttc ttcggggcga aaactctcaa ggatcttacc gctgttgaga tccagttcga 5940

tgtaacccac tcgtgcaccc aactgatctt cagcatcttt tactttcacc agcgtttctg 6000

ggtgagcaaa aacaggaagg caaaatgccg caaaaaaggg aataagggcg acacggaaat 6060

gttgaatact catactcttc ctttttcaat attattgaag catttatcag ggttattgtc 6120

tcatgagcgg atacatattt gaatgtattt agaaaaataa acaaataggg gttccgcgca 6180

catttccccg aaaagtgcca cctgatgcgg tgtgaaatac cgcacagatg cgtaaggaga 6240

aaataccgca tcaggaaatt gtaagcgtta atattttgtt aaaattcgcg ttaaattttt 6300

gttaaatcag ctcatttttt aaccaatagg ccgaaatcgg caaaatccct tataaatcaa 6360

aagaatagac cgagataggg ttgagtgttg ttccagtttg gaacaagagt ccactattaa 6420

agaacgtgga ctccaacgtc aaagggcgaa aaaccgtcta tcagggcgat ggcccactac 6480

gtgaaccatc accctaatca agtttttttgg ggtcgaggtg ccgtaaagca ctaaatcgga 6540

accctaaagg gagcccccga tttagagctt gacggggaaa gccggcgaac gtggcgagaa 6600

aggaagggaa gaaagcgaaa ggagcgggcg ctagggcgct ggcaagtgta gcggtcacgc 6660

tgcgcgtaac caccacaccc gccgcgctta atgcgccgct acagggcgcg tccattcgcc 6720

attcaggctg cgcaactgtt gggaagggcg atcggtgcgg gcctcttcgc tattacgcca 6780

gctggcgaaa gggggatgtg ctgcaaggcg attaagttgg gtaacgccag ggttttccca 6840

gtcacgacgt tgtaaaacga cggccagtga attgtaatac gactcactat agggcgaatt 6900

gggcccgacg tcgcatgctt agaagtgagg attacaagaa gcctctggat atcaatgatg 6960

aacgtactca gcggctggtc aagcatttcg accgtcgaat cgacgaggtg ttcacctttg 7020

acaagcgagg gttcccaatt gatcacgttc tcgagttgtt caaatcttct ctcaacatct 7080

ctctgcatga actatctctg ttgacgaacg tgtcacccac tgttcctcga acgcccttct 7140

ccgagtttgg tctgaacatc ttcgatctca aactgacccc cgcagtgatc aatagtgcca 7200

tgccactgcc gatgcggtgc gaacatccct ggagggattc tcggagctct acacaatgca 7260

gattctgtcg tcgagtactc tctaccttgc tcgaatgact tattgtgcta ctactgcact 7320

catgcttcga tcatgtgccc tactgcaccc caaatttggt gatctgattg agacagagta 7380

ccctcttcag ctgattcaga agatcatcag caacatgaat gatgtggttg accaggcagg 7440

ctgttgtagt cacgtccttc acttcaagtt cattcttcat ctgcttctgt tttactttga 7500

caggcaaatg aagacatggt acgacttgat ggaggccaag aacgccattt caccccgaga 7560

caccgaagtg cctgaaatcc tggctgcccc cattgataac atcggaaact acggtattcc 7620

ggaaagtgta tatagaacct ttccccagct tgtgtctgtg gatatggatg gtgtaatccc 7680

cttaattaac gaggaagagg acaagcggct gcttcttaag tttgtgacat cagtatccaa 7740

ggcaccattg caaggattca aggctttgaa cccgtcattt gccattcgta acgctggtag 7800

acaggttgat cggttcccta cggcctccac ctgtgtcaat cttctcaagc tgcctgacta 7860

tcaggacatt gatcaacttc ggaagaaact tttgtatgcc attcgatcac atgctggttt 7920

cgatttgtct tagaggaacg catatacagt aatcatagag aataaacgat attcatttat 7980

taaagtagat agttgaggta gaagttgtaa agagtgataa atagcggccg cgcctactta 8040

agcaacgggc ttgataacag cgggggggt gcccacgttg ttgcggttgc ggaagaacag 8100

aacaccctta ccagcaccct cggcaccagc gctgggctca acccactggc acatacgcgc 8160

actgcggtac atggcgcgga tgaagccacg aggaccatcc tggacatcag cccggtagtg 8220

cttgcccatg atgggcttaa tggcctcggt ggcctcgtcc gcgttgtaga aggggatgct 8280

gctgacgtag tggtggagga catgagtctc gatgatgccg tggagaaggt ggcggccgat 8340

gaagcccatc tcacggtcaa tggtagcagc ggcaccacgg acgaagttcc actcgtcgtt 8400

ggtgtagtgg ggaagggtag ggtcggtgtg ctggaggaag gtgatggcaa cgagccagtg 8460

gttaacccag aggtagggaa caaagtacca gatggccatg ttgtagaaac cgaacttctg 8520

aacgaggaag tacagagcag tggccatcag accgatacca atatcgctga ggacgatgag 8580

cttagcgtca ctgttctcgt acagagggct gcggggatcg aagtggttaa caccaccgcc 8640

gaggccgtta tgcttgccct tgccgcgacc ctcacgctgg cgctcgtggt agttgtggcc 8700

ggtaacattg gtgatgaggt agttgggcca gccnacgann nnctcagtaa gatgagcgag 8760

ctcgtgggtc atctttccga gacgagtagc ctgctgctcg cgggttcggg gaacgaagac 8820

catgtcacgc tccatgttgc cagtggcctt gtggtgcttt cggtgggaga tttgccagct 8880

gaagtagggg acaaggaggg aagagtgaag aacccagcca gtaatgtcgt tgatgatgcg 8940

agaatcggag aaagcaccgt gaccgcactc atgggcaata acccagagac cagtaccgaa 9000

aagaccctga agaacggtgt acacggccca cagaccagcg cgggcggggg tggaggggat 9060

atattcgggg gtcacaaagt tgtaccagat gctgaaagtg gtagtcagga ggacaatgtc 9120

gcggaggata taaccgtatc ccttgagagc ggagcgcttg aagcagtgct tagggatggc 9180

attgtagatg tccttgatgg taaagtcggg aacctcgaac tggttgccgt aggtgtcgag 9240

catgacacca tactcggact tgggcttggc gatatcaacc tcggacatgg acgagagcga 9300

tgtggaagag gccgagtggc ggggagagtc tgaaggagag acggcggcag actcagaatc 9360

cgtcacagta gttgaggtga cggtgcgtct aagcgcaggg ttctgcttgg gcagagccga 9420

agtggacgcc atggagagct gggttagttt gtgtagagag tgtgtgttgc tagcgacttt 9480

cggattgtgt cattacacaa aacgcgtcgt ctcgacactg atcttgtcgt ggatactcac 9540

ggctcggaca tcgtcgccga cgatgacacc ggactttcgc ttaaggacgt cagtaacagg 9600

cattgtgtga tgtgtagttt agatttcgaa tctgtgggga aagaaaggaa aaaagagact 9660

ggcaaccgat tgggagagcc actgtttata tataccctag acaagccccc cgcttgtaag 9720

atgttggtca atgtaaacca gtattaaggt tggcaagtgc aggagaagca aggtgtgggt 9780

accgagcaat ggaaatgtgc ggaaggcaaa aaaatgaggc cacggcctat tgtcggggct 9840

atatccaggg ggcgattgaa gtacactaac atgacatgtg tccacagacc ctcaatctgg 9900

cctgatgagc caaatccata cgcgctttcg cagctctaaa ggctataaca agtcacacca 9960

ccctgctcga cctcagcgcc ctcacttttt gttaagacaa actgtacacg ctgttccagc 10020

gttttctgcc tgcacctggt gggacatttg gtgcaaccta aagtgctcgg aacctctgtg 10080

gtgtccagat cagcgcagca gttccgaggt agttttgagg cccttagatg atgcaatggt 10140

gtcagtcgct ggatcacgag tcttaatggc agtattcgtt cttatttgtg ccattgagcc 10200

ccgttatcct cgtatcttct acccccccatc ccatcccttt gttggtgcaa ccctacccat 10260

ttattgttgg gtgcagccca accgacgtgg agagcttggc ttggccatat aaaaaggccc 10320

ccccctagtg gcaatggcag aaagtcagct gtgagttgtt gaatttgtca tctaggcggc 10380

ctggccgtct tctccggggc aattgttcct ctatagtact gcgt 10424

<210> 180
<211> 10485
<212> DNA

<213> Artificial Sequence

<220>
<223> Plasmid pY94

<400> 180
taaaaggcgt tgaaacagaa tgagccagac agcaaggaca aggtggccaa cagcaaggag 60

tccaaaaagc cctctattga cgagatccac gatgttattg ctcatgaggt ttccgagctc 120

gatgctggga agaagaagtg atttgtatat aagaaataaa tgagatatag taaaggagtg 180

caagagaatg gcaaggtggt caaattctat attacttgca gtcactggtt cctcgttgac 240

atgaatgaag ttaccgttgg catagctgat ttaatatata actgtccaac taactctcac 300

ctagatataa cccatgtgtg tgtttccaat catcaatgcg gccgcttaga gagccttgcc 360

agcgggctgc ttctcggcca ttcgagcgaa cacggccagg tatcggagca ggatgacgag 420

accttcatgg ggcagagggt ttcggtaggg gaggttgtgc ttctggcaca gctgttccac 480

ctggtaggaa acggcagtga ggttgtgtcg aggcagggtg ggccagagat ggtgctcgat 540

ctggtagttc aggcctccaa agaaccagtc agtaatgatg cctcgtcgaa tgttcatggt 600

ctcatggatc tgacccacag agaagccatg tccgtcccag acggaatcac cgatcttctc 660

cagagggtag tggttcatga agaccacgat ggcaattccg aagccaccga cgagctcgga 720

aacaaagaac accaggagcg aggtcaggat ggagggcata aagaagaggt ggaacagggt 780

cttgagagtc cagtgcagag cgagtccaat ggcctctttc ttgtactgag atcggtagaa 840

ctggttgtct cggtccttga gggatcgaac ggtcagcaca gactggaaac accagatgaa 900

tcgcaggaga atacagatga ccaggaaata gtactgttgg aactgaatga gctttcggga 960

gatgggagaa gctcgagtga catcgtcctc ggaccaggcg agcagaggca ggttatcaat 1020

gtcgggatcg tgaccctgaa cgttggtagc agaatgatgg gcgttgtgtc tgtccttcca 1080

ccaggtcacg gagaagccct ggagtccgtt gccaaagacc agacccacga ggttattcca 1140

gtttcggttc ttgaaggtct ggtggtggca gatgtcatga gacagccatc ccatttgctg 1200

gtagtgcata ccgagcagca cagcaccaat gaagtacatc tggtactgga ccatcaggaa 1260

gtatccaagc acgccaagac ccagggtggt ggagatcttg tacgagtacc agaggggaga 1320

ggcgtcaaac atgccagtgg cgatcagctc ttctcggagc tttcggaaat cctcctgagc 1380

ttcgttgacg gcagcctggg gaggcagctc ggaggagggg ttgatcttgg gcattcgctt 1440

gagcttgtcg aaggcttcct gagagtgcat aaccatgaag gcgtcagtag catctcgtcc 1500

ctggtagttc tcaatgattt cagctccacc agggtggaag ttcacccaag cggagacgtc 1560

gtaggtagtt ccgtcgatgg tgaggggcag agcctgtcgc ttggacttca ccatggagag 1620

ctgggttagt ttgtgtagag agtgtgtgtt gctagcgact ttcggattgt gtcattacac 1680

aaaacgcgtc gtctcgacac tgatcttgtc gtggatactc acggctcgga catcgtcgcc 1740

gacgatgaca ccggactttc gcttaaggac gtcagtaaca ggcattgtgt gatgtgtagt 1800

ttagatttcg aatctgtggg gaaagaaagg aaaaaagaga ctggcaaccg attgggagag 1860

ccactgttta tatataccct agacaagccc cccgcttgta agatgttggt caatgtaaac 1920

cagtattaag gttggcaagt gcaggagaag caaggtgtgg gtaccgagca atggaaatgt 1980

gcggaaggca aaaaaatgag gccacggcct attgtcgggg ctatatccag ggggcgattg 2040

aagtacacta acatgacatg tgtccacaga ccctcaatct ggcctgatga gccaaatcca 2100

tacgcgcttt cgcagctcta aaggctataa caagtcacac caccctgctc gacctcagcg 2160

ccctcacttt ttgttaagac aaactgtaca cgctgttcca gcgttttctg cctgcacctg 2220

gtgggacatt tggtgcaacc taaagtgctc ggaacctctg tggtgtccag atcagcgcag 2280

cagttccgag gtagttttga ggcccttaga tgatgcaatg gtgtcagtcg ctggatcacg 2340

agtcttaatg gcagtattcg ttcttatttg tgccattgag ccccgttatc ctcgtatctt 2400

ctacccccca tcccatccct ttgttggtgc aaccctaccc atttattgtt gggtgcagcc 2460

caaccgacgt ggagagcttg gcttggccat ataaaaaggc ccccccctag tggcaatggc 2520

agaaagtcag ctgtgagttg ttgaatttgt catctaggcg gcctggccgt cttctccggg 2580

gcaatttaaa ttccttcact tcaagttcat tcttcatctg cttctgtttt actttgacag 2640

gcaaatgaag acatggtacg acttgatgga ggccaagaac gccatttcac cccgagacac 2700

cgaagtgcct gaaatcctgg ctgcccccat tgataacatc ggaaactacg gtattccgga 2760

aagtgtatat agaacctttc cccagcttgt gtctgtggat atggatggtg taatcccctt 2820

tgagtactcg tcttggcttc tctccgagca gtatgaggct ctctaatcta gcgcatttaa 2880

tatctcaatg tatttatata tttatcttct catgcggccg ctaaagctgc ttaccagcct 2940

tagcggattt cttggtggcc aggttgtcct ggtaaaagaa gtgacagaac aggagaaaga 3000

cagatccgac gtaggcgtag ttgaaagccc aggagaacag cttgcccttg tcagagttga 3060

agcagggaac gttgatgtag tcccagacca ggagaaagcc accgacgaac tggcaaatct 3120

gcatggcagt gatcagaggc ttggccttga acttgtagcc agcggcagtc agtccatagt 3180

aggtgtacat gatggtgtga atgaacgagt taaagaacat gaagatccac acaccctcgt 3240

tgtgcagtcg aatgccgagg tagacgtccc agggagctcc aaagtgatgg aaggcctgca 3300

gaaaggacac tcgcttgccc ttgaggacca gccaagcggt gtcgaggtac tccacgtact 3360

tagaatagta gaaggccttg gcagtccagg tgaacagctt ggagtcccag acaggagagg 3420

gacactgaaa gagaggctgg ggagtatcac cggtctgtct tcgcagccag gctccagtac 3480

cgtagtccca gccgagagcg gtggcagtca cgtagaagga cagggcagag aagagagcca 3540

ggaggacgtt gtaccagatc atggaggttc ggtaggctcc tttcttctcg tccacgagac 3600

cagagtttcg caggagaggc ttcaggagca ggtaggagaa ggtgccaatg aggatttcgg 3660

gatcggtgac ggcagcccag attcgctcgc cagcgtcgtt ggccagagcc atggttagcg 3720

tgtcgtgttt ttgttgtgct ggaagaacca aagggtggcg caatgtgtgt agatatatat 3780

gtcgtgaccc acaagtcaca caaacaagta tcgggaggag tggtgcacct ctatgcggag 3840

aaaccttata ccgctgtaga ccaactgggg cagaggtgtg agttgaagtc agctggagga 3900

gatgtgtgac agaagcacaa gaagtgagat tgtgagatgt atgtctaggg ggggaagttt 3960

tgtgtcaaat atatgggaat tattatcagc accacgaaat tatacgcctc atatgaccca 4020

tttaggtgga tagatcatgg acactgttga cagctgcgaa gaaaaagcgt attggggatg 4080

atccgaaatt agtccggtac cgaggcgcaa atacgtaaga cagccgatwa aatatatgcg 4140

agaaacacca aagagactct agatgtttgt ttggcacagt tttgacttct gcgaaggcct 4200

tacaccacct tgttgaccct tgtcgcgggt cgggcaatat cggctgacag agttttactt 4260

gctcaataag atacgagctg catagagttg aactacagga caatattggg gctggccaca 4320

tgaagggcat tgtttggagg tgtattgatg gtgaaaacac gatatgaaat gacaacgccc 4380

cctgttttat tattattctt attattttgg gtgcttctct atccatacaa gcacctccta 4440

acatgcttca taagtgacct cctcatcaca aggcctgagg tctcatttat ccagtggcgc 4500

caagctaaac taaaactggt ccgagtagac taaggcgaag agagaaggag agaagacagt 4560

ttttttgtgg ccgcctgtga acaatgaaaa cgatgagggt gagatggagc aaaccatatg 4620

gacagtcaga ggagtacgta cgctgcttac ataatggcgc aacgaccaca tgtcccacag 4680

atacgcatcc gtacggtgag atctatacct aggataactt cgtatagcat acattatacg 4740

aagttatatc gatgagtcga cgagtatctg tctgactcgt cattgccgcc tttggagtac 4800

gactccaact atgagtgtgc ttggatcact ttgacgatac attcttcgtt ggaggctgtg 4860

ggtctgacag ctgcgttttc ggcgcggttg gccgacaaca atatcagctg caacgtcatt 4920

gctggctttc atcatgatca catttttgtc ggcaaaggcg acgcccagag agccattgac 4980

gttctttcta atttggaccg atagccgtat agtccagtct atctataagt tcaactaact 5040

cgtaactatt accataacat atacttcact gccccagata aggttccgat aaaaagttct 5100

gcagactaaa tttatttcag tctcctcttc accaccaaaa tgccctccta cgaagctcga 5160

gctaacgtcc acaagtccgc ctttgccgct cgagtgctca agctcgtggc agccaagaaa 5220

accaacctgt gtgcttctct ggatgttacc accaccaagg agctcattga gcttgccgat 5280

aaggtcggac cttatgtgtg catgatcaaa acccatatcg acatcattga cgacttcacc 5340

tacgccggca ctgtgctccc cctcaaggaa cttgctctta agcacggttt cttcctgttc 5400

gaggacagaa agttcgcaga tattggcaac actgtcaagc accagtaccg gtgtcaccga 5460

atcgccgagt ggtccgatat caccaacgcc cacggtgtac ccggaaccgg aatcattgct 5520

ggcctgcgag ctggtgccga ggaaactgtc tctgaacaga agaaggagga cgtctctgac 5580

tacgagaact cccagtacaa ggagttccta gtcccctctc ccaacgagaa gctggccaga 5640

ggtctgctca tgctggccga gctgtcttgc aagggctctc tggccactgg cgagtactcc 5700

aagcagacca ttgagcttgc ccgatccgac cccgagtttg tggttggctt cattgcccag 5760

aaccgaccta agggcgactc tgaggactgg cttattctga cccccggggt gggtcttgac 5820

gacaagggag acgctctcgg acagcagtac cgaactgttg aggatgtcat gtctaccgga 5880

acggatatca taattgtcgg ccgaggtctg tacggccaga accgagatcc tattgaggag 5940

gccaagcgat accagaaggc tggctgggag gcttaccaga agattaactg ttagaggtta 6000

gactatggat atgtaattta actgtgtata tagagagcgt gcaagtatgg agcgcttgtt 6060

cagcttgtat gatggtcaga cgacctgtct gatcgagtat gtatgatact gcacaacctg 6120

tgtatccgca tgatctgtcc aatggggcat gttgttgtgt ttctcgatac ggagatgctg 6180

ggtacagtgc taatacgttg aactacttat acttatatga ggctcgaaga aagctgactt 6240

gtgtatgact tattctcaac tacatcccca gtcacaatac caccactgca ctaccactac 6300

accaaaacca tgatcaaacc acccatggac ttcctggagg cagaagaact tgttatggaa 6360

aagctcaagg agagaattca taacttcgta tagcatacat tatacgaagt tatcctgcag 6420

gtaaaggccg gccgatgtat atcgtattca ttcatgttag ttgcgtacga agtcgtcaat 6480

gatgtcgata tgggttttga tcatgcacac ataaggtccg accttatcgg caagctcaat 6540

gagctccttg gtggtggtaa catccagaga agcacacagg ttggttttct tggctgccac 6600

gagcttgagc actcgagcgg caaaggcgga cttgtggacg ttagctcgag cttcgtagga 6660

gggcattttg gtggtgaaga ggagactgaa ataaatttag tctgcagaac tttttatcgg 6720

aaccttatct ggggcagtga agtatatgtt atggtaatag ttacgagtta gttgaactta 6780

tagatagact ggactatacg gctatcggtc caaattagaa agaacgtcaa tggctctctg 6840

ggcgtcgcct ttgccgacaa aaatgtgatc atgatgaaag ccagcaatga cgttgcagct 6900

gatattgttg tcggccaacc gcgccgaaaa cgcagctgtc agacccacag cctccaacga 6960

agaatgtatc gtcaaagtga tccaagcaca ctcatagttg gagtcgtact ccaaaggcgg 7020

caatgacgag tcagacagat actcgtcgac cttttccttg ggaaccacca ccgtcagccc 7080

ttctgactca cgtattgtag ccaccgacac aggcaacagt ccgtggatag cagaatatgt 7140

cttgtcggtc catttctcac caactttagg cgtcaagtga atgttgcaga agaagtatgt 7200

gccttcattg agaatcggtg ttgctgattt caataaagtc ttgagatcag tttggcgcgc 7260

cagctgcatt aatgaatcgg ccaacgcgcg gggagaggcg gtttgcgtat tgggcgctct 7320

tccgcttcct cgctcactga ctcgctgcgc tcggtcgttc ggctgcggcg agcggtatca 7380

gctcactcaa aggcggtaat acggttatcc acagaatcag gggataacgc aggaaagaac 7440

atgtgagcaa aaggccagca aaaggccagg aaccgtaaaa aggccgcgtt gctggcgttt 7500

ttccataggc tccgcccccc tgacgagcat cacaaaaatc gacgctcaag tcagaggtgg 7560

cgaaacccga caggactata aagataccag gcgtttcccc ctggaagctc cctcgtgcgc 7620

tctcctgttc cgaccctgcc gcttaccgga tacctgtccg cctttctccc ttcgggaagc 7680

gtggcgcttt ctcatagctc acgctgtagg tatctcagtt cggtgtaggt cgttcgctcc 7740

aagctgggct gtgtgcacga acccccccgtt cagcccgacc gctgcgcctt atccggtaac 7800

tatcgtcttg agtccaaccc ggtaagacac gacttatcgc cactggcagc agccactggt 7860

aacaggatta gcagagcgag gtatgtaggc ggtgctacag agttcttgaa gtggtggcct 7920

aactacggct acactagaag aacagtattt ggtatctgcg ctctgctgaa gccagttacc 7980

ttcggaaaaa gagttggtag ctcttgatcc ggcaaacaaa ccaccgctgg tagcggtggt 8040

ttttttgttt gcaagcagca gattacgcgc agaaaaaaag gatctcaaga agatcctttg 8100

atcttttcta cggggtctga cgctcagtgg aacgaaaact cacgttaagg gattttggtc 8160

atgagattat caaaaaggat cttcacctag atccttttaa attaaaaatg aagttttaaa 8220

tcaatctaaa gtatatatga gtaaacttgg tctgacagtt accaatgctt aatcagtgag 8280

gcacctatct cagcgatctg tctatttcgt tcatccatag ttgcctgact ccccgtcgtg 8340

tagataacta cgatacggga gggcttacca tctggcccca gtgctgcaat gataccgcga 8400

gacccacgct caccggctcc agatttatca gcaataaacc agccagccgg aagggccgag 8460

cgcagaagtg gtcctgcaac tttatccgcc tccatccagt ctattaattg ttgccgggaa 8520

gctagagtaa gtagttcgcc agttaatagt ttgcgcaacg ttgttgccat tgctacaggc 8580

atcgtggtgt cacgctcgtc gtttggtatg gcttcattca gctccggttc ccaacgatca 8640

aggcgagtta catgatcccc catgttgtgc aaaaaagcgg ttagctcctt cggtcctccg 8700

atcgttgtca gaagtaagtt ggccgcagtg ttatcactca tggttatggc agcactgcat 8760

aattctctta ctgtcatgcc atccgtaaga tgcttttctg tgactggtga gtactcaacc 8820

aagtcattct gagaatagtg tatgcggcga ccgagttgct cttgcccggc gtcaatacgg 8880

gataataccg cgccacatag cagaacttta aaagtgctca tcattggaaa acgttcttcg 8940

gggcgaaaac tctcaaggat cttaccgctg ttgagatcca gttcgatgta acccactcgt 9000

gcacccaact gatcttcagc atcttttact ttcaccagcg tttctgggtg agcaaaaaca 9060

ggaaggcaaa atgccgcaaa aaagggaata agggcgacac ggaaatgttg aatactcata 9120

ctcttccttt ttcaatatta ttgaagcatt tatcagggtt attgtctcat gagcggatac 9180

atatttgaat gtatttagaa aaataaacaa ataggggttc cgcgcacatt tccccgaaaa 9240

gtgccacctg atgcggtgtg aaataccgca cagatgcgta aggagaaaat accgcatcag 9300

gaaattgtaa gcgttaatat tttgttaaaa ttcgcgttaa atttttgtta aatcagctca 9360

ttttttaacc aataggccga aatcggcaaa atcccttata aatcaaaaga atagaccgag 9420

atagggttga gtgttgttcc agtttggaac aagagtccac tattaaagaa cgtggactcc 9480

aacgtcaaag ggcgaaaaac cgtctatcag ggcgatggcc cactacgtga accatcaccc 9540

taatcaagtt ttttggggtc gaggtgccgt aaagcactaa atcggaaccc taaagggagc 9600

ccccgattta gagcttgacg gggaaagccg gcgaacgtgg cgagaaagga agggaagaaa 9660

gcgaaaggag cgggcgctag ggcgctggca agtgtagcgg tcacgctgcg cgtaaccacc 9720

acacccgccg cgcttaatgc gccgctacag ggcgcgtcca ttcgccattc aggctgcgca 9780

actgttggga agggcgatcg gtgcgggcct cttcgctatt acgccagctg gcgaaagggg 9840

gatgtgctgc aaggcgatta agttgggtaa cgccagggtt ttcccagtca cgacgttgta 9900

aaacgacggc cagtgaattg taatacgact cactataggg cgaattgggc ccgacgtcgc 9960

atgcagtggt ggtattgtga ctggggatgt agttgagaat aagtcataca caagtcagct 10020

ttcttcgagc ctcatataag tataagtagt tcaacgtatt agcactgtac ccagcatctc 10080

cgtatcgaga aacacaacaa catgccccat tggacagatc atgcggatac acaggttgtg 10140

cagtatcata catactcgat cagacaggtc gtctgaccat catacaagct gaacaagcgc 10200

tccatacttg cacgctctct atatacacag ttaaattaca tatccatagt ctaacctcta 10260

acagttaatc ttctggtaag cctcccagcc agccttctgg tatcgcttgg cctcctcaat 10320

aggatctcgg ttctggccgt acagacctcg gccgacaatt atgatatccg ttccggtaga 10380

catgacatcc tcaacagttc ggtactgctg tccgagagcg tctcccttgt cgtcaagacc 10440

caccccgggg gtcagaataa gccagtcctc agagtcgccc ttaat 10485

<210> 181
<211> 8423
<212> DNA
<213> Artificial Sequence

<220>
<223> Plasmid pY91M

<400> 181

gtacgagccg gaagcataaa gtgtaaagcc tggggtgcct aatgagtgag ctaactcaca 60

ttaattgcgt tgcgctcact gcccgctttc cagtcgggaa acctgtcgtg ccagctgcat 120

taatgaatcg gccaacgcgc ggggagaggc ggtttgcgta ttgggcgctc ttccgcttcc 180

tcgctcactg actcgctgcg ctcggtcgtt cggctgcggc gagcggtatc agctcactca 240

aaggcggtaa tacggttatc cacagaatca ggggataacg caggaaagaa catgtgagca 300

aaaggccagc aaaaggccag gaaccgtaaa aaggccgcgt tgctggcgtt tttccatagg 360

ctccgccccc ctgacgagca tcacaaaaat cgacgctcaa gtcagaggtg gcgaaacccg 420

acaggactat aaagatacca ggcgtttccc cctggaagct ccctcgtgcg ctctcctgtt 480

ccgaccctgc cgcttaccgg atacctgtcc gcctttctcc cttcgggaag cgtggcgctt 540

tctcatagct cacgctgtag gtatctcagt tcggtgtagg tcgttcgctc caagctgggc 600

tgtgtgcacg aaccccccgt tcagcccgac cgctgcgcct tatccggtaa ctatcgtctt 660

gagtccaacc cggtaagaca cgacttatcg ccactggcag cagccactgg taacaggatt 720

agcagagcga ggtatgtagg cggtgctaca gagttcttga agtggtggcc taactacggc 780

tacactagaa ggacagtatt tggtatctgc gctctgctga agccagttac cttcggaaaa 840

agagttggta gctcttgatc cggcaaacaa accaccgctg gtagcggtgg ttttttttgtt 900

tgcaagcagc agattacgcg cagaaaaaaa ggatctcaag aagatccttt gatcttttct 960

acggggtctg acgctcagtg gaacgaaaac tcacgttaag ggattttggt catgagatta 1020

tcaaaaagga tcttcaccta gatcctttta aattaaaaat gaagttttaa atcaatctaa 1080

agtatatatg agtaaacttg gtctgacagt taccaatgct taatcagtga ggcacctatc 1140

tcagcgatct gtctatttcg ttcatccata gttgcctgac tccccgtcgt gtagataact 1200

acgatacggg agggcttacc atctggcccc agtgctgcaa tgataccgcg agacccacgc 1260

tcaccggctc cagatttatc agcaataaac cagccagccg gaagggccga gcgcagaagt 1320

ggtcctgcaa ctttatccgc ctccatccag tctattaatt gttgccggga agctagagta 1380

agtagttcgc cagttaatag tttgcgcaac gttgttgcca ttgctacagg catcgtggtg 1440

tcacgctcgt cgtttggtat ggcttcattc agctccggtt cccaacgatc aaggcgagtt 1500

acatgatccc ccatgttgtg caaaaaagcg gttagctcct tcggtcctcc gatcgttgtc 1560

agaagtaagt tggccgcagt gttatcactc atggttatgg cagcactgca taattctctt 1620

actgtcatgc catccgtaag atgcttttct gtgactggtg agtactcaac caagtcattc 1680

tgagaatagt gtatgcggcg accgagttgc tcttgcccgg cgtcaatacg ggataatacc 1740

gcgccacata gcagaacttt aaaagtgctc atcattggaa aacgttcttc ggggcgaaaa 1800

ctctcaagga tcttaccgct gttgagatcc agttcgatgt aacccactcg tgcacccaac 1860

tgatcttcag catcttttac tttcaccagc gtttctgggt gagcaaaaac aggaaggcaa 1920

aatgccgcaa aaaagggaat aagggcgaca cggaaatgtt gaatactcat actcttcctt 1980

tttcaatatt attgaagcat ttatcagggt tattgtctca tgagcggata catatttgaa 2040

tgtatttaga aaaataaaca aatagggttt ccgcgcacat ttccccgaaa agtgccacct 2100

gacgcgccct gtagcggcgc attaagcgcg gcgggtgtgg tggttacgcg cagcgtgacc 2160

gctacacttg ccagcgccct agcgcccgct cctttcgctt tcttcccttc ctttctcgcc 2220

acgttcgccg gctttccccg tcaagctcta aatcgggggc tcccttttagg gttccgattt 2280

agtgctttac ggcacctcga ccccaaaaaa cttgattagg gtgatggttc acgtagtggg 2340

ccatcgccct gatagacggt ttttcgccct ttgacgttgg agtccacgtt ctttaatagt 2400

ggactcttgt tccaaactgg aacaacactc aaccctatct cggtctattc ttttgattta 2460

taagggattt tgccgatttc ggcctattgg ttaaaaaatg agctgattta acaaaaattt 2520

aacgcgaatt ttaacaaaat attaacgctt acaatttcca ttcgccattc aggctgcgca 2580

actgttggga agggcgatcg gtgcgggcct cttcgctatt acgccagctg gcgaaagggg 2640

gatgtgctgc aaggcgatta agttgggtaa cgccagggtt ttcccagtca cgacgttgta 2700

aaacgacggc cagtgaattg taatacgact cactataggg cgaattgggt accgggcccc 2760

ccctcgaggt cgatggtgtc gataagcttg atatcgaatt catgtcacac aaaccgatct 2820

tcgcctcaag gaaacctaat tctacatccg agagactgcc gagatccagt ctacactgat 2880

taattttcgg gccaataatt taaaaaaatc gtgttatata atattatatg tattatatat 2940

atacatcatg atgatactga cagtcatgtc ccattgctaa atagacagac tccatctgcc 3000

gcctccaact gatgttctca atatttaagg ggtcatctcg cattgtttaa taataaacag 3060

actccatcta ccgcctccaa atgatgttct caaaatatat tgtatgaact tatttttatt 3120

acttagtatt attagacaac ttacttgctt tatgaaaaac acttcctatt taggaaacaa 3180

tttataatgg cagttcgttc atttaacaat ttatgtagaa taaatgttat aaatgcgtat 3240

gggaaatctt aaatatggat agcataaatg atatctgcat tgcctaattc gaaatcaaca 3300

gcaacgaaaa aaatcccttg tacaacataa atagtcatcg agaaatatca actatcaaag 3360

aacagctatt cacacgttac tattgagatt attattggac gagaatcaca cactcaactg 3420

tctttctctc ttctagaaat acaggtacaa gtatgtacta ttctcattgt tcatacttct 3480

agtcatttca tcccacatat tccttggatt tctctccaat gaatgacatt ctatcttgca 3540

aattcaacaa ttataataag atataccaaa gtagcggtat agtggcaatc aaaaagcttc 3600

tctggtgtgc ttctcgtatt tatttttatt ctaatgatcc attaaaggta tatatttatt 3660

tcttgttata taatcctttt gtttattaca tgggctggat acataaaggt attttgattt 3720

aatttttgc ttaaattcaa tccccccctcg ttcagtgtca actgtaatgg taggaaatta 3780

ccatactttt gaagaagcaa aaaaaatgaa agaaaaaaaa aatcgtattt ccaggttaga 3840

cgttccgcag aatctagaat gcggtatgcg gtacattgtt cttcgaacgt aaaagttgcg 3900

ctccctgaga tattgtacat ttttgctttt acaagtacaa gtacatcgta caactatgta 3960

ctactgttga tgcatccaca acagtttgtt ttgttttttt ttgttttttt tttttctaat 4020

gattcattac cgctatgtat acctacttgt acttgtagta agccgggtta ttggcgttca 4080

attaatcata gacttatgaa tctgcacggt gtgcgctgcg agttactttt agcttatgca 4140

tgctacttgg gtgtaatatt gggatctgtt cggaaatcaa cggatgctca atcgatttcg 4200

acagtaatta attaagtcat acacaagtca gctttcttcg agcctcatat aagtataagt 4260

agttcaacgt attagcactg tacccagcat ctccgtatcg agaaacacaa caacatgccc 4320

cattggacag atcatgcgga tacacaggtt gtgcagtatc atacatactc gatcagacag 4380

gtcgtctgac catcatacaa gctgaacaag cgctccatac ttgcacgctc tctatataca 4440

cagttaaatt acatatccat agtctaacct ctaacagtta atcttctggt aagcctccca 4500

gccagccttc tggtatcgct tggcctcctc aataggatct cggttctggc cgtacagacc 4560

tcggccgaca attatgatat ccgttccggt agacatgaca tcctcaacag ttcggtactg 4620

ctgtccgaga gcgtctccct tgtcgtcaag acccaccccg ggggtcagaa taagccagtc 4680

ctcagagtcg cccttaggtc ggttctgggc aatgaagcca accacaaact cggggtcgga 4740

tcgggcaagc tcaatggtct gcttggagta ctcgccagtg gccagagagc ccttgcaaga 4800

cagctcggcc agcatgagca gacctctggc cagcttctcg ttgggagagg ggactaggaa 4860

ctccttgtac tgggagttct cgtagtcaga gacgtcctcc ttcttctgtt cagagacagt 4920

ttcctcggca ccagctcgca ggccagcaat gattccggtt ccgggtacac cgtgggcgtt 4980

ggtgatatcg gaccactcgg cgattcggtg acaccggtac tggtgcttga cagtgttgcc 5040

aatatctgcg aactttctgt cctcgaacag gaagaaaccg tgcttaagag caagttcctt 5100

gaggggggagc acagtgccgg cgtaggtgaa gtcgtcaatg atgtcgatat gggttttgat 5160

catgcacaca taaggtccga ccttatcggc aagctcaatg agctccttgg tggtggtaac 5220

atccagagaa gcacacaggt tggttttctt ggctgccacg agcttgagca ctcgagcggc 5280

aaaggcggac ttgtggacgt tagctcgagc ttcgtaggag ggcattttgg tggtgaagag 5340

gagactgaaa taaatttagt ctgcagaact ttttatcgga accttatctg gggcagtgaa 5400

gtatatgtta tggtaatagt tacgagttag ttgaacttat agatagactg gactatacgg 5460

ctatcggtcc aaattagaaa gaacgtcaat ggctctctgg gcgtcgcctt tgccgacaaa 5520

aatgtgatca tgatgaaagc cagcaatgac gttgcagctg atattgttgt cggccaaccg 5580

cgccgaaaac gcagctgtca gacccacagc ctccaacgaa gaatgtatcg tcaaagtgat 5640

ccaagcacac tcatagttgg agtcgtactc caaaggcggc aatgacgagt cagacagata 5700

ctcgtcgact caggcgacga cggaattcct gcagcccatc tgcagaattc aggagagacc 5760

gggttggcgg cgtatttgtg tcccaaaaaa cagccccaat tgccccggag aagacggcca 5820

ggccgcctag atgacaaatt caacaactca cagctgactt tctgccattg ccactagggg 5880

ggggcctttt tatatggcca agccaagctc tccacgtcgg ttgggctgca cccaacaata 5940

aatgggtagg gttgcaccaa caaagggatg ggatgggggg tagaagatac gaggataacg 6000

gggctcaatg gcacaaataa gaacgaatac tgccattaag actcgtgatc cagcgactga 6060

caccattgca tcatctaagg gcctcaaaac tacctcggaa ctgctgcgct gatctggaca 6120

ccacagaggt tccgagcact ttaggttgca ccaaatgtcc caccaggtgc aggcagaaaa 6180

cgctggaaca gcgtgtacag tttgtcttaa caaaaagtga gggcgctgag gtcgagcagg 6240

gtggtgtgac ttgttatagc ctttagagct gcgaaagcgc gtatggattt ggctcatcag 6300

gccagattga gggtctgtgg acacatgtca tgttagtgta cttcaatcgc cccctggata 6360

tagccccgac aataggccgt ggcctcattt ttttgccttc cgcacatttc cattgctcgg 6420

tacccacacc ttgcttctcc tgcacttgcc aaccttaata ctggtttaca ttgaccaaca 6480

tcttacaagc ggggggcttg tctagggtat atataaacag tggctctccc aatcggttgc 6540

cagtctcttt tttcctttct ttccccacag attcgaaatc taaactacac atcacacaat 6600

gcctgttact gacgtcctta agcgaaagtc cggtgtcatc gtcggcgacg atgtccgagc 6660

cgtgagtatc cacgacaaga tcagtgtcga gacgacgcgt tttgtgtaat gacacaatcc 6720

gaaagtcgct agcaacacac actctctaca caaactaacc cagctctcca tgggtggcgg 6780

aggacagcag acagaccgaa tcaccgacac caacggcaga ttcagcagct acacctggga 6840

ggaggtgcag aaacacacca aacatggaga tcagtgggtg gtggtggaga ggaaggttta 6900

taacgtcagc cagtgggtga agagacaccc cggaggactg aggatcctcg gacactatgc 6960

tggagaagac gccacggagg cgttcactgc gtttcatcca aaccttcagc tggtgaggaa 7020

atacctgaag ccgctgctaa tcggagagct ggaggcgtct gaacccagtc aggaccggca 7080

gaaaaacgct gctctcgtgg aggatttccg agccctgcgt gagcgtctgg aggctgaagg 7140

ctgttttaaa acgcagccgc tgtttttcgc tctgcatttg ggccacattc tgctcctgga 7200

ggccatcgct ttcatgatgg tgtggtattt cggcaccggt tggatcaaca cgctcatcgt 7260

cgctgttatt ctggctactg cacagtcaca agctggatgg ttgcagcatg acttcggtca 7320

tctgtccgtg tttaaaacct ctggaatgaa tcatttggtg cacaaatttg tcatcggaca 7380

cctgaaggga gcgtctgcgg gctggtggaa ccatcggcac ttccagcatc acgctaaacc 7440

caacatcttc aagaaggacc cggacgtcaa catgctgaac gcctttgtgg tgggaaacgt 7500

gcagcccgtg gagtatggcg ttaagaagat caagcatctg ccctacaacc atcagcacaa 7560

gtacttcttc ttcattggtc ctcccctgct catcccagtg tatttccagt tccaaatctt 7620

tcacaatatg atcagtcatg gcatgtgggt ggacctgctg tggtgtatca gctactacgt 7680

ccgatacttc ctttgttaca cgcagttcta cggcgtcttt tgggctatta tcctctttaa 7740

tttcgtcagg tttatggaga gccactggtt tgtttgggtc acacagatga gccacatccc 7800

catgaacatt gactatgaga aaaatcagga ctggctcagc atgcagctgg tcgcgacctg 7860

taacatcgag cagtctgcct tcaacgactg gttcagcgga cacctcaact tccagatcga 7920

gcatcatctc tttcccacaa tgcctcggca caactactgg cgcgccgctc cacgggtgcg 7980

agcgttgtgt gagaaatacg gagtcaaata ccaagagaag accttgtacg gagcatttgc 8040

ggatatcatt aggtctttgg agaaatctgg cgagctctgg ctggatgcgt atctcaacaa 8100

ataagcggcc gcaagtgtgg atggggaagt gagtgcccgg ttctgtgtgc acaattggca 8160

atccaagatg gatggattca acacagggat atagcgagct acgtggtggt gcgaggatat 8220

agcaacggat atttatgttt gacacttgag aatgtacgat acaagcactg tccaagtaca 8280

atactaaaca tactgtacat actcatactc gtacccgggc aacggtttca cttgagtgca 8340

gtggctagtg ctcttactcg tacagtgtgc aatactgcgt atcatagtct ttgatgtata 8400

tcgtattcat tcatgttagt tgc 8423

<210> 182
<211> 10945
<212> DNA
<213> Artificial Sequence

<220>
<223> Plasmid pDMW232

<400> 182
aattcctgca gcccatcgat caggagagac cgggttggcg gcgtatttgt gtcccaaaaa 60

acagccccaa ttgccccaat tgaccccaaa ttgacccagt agcgggccca accccggcga 120

gagccccctt caccccacat atcaaacctc ccccggttcc cacacttgcc gttaagggcg 180

tagggtactg cagtctggaa tctacgcttg ttcagacttt gtactagttt ctttgtctgg 240

ccatccgggt aacccatgcc ggacgcaaaa tagactactg aaaatttttt tgctttgtgg 300

ttgggacttt agccaagggt ataaaagacc accgtccccg aattaccttt cctcttcttt 360

tctctctctc cttgtcaact cacacccgaa atcgttaagc atttccttct gagtataaga 420

atcattcacc atgggaacgg accaaggaaa aaccttcacc tgggaagagc tggcggccca 480

taacaccaag gacgacctac tcttggccat ccgcggcagg gtgtacgatg tcacaaagtt 540

cttgagccgc catcctggtg gagtggacac tctcctgctc ggagctggcc gagatgttac 600

tccggtcttt gagatgtatc acgcgtttgg ggctgcagat gccattatga agaagtacta 660

tgtcggtaca ctggtctcga atgagctgcc catcttcccg gagccaacgg tgttccacaa 720

aaccatcaag acgagagtcg agggctactt tacggatcgg aacattgatc ccaagaatag 780

accagagatc tggggacgat acgctcttat ctttggatcc ttgatcgctt cctactacgc 840

gcagctcttt gtgcctttcg ttgtcgaacg cacatggctt caggtggtgt ttgcaatcat 900

catgggattt gcgtgcgcac aagtcggact caaccctctt catgatgcgt ctcacttttc 960

agtgacccac aaccccactg tctggaagat tctgggagcc acgcacgact ttttcaacgg 1020

agcatcgtac ctggtgtgga tgtaccaaca tatgctcggc catcacccct acaccaacat 1080

tgctggagca gatcccgacg tgtcgacgtc tgagcccgat gttcgtcgta tcaagcccaa 1140

ccaaaagtgg tttgtcaacc acatcaacca gcacatgttt gttcctttcc tgtacggact 1200

gctggcgttc aaggtgcgca ttcaggacat caacattttg tactttgtca agaccaatga 1260

cgctattcgt gtcaatccca tctcgacatg gcacactgtg atgttctggg gcggcaaggc 1320

tttctttgtc tggtatcgcc tgattgttcc cctgcagtat ctgcccctgg gcaaggtgct 1380

gctcttgttc acggtcgcgg acatggtgtc gtcttactgg ctggcgctga ccttccaggc 1440

gaaccacgtt gttgaggaag ttcagtggcc gttgcctgac gagaacggga tcatccaaaa 1500

ggactgggca gctatgcagg tcgagactac gcaggattac gcacacgatt cgcacctctg 1560

gaccagcatc actggcagct tgaactacca ggctgtgcac catctgttcc ccaacgtgtc 1620

gcagcaccat tatcccgata ttctggccat catcaagaac acctgcagcg agtacaaggt 1680

tccatacctt gtcaaggata cgttttggca agcatttgct tcacatttgg agcacttgcg 1740

tgttcttgga ctccgtccca aggaagagta ggcagctaag cggccgcatg agaagataaa 1800

tatataaata cattgagata ttaaatgcgc tagattagag agcctcatac tgctcggaga 1860

gaagccaaga cgagtactca aaggggatta caccatccat atccacagac acaagctggg 1920

gaaaggttct atatacactt tccggaatac cgtagtttcc gatgttatca atgggggcag 1980

ccaggatttc aggcacttcg gtgtctcggg gtgaaatggc gttcttggcc tccatcaagt 2040

cgtaccatgt cttcatttgc ctgtcaaagt aaaacagaag cagatgaaga atgaacttga 2100

agtgaaggaa tttaaattgc cccggagaag acggccaggc cgcctagatg acaaattcaa 2160

caactcacag ctgactttct gccattgcca ctaggggggg gccttttttat atggccaagc 2220

caagctctcc acgtcggttg ggctgcaccc aacaataaat gggtagggtt gcaccaacaa 2280

agggatggga tgggggggtag aagatacgag gataacgggg ctcaatggca caaataagaa 2340

cgaatactgc cattaagact cgtgatccag cgactgacac cattgcatca tctaagggcc 2400

tcaaaactac ctcggaactg ctgcgctgat ctggacacca cagaggttcc gagcacttta 2460

ggttgcacca aatgtcccac caggtgcagg cagaaaacgc tggaacagcg tgtacagttt 2520

gtcttaacaa aaagtgaggg cgctgaggtc gagcagggtg gtgtgacttg ttatagcctt 2580

tagagctgcg aaagcgcgta tggatttggc tcatcaggcc agattgaggg tctgtggaca 2640

catgtcatgt tagtgtactt caatcgcccc ctggatatag ccccgacaat aggccgtggc 2700

ctcatttttt tgccttccgc acatttccat tgctcggtac ccacaccttg cttctcctgc 2760

acttgccaac cttaatactg gtttacattg accaacatct tacaagcggg gggcttgtct 2820

agggtatata taaacagtgg ctctcccaat cggttgccag tctcttttttt cctttctttc 2880

cccacagatt cgaaatctaa actacacatc acacaatgcc tgttactgac gtccttaagc 2940

gaaagtccgg tgtcatcgtc ggcgacgatg tccgagccgt gagtatccac gacaagatca 3000

gtgtcgagac gacgcgtttt gtgtaatgac acaatccgaa agtcgctagc aacacacact 3060

ctctacacaa actaacccag ctctccatgg gaacggacca aggaaaaacc ttcacctggg 3120

aagagctggc ggcccataac accaaggacg acctactctt ggccatccgc ggcagggtgt 3180

acgatgtcac aaagttcttg agccgccatc ctggtggagt ggacactctc ctgctcggag 3240

ctggccgaga tgttactccg gtctttgaga tgtatcacgc gtttggggct gcagatgcca 3300

ttatgaagaa gtactatgtc ggtacactgg tctcgaatga gctgcccatc ttcccggagc 3360

caacggtgtt ccacaaaacc atcaagacga gagtcgaggg ctactttacg gatcggaaca 3420

ttgatcccaa gaatagacca gagatctggg gacgatacgc tcttatcttt ggatccttga 3480

tcgcttccta ctacgcgcag ctctttgtgc ctttcgttgt cgaacgcaca tggcttcagg 3540

tggtgtttgc aatcatcatg ggatttgcgt gcgcacaagt cggactcaac cctcttcatg 3600

atgcgtctca cttttcagtg acccacaacc ccactgtctg gaagattctg ggagccacgc 3660

acgacttttt caacggagca tcgtacctgg tgtggatgta ccaacatatg ctcggccatc 3720

acccctacac caacattgct ggagcagatc ccgacgtgtc gacgtctgag cccgatgttc 3780

gtcgtatcaa gcccaaccaa aagtggtttg tcaaccacat caaccagcac atgtttgttc 3840

ctttcctgta cggactgctg gcgttcaagg tgcgcattca ggacatcaac attttgtact 3900

ttgtcaagac caatgacgct attcgtgtca atcccatctc gacatggcac actgtgatgt 3960

tctggggcgg caaggctttc tttgtctggt atcgcctgat tgttcccctg cagtatctgc 4020

ccctgggcaa ggtgctgctc ttgttcacgg tcgcggacat ggtgtcgtct tactggctgg 4080

cgctgacctt ccaggcgaac cacgttgttg aggaagttca gtggccgttg cctgacgaga 4140

acgggatcat ccaaaaggac tgggcagcta tgcaggtcga gactacgcag gattacgcac 4200

acgattcgca cctctggacc agcatcactg gcagcttgaa ctaccaggct gtgcaccatc 4260

tgttccccaa cgtgtcgcag caccattatc ccgatattct ggccatcatc aagaacacct 4320

gcagcgagta caaggttcca taccttgtca aggatacgtt ttggcaagca tttgcttcac 4380

atttggagca cttgcgtgtt cttggactcc gtcccaagga agagtaggca gctaagcggc 4440

cgcaagtgtg gatggggaag tgagtgcccg gttctgtgtg cacaattggc aatccaagat 4500

ggatggattc aacacaggga tatagcgagc tacgtggtgg tgcgaggata tagcaacgga 4560

tatttatgtt tgacacttga gaatgtacga tacaagcact gtccaagtac aatactaaac 4620

atactgtaca tactcatact cgtacccggg caacggtttc acttgagtgc agtggctagt 4680

gctcttactc gtacagtgtg caatactgcg tatcatagtc tttgatgtat atcgtattca 4740

ttcatgttag ttgcgtacgc caccattctg tctgccgcca tgatgctcaa gttctctctt 4800

aacatgaagc ccgccggtga cgctgttgag gctgccgtca aggagtccgt cgaggctggt 4860

atcactaccg ccgatatcgg aggctcttcc tccacctccg aggtcggaga cttgttgcca 4920

acaaggtcaa ggagctgctc aagaaggagt aagtcgtttc tacgacgcat tgatggaagg 4980

agcaaactga cgcgcctgcg ggttggtcta ccggcagggt ccgctagtgt ataagactct 5040

ataaaaaggg ccctgccctg ctaatgaaat gatgatttat aatttaccgg tgtagcaacc 5100

ttgactagaa gaagcagatt gggtgtgttt gtagtggagg acagtggtac gttttggaaa 5160

cagtcttctt gaaagtgtct tgtctacagt atattcactc ataacctcaa tagccaaggg 5220

tgtagtcggt ttattaaagg aagggagttg tggctgatgt ggatagatat ctttaagctg 5280

gcgactgcac ccaacgagtg tggtggtagc ttgttactgt atattcggta agatatattt 5340

tgtggggttt tagtggtgtt tggtaggtta gtgcttggta tatgagttgt aggcatgaca 5400

atttggaaag gggtggactt tgggaatatt gtgggatttc aataccttag tttgtacagg 5460

gtaattgtta caaatgatac aaagaactgt atttcttttc atttgtttta attggttgta 5520

tatcaagtcc gttagacgag ctcagtgggc gcgccagctg cattaatgaa tcggccaacg 5580

cgcggggaga ggcggtttgc gtattgggcg ctcttccgct tcctcgctca ctgactcgct 5640

gcgctcggtc gttcggctgc ggcgagcggt atcagctcac tcaaaggcgg taatacggtt 5700

atccacagaa tcaggggata acgcaggaaa gaacatgtga gcaaaaggcc agcaaaaggc 5760

caggaaccgt aaaaaggccg cgttgctggc gtttttccat aggctccgcc cccctgacga 5820

gcatcacaaa aatcgacgct caagtcagag gtggcgaaac ccgacaggac tataaagata 5880

ccaggcgttt ccccctggaa gctccctcgt gcgctctcct gttccgaccc tgccgcttac 5940

cggatacctg tccgcctttc tcccttcggg aagcgtggcg ctttctcata gctcacgctg 6000

taggtatctc agttcggtgt aggtcgttcg ctccaagctg ggctgtgtgc acgaaccccc 6060

cgttcagccc gaccgctgcg ccttatccgg taactatcgt cttgagtcca acccggtaag 6120

acacgactta tcgccactgg cagcagccac tggtaacagg attagcagag cgaggtatgt 6180

aggcggtgct acagagttct tgaagtggtg gcctaactac ggctacacta gaagaacagt 6240

atttggtatc tgcgctctgc tgaagccagt taccttcgga aaaagagttg gtagctcttg 6300

atccggcaaa caaaccaccg ctggtagcgg tggttttttt gtttgcaagc agcagattac 6360

gcgcagaaaa aaaggatctc aagaagatcc tttgatcttt tctacggggt ctgacgctca 6420

gtggaacgaa aactcacgtt aagggatttt ggtcatgaga ttatcaaaaa ggatcttcac 6480

ctagatcctt ttaaattaaa aatgaagttt taaatcaatc taaagtatat atgagtaaac 6540

ttggtctgac agttaccaat gcttaatcag tgaggcacct atctcagcga tctgtctatt 6600

tcgttcatcc atagttgcct gactccccgt cgtgtagata actacgatac gggagggctt 6660

accatctggc cccagtgctg caatgatacc gcgagaccca cgctcaccgg ctccagattt 6720

atcagcaata aaccagccag ccggaagggc cgagcgcaga agtggtcctg caactttatc 6780

cgcctccatc cagtctatta attgttgccg ggaagctaga gtaagtagtt cgccagttaa 6840

tagtttgcgc aacgttgttg ccattgctac aggcatcgtg gtgtcacgct cgtcgtttgg 6900

tatggcttca ttcagctccg gttcccaacg atcaaggcga gttacatgat cccccatgtt 6960

gtgcaaaaaa gcggttagct ccttcggtcc tccgatcgtt gtcagaagta agttggccgc 7020

agtgttatca ctcatggtta tggcagcact gcataattct cttactgtca tgccatccgt 7080

aagatgcttt tctgtgactg gtgagtactc aaccaagtca ttctgagaat agtgtatgcg 7140

gcgaccgagt tgctcttgcc cggcgtcaat acgggataat accgcgccac atagcagaac 7200

tttaaaagtg ctcatcattg gaaaacgttc ttcggggcga aaactctcaa ggatcttacc 7260

gctgttgaga tccagttcga tgtaacccac tcgtgcaccc aactgatctt cagcatcttt 7320

tactttcacc agcgtttctg ggtgagcaaa aacaggaagg caaaatgccg caaaaaaggg 7380

aataagggcg acacggaaat gttgaatact catactcttc ctttttcaat attattgaag 7440

catttatcag ggttattgtc tcatgagcgg atacatattt gaatgtattt agaaaaataa 7500

acaaataggg gttccgcgca catttccccg aaaagtgcca cctgatgcgg tgtgaaatac 7560

cgcacagatg cgtaaggaga aaataccgca tcaggaaatt gtaagcgtta atattttgtt 7620

aaaattcgcg ttaaattttt gttaaatcag ctcatttttt aaccaatagg ccgaaatcgg 7680

caaaatccct tataaatcaa aagaatagac cgagataggg ttgagtgttg ttccagtttg 7740

gaacaagagt ccactattaa agaacgtgga ctccaacgtc aaagggcgaa aaaccgtcta 7800

tcagggcgat ggcccactac gtgaaccatc accctaatca agtttttttgg ggtcgaggtg 7860

ccgtaaagca ctaaatcgga accctaaagg gagcccccga tttagagctt gacggggaaa 7920

gccggcgaac gtggcgagaa aggaagggaa gaaagcgaaa ggagcgggcg ctagggcgct 7980

ggcaagtgta gcggtcacgc tgcgcgtaac caccacaccc gccgcgctta atgcgccgct 8040

acagggcgcg tccattcgcc attcaggctg cgcaactgtt gggaagggcg atcggtgcgg 8100

gcctcttcgc tattacgcca gctggcgaaa gggggatgtg ctgcaaggcg attaagttgg 8160

gtaacgccag ggttttccca gtcacgacgt tgtaaaacga cggccagtga attgtaatac 8220

gactcactat agggcgaatt gggcccgacg tcgcatgcta tcggcatcga caaggtttgg 8280

gtccctagcc gataccgcac tacctgagtc acaatcttcg gaggtttagt cttccacata 8340

gcacgggcaa aagtgcgtat atatacaaga gcgtttgcca gccacagatt ttcactccac 8400

acaccacatc acacatacaa ccacacacat ccacaatgga acccgaaact aagaagacca 8460

agactgactc caagaagatt gttcttctcg gcggcgactt ctgtggcccc gaggtgattg 8520

ccgaggccgt caaggtgctc aagtctgttg ctgaggcctc cggcaccgag tttgtgtttg 8580

aggaccgact cattggagga gctgccattg agaaggaggg cgagcccatc accgacgcta 8640

ctctcgacat ctgccgaaag gctgactcta ttatgctcgg tgctgtcgga ggcgctgcca 8700

acaccgtatg gaccactccc gacggacgaa ccgacgtgcg acccgagcag ggtctcctca 8760

agctgcgaaa ggacctgaac ctgtacgcca acctgcgacc ctgccagctg ctgtcgccca 8820

agctcgccga tctctccccc atccgaaacg ttgagggcac cgacttcatc attgtccgag 8880

agctcgtcgg aggtatctac tttggagagc gaaaggagga tgacggatct ggcgtcgctt 8940

ccgacaccga gacctactcc gttaattaat ttgaatcgaa tcgatgagcc taaaatgaac 9000

ccgagtatat ctcataaaat tctcggtgag aggtctgtga ctgtcagtac aaggtgcctt 9060

cattatgccc tcaaccttac catacctcac tgaatgtagt gtacctctaa aaatgaaata 9120

cagtgccaaa agccaaggca ctgagctcgt ctaacggact tgatatacaa ccaattaaaa 9180

caaatgaaaa gaaatacagt tctttgtatc atttgtaaca attaccctgt acaaactaag 9240

gtattgaaat cccacaatat tcccaaagtc cacccctttc caaattgtca tgcctacaac 9300

tcatatacca agcactaacc taccgtttaa acagtgtacg cagatctggt gtagtggtag 9360

tgcagtggtg gtattgtgac tggggatgta gttgagaata agtcatacac aagtcagctt 9420

tcttcgagcc tcatataagt ataagtagtt caacgtatta gcactgtacc cagcatctcc 9480

gtatcgagaa acacaacaac atgccccatt ggacagatca tgcggataca caggttgtgc 9540

agtatcatac atactcgatc agacaggtcg tctgaccatc atacaagctg aacaagcgct 9600

ccatacttgc acgctctcta tatacacagt taaattacat atccatagtc taacctctaa 9660

cagttaatct tctggtaagc ctcccagcca gccttctggt atcgcttggc ctcctcaata 9720

ggatctcggt tctggccgta cagacctcgg ccgacaatta tgatatccgt tccggtagac 9780

atgacatcct caacagttcg gtactgctgt ccgagagcgt ctcccttgtc gtcaagaccc 9840

accccggggg tcagaataag ccagtcctca gagtcgccct taggtcggtt ctgggcaatg 9900

aagccaacca caaactcggg gtcggatcgg gcaagctcaa tggtctgctt ggagtactcg 9960

ccagtggcca gagagcccct gcaagacagc tcggccagca tgagcagacc tctggccagc 10020

ttctcgttgg gagaggggac taggaactcc ttgtactggg agttctcgta gtcagagacg 10080

tcctccttct tctgttcaga gacagtttcc tcggcaccag ctcgcaggcc agcaatgatt 10140

ccggttccgg gtacaccgtg ggcgttggtg atatcggacc actcggcgat tcggtgacac 10200

cggtactggt gcttgacagt gttgccaata tctgcgaact ttctgtcctc gaacaggaag 10260

aaaccgtgct taagagcaag ttccttgagg gggagcacag tgccggcgta ggtgaagtcg 10320

tcaatgatgt cgatatgggt tttgatcatg cacacataag gtccgacctt atcggcaagc 10380

tcaatgagct ccttggtggt ggtaacatcc agagaagcac acaggttggt tttcttggct 10440

gccacgagct tgagcactcg agcggcaaag gcggacttgt ggacgttagc tcgagcttcg 10500

taggagggca ttttggtggt gaagaggaga ctgaaataaa tttagtctgc agaacttttt 10560

atcggaacct tatctggggc agtgaagtat atgttatggt aatagttacg agttagttga 10620

acttatagat agactggact atacggctat cggtccaaat tagaaagaac gtcaatggct 10680

ctctgggcgt cgcctttgcc gacaaaaatg tgatcatgat gaaagccagc aatgacgttg 10740

cagctgatat tgttgtcggc caaccgcgcc gaaaacgcag ctgtcagacc cacagcctcc 10800

aacgaagaat gtatcgtcaa agtgatccaa gcacactcat agttggagtc gtactccaaa 10860

ggcggcaatg acgagtcaga cagatactcg tcgacctttt ccttgggaac caccaccgtc 10920

agcccttctg actcacgtat tgtag 10945

<210> 183
<211> 8194
<212> DNA
<213> Artificial Sequence

<220>
<223> Plasmid pY37/F15

<400> 183
ggccgcacaa tggcgactcg acagcgaact gccaccactg ttgtggtcga ggaccttccc 60

aaggtcactc ttgaggccaa gtctgaacct gtgttccccg atatcaagac catcaaggat 120

gccattcccg cgcactgctt ccagccctcg ctcgtcacct cattctacta cgtcttccgc 180

gattttgcca tggtctctgc cctcgtctgg gctgctctca cctacatccc cagcatcccc 240

gaccagaccc tccgcgtcgc agcttggatg gtctacggct tcgtccaggg tctgttctgc 300

accggtgtct ggattctcgg ccatgagtgc ggccacggtg ctttctctct ccacggaaag 360

gtcaacaatg tgaccggctg gttcctccac tcgttcctcc tcgtccccta cttcagctgg 420

aagtactctc accaccgcca ccaccgcttc accggccaca tggatctcga catggctttc 480

gtccccaaga ctgagcccaa gccctccaag tcgctcatga ttgctggcat tgacgtcgcc 540

gagcttgttg aggacacccc cgctgctcag atggtcaagc tcatcttcca ccagcttttc 600

ggatggcagg cgtacctctt cttcaacgct agctctggca agggcagcaa gcagtgggag 660

cccaagactg gcctctccaa gtggttccga gtcagtcact tcgagcctac cagcgctgtc 720

ttccgcccca acgaggccat cttcatcctc atctccgata tcggtcttgc tctaatggga 780

actgctctgt actttgcttc caagcaagtt ggtgtttcga ccattctctt cctctacctt 840

gttccctacc tgtgggttca ccactggctc gttgccatta cctacctcca ccaccaccac 900

accgagctcc ctcactacac cgctgagggc tggacctacg tcaagggagc tctcgccact 960

gtcgaccgtg agtttggctt catcggaaag cacctcttcc acggtatcat tgagaagcac 1020

gttgttcacc atctcttccc taagatcccc ttctacaagg ctgacgaggc caccgaggcc 1080

atcaagcccg tcattggcga ccactactgc cacgacgacc gaagcttcct gggccagctg 1140

tggaccatct tcggcacgct caagtacgtc gagcacgacc ctgcccgacc cggtgccatg 1200

cgatggaaca aggactaggc taggcggccg ccaccgcggc ccgaattccg gcctcttcgg 1260

ccgccaagcg acccgggtgg acgtctagag gtacctagca attaacagat agtttgccgg 1320

tgataattct cttaacctcc cacactcctt tgacataacg atttatgtaa cgaaactgaa 1380

atttgaccag atattgtgtc cgcggtggag ctccagcttt tgttcccttt agtgagggtt 1440

aatttcgagc ttggcgtaat catggtcata gctgtttcct gtgtgaaatt gttatccgct 1500

cacaattcca cacaacgtac gagccggaag cataaagtgt aaagcctggg gtgcctaatg 1560

agtgagctaa ctcacattaa ttgcgttgcg ctcactgccc gctttccagt cgggaaacct 1620

gtcgtgccag ctgcattaat gaatcggcca acgcgcgggg agaggcggtt tgcgtattgg 1680

gcgctcttcc gcttcctcgc tcactgactc gctgcgctcg gtcgttcggc tgcggcgagc 1740

ggtatcagct cactcaaagg cggtaatacg gttatccaca gaatcagggg ataacgcagg 1800

aaagaacatg tgagcaaaag gccagcaaaa ggccaggaac cgtaaaaagg ccgcgttgct 1860

ggcgtttttc cataggctcc gccccccctga cgagcatcac aaaaatcgac gctcaagtca 1920

gaggtggcga aacccgacag gactataaag ataccaggcg tttcccccctg gaagctccct 1980

cgtgcgctct cctgttccga ccctgccgct taccggatac ctgtccgcct ttctcccttc 2040

gggaagcgtg gcgctttctc atagctcacg ctgtaggtat ctcagttcgg tgtaggtcgt 2100

tcgctccaag ctgggctgtg tgcacgaacc ccccgttcag cccgaccgct gcgccttatc 2160

cggtaactat cgtcttgagt ccaacccggt aagacacgac ttatcgccac tggcagcagc 2220

cactggtaac aggattagca gagcgaggta tgtaggcggt gctacagagt tcttgaagtg 2280

gtggcctaac tacggctaca ctagaaggac agtatttggt atctgcgctc tgctgaagcc 2340

agttaccttc ggaaaaagag ttggtagctc ttgatccggc aaacaaacca ccgctggtag 2400

cggtggtttt tttgtttgca agcagcagat tacgcgcaga aaaaaaggat ctcaagaaga 2460

tcctttgatc ttttctacgg ggtctgacgc tcagtggaac gaaaactcac gttaagggat 2520

tttggtcatg agattatcaa aaaggatctt cacctagatc cttttaaatt aaaaatgaag 2580

ttttaaatca atctaaagta tatatgagta aacttggtct gacagttacc aatgcttaat 2640

cagtgaggca cctatctcag cgatctgtct atttcgttca tccatagttg cctgactccc 2700

cgtcgtgtag ataactacga tacgggaggg cttaccatct ggccccagtg ctgcaatgat 2760

accgcgagac ccacgctcac cggctccaga tttatcagca ataaaccagc cagccggaag 2820

ggccgagcgc agaagtggtc ctgcaacttt atccgcctcc atccagtcta ttaattgttg 2880

ccgggaagct agagtaagta gttcgccagt taatagtttg cgcaacgttg ttgccattgc 2940

tacaggcatc gtggtgtcac gctcgtcgtt tggtatggct tcattcagct ccggttccca 3000

acgatcaagg cgagttacat gatcccccat gttgtgcaaa aaagcggtta gctccttcgg 3060

tcctccgatc gttgtcagaa gtaagttggc cgcagtgtta tcactcatgg ttatggcagc 3120

actgcataat tctcttactg tcatgccatc cgtaagatgc ttttctgtga ctggtgagta 3180

ctcaaccaag tcattctgag aatagtgtat gcggcgaccg agttgctctt gcccggcgtc 3240

aatacgggat aataccgcgc cacatagcag aactttaaaa gtgctcatca ttggaaaacg 3300

ttcttcgggg cgaaaactct caaggatctt accgctgttg agatccagtt cgatgtaacc 3360

cactcgtgca cccaactgat cttcagcatc ttttactttc accagcgttt ctgggtgagc 3420

aaaaacagga aggcaaaatg ccgcaaaaaa gggaataagg gcgacacgga aatgttgaat 3480

actcatactc ttccttttc aatattattg aagcatttat cagggttatt gtctcatgag 3540

cggatacata tttgaatgta tttagaaaaa taaacaaata ggggttccgc gcacatttcc 3600

ccgaaaagtg ccacctgacg cgccctgtag cggcgcatta agcgcggcgg gtgtggtggt 3660

tacgcgcagc gtgaccgcta cacttgccag cgccctagcg cccgctcctt tcgctttctt 3720

cccttccttt ctcgccacgt tcgccggctt tccccgtcaa gctctaaatc gggggctccc 3780

tttagggttc cgatttagtg ctttacggca cctcgacccc aaaaaacttg attagggtga 3840

tggttcacgt agtgggccat cgccctgata gacggttttt cgccctttga cgttggagtc 3900

cacgttcttt aatagtggac tcttgttcca aactggaaca acactcaacc ctatctcggt 3960

ctattctttt gatttataag ggattttgcc gatttcggcc tattggttaa aaaatgagct 4020

gatttaacaa aaatttaacg cgaattttaa caaaatatta acgcttacaa tttccattcg 4080

ccattcaggc tgcgcaactg ttgggaaggg cgatcggtgc gggcctcttc gctattacgc 4140

cagctggcga aaggggggatg tgctgcaagg cgattaagtt gggtaacgcc agggttttcc 4200

cagtcacgac gttgtaaaac gacggccagt gaattgtaat acgactcact atagggcgaa 4260

ttgggtaccg ggcccccccct cgaggtcgat ggtgtcgata agcttgatat cgaattcatg 4320

tcacacaaac cgatcttcgc ctcaaggaaa cctaattcta catccgagag actgccgaga 4380

tccagtctac actgattaat tttcgggcca ataatttaaa aaaatcgtgt tatataatat 4440

tatatgtatt atatatatac atcatgatga tactgacagt catgtcccat tgctaaatag 4500

acagactcca tctgccgcct ccaactgatg ttctcaatat ttaaggggtc atctcgcatt 4560

gtttaataat aaacagactc catctaccgc ctccaaatga tgttctcaaa atatattgta 4620

tgaacttatt tttattactt agtattatta gacaacttac ttgctttatg aaaaacactt 4680

cctatttagg aaacaattta taatggcagt tcgttcattt aacaatttat gtagaataaa 4740

tgttataaat gcgtatggga aatcttaaat atggatagca taaatgatat ctgcattgcc 4800

taattcgaaa tcaacagcaa cgaaaaaaat cccttgtaca acataaatag tcatcgagaa 4860

atatcaacta tcaaagaaca gctattcaca cgttactatt gagattatta ttggacgaga 4920

atcacacact caactgtctt tctctcttct agaaatacag gtacaagtat gtactattct 4980

cattgttcat acttctagtc atttcatccc acatattcct tggatttctc tccaatgaat 5040

gacattctat cttgcaaatt caacaattat aataagatat accaaagtag cggtatagtg 5100

gcaatcaaaa agcttctctg gtgtgcttct cgtatttatt tttattctaa tgatccatta 5160

aaggtatata tttatttctt gttatataat ccttttgttt attacatggg ctggatacat 5220

aaaggtattt tgatttaatt ttttgcttaa attcaatccc ccctcgttca gtgtcaactg 5280

taatggtagg aaattaccat acttttgaag aagcaaaaaa aatgaaagaa aaaaaaaatc 5340

gtatttccag gttagacgtt ccgcagaatc tagaatgcgg tatgcggtac attgttcttc 5400

gaacgtaaaa gttgcgctcc ctgagatatt gtacattttt gcttttacaa gtacaagtac 5460

atcgtacaac tatgtactac tgttgatgca tccacaacag tttgttttgt tttttttttgt 5520

ttttttttttt tctaatgatt cattaccgct atgtatacct acttgtactt gtagtaagcc 5580

gggttattgg cgttcaatta atcatagact tatgaatctg cacggtgtgc gctgcgagtt 5640

acttttagct tatgcatgct acttgggtgt aatattggga tctgttcgga aatcaacgga 5700

tgctcaaccg atttcgacag taattaatta agtcatacac aagtcagctt tcttcgagcc 5760

tcatataagt ataagtagtt caacgtatta gcactgtacc cagcatctcc gtatcgagaa 5820

acacaacaac atgccccatt ggacagatca tgcggataca caggttgtgc agtatcatac 5880

atactcgatc agacaggtcg tctgaccatc atacaagctg aacaagcgct ccatacttgc 5940

acgctctcta tatacacagt taaattacat atccatagtc taacctctaa cagttaatct 6000

tctggtaagc ctcccagcca gccttctggt atcgcttggc ctcctcaata ggatctcggt 6060

tctggccgta cagacctcgg ccgacaatta tgatatccgt tccggtagac atgacatcct 6120

caacagttcg gtactgctgt ccgagagcgt ctcccttgtc gtcaagaccc accccggggg 6180

tcagaataag ccagtcctca gagtcgccct taggtcggtt ctgggcaatg aagccaacca 6240

caaactcggg gtcggatcgg gcaagctcaa tggtctgctt ggagtactcg ccagtggcca 6300

gagagccctt gcaagacagc tcggccagca tgagcagacc tctggccagc ttctcgttgg 6360

gagaggggac taggaactcc ttgtactggg agttctcgta gtcagagacg tcctccttct 6420

tctgttcaga gacagtttcc tcggcaccag ctcgcaggcc agcaatgatt ccggttccgg 6480

gtacaccgtg ggcgttggtg atatcggacc actcggcgat tcggtgacac cggtactggt 6540

gcttgacagt gttgccaata tctgcgaact ttctgtcctc gaacaggaag aaaccgtgct 6600

taagagcaag ttccttgagg gggagcacag tgccggcgta ggtgaagtcg tcaatgatgt 6660

cgatatgggt tttgatcatg cacacataag gtccgacctt atcggcaagc tcaatgagct 6720

ccttggtggt ggtaacatcc agagaagcac acaggttggt tttcttggct gccacgagct 6780

tgagcactcg agcggcaaag gcggacttgt ggacgttagc tcgagcttcg taggagggca 6840

ttttggtggt gaagaggaga ctgaaataaa tttagtctgc agaacttttt atcggaacct 6900

tatctggggc agtgaagtat atgttatggt aatagttacg agttagttga acttatagat 6960

agactggact atacggctat cggtccaaat tagaaagaac gtcaatggct ctctgggcgt 7020

cgcctttgcc gacaaaaatg tgatcatgat gaaagccagc aatgacgttg cagctgatat 7080

tgttgtcggc caaccgcgcc gaaaacgcag ctgtcagacc cacagcctcc aacgaagaat 7140

gtatcgtcaa agtgatccaa gcacactcat agttggagtc gtactccaaa ggcggcaatg 7200

acgagtcaga cagatactcg tcgacgcagt aggatgtcct gcacgggtct ttttgtgggg 7260

tgtggagaaa ggggtgcttg gagatggaag ccggtagaac cgggctgctt gtgcttggag 7320

atggaagccg gtagaaccgg gctgcttggg gggatttggg gccgctgggc tccaaagagg 7380

ggtaggcatt tcgttggggt tacgtaattg cggcatttgg gtcctgcgcg catgtcccat 7440

tggtcagaat tagtccggat aggagactta tcagccaatc acagcgccgg atccacctgt 7500

aggttgggtt gggtgggagc acccctccac agagtagagt caaacagcag cagcaacatg 7560

atagttgggg gtgtgcgtgt taaaggaaaa aaaagaagct tgggttatat tcccgctcta 7620

tttagaggtt gcgggataga cgccgacgga gggcaatggc gccatggaac cttgcggata 7680

tcgatacgcc gcggcggact gcgtccgaac cagctccagc agcgtttttt ccgggccatt 7740

gagccgactg cgaccccgcc aacgtgtctt ggcccacgca ctcatgtcat gttggtgttg 7800

ggaggccact ttttaagtag cacaaggcac ctagctcgca gcaaggtgtc cgaaccaaag 7860

aagcggctgc agtggtgcaa acggggcgga aacggcggga aaaagccacg ggggcacgaa 7920

ttgaggcacg ccctcgaatt tgagacgagt cacggcccca ttcgcccgcg caatggctcg 7980

ccaacgcccg gtcttttgca ccacatcagg ttaccccaag ccaaaccttt gtgttaaaaa 8040

gcttaacata ttataccgaa cgtaggtttg ggcgggcttg ctccgtctgt ccaaggcaac 8100

atttatataa gggtctgcat cgccggctca attgaatctt ttttcttctt ctcttctcta 8160

tattcattct tgaattaaac acacatcaat ccgc 8194

<210> 184
<211> 10838
<212> DNA
<213> Artificial Sequence

<220>
<223> Plasmid pKO2UF2PE

<220>
<221> misc_feature
<222> (1615)..(1618)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (1623)..(1623)
<223> n is a, c, g, or t

<400> 184
aaattgcccc ggagaagacg gccaggccgc ctagatgaca aattcaacaa ctcacagctg 60

actttctgcc attgccacta gggggggccc tttttatatg gccaagccaa gctctccacg 120

tcggttgggc tgcacccaac aataaatggg tagggttgca ccaacaaagg gatgggatgg 180

ggggtagaag atacgaggat aacggggctc aatggcacaa ataagaacga atactgccat 240

taagactcgt gatccagcga ctgacaccat tgcatcatct aagggcctca aaactacctc 300

ggaactgctg cgctgatctg gacaccacag aggttccgag cactttaggt tgcaccaaat 360

gtcccaccag gtgcaggcag aaaacgctgg aacagcgtgt acagtttgtc ttaacaaaaa 420

gtgagggcgc tgaggtcgag cagggtggtg tgacttgtta tagcctttag agctgcgaaa 480

gcgcgtatgg atttggctca tcaggccaga ttgagggtct gtggacacat gtcatgttag 540

tgtacttcaa tcgccccctg gatatagccc cgacaatagg ccgtggcctc atttttttgc 600

cttccgcaca tttccattgc tcgataccca caccttgctt ctcctgcact tgccaacctt 660

aatactggtt tacattgacc aacatcttac aagcgggggg cttgtctagg gtatatataa 720

acagtggctc tcccaatcgg ttgccagtct ctttttttcct ttctttcccc acagattcga 780

aatctaaact acacatcaca gaattccgag ccgtgagtat ccacgacaag atcagtgtcg 840

agacgacgcg ttttgtgtaa tgacacaatc cgaaagtcgc tagcaacaca cactctctac 900

acaaactaac ccagctctgg taccatggcg tccacttcgg ctctgcccaa gcagaaccct 960

gcgcttagac gcaccgtcac ctcaactact gtgacggatt ctgagtctgc cgccgtctct 1020

ccttcagact ctccccgcca ctcggcctct tccacatcgc tctcgtccat gtccgaggtt 1080

gatatcgcca agcccaagtc cgagtatggt gtcatgctcg acacctacgg caaccagttc 1140

gaggttcccg actttaccat caaggacatc tacaatgcca tccctaagca ctgcttcaag 1200

cgctccgctc tcaagggata cggttatatc ctccgcgaca ttgtcctcct gactaccact 1260

ttcagcatct ggtacaactt tgtgaccccc gaatatatcc cctccacccc cgcccgcgct 1320

ggtctgtggg ccgtgtacac cgttcttcag ggtcttttcg gtactggtct ctgggttatt 1380

gcccatgagt gcggtcacgg tgctttctcc gattctcgca tcatcaacga cattactggc 1440

tgggttcttc actcttccct ccttgtcccc tacttcagct ggcaaatctc ccaccgaaag 1500

caccacaagg ccactggcaa catggagcgt gacatggtct tcgttccccg aacccgcgag 1560

cagcaggcta ctcgtctcgg aaagatgacc cacgagctcg ctcatcttac tgagnnnntc 1620

gtnggctggc ccaactacct catcaccaat gttaccggcc acaactacca cgagcgccag 1680

cgtgagggtc gcggcaaggg caagcataac ggcctcggcg gtggtgttaa ccacttcgat 1740

ccccgcagcc ctctgtacga gaacagtgac gctaagctca tcgtcctcag cgatattggt 1800

atcggtctga tggccactgc tctgtacttc ctcgttcaga agttcggttt ctacaacatg 1860

gccatctggt actttgttcc ctacctctgg gttaaccact ggctcgttgc catcaccttc 1920

ctccagcaca ccgaccctac ccttccccac tacaccaacg acgagtggaa cttcgtccgt 1980

ggtgccgctg ctaccattga ccgtgagatg ggcttcatcg gccgccacct tctccacggc 2040

atcatcgaga ctcatgtcct ccaccactac gtcagcagca tccccttcta caacgcggac 2100

gaggccaccg aggccattaa gcccatcatg ggcaagcact accgggctga tgtccaggat 2160

ggtcctcgtg gcttcatccg cgccatgtac cgcagtgcgc gtatgtgcca gtgggttgag 2220

cccagcgctg gtgccgaggg tgctggtaag ggtgttctgt tcttccgcaa ccgcaacaac 2280

gtgggcaccc cccccgctgt tatcaagccc gttgcttaag taggcgcggc cgcaagtgtg 2340

gatggggaag tgagtgcccg gttctgtgtg cacaattggc aatccaagat ggatggattc 2400

aacacaggga tatagcgagc tacgtggtgg tgcgaggata tagcaacgga tatttatgtt 2460

tgacacttga gaatgtacga tacaagcact gtccaagtac aatactaaac atactgtaca 2520

tactcatact cgtacccggg caacggtttc acttgagtgc agtggctagt gctcttactc 2580

gtacagtgtg caatactgcg tatcatagtc tttgatgtat atcgtattca ttcatgttag 2640

ttgcgtacgg gtgaagcttc cactggtcgg cgtggtagtg gggcagagtg gggtcggtgt 2700

gctgcaggta ggtgatggcc acgagccagt ggttgaccca caggtagggg atcaggtagt 2760

agagggtgac ggaagccagg ccccatcggt tgatggagta tgcgatgacg gacatggtga 2820

taccaatacc gacgttagag atccagatgt tgaaccagtc cttcttctca aacagcgggg 2880

cgttggggtt gaagtggttg acagcccatt tgttgagctt ggggtacttc tgtccggtaa 2940

cgtaagacag cagatacaga ggccatccaa acacctgctg ggtgatgagg ccgtagaggg 3000

tcatgagggg agcgtcctca gcaagctcag accagtcatg ggcgcctcgg ttctccataa 3060

actcctttcg gtccttgggc acaaacacca tatcacgggt gaggtgacca gtggacttgt 3120

ggtgcatgga gtgggtcagc ttccaggcgt agtaagggac cagcatggag gagtgcagaa 3180

cccatccggt gacgttgttg acggtgttag agtcggagaa agcagagtgg ccacactcgt 3240

gggcaagaac ccacagaccg gtgccaaaca gaccctggac aatggagtac atggcccagg 3300

ccacagctcg gccggaagcc gagggaataa gaggcaggta cgcgtaggcc atgtaggcaa 3360

aaacggcgat aaagaagcag gcgcgccagc tgcattaatg aatcggccaa cgcgcgggga 3420

gaggcggttt gcgtattggg cgctcttccg cttcctcgct cactgactcg ctgcgctcgg 3480

tcgttcggct gcggcgagcg gtatcagctc actcaaaggc ggtaatacgg ttatccacag 3540

aatcagggga taacgcagga aagaacatgt gagcaaaagg ccagcaaaag gccaggaacc 3600

gtaaaaaggc cgcgttgctg gcgtttttcc ataggctccg ccccccctgac gagcatcaca 3660

aaaatcgacg ctcaagtcag aggtggcgaa acccgacagg actataaaga taccaggcgt 3720

ttcccccctgg aagctccctc gtgcgctctc ctgttccgac cctgccgctt accggatacc 3780

tgtccgcctt tctcccttcg ggaagcgtgg cgctttctca tagctcacgc tgtaggtatc 3840

tcagttcggt gtaggtcgtt cgctccaagc tgggctgtgt gcacgaaccc cccgttcagc 3900

ccgaccgctg cgccttatcc ggtaactatc gtcttgagtc caacccggta agacacgact 3960

tatcgccact ggcagcagcc actggtaaca ggattagcag agcgaggtat gtaggcggtg 4020

ctacagagtt cttgaagtgg tggcctaact acggctacac tagaagaaca gtatttggta 4080

tctgcgctct gctgaagcca gttaccttcg gaaaaagagt tggtagctct tgatccggca 4140

aacaaaccac cgctggtagc ggtggttttt ttgtttgcaa gcagcagatt acgcgcagaa 4200

aaaaaggatc tcaagaagat cctttgatct tttctacggg gtctgacgct cagtggaacg 4260

aaaactcacg ttaagggatt ttggtcatga gattatcaaa aaggatcttc acctagatcc 4320

ttttaaatta aaaatgaagt tttaaatcaa tctaaagtat atatgagtaa acttggtctg 4380

acagttacca atgcttaatc agtgaggcac ctatctcagc gatctgtcta tttcgttcat 4440

ccatagttgc ctgactcccc gtcgtgtaga taactacgat acgggagggc ttaccatctg 4500

gccccagtgc tgcaatgata ccgcgagacc cacgctcacc ggctccagat ttatcagcaa 4560

taaaccagcc agccggaagg gccgagcgca gaagtggtcc tgcaacttta tccgcctcca 4620

tccagtctat taattgttgc cgggaagcta gagtaagtag ttcgccagtt aatagtttgc 4680

gcaacgttgt tgccattgct acaggcatcg tggtgtcacg ctcgtcgttt ggtatggctt 4740

cattcagctc cggttcccaa cgatcaaggc gagttacatg atcccccatg ttgtgcaaaa 4800

aagcggttag ctccttcggt cctccgatcg ttgtcagaag taagttggcc gcagtgttat 4860

cactcatggt tatggcagca ctgcataatt ctcttactgt catgccatcc gtaagatgct 4920

tttctgtgac tggtgagtac tcaaccaagt cattctgaga atagtgtatg cggcgaccga 4980

gttgctcttg cccggcgtca atacgggata ataccgcgcc acatagcaga actttaaaag 5040

tgctcatcat tggaaaacgt tcttcggggc gaaaactctc aaggatctta ccgctgttga 5100

gatccagttc gatgtaaccc actcgtgcac ccaactgatc ttcagcatct tttactttca 5160

ccagcgtttc tgggtgagca aaaacaggaa ggcaaaatgc cgcaaaaaag ggaataaggg 5220

cgacacggaa atgttgaata ctcatactct tccttttttca atattattga agcatttatc 5280

agggttattg tctcatgagc ggatacatat ttgaatgtat ttagaaaaat aaacaaatag 5340

gggttccgcg cacatttccc cgaaaagtgc cacctgatgc ggtgtgaaat accgcacaga 5400

tgcgtaagga gaaaataccg catcaggaaa ttgtaagcgt taatattttg ttaaaattcg 5460

cgttaaattt ttgttaaatc agctcatttt ttaaccaata ggccgaaatc ggcaaaatcc 5520

cttataaatc aaaagaatag accgagatag ggttgagtgt tgttccagtt tggaacaaga 5580

gtccactatt aaagaacgtg gactccaacg tcaaagggcg aaaaaccgtc tatcagggcg 5640

atggcccact acgtgaacca tcaccctaat caagtttttt ggggtcgagg tgccgtaaag 5700

cactaaatcg gaaccctaaa gggagccccc gatttagagc ttgacgggga aagccggcga 5760

acgtggcgag aaaggaaggg aagaaagcga aaggagcggg cgctagggcg ctggcaagtg 5820

tagcggtcac gctgcgcgta accaccacac ccgccgcgct taatgcgccg ctacagggcg 5880

cgtccattcg ccattcaggc tgcgcaactg ttgggaaggg cgatcggtgc gggcctcttc 5940

gctattacgc cagctggcga aaggggggatg tgctgcaagg cgattaagtt gggtaacgcc 6000

agggttttcc cagtcacgac gttgtaaaac gacggccagt gaattgtaat acgactcact 6060

atagggcgaa ttgggcccga cgtcgcatgc ttgaatctac aagtaggagg gttggagtga 6120

ttaagtgaaa cttctttaac ggctctatgc cagttctatt gatatccgaa acatcagtat 6180

gaaggtctga taagggtgac ttcttcccac agattcgtat cagtacgagt acgagaccgg 6240

tacttgtaac agtattgata ctaaagggaa actacaacgg ttgtcagcgt aatgtgactt 6300

cgcccatgaa cgcagacacg cagtgccgag tgcggtgata tcgcctactc gttacgtcca 6360

tggactacac aacccctcgg cttcgcttgg cttagcctcg ggctcggtgc tgttcagtta 6420

aaacacaatc aaataacatt tctacttttt agaaggcagg ccgtcaggag caactccgac 6480

tccattgacg tttctaaaca tctgaatgcc ttccttacct tcaacaaact ggcaggttcg 6540

ggcgacagtg taaagagact tgatgaagtt ggtgtcgtcg tgtcggtagt gcttgcccat 6600

gaccttcttg atcttctcag tggcgattcg ggcgttgtag aagggaattc cgtcgtcgcc 6660

tgagtcgacg agtatctgtc tgactcgtca ttgccgcctt tggagtacga ctccaactat 6720

gagtgtgctt ggatcacttt gacgatacat tcttcgttgg aggctgtggg tctgacagct 6780

gcgttttcgg cgcggttggc cgacaacaat atcagctgca acgtcattgc tggctttcat 6840

catgatcaca tttttgtcgg caaaggcgac gcccagagag ccattgacgt tctttctaat 6900

ttggaccgat agccgtatag tccagtctat ctataagttc aactaactcg taactattac 6960

cataacatat acttcactgc cccagataag gttccgataa aaagttctgc agactaaatt 7020

tatttcagtc tcctcttcac caccaaaatg ccctcctacg aagctcgagc taacgtccac 7080

aagtccgcct ttgccgctcg agtgctcaag ctcgtggcag ccaagaaaac caacctgtgt 7140

gcttctctgg atgttaccac caccaaggag ctcattgagc ttgccgataa ggtcggacct 7200

tatgtgtgca tgatcaaaac ccatatcgac atcattgacg acttcaccta cgccggcact 7260

gtgctccccc tcaaggaact tgctcttaag cacggtttct tcctgttcga ggacagaaag 7320

ttcgcagata ttggcaacac tgtcaagcac cagtaccggt gtcaccgaat cgccgagtgg 7380

tccgatatca ccaacgccca cggtgtaccc ggaaccggaa tcattgctgg cctgcgagct 7440

ggtgccgagg aaactgtctc tgaacagaag aaggaggacg tctctgacta cgagaactcc 7500

cagtacaagg agttcctagt ccctctccc aacgagaagc tggccagagg tctgctcatg 7560

ctggccgagc tgtcttgcaa gggctctctg gccactggcg agtactccaa gcagaccatt 7620

gagcttgccc gatccgaccc cgagtttgtg gttggcttca ttgcccagaa ccgacctaag 7680

ggcgactctg aggactggct tattctgacc cccggggtgg gtcttgacga caagggagac 7740

gctctcggac agcagtaccg aactgttgag gatgtcatgt ctaccggaac ggatatcata 7800

attgtcggcc gaggtctgta cggccagaac cgagatccta ttgaggaggc caagcgatac 7860

cagaaggctg gctgggaggc ttaccagaag attaactgtt agaggttaga ctatggatat 7920

gtaatttaac tgtgtatata gagagcgtgc aagtatggag cgcttgttca gcttgtatga 7980

tggtcagacg acctgtctga tcgagtatgt atgatactgc acaacctgtg tatccgcatg 8040

atctgtccaa tggggcatgt tgttgtgttt ctcgatacgg agatgctggg tacagtgcta 8100

atacgttgaa ctacttatac ttatatgagg ctcgaagaaa gctgacttgt gtatgactta 8160

attaatttga atcgaatcga tgagcctaaa atgaacccga gtatatctca taaaattctc 8220

ggtgagaggt ctgtgactgt cagtacaagg tgccttcatt atgccctcaa ccttaccata 8280

cctcactgaa tgtagtgtac ctctaaaaat gaaatacagt gccaaaagcc aaggcactga 8340

gctcgtctaa cggacttgat atacaaccaa ttaaaacaaa tgaaaagaaa tacagttctt 8400

tgtatcattt gtaacaatta ccctgtacaa actaaggtat tgaaatccca caatattccc 8460

aaagtccacc cctttccaaa ttgtcatgcc tacaactcat ataccaagca ctaacctacc 8520

gtttaaacag tgtacgcaga tcccgtcaac agttttatat atcgtagtta caaccatcaa 8580

cacttttgg taagtgtacc attctatact ccaactggtc tgcaactgta caagtagaca 8640

tgttaatggt agttaataac atctacagca gaacctatgg taaagacatt gcatttttac 8700

aggaagtatc gtcctacacg ttgataaatc caaagatgcg gaacttcttc cacttttatc 8760

atcatcccct actcgtacac tcgtactctt tgttcgatcg cgattcattt ctataaataa 8820

tcttgtatgt acatgcggcc gcttactgga gctttctggc cttctccttg gcagcgtcag 8880

ccttggcctg cttggcgagc ttggcgttct ttcggtaaaa gttgtagaag agaccgagca 8940

tggtccacat gtagaaccag agcagagcgg tgatgaagaa ggggtatcca ggtcggccaa 9000

ggaccttcat ggcgtacatg tcccaggaag actggacaga catcatgcag aactgggtca 9060

tctgggatcg agtgatgtag aacttgatga acgacacctg cttgaagccc agggcagaca 9120

gaaagtagta gccgtacatg atgacgtgga tgaaggagtt cagggcagca gagaagtagg 9180

cttcaccgtt gggagcaacg aaggtgacca gccaccagat ggtgaagatg gaagagtggt 9240

ggtacacgtg cagaaaggaa atctgtcggt tgttcttctt gaggaccatg atcatggtgt 9300

cgacaaactc catgatcttg gagaagtaga agagccagat catcttagcc atagggagac 9360

ccttgaaggt gtgatcggca gcgttctcaa acagtccata gttggcctga taagcctcgt 9420

acaggatgcc accgcacatg taggcggaga tggagaccag acagaagttg tgcaggaggg 9480

agaaggtctt gacctcgaat cgttcaaagt tcttcatgat ctgcataccc acaaacacgg 9540

tgaccaggta ggcgagcacg atcaggagca cgtggaaggg gttcatcaga ggcagctctc 9600

gagccagggg agactccacg gcaaccagga agcctcgagt gtgatggaca atggtgggaa 9660

tgtacttctc ggcctgggca accagggcag cctccagggg atcgacgtag ggagcagctc 9720

ggacaccgat agcgctggcg aggtccatga acaggtcctg aggcatcttg gagggcagga 9780

agggagcaat ggactccatg gttagcgtgt cgtgtttttg ttgtgctgga agaaccaaag 9840

ggtggcgcaa tgtgtgtaga tatatatgtc gtgacccaca agtcacacaa acaagtatcg 9900

ggaggagtgg tgcacctcta tgcggagaaa ccttataccg ctgtagacca actggggcag 9960

aggtgtgagt tgaagtcagc tggaggagat gtgtgacaga agcacaagaa gtgagattgt 10020

gagatgtatg tctaggggggg gaagttttgt gtcaaatata tgggaattat tatcagcacc 10080

acgaaattat acgcctcata tgacccattt aggtggatag atcatggaca ctgttgacag 10140

ctgcgaagaa aaagcgtatt ggggatgatc cgaaattagt ccggtaccga ggcgcaaata 10200

cgtaagacag ccgatwaaat atatgcgaga aacaccaaag agactctaga tgtttgtttg 10260

gcacagtttt gacttctgcg aaggccttac accaccttgt tgacccttgt cgcgggtcgg 10320

gcaatatcgg ctgacagagt tttacttgct caataagata cgagctgcat agagttgaac 10380

tacaggacaa tattggggct ggccacatga agggcattgt ttggaggtgt attgatggtg 10440

aaaacacgat atgaaatgac aacgcccccc gttttattat tattcttatt attttgggtg 10500

cttctctatc catacaagca cctcctaaca tgcttcataa gtgacctcct catcacaagg 10560

cctgaggtct catttatcca gtggcgccaa gctaaactaa aactggtccg agtagactaa 10620

ggcgaagaga gaaggagaga agacagtttt tttgtggccg cctgtgaaca atgaaaacga 10680

tgagggtgag atggagcaaa ccatatggac agtcagagga gtacacgctg cttacataat 10740

ggcgcaacga ccacatgtcc cacagatacg cattatgcct gtacatattc cgggggaggt 10800

atgtaccagt agttcgcctg ctaccgttag ctacattt 10838

<210> 185
<211> 6912
<212> DNA
<213> Artificial Sequence

<220>
<223> Plasmid pZKUGPI5S

<400> 185
ggccgcaagt gtggatgggg aagtgagtgc ccggttctgt gtgcacaatt ggcaatccaa 60

gatggatgga ttcaacacag ggatatagcg agctacgtgg tggtgcgagg atatagcaac 120

ggatatttat gtttgacact tgagaatgta cgatacaagc actgtccaag tacaatacta 180

aacatactgt acatactcat actcgtaccc gggcaacggt ttcacttgag tgcagtggct 240

agtgctctta ctcgtacagt gtgcaatact gcgtatcata gtctttgatg tatatcgtat 300

tcattcatgt tagttgcgta cgaggaaact gtctctgaac agaagaagga ggacgtctct 360

gactacgaga actcccagta caaggagttc ctagtcccct ctcccaacga gaagctggcc 420

agaggtctgc tcatgctggc cgagctgtct tgcaagggct ctctggccac tggcgagtac 480

tccaagcaga ccattgagct tgcccgatcc gaccccgagt ttgtggttgg cttcattgcc 540

cagaaccgac ctaagggcga ctctgaggac tggcttattc tgacccccgg ggtgggtctt 600

gacgacaagg gagacgctct cggacagcag taccgaactg ttgaggatgt catgtctacc 660

ggaacggata tcataattgt cggccgaggt ctgtacggcc agaaccgaga tcctattgag 720

gaggccaagc gataccagaa ggctggctgg gaggcttacc agaagattaa ctgttagagg 780

ttagactatg gatatgtaat ttaactgtgt atatagagag cgtgcaagta tggagcgctt 840

gttcagcttg tatgatggtc agacgacctg tctgatcgag tatgtatgat actgcacaac 900

ctgtgtatcc gcatgatctg tccaatgggg catgttgttg tgtttctcga tacggagatg 960

ctgggtacag tgctaatacg ttgaactact tatacttata tgaggctcga agaaagctga 1020

cttgtgtatg acttaattaa tcgagcttgg cgtaatcatg gtcatagctg tttcctgtgt 1080

gaaattgtta tccgctcaca attccacaca acatacgagc cggaagcata aagtgtaaag 1140

cctggggtgc ctaatgagtg agctaactca cattaattgc gttgcgctca ctgcccgctt 1200

tccagtcggg aaacctgtcg tgccagctgc attaatgaat cggccaacgc gcggggagag 1260

gcggtttgcg tattgggcgc tcttccgctt cctcgctcac tgactcgctg cgctcggtcg 1320

ttcggctgcg gcgagcggta tcagctcact caaaggcggt aatacggtta tccacagaat 1380

caggggataa cgcaggaaag aacatgtgag caaaaggcca gcaaaaggcc aggaaccgta 1440

aaaaggccgc gttgctggcg tttttccata ggctccgccc ccctgacgag catcacaaaa 1500

atcgacgctc aagtcagagg tggcgaaacc cgacaggact ataaagatac caggcgtttc 1560

cccctggaag ctccctcgtg cgctctcctg ttccgaccct gccgcttacc ggatacctgt 1620

ccgcctttct cccttcggga agcgtggcgc tttctcatag ctcacgctgt aggtatctca 1680

gttcggtgta ggtcgttcgc tccaagctgg gctgtgtgca cgaacccccc gttcagcccg 1740

accgctgcgc cttatccggt aactatcgtc ttgagtccaa cccggtaaga cacgacttat 1800

cgccactggc agcagccact ggtaacagga ttagcagagc gaggtatgta ggcggtgcta 1860

cagagttctt gaagtggtgg cctaactacg gctacactag aaggacagta tttggtatct 1920

gcgctctgct gaagccagtt accttcggaa aaagagttgg tagctcttga tccggcaaac 1980

aaaccaccgc tggtagcggt ggtttttttg tttgcaagca gcagattacg cgcagaaaaa 2040

aaggatctca agaagatcct ttgatctttt ctacggggtc tgacgctcag tggaacgaaa 2100

actcacgtta agggattttg gtcatgagat tatcaaaaag gatcttcacc tagatccttt 2160

taaattaaaa atgaagtttt aaatcaatct aaagtatata tgagtaaact tggtctgaca 2220

gttaccaatg cttaatcagt gaggcaccta tctcagcgat ctgtctattt cgttcatcca 2280

tagttgcctg actccccgtc gtgtagataa ctacgatacg ggagggctta ccatctggcc 2340

ccagtgctgc aatgataccg cgagacccac gctcaccggc tccagattta tcagcaataa 2400

accagccagc cggaagggcc gagcgcagaa gtggtcctgc aactttatcc gcctccatcc 2460

agtctattaa ttgttgccgg gaagctagag taagtagttc gccagttaat agtttgcgca 2520

acgttgttgc cattgctaca ggcatcgtgg tgtcacgctc gtcgtttggt atggcttcat 2580

tcagctccgg ttcccaacga tcaaggcgag ttacatgatc ccccatgttg tgcaaaaaag 2640

cggttagctc cttcggtcct ccgatcgttg tcagaagtaa gttggccgca gtgttatcac 2700

tcatggttat ggcagcactg cataattctc ttactgtcat gccatccgta agatgctttt 2760

ctgtgactgg tgagtactca accaagtcat tctgagaata gtgtatgcgg cgaccgagtt 2820

gctcttgccc ggcgtcaata cgggataata ccgcgccaca tagcagaact ttaaaagtgc 2880

tcatcattgg aaaacgttct tcggggcgaa aactctcaag gatcttaccg ctgttgagat 2940

ccagttcgat gtaacccact cgtgcaccca actgatcttc agcatctttt actttcacca 3000

gcgtttctgg gtgagcaaaa acaggaaggc aaaatgccgc aaaaaaggga ataagggcga 3060

cacggaaatg ttgaatactc atactcttcc tttttcaata ttattgaagc atttatcagg 3120

gttattgtct catgagcgga tacatatttg aatgtattta gaaaaataaa caaatagggg 3180

ttccgcgcac atttccccga aaagtgccac ctgacgcgcc ctgtagcggc gcattaagcg 3240

cggcgggtgt ggtggttacg cgcagcgtga ccgctacact tgccagcgcc ctagcgcccg 3300

ctcctttcgc tttcttccct tcctttctcg ccacgttcgc cggctttccc cgtcaagctc 3360

taaatcgggg gctcccttta gggttccgat ttagtgcttt acggcacctc gaccccaaaa 3420

aacttgatta gggtgatggt tcacgtagtg ggccatcgcc ctgatagacg gtttttcgcc 3480

ctttgacgtt ggagtccacg ttctttaata gtggactctt gttccaaact ggaacaacac 3540

tcaaccctat ctcggtctat tcttttgatt tataagggat tttgccgatt tcggcctatt 3600

ggttaaaaaa tgagctgatt taacaaaaat ttaacgcgaa ttttaacaaa atattaacgc 3660

ttacaatttc cattcgccat tcaggctgcg caactgttgg gaagggcgat cggtgcgggc 3720

ctcttcgcta ttacgccagc tggcgaaagg gggatgtgct gcaaggcgat taagttgggt 3780

aacgccaggg ttttcccagt cacgacgttg taaaacgacg gccagtgaat tgtaatacga 3840

ctcactatag ggcgaattgg gtaccgggcc ccccctcgag gtcgacgagt atctgtctga 3900

ctcgtcattg ccgcctttgg agtacgactc caactatgag tgtgcttgga tcactttgac 3960

gatacattct tcgttggagg ctgtgggtct gacagctgcg ttttcggcgc ggttggccga 4020

caacaatatc agctgcaacg tcattgctgg ctttcatcat gatcacattt ttgtcggcaa 4080

aggcgacgcc cagagagcca ttgacgttct ttctaatttg gaccgatagc cgtatagtcc 4140

agtctatcta taagttcaac taactcgtaa ctattaccat aacatatact tcactgcccc 4200

agataaggtt ccgataaaaa gttctgcaga ctaaatttat ttcagtctcc tcttcaccac 4260

caaaatgccc tcctacgaag ctcgagtgct caagctcgtg gcagccaaga aaaccaacct 4320

gtgtgcttct ctggatgtta ccaccaccaa ggagctcatt gagcttgccg ataaggtcgg 4380

accttatgtg tgcatgatca aaacccatat cgacatcatt gacgacttca cctacgccgg 4440

cactgtgctc cccctcaagg aacttgctct taagcacggt ttcttcctgt tcgaggacag 4500

aaagttcgca gatattggca acactgtcaa gcaccagtac cggtgtcacc gaatcgccga 4560

gtggtccgat atcaccaacg cccacggtgt acccggaacc ggaatcgatg cgtatctgtg 4620

ggacatgtgg tcgttgcgcc attatgtaag cagcgtgtac tcctctgact gtccatatgg 4680

tttgctccat ctcaccctca tcgttttcat tgttcacagg cggccacaaa aaaactgtct 4740

tctctccttc tctcttcgcc ttagtctact cggaccagtt ttagtttagc ttggcgccac 4800

tggataaatg agacctcagg ccttgtgatg aggaggtcac ttatgaagca tgttaggagg 4860

tgcttgtatg gatagagaag cacccaaaat aataagaata ataataaaac agggggcgtt 4920

gtcatttcat atcgtgtttt caccatcaat acacctccaa acaatgccct tcatgtggcc 4980

agccccaata ttgtcctgta gttcaactct atgcagctcg tatcttattg agcaagtaaa 5040

actctgtcag ccgatattgc ccgacccgcg acaagggtca acaaggtggt gtaaggcctt 5100

cgcagaagtc aaaactgtgc caaacaaaca tctagagtct ctttggtgtt tctcgcatat 5160

atttwatcgg ctgtcttacg tatttgcgcc tcggtaccgg actaatttcg gatcatcccc 5220

aatacgcttt ttcttcgcag ctgtcaacag tgtccatgat ctatccacct aaatgggtca 5280

tatgaggcgt ataatttcgt ggtgctgata ataattccca tatatttgac acaaaacttc 5340

ccccccctaga catacatctc acaatctcac ttcttgtgct tctgtcacac atctcctcca 5400

gctgacttca actcacacct ctgccccagt tggtctacag cggtataagg tttctccgca 5460

tagaggtgca ccactcctcc cgatacttgt ttgtgtgact tgtgggtcac gacatatata 5520

tctacacaca ttgcgccacc ctttggttct tccagcacaa caaaaacacg acacgctaac 5580

catggtcgct ggcaagtccg gagctgcagc ccacgtgacc cactcttcca ctctccctcg 5640

agagtaccac ggtgctacca acgactcccg atctgaggct gccgatgtca ccgtctcgtc 5700

tatcgacgcc gagaaggaga tgattatcaa cggacgagtg tacgacgtct cctcgttcgt 5760

gaagcgacac cctggtggct ccgtcatcaa gttccagctc ggagcagatg cttctgacgc 5820

ttacaacaac ttccacgtcc gatcgaagaa ggctgacaag atgctgtact ctcttccctc 5880

ccgacctgcc gaggctggct atgcacagga cgacatctct agagactttg agaagctgcg 5940

actggaactg aaagaggaag gttacttcga gcccaatctg gtgcacgtgt cctaccgatg 6000

tgtcgaggtg cttgccatgt actgggctgg cgtccagctg atctggtccg gatactggtt 6060

cctcggtgcc atcgttgctg gaattgctca aggtcgatgc ggatggctcc agcatgaagg 6120

cggacactac tcgctcactg gcaacatcaa gattgaccga catctccaga tggccatcta 6180

tggactgggc tgtggtatgt ctggctgcta ctggagaaac cagcacaaca aacatcacgc 6240

cactcctcag aagctcggaa ccgatcccga cctgcagacc atgcctctcg ttgccttcca 6300

caagattgtc ggagccaagg cacgaggcaa gggtaaagcc tggcttgctt ggcaagctcc 6360

cctcttcttt ggaggcatca tttgctccct ggtctctttc ggctggcagt tcgttctcca 6420

ccccaatcat gcactgcgag tgcacaacca tctcgaactg gcctacatgg gtctccgata 6480

cgttctctgg caccttgcct ttggccatct gggactcctg tcctctcttc gactgtatgc 6540

cttctacgtg gctgtcggtg gcacctacat cttcaccaac ttcgccgtct cccatactca 6600

caaggatgtc gttcctccca ccaagcacat ttcgtgggct ctgtactctg ccaaccacac 6660

taccaactgt tccgactctc cctttgtcaa ctggtggatg gcctacctca acttccagat 6720

cgagcaccat ctgttcccct ccatgcctca gtacaaccac cccaagattg ctcctcgagt 6780

gcgagcactc ttcgagaagc acggagtcga gtacgacgtc cgaccctatc tggaatgctt 6840

tcgagtgacc tacgtcaacc tccttgctgt tggcaaccct gagcactcct accacgagca 6900

tactcactaa gc 6912

<210> 186
<211> 14864
<212> DNA
<213> Artificial Sequence

<220>
<223> Plasmid pDMW302T16

<220>
<221> misc_feature
<222> (5341)..(5341)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (5346)..(5349)
<223> n is a, c, g, or t

<400> 186
gtacgatagt tagtagacaa caatcagaac atctccctcc ttatataatc acacaggcca 60

gaacgcgcta aactaaagcg ctttggacac tatgttacat tggcattgat tgaactgaaa 120

ccacagtctc cctcgcctga atcgagcaat ggatgttgtc ggaagtcaac ttcactagaa 180

gagcggttct atgccttgtc aagatcatat cataaactca ctctgtatta ccccatctat 240

agaacacttg ttatgaatgg gcggaaacat tccgctatat gcacctttcc acactaatgc 300

aaagatgtgc atcttcaacg ggtagtaaga ctggttccga cttccgttgc atggagagca 360

atgacctcga taatgcgaac atcccccaca tatacactct tacacaggcc aatataatct 420

gtgcatttac taaatattta agtctatgca cctgcttgat gaaaagcggc acggatggta 480

tcatctagtt tccgccaatc caagaaccaa ctgtgttggc agtggtgtag cccatggcac 540

acagaccaaa gatgaaaata cagacatcgg cggttcgagc cgtggtgcct cgagcaacac 600

ccttgtaatg caaaagagga gggtaaatgt acaccagagg cacacatgca aacgatccgg 660

tgagagcgac gaaccgatcg agatcgtcgg cacctcccca tgcaacaaag gcggtgacaa 720

acacaaggaa gaaccggaaa atgttcttct gccacttgat ggtagagttg tacttgcctg 780

atcgggtgaa gagaccattc tcgatgattc ggatggcgcg ccagctgcat taatgaatcg 840

gccaacgcgc ggggagaggc ggtttgcgta ttgggcgctc ttccgcttcc tcgctcactg 900

actcgctgcg ctcggtcgtt cggctgcggc gagcggtatc agctcactca aaggcggtaa 960

tacggttatc cacagaatca ggggataacg caggaaagaa catgtgagca aaaggccagc 1020

aaaaggccag gaaccgtaaa aaggccgcgt tgctggcgtt tttccatagg ctccgccccc 1080

ctgacgagca tcacaaaaat cgacgctcaa gtcagaggtg gcgaaacccg acaggactat 1140

aaagatacca ggcgtttccc cctggaagct ccctcgtgcg ctctcctgtt ccgaccctgc 1200

cgcttaccgg atacctgtcc gcctttctcc cttcgggaag cgtggcgctt tctcatagct 1260

cacgctgtag gtatctcagt tcggtgtagg tcgttcgctc caagctgggc tgtgtgcacg 1320

aacccccccgt tcagcccgac cgctgcgcct tatccggtaa ctatcgtctt gagtccaacc 1380

cggtaagaca cgacttatcg ccactggcag cagccactgg taacaggatt agcagagcga 1440

ggtatgtagg cggtgctaca gagttcttga agtggtggcc taactacggc tacactagaa 1500

gaacagtatt tggtatctgc gctctgctga agccagttac cttcggaaaa agagttggta 1560

gctcttgatc cggcaaacaa accaccgctg gtagcggtgg ttttttttgtt tgcaagcagc 1620

agattacgcg cagaaaaaaa ggatctcaag aagatccttt gatcttttct acggggtctg 1680

acgctcagtg gaacgaaaac tcacgttaag ggattttggt catgagatta tcaaaaagga 1740

tcttcaccta gatcctttta aattaaaaat gaagttttaa atcaatctaa agtatatatg 1800

agtaaacttg gtctgacagt taccaatgct taatcagtga ggcacctatc tcagcgatct 1860

gtctatttcg ttcatccata gttgcctgac tccccgtcgt gtagataact acgatacggg 1920

agggcttacc atctggcccc agtgctgcaa tgataccgcg agacccacgc tcaccggctc 1980

cagatttatc agcaataaac cagccagccg gaagggccga gcgcagaagt ggtcctgcaa 2040

ctttatccgc ctccatccag tctattaatt gttgccggga agctagagta agtagttcgc 2100

cagttaatag tttgcgcaac gttgttgcca ttgctacagg catcgtggtg tcacgctcgt 2160

cgtttggtat ggcttcattc agctccggtt cccaacgatc aaggcgagtt acatgatccc 2220

ccatgttgtg caaaaaagcg gttagctcct tcggtcctcc gatcgttgtc agaagtaagt 2280

tggccgcagt gttatcactc atggttatgg cagcactgca taattctctt actgtcatgc 2340

catccgtaag atgcttttct gtgactggtg agtactcaac caagtcattc tgagaatagt 2400

gtatgcggcg accgagttgc tcttgcccgg cgtcaatacg ggataatacc gcgccacata 2460

gcagaacttt aaaagtgctc atcattggaa aacgttcttc ggggcgaaaa ctctcaagga 2520

tcttaccgct gttgagatcc agttcgatgt aacccactcg tgcacccaac tgatcttcag 2580

catcttttac tttcaccagc gtttctgggt gagcaaaaac aggaaggcaa aatgccgcaa 2640

aaaagggaat aagggcgaca cggaaatgtt gaatactcat actcttcctt tttcaatatt 2700

attgaagcat ttatcagggt tattgtctca tgagcggata catatttgaa tgtatttaga 2760

aaaataaaca aatagggttt ccgcgcacat ttccccgaaa agtgccacct gatgcggtgt 2820

gaaataccgc acagatgcgt aaggagaaaa taccgcatca ggaaattgta agcgttaata 2880

ttttgttaaa attcgcgtta aatttttgtt aaatcagctc atttttaac caataggccg 2940

aaatcggcaa aatcccttat aaatcaaaag aatagaccga gatagggttg agtgttgttc 3000

cagtttggaa caagagtcca ctattaaaga acgtggactc caacgtcaaa gggcgaaaaa 3060

ccgtctatca gggcgatggc ccactacgtg aaccatcacc ctaatcaagt tttttggggt 3120

cgaggtgccg taaagcacta aatcggaacc ctaaagggag cccccgattt agagcttgac 3180

ggggaaagcc ggcgaacgtg gcgagaaagg aagggaagaa agcgaaagga gcgggcgcta 3240

gggcgctggc aagtgtagcg gtcacgctgc gcgtaaccac cacacccgcc gcgcttaatg 3300

cgccgctaca gggcgcgtcc attcgccatt caggctgcgc aactgttggg aagggcgatc 3360

ggtgcgggcc tcttcgctat tacgccagct ggcgaaaggg ggatgtgctg caaggcgatt 3420

aagttgggta acgccagggt tttcccagtc acgacgttgt aaaacgacgg ccagtgaatt 3480

gtaatacgac tcactatagg gcgaattggg cccgacgtcg catgcttaga agtgaggatt 3540

acaagaagcc tctggatatc aatgatgaac gtactcagcg gctggtcaag catttcgacc 3600

gtcgaatcga cgaggtgttc acctttgaca agcgagggtt cccaattgat cacgttctcg 3660

agttgttcaa atcttctctc aacatctctc tgcatgaact atctctgttg acgaacgtgt 3720

cacccactgt tcctcgaacg cccttctccg agtttggtct gaacatcttc gatctcaaac 3780

tgaccccccgc agtgatcaat agtgccatgc cactgccgat gcggtgcgaa catccctgga 3840

gggattctcg gagctctaca caatgcagat tctgtcgtcg agtactctct accttgctcg 3900

aatgacttat tgtgctacta ctgcactcat gcttcgatca tgtgccctac tgcaccccaa 3960

atttggtgat ctgattgaga cagagtaccc tcttcagctg attcagaaga tcatcagcaa 4020

catgaatgat gtggttgacc aggcaggctg ttgtagtcac gtccttcact tcaagttcat 4080

tcttcatctg cttctgtttt actttgacag gcaaatgaag acatggtacg acttgatgga 4140

ggccaagaac gccatttcac cccgagacac cgaagtgcct gaaatcctgg ctgcccccat 4200

tgataacatc ggaaactacg gtattccgga aagtgtatat agaacctttc cccagcttgt 4260

gtctgtggat atggatggtg taatcccctt aattaaaagg cgttgaaaca gaatgagcca 4320

gacagcaagg acaaggtggc caacagcaag gagtccaaaa agccctctat tgacgagatc 4380

cacgatgtta ttgctcatga ggtttccgag ctcgatgctg ggaagaagaa gtgatttgta 4440

tataagaaat aaatgagata tagtaaagga gtgcaagaga atggcaaggt ggtcaaattc 4500

tatattactt gcagtcactg gttcctcgtt gacatgaatg aagttaccgt tggcatagct 4560

gatttaatat ataactgtcc aactaactct cacctagata taacccatgt gtgtgtttcc 4620

aatcatcaat gcggccgcgc ctacttaagc aacgggcttg ataacagcgg ggggggtgcc 4680

cacgttgttg cggttgcgga agaacagaac acccttacca gcaccctcgg caccagcgct 4740

gggctcaacc cactggcaca tacgcgcact gcggtacatg gcgcggatga agccacgagg 4800

accatcctgg acatcagccc ggtagtgctt gcccatgatg ggcttaatgg cctcggtggc 4860

ctcgtccgcg ttgtagaagg ggatgctgct gacgtagtgg tggaggacat gagtctcgat 4920

gatgccgtgg agaaggtggc ggccgatgaa gcccatctca cggtcaatgg tagcagcggc 4980

accacggacg aagttccact cgtcgttggt gtagtgggga agggtagggt cggtgtgctg 5040

gaggaaggtg atggcaacga gccagtggtt aacccagagg tagggaacaa agtaccagat 5100

ggccatgttg tagaaaccga acttctgaac gaggaagtac agagcagtgg ccatcagacc 5160

gataccaata tcgctgagga cgatgagctt agcgtcactg ttctcgtaca gagggctgcg 5220

gggatcgaag tggttaacac caccgccgag gccgttatgc ttgcccttgc cgcgaccctc 5280

acgctggcgc tcgtggtagt tgtggccggt aacattggtg atgaggtagt tgggccagcc 5340

nacgannnnc tcagtaagat gagcgagctc gtgggtcatc tttccgagac gagtagcctg 5400

ctgctcgcgg gttcggggaa cgaagaccat gtcacgctcc atgttgccag tggccttgtg 5460

gtgctttcgg tgggagattt gccagctgaa gtaggggaca aggagggaag agtgaagaac 5520

ccagccagta atgtcgttga tgatgcgaga atcggagaaa gcaccgtgac cgcactcatg 5580

ggcaataacc cagagaccag taccgaaaag accctgaaga acggtgtaca cggcccacag 5640

accagcgcgg gcgggggtgg aggggatata ttcgggggtc acaaagttgt accagatgct 5700

gaaagtggta gtcaggagga caatgtcgcg gaggatataa ccgtatccct tgagagcgga 5760

gcgcttgaag cagtgcttag ggatggcatt gtagatgtcc ttgatggtaa agtcgggaac 5820

ctcgaactgg ttgccgtagg tgtcgagcat gacaccatac tcggacttgg gcttggcgat 5880

atcaacctcg gacatggacg agagcgatgt ggaagaggcc gagtggcggg gagagtctga 5940

aggagagacg gcggcagact cagaatccgt cacagtagtt gaggtgacgg tgcgtctaag 6000

cgcagggttc tgcttgggca gagccgaagt ggacgccatg gtgaatgatt cttatactca 6060

gaaggaaatg cttaacgatt tcgggtgtga gttgacaagg agagagagaa aagaagagga 6120

aaggtaattc ggggacggtg gtcttttata cccttggcta aagtcccaac cacaaagcaa 6180

aaaaattttc agtagtctat tttgcgtccg gcatgggtta cccggatggc cagacaaaga 6240

aactagtaca aagtctgaac aagcgtagat tccagactgc agtaccctac gcccttaacg 6300

gcaagtgtgg gaaccggggg aggtttgata tgtggggtga agggggctct cgccggggtt 6360

gggcccgcta ctgggtcaat ttggggtcaa ttggggcaat tggggctgtt ttttgggaca 6420

caaatacgcc gccaacccgg tctctcctga attctctctc ttgagctttt ccataacaag 6480

ttcttctgcc tccaggaagt ccatgggtgg tttgatcatg gttttggtgt agtggtagtg 6540

cagtggtggt attgtgactg gggatgtagt tgagaataag tcatacacaa gtcagctttc 6600

ttcgagcctc atataagtat aagtagttca acgtattagc actgtaccca gcatctccgt 6660

atcgagaaac acaacaacat gccccattgg acagatcatg cggatacaca ggttgtgcag 6720

tatcatacat actcgatcag acaggtcgtc tgaccatcat acaagctgaa caagcgctcc 6780

atacttgcac gctctctata tacacagtta aattacatat ccatagtcta acctctaaca 6840

gttaatcttc tggtaagcct cccagccagc cttctggtat cgcttggcct cctcaatagg 6900

atctcggttc tggccgtaca gacctcggcc gacaattatg atatccgttc cggtagacat 6960

gacatcctca acagttcggt actgctgtcc gagagcgtct cccttgtcgt caagacccac 7020

cccgggggtc agaataagcc agtcctcaga gtcgccctta ggtcggttct gggcaatgaa 7080

gccaaccaca aactcggggt cggatcgggc aagctcaatg gtctgcttgg agtactcgcc 7140

agtggccaga gagcccttgc aagacagctc ggccagcatg agcagacctc tggccagctt 7200

ctcgttggga gaggggacta ggaactcctt gtactgggag ttctcgtagt cagagacgtc 7260

ctccttcttc tgttcagaga cagtttcctc ggcaccagct cgcaggccag caatgattcc 7320

ggttccgggt acaccgtggg cgttggtgat atcggaccac tcggcgattc ggtgacaccg 7380

gtactggtgc ttgacagtgt tgccaatatc tgcgaacttt ctgtcctcga acaggaagaa 7440

accgtgctta agagcaagtt ccttgagggg gagcacagtg ccggcgtagg tgaagtcgtc 7500

aatgatgtcg atatgggttt tgatcatgca cacataaggt ccgaccttat cggcaagctc 7560

aatgagctcc ttggtggtgg taacatccag agaagcacac aggttggttt tcttggctgc 7620

cacgagcttg agcactcgag cggcaaaggc ggacttgtgg acgttagctc gagcttcgta 7680

ggagggcatt ttggtggtga agaggagact gaaataaatt tagtctgcag aacttttat 7740

cggaacctta tctggggcag tgaagtatat gttatggtaa tagttacgag ttagttgaac 7800

ttatagatag actggactat acggctatcg gtccaaatta gaaagaacgt caatggctct 7860

ctgggcgtcg cctttgccga caaaaatgtg atcatgatga aagccagcaa tgacgttgca 7920

gctgatattg ttgtcggcca accgcgccga aaacgcagct gtcagaccca cagcctccaa 7980

cgaagaatgt atcgtcaaag tgatccaagc acactcatag ttggagtcgt actccaaagg 8040

cggcaatgac gagtcagaca gatactcgtc gactcatcga tcgaggaaga ggacaagcgg 8100

ctgcttctta agtttgtgac atcagtatcc aaggcaccat tgcaaggatt caaggctttg 8160

aacccgtcat ttgccattcg taacgctggt agacaggttg atcggttccc tacggcctcc 8220

acctgtgtca atcttctcaa gctgcctgac tatcaggaca ttgatcaact tcggaagaaa 8280

cttttgtatg ccattcgatc acatgctggt ttcgatttgt cttagaggaa cgcatataca 8340

gtaatcatag agaataaacg atattcattt attaaagtag atagttgagg tagaagttgt 8400

aaagagtgat aaatagcggc cgcttactgg agctttctgg ccttctcctt ggcagcgtca 8460

gccttggcct gcttggcgag cttggcgttc tttcggtaaa agttgtagaa gagaccgagc 8520

atggtccaca tgtagaacca gagcagagcg gtgatgaaga aggggtatcc aggtcggcca 8580

aggaccttca tggcgtacat gtcccaggaa gactggacag acatcatgca gaactgggtc 8640

atctgggatc gagtgatgta gaacttgatg aacgacacct gcttgaagcc cagggcagac 8700

agaaagtagt agccgtacat gatgacgtgg atgaaggagt tcagggcagc agagaagtag 8760

gcttcaccgt tgggagcaac gaaggtgacc agccaccaga tggtgaagat ggaagagtgg 8820

tggtacacgt gcagaaagga aatctgtcgg ttgttcttct tgaggaccat gatcatggtg 8880

tcgacaaact ccatgatctt ggagaagtag aagagccaga tcatcttagc catagggaga 8940

cccttgaagg tgtgatcggc agcgttctca aacagtccat agttggcctg ataagcctcg 9000

tacaggatgc caccgcacat gtaggcggag atggagacca gacagaagtt gtgcaggagg 9060

gagaaggtct tgacctcgaa tcgttcaaag ttcttcatga tctgcatacc cacaaacacg 9120

gtgaccaggt aggcgagcac gatcaggagc acgtggaagg ggttcatcag aggcagctct 9180

cgagccaggg gagactccac ggcaaccagg aagcctcgag tgtgatggac aatggtggga 9240

atgtacttct cggcctgggc aaccagggca gcctccaggg gatcgacgta gggagcagct 9300

cggacaccga tagcgctggc gaggtccatg aacaggtcct gaggcatctt ggagggcagg 9360

aagggagcaa tggactccat gggcaggacc tgtgttagta cattgtcggg gagtcatcaa 9420

ttggttcgac aggttgtcga ctgttagtat gagctcaatt gggctctggt gggtcgatga 9480

cacttgtcat ctgtttctgt tgggtcatgt ttccatcacc ttctatggta ctcacaattc 9540

gtccgattcg cccgaatccg ttaataccga ctttgatggc catgttgatg tgtgtttaat 9600

tcaagaatga atatagagaa gagaagaaga aaaaagattc aattgagccg gcgatgcaga 9660

cccttatata aatgttgcct tggacagacg gagcaagccc gcccaaacct acgttcggta 9720

taatatgtta agctttttaa cacaaaggtt tggcttgggg taacctgatg tggtgcaaaa 9780

gaccgggcgt tggcgagcca ttgcgcgggc gaatggggcc gtgactcgtc tcaaattcga 9840

gggcgtgcct caattcgtgc ccccgtggct ttttcccgcc gtttccgccc cgtttgcacc 9900

actgcagccg cttctttggt tcggacacct tgctgcgagc taggtgcctt gtgctactta 9960

aaaagtggcc tcccaacacc aacatgacat gagtgcgtgg gccaagacac gttggcgggg 10020

tcgcagtcgg ctcaatggcc cggaaaaaac gctgctggag ctggttcgga cgcagtccgc 10080

cgcggcgtat ggatatccgc aaggttccat agcgccattg ccctccgtcg gcgtctatcc 10140

cgcaacctct aaatagagcg ggaatataac ccaagcttct ttttttttcct ttaacacgca 10200

caccccccaac tatcatgttg ctgctgctgt ttgactctac tctgtggagg ggtgctccca 10260

cccaacccaa cctacaggtg gatccggcgc tgtgattggc tgataagtct cctatccgga 10320

ctaattctga ccaatgggac atgcgcgcag gacccaaatg ccgcaattac gtaaccccaa 10380

cgaaatgcct acccctcttt ggagcccagc ggccccaaat ccccccaagc agcccggttc 10440

taccggcttc catctccaag cacaagcagc ccggttctac cggcttccat ctccaagcac 10500

ccctttctcc acaccccaca aaaagacccg tgcaggacat cctactgcgt gtttaaacac 10560

cactaaaacc ccacaaaata tatcttaccg aatatacaga tctactatag aggaacaatt 10620

gccccggaga agacggccag gccgcctaga tgacaaattc aacaactcac agctgacttt 10680

ctgccattgc cactaggggg gggccttttt atatggccaa gccaagctct ccacgtcggt 10740

tgggctgcac ccaacaataa atgggtaggg ttgcaccaac aaagggatgg gatgggggt 10800

agaagatacg aggataacgg ggctcaatgg cacaaataag aacgaatact gccattaaga 10860

ctcgtgatcc agcgactgac accattgcat catctaaggg cctcaaaact acctcggaac 10920

tgctgcgctg atctggacac cacagaggtt ccgagcactt taggttgcac caaatgtccc 10980

accaggtgca ggcagaaaac gctggaacag cgtgtacagt ttgtcttaac aaaaagtgag 11040

ggcgctgagg tcgagcaggg tggtgtgact tgttatagcc tttagagctg cgaaagcgcg 11100

tatggatttg gctcatcagg ccagattgag ggtctgtgga cacatgtcat gttagtgtac 11160

ttcaatcgcc ccctggatat agccccgaca ataggccgtg gcctcatttt tttgccttcc 11220

gcacatttcc attgctcggt acccacacct tgcttctcct gcacttgcca accttaatac 11280

tggtttacat tgaccaacat cttacaagcg gggggcttgt ctagggtata tataaacagt 11340

ggctctccca atcggttgcc agtctctttt ttcctttctt tccccacaga ttcgaaatct 11400

aaactacaca tcacacaatg cctgttactg acgtccttaa gcgaaagtcc ggtgtcatcg 11460

tcggcgacga tgtccgagcc gtgagtatcc acgacaagat cagtgtcgag acgacgcgtt 11520

ttgtgtaatg acacaatccg aaagtcgcta gcaacacaca ctctctacac aaactaaccc 11580

agctctccat ggctgccgct ccctctgtgc gaacctttac ccgagccgag gttctgaacg 11640

ctgaggctct gaacgagggc aagaaggacg ctgaggctcc cttcctgatg atcatcgaca 11700

acaaggtgta cgacgtccga gagttcgtcc ctgaccatcc tggaggctcc gtgattctca 11760

cccacgttgg caaggacggc accgacgtct ttgacacctt tcatcccgag gctgcttggg 11820

agactctcgc caacttctac gttggagaca ttgacgagtc cgaccgagac atcaagaacg 11880

atgactttgc cgctgaggtc cgaaagctgc gaaccctgtt ccagtctctc ggctactacg 11940

actcctctaa ggcctactac gccttcaagg tctccttcaa cctctgcatc tggggactgt 12000

ccaccgtcat tgtggccaag tggggtcaga cctccaccct cgccaacgtg ctctctgctg 12060

ccctgctcgg cctgttctgg cagcagtgcg gatggctggc tcacgacttt ctgcaccacc 12120

aggtcttcca ggaccgattc tggggtgatc tcttcggagc cttcctggga ggtgtctgcc 12180

agggcttctc ctcttcctgg tggaaggaca agcacaacac tcaccatgcc gctcccaacg 12240

tgcatggcga ggatcctgac attgacaccc accctctcct gacctggtcc gagcacgctc 12300

tggagatgtt ctccgacgtc cccgatgagg agctgacccg aatgtggtct cgattcatgg 12360

tcctgaacca gacctggttc tacttcccca ttctctcctt cgctcgactg tcttggtgcc 12420

tccagtccat tctctttgtg ctgcccaacg gtcaggctca caagccctcc ggagctcgag 12480

tgcccatctc cctggtcgag cagctgtccc tcgccatgca ctggacctgg tacctcgcta 12540

ccatgttcct gttcatcaag gatcctgtca acatgctcgt gtacttcctg gtgtctcagg 12600

ctgtgtgcgg aaacctgctc gccatcgtgt tctccctcaa ccacaacggt atgcctgtga 12660

tctccaagga ggaggctgtc gacatggatt tctttaccaa gcagatcatc actggtcgag 12720

atgtccatcc tggactgttc gccaactggt tcaccggtgg cctgaactac cagatcgagc 12780

atcacctgtt cccttccatg cctcgacaca acttctccaa gatccagcct gccgtcgaga 12840

ccctgtgcaa gaagtacaac gtccgatacc acaccactgg tatgatcgag ggaactgccg 12900

aggtcttctc ccgactgaac gaggtctcca aggccacctc caagatgggc aaggctcagt 12960

aagcggccgc atgagaagat aaatatataa atacattgag atattaaatg cgctagatta 13020

gagagcctca tactgctcgg agagaagcca agacgagtac tcaaagggga ttacaccatc 13080

catatccaca gacacaagct ggggaaaggt tctatataca ctttccggaa taccgtagtt 13140

tccgatgtta tcaatggggg cagccaggat ttcaggcact tcggtgtctc ggggtgaaat 13200

ggcgttcttg gcctccatca agtcgtacca tgtcttcatt tgcctgtcaa agtaaaacag 13260

aagcagatga agaatgaact tgaagtgaag gaatttaaac gacggaattc ctgcagccca 13320

tctgcagaat tcaggagaga ccgggttggc ggcgtatttg tgtcccaaaa aacagcccca 13380

attgccccaa ttgaccccaa attgacccag tagcgggccc aaccccggcg agagcccccc 13440

tcaccccaca tatcaaacct cccccggttc ccacacttgc cgttaagggc gtagggtact 13500

gcagtctgga atctacgctt gttcagactt tgtactagtt tctttgtctg gccatccggg 13560

taacccatgc cggacgcaaa atagactact gaaaattttt ttgctttgtg gttgggactt 13620

tagccaaggg tataaaagac caccgtcccc gaattacctt tcctcttctt ttctctctct 13680

ccttgtcaac tcacacccga aatcgttaag catttccttc tgagtataag aatcattcac 13740

catggacatg tccgtcctga ctctccaaga gtacgagttc gagaagcagt tcaacgagaa 13800

tgaagccatc caatggatgc aggaaaactg gaagaaatcc ttcctgtttt ctgccctcta 13860

cgctgccttt atctttggtg gacgacatct gatgaacaag cgagccaagt ttgagctgcg 13920

aaaacctctc gtgctctggt ccctgaccct cgctgtcttc tctatcttcg gtgctctgcg 13980

aactggagcc tacatgctct acatcctgat gaccaaaggc ctgaaacagt ctgtttgtga 14040

ccagtccttt tacaacggac ccgtctcgaa attctgggct tacgcctttg tgctctccaa 14100

agctcccgaa cttggcgata ccatcttcat cattctgcga aagcagaaac tcatcttcct 14160

gcactggtat caccacatca ccgtcctcct gtactcttgg tactcctaca aggacatggt 14220

ggctggaggt ggctggttca tgactatgaa ctacggtgtc cacgccgtga tgtactccta 14280

ctacgccctc cgagctgccg gtttccgagt ctctcgaaag tttgccatgt tcatcaccct 14340

gtcgcagatc actcagatgc tcatgggctg tgtcattaac tacctggtct tcaactggat 14400

gcagcatgac aatgaccagt gctactccca ctttcagaac atcttctggt cctctctcat 14460

gtacctctcc taccttctgc tcttctgcca tttcttcttt gaggcctaca ttggcaaagt 14520

gaagaaagcc accaaggctg agtaagcggc cgcaagtgtg gatggggaag tgagtgcccg 14580

gttctgtgtg cacaattggc aatccaagat ggatggattc aacacaggga tatagcgagc 14640

tacgtggtgg tgcgaggata tagcaacgga tatttatgtt tgacacttga gaatgtacga 14700

tacaagcact gtccaagtac aatactaaac atactgtaca tactcatact cgtacccggg 14760

caacggtttc acttgagtgc agtggctagt gctcttactc gtacagtgtg caatactgcg 14820

tatcatagtc tttgatgtat atcgtattca ttcatgttag ttgc 14864

<210> 187
<211> 6540
<212> DNA
<213> Artificial Sequence

<220>
<223> Plasmid pZKUGPE1S

<400> 187
ggccgcaagt gtggatgggg aagtgagtgc ccggttctgt gtgcacaatt ggcaatccaa 60

gatggatgga ttcaacacag ggatatagcg agctacgtgg tggtgcgagg atatagcaac 120

ggatatttat gtttgacact tgagaatgta cgatacaagc actgtccaag tacaatacta 180

aacatactgt acatactcat actcgtaccc gggcaacggt ttcacttgag tgcagtggct 240

agtgctctta ctcgtacagt gtgcaatact gcgtatcata gtctttgatg tatatcgtat 300

tcattcatgt tagttgcgta cgaggaaact gtctctgaac agaagaagga ggacgtctct 360

gactacgaga actcccagta caaggagttc ctagtcccct ctcccaacga gaagctggcc 420

agaggtctgc tcatgctggc cgagctgtct tgcaagggct ctctggccac tggcgagtac 480

tccaagcaga ccattgagct tgcccgatcc gaccccgagt ttgtggttgg cttcattgcc 540

cagaaccgac ctaagggcga ctctgaggac tggcttattc tgaccccccgg ggtgggtctt 600

gacgacaagg gagacgctct cggacagcag taccgaactg ttgaggatgt catgtctacc 660

ggaacggata tcataattgt cggccgaggt ctgtacggcc agaaccgaga tcctattgag 720

gaggccaagc gataccagaa ggctggctgg gaggcttacc agaagattaa ctgttagagg 780

ttagactatg gatatgtaat ttaactgtgt atatagagag cgtgcaagta tggagcgctt 840

gttcagcttg tatgatggtc agacgacctg tctgatcgag tatgtatgat actgcacaac 900

ctgtgtatcc gcatgatctg tccaatgggg catgttgttg tgtttctcga tacggagatg 960

ctgggtacag tgctaatacg ttgaactact tatacttata tgaggctcga agaaagctga 1020

cttgtgtatg acttaattaa tcgagcttgg cgtaatcatg gtcatagctg tttcctgtgt 1080

gaaattgtta tccgctcaca attccacaca acatacgagc cggaagcata aagtgtaaag 1140

cctggggtgc ctaatgagtg agctaactca cattaattgc gttgcgctca ctgcccgctt 1200

tccagtcggg aaacctgtcg tgccagctgc attaatgaat cggccaacgc gcggggagag 1260

gcggtttgcg tattgggcgc tcttccgctt cctcgctcac tgactcgctg cgctcggtcg 1320

ttcggctgcg gcgagcggta tcagctcact caaaggcggt aatacggtta tccacagaat 1380

caggggataa cgcaggaaag aacatgtgag caaaaggcca gcaaaaggcc aggaaccgta 1440

aaaaggccgc gttgctggcg tttttccata ggctccgccc ccctgacgag catcacaaaa 1500

atcgacgctc aagtcagagg tggcgaaacc cgacaggact ataaagatac caggcgtttc 1560

cccctggaag ctccctcgtg cgctctcctg ttccgaccct gccgcttacc ggatacctgt 1620

ccgcctttct cccttcggga agcgtggcgc tttctcatag ctcacgctgt aggtatctca 1680

gttcggtgta ggtcgttcgc tccaagctgg gctgtgtgca cgaacccccc gttcagcccg 1740

accgctgcgc cttatccggt aactatcgtc ttgagtccaa cccggtaaga cacgacttat 1800

cgccactggc agcagccact ggtaacagga ttagcagagc gaggtatgta ggcggtgcta 1860

cagagttctt gaagtggtgg cctaactacg gctacactag aaggacagta tttggtatct 1920

gcgctctgct gaagccagtt accttcggaa aaagagttgg tagctcttga tccggcaaac 1980

aaaccaccgc tggtagcggt ggtttttttg tttgcaagca gcagattacg cgcagaaaaa 2040

aaggatctca agaagatcct ttgatctttt ctacggggtc tgacgctcag tggaacgaaa 2100

actcacgtta agggattttg gtcatgagat tatcaaaaag gatcttcacc tagatccttt 2160

taaattaaaa atgaagtttt aaatcaatct aaagtatata tgagtaaact tggtctgaca 2220

gttaccaatg cttaatcagt gaggcaccta tctcagcgat ctgtctattt cgttcatcca 2280

tagttgcctg actccccgtc gtgtagataa ctacgatacg ggagggctta ccatctggcc 2340

ccagtgctgc aatgataccg cgagacccac gctcaccggc tccagattta tcagcaataa 2400

accagccagc cggaagggcc gagcgcagaa gtggtcctgc aactttatcc gcctccatcc 2460

agtctattaa ttgttgccgg gaagctagag taagtagttc gccagttaat agtttgcgca 2520

acgttgttgc cattgctaca ggcatcgtgg tgtcacgctc gtcgtttggt atggcttcat 2580

tcagctccgg ttcccaacga tcaaggcgag ttacatgatc ccccatgttg tgcaaaaaag 2640

cggttagctc cttcggtcct ccgatcgttg tcagaagtaa gttggccgca gtgttatcac 2700

tcatggttat ggcagcactg cataattctc ttactgtcat gccatccgta agatgctttt 2760

ctgtgactgg tgagtactca accaagtcat tctgagaata gtgtatgcgg cgaccgagtt 2820

gctcttgccc ggcgtcaata cgggataata ccgcgccaca tagcagaact ttaaaagtgc 2880

tcatcattgg aaaacgttct tcggggcgaa aactctcaag gatcttaccg ctgttgagat 2940

ccagttcgat gtaacccact cgtgcaccca actgatcttc agcatctttt actttcacca 3000

gcgtttctgg gtgagcaaaa acaggaaggc aaaatgccgc aaaaaaggga ataagggcga 3060

cacggaaatg ttgaatactc atactcttcc tttttcaata ttattgaagc atttatcagg 3120

gttattgtct catgagcgga tacatatttg aatgtattta gaaaaataaa caaatagggg 3180

ttccgcgcac atttccccga aaagtgccac ctgacgcgcc ctgtagcggc gcattaagcg 3240

cggcgggtgt ggtggttacg cgcagcgtga ccgctacact tgccagcgcc ctagcgcccg 3300

ctcctttcgc tttcttccct tcctttctcg ccacgttcgc cggctttccc cgtcaagctc 3360

taaatcgggg gctcccttta gggttccgat ttagtgcttt acggcacctc gaccccaaaa 3420

aacttgatta gggtgatggt tcacgtagtg ggccatcgcc ctgatagacg gtttttcgcc 3480

ctttgacgtt ggagtccacg ttctttaata gtggactctt gttccaaact ggaacaacac 3540

tcaaccctat ctcggtctat tcttttgatt tataagggat tttgccgatt tcggcctatt 3600

ggttaaaaaa tgagctgatt taacaaaaat ttaacgcgaa ttttaacaaa atattaacgc 3660

ttacaatttc cattcgccat tcaggctgcg caactgttgg gaagggcgat cggtgcgggc 3720

ctcttcgcta ttacgccagc tggcgaaagg gggatgtgct gcaaggcgat taagttgggt 3780

aacgccaggg ttttcccagt cacgacgttg taaaacgacg gccagtgaat tgtaatacga 3840

ctcactatag ggcgaattgg gtaccgggcc ccccctcgag gtcgacgagt atctgtctga 3900

ctcgtcattg ccgcctttgg agtacgactc caactatgag tgtgcttgga tcactttgac 3960

gatacattct tcgttggagg ctgtgggtct gacagctgcg ttttcggcgc ggttggccga 4020

caacaatatc agctgcaacg tcattgctgg ctttcatcat gatcacattt ttgtcggcaa 4080

aggcgacgcc cagagagcca ttgacgttct ttctaatttg gaccgatagc cgtatagtcc 4140

agtctatcta taagttcaac taactcgtaa ctattaccat aacatatact tcactgcccc 4200

agataaggtt ccgataaaaa gttctgcaga ctaaatttat ttcagtctcc tcttcaccac 4260

caaaatgccc tcctacgaag ctcgagtgct caagctcgtg gcagccaaga aaaccaacct 4320

gtgtgcttct ctggatgtta ccaccaccaa ggagctcatt gagcttgccg ataaggtcgg 4380

accttatgtg tgcatgatca aaacccatat cgacatcatt gacgacttca cctacgccgg 4440

cactgtgctc cccctcaagg aacttgctct taagcacggt ttcttcctgt tcgaggacag 4500

aaagttcgca gatattggca acactgtcaa gcaccagtac cggtgtcacc gaatcgccga 4560

gtggtccgat atcaccaacg cccacggtgt acccggaacc ggaatcgatg cgtatctgtg 4620

ggacatgtgg tcgttgcgcc attatgtaag cagcgtgtac tcctctgact gtccatatgg 4680

tttgctccat ctcaccctca tcgttttcat tgttcacagg cggccacaaa aaaactgtct 4740

tctctccttc tctcttcgcc ttagtctact cggaccagtt ttagtttagc ttggcgccac 4800

tggataaatg agacctcagg ccttgtgatg aggaggtcac ttatgaagca tgttaggagg 4860

tgcttgtatg gatagagaag cacccaaaat aataagaata ataataaaac agggggcgtt 4920

gtcatttcat atcgtgtttt caccatcaat acacctccaa acaatgccct tcatgtggcc 4980

agccccaata ttgtcctgta gttcaactct atgcagctcg tatcttattg agcaagtaaa 5040

actctgtcag ccgatattgc ccgacccgcg acaagggtca acaaggtggt gtaaggcctt 5100

cgcagaagtc aaaactgtgc caaacaaaca tctagagtct ctttggtgtt tctcgcatat 5160

atttwatcgg ctgtcttacg tatttgcgcc tcggtaccgg actaatttcg gatcatcccc 5220

aatacgcttt ttcttcgcag ctgtcaacag tgtccatgat ctatccacct aaatgggtca 5280

tatgaggcgt ataatttcgt ggtgctgata ataattccca tatatttgac acaaaacttc 5340

ccccccctaga catacatctc acaatctcac ttcttgtgct tctgtcacac atctcctcca 5400

gctgacttca actcacacct ctgccccagt tggtctacag cggtataagg tttctccgca 5460

tagaggtgca ccactcctcc cgatacttgt ttgtgtgact tgtgggtcac gacatatata 5520

tctacacaca ttgcgccacc ctttggttct tccagcacaa caaaaacacg acacgctaac 5580

catggagtcc attgctccct tcctgccctc caagatgcct caggacctgt tcatggacct 5640

cgccagcgct atcggtgtcc gagctgctcc ctacgtcgat cccctggagg ctgccctggt 5700

tgcccaggcc gagaagtaca ttcccaccat tgtccatcac actcgaggct tcctggttgc 5760

cgtggagtct cccctggctc gagagctgcc tctgatgaac cccttccacg tgctcctgat 5820

cgtgctcgcc tacctggtca ccgtgtttgt gggtatgcag atcatgaaga actttgaacg 5880

attcgaggtc aagaccttct ccctcctgca caacttctgt ctggtctcca tctccgccta 5940

catgtgcggt ggcatcctgt acgaggctta tcaggccaac tatggactgt ttgagaacgc 6000

tgccgatcac accttcaagg gtctccctat ggctaagatg atctggctct tctacttctc 6060

caagatcatg gagtttgtcg acaccatgat catggtcctc aagaagaaca accgacagat 6120

ttcctttctg cacgtgtacc accactcttc catcttcacc atctggtggc tggtcacctt 6180

cgttgctccc aacggtgaag cctacttctc tgctgccctg aactccttca tccacgtcat 6240

catgtacggc tactactttc tgtctgccct gggcttcaag caggtgtcgt tcatcaagtt 6300

ctacatcact cgatcccaga tgacccagtt ctgcatgatg tctgtccagt cttcctggga 6360

catgtacgcc atgaaggtcc ttggccgacc tggatacccc ttcttcatca ccgctctgct 6420

ctggttctac atgtggacca tgctcggtct cttctacaac ttttaccgaa agaacgccaa 6480

gctcgccaag caggccaagg ctgacgctgc caaggagaag gccagaaagc tccagtaagc 6540

<210> 188
<211> 13321
<212> DNA
<213> Artificial Sequence

<220>
<223> Plasmid pKO2UM26E

<400> 188
tgggtgagca aaaacaggaa ggcaaaatgc cgcaaaaaag ggaataaggg cgacacggaa 60

atgttgaata ctcatactct tccttttca atattattga agcatttatc agggttattg 120

tctcatgagc ggatacatat ttgaatgtat ttagaaaaat aaacaaatag gggttccgcg 180

cacatttccc cgaaaagtgc cacctgatgc ggtgtgaaat accgcacaga tgcgtaagga 240

gaaaataccg catcaggaaa ttgtaagcgt taatattttg ttaaaattcg cgttaaattt 300

ttgttaaatc agctcatttt ttaaccaata ggccgaaatc ggcaaaatcc cttataaatc 360

aaaagaatag accgagatag ggttgagtgt tgttccagtt tggaacaaga gtccactatt 420

aaagaacgtg gactccaacg tcaaagggcg aaaaaccgtc tatcagggcg atggcccact 480

acgtgaacca tcaccctaat caagtttttt ggggtcgagg tgccgtaaag cactaaatcg 540

gaaccctaaa gggagccccc gatttagagc ttgacgggga aagccggcga acgtggcgag 600

aaaggaaggg aagaaagcga aaggagcggg cgctagggcg ctggcaagtg tagcggtcac 660

gctgcgcgta accaccacac ccgccgcgct taatgcgccg ctacagggcg cgtccattcg 720

ccattcaggc tgcgcaactg ttgggaaggg cgatcggtgc gggcctcttc gctattacgc 780

cagctggcga aaggggggatg tgctgcaagg cgattaagtt gggtaacgcc agggttttcc 840

cagtcacgac gttgtaaaac gacggccagt gaattgtaat acgactcact atagggcgaa 900

ttgggcccga cgtcgcatgc ttgaatctac aagtaggagg gttggagtga ttaagtgaaa 960

cttctttaac ggctctatgc cagttctatt gatatccgaa acatcagtat gaaggtctga 1020

taagggtgac ttcttcccac agattcgtat cagtacgagt acgagaccgg tacttgtaac 1080

agtattgata ctaaagggaa actacaacgg ttgtcagcgt aatgtgactt cgcccatgaa 1140

cgcagacacg cagtgccgag tgcggtgata tcgcctactc gttacgtcca tggactacac 1200

aacccctcgg cttcgcttgg cttagcctcg ggctcggtgc tgttcagtta aaacacaatc 1260

aaataacatt tctacttttt agaaggcagg ccgtcaggag caactccgac tccattgacg 1320

tttctaaaca tctgaatgcc ttccttacct tcaacaaact ggcaggttcg ggcgacagtg 1380

taaagagact tgatgaagtt ggtgtcgtcg tgtcggtagt gcttgcccat gaccttcttg 1440

atcttctcag tggcgattcg ggcgttgtag aagggaattc aggagagacc gggttggcgg 1500

cgtatttgtg tcccaaaaaa cagccccaat tgccccggag aagacggcca ggccgcctag 1560

atgacaaatt caacaactca cagctgactt tctgccattg ccactagggg ggggccttttt 1620

tatatggcca agccaagctc tccacgtcgg ttgggctgca cccaacaata aatgggtagg 1680

gttgcaccaa caaagggatg ggatgggggg tagaagatac gaggataacg gggctcaatg 1740

gcacaaataa gaacgaatac tgccattaag actcgtgatc cagcgactga caccattgca 1800

tcatctaagg gcctcaaaac tacctcggaa ctgctgcgct gatctggaca ccacagaggt 1860

tccgagcact ttaggttgca ccaaatgtcc caccaggtgc aggcagaaaa cgctggaaca 1920

gcgtgtacag tttgtcttaa caaaaagtga gggcgctgag gtcgagcagg gtggtgtgac 1980

ttgttatagc ctttagagct gcgaaagcgc gtatggattt ggctcatcag gccagattga 2040

gggtctgtgg acacatgtca tgttagtgta cttcaatcgc cccctggata tagccccgac 2100

aataggccgt ggcctcattt ttttgccttc cgcacatttc cattgctcgg tacccacacc 2160

ttgcttctcc tgcacttgcc aaccttaata ctggtttaca ttgaccaaca tcttacaagc 2220

ggggggcttg tctagggtat atataaacag tggctctccc aatcggttgc cagtctcttt 2280

tttcctttct ttccccacag attcgaaatc taaactacac atcacacaat gcctgttact 2340

gacgtcctta agcgaaagtc cggtgtcatc gtcggcgacg atgtccgagc cgtgagtatc 2400

cacgacaaga tcagtgtcga gacgacgcgt tttgtgtaat gacacaatcc gaaagtcgct 2460

agcaacacac actctctaca caaactaacc cagctctcca tggctacctc tgacccttcc 2520

gtccgagctt tcacacgatc tgaagtcttg cacgccgatg ccttgaacga gggcaaaaag 2580

aatgctgaag ccccgttcct catgatcatt gacaacaaag tctacgatgt gcgcgagttt 2640

gtccccgacc atcctggcgg tagtgtcatt ttgacccacg taggcaagga cggcacggac 2700

gtcttcgaga ccttccatcc cgaggccgcg tgggaaacgc tcgccaactt ttatgtcggc 2760

gacattgcag aatccgatcg tgcgattgag aatgacgagt ttgcagccga agtccgcaag 2820

ctgcgtactt tgttccagtc tctgggatat tacaactcat ccaaaaccta ctacgccttc 2880

aaggtctcat tcaacctctg catctggggc ctgtcggcgt ttatggttgc caagtggggc 2940

cagacctcga ccctcgccaa cgtggtgtct gcgtcgcttc tgggtgtctt ttggcaacag 3000

tgtggttggc tcgcccatga cttcttacac caccaggtct ttcacgatcg gttctggggt 3060

gatttgtttg gtgccttcct cggcggagtc tgtcaggggt tctcctcttc ctggtggaag 3120

gacaaacaca acacccacca cgcggcgccc aatgtccacg gagaggatcc cgatatcgac 3180

acgcacccac tgttgacatg gagcgaacat gcgctcgaga tgttttcgga tgtgcccgat 3240

gaggagttgg cccaaatgtg gtcccggttt atggttgtga accagacctg gttttacttt 3300

cccattctgt cgttcgcccg gttgtcgtgg tgcattcagt ccattctttt tgtgctcccg 3360

aacggacagg cacacaagcc tgcgggagcc cgggtcccca tctcgctggt ggagcaattg 3420

tcgctggcga tgcattggac ctggtatctg gcgaccatgt tcttgtttat caaggatcct 3480

gtcaacataa tggtgtattt cttggtctcg caagcggtct gtggcaacct gttggcgatt 3540

gtgttctctc tgaaccataa cggaatgcct gtgatctctc aggaagaagc ggtcgagatg 3600

gatttcttta caaagcagat cattacgggc cgtgatgtcc acccaggctg gttcgcagac 3660

tggttcacgg gcggattgaa ctatcaaatt gagcaccacc tgttcccgtc gatgcctcgc 3720

cactatttct caaagatcca accaacggtc gaatcattgt gcaaaaagta cggggtccga 3780

tatcacacga cggggatgat agatggcacc gcagaggtct tttcgcgact gaacgaggtc 3840

tctcaggctg cctccaagct cggcaagtct gcttgagcgg ccgcaagtgt ggatggggaa 3900

gtgagtgccc ggttctgtgt gcacaattgg caatccaaga tggatggatt caacacaggg 3960

atatagcgag ctacgtggtg gtgcgaggat atagcaacgg atatttatgt ttgacacttg 4020

agaatgtacg atacaagcac tgtccaagta caatactaaa catactgtac atactcatac 4080

tcgtacccgg gcaacggttt cacttgagtg cagtggctag tgctcttact cgtacagtgt 4140

gcaatactgc gtatcatagt ctttgatgta tatcgtattc attcatgtcg acgagtatct 4200

gtctgactcg tcattgccgc ctttggagta cgactccaac tatgagtgtg cttggatcac 4260

tttgacgata cattcttcgt tggaggctgt gggtctgaca gctgcgtttt cggcgcggtt 4320

ggccgacaac aatatcagct gcaacgtcat tgctggcttt catcatgatc acatttttgt 4380

cggcaaaggc gacgcccaga gagccattga cgttctttct aatttggacc gatagccgta 4440

tagtccagtc tatctataag ttcaactaac tcgtaactat taccataaca tatacttcac 4500

tgccccagat aaggttccga taaaaagttc tgcagactaa atttatttca gtctcctctt 4560

caccaccaaa atgccctcct acgaagctcg agctaacgtc cacaagtccg cctttgccgc 4620

tcgagtgctc aagctcgtgg cagccaagaa aaccaacctg tgtgcttctc tggatgttac 4680

caccaccaag gagctcattg agcttgccga taaggtcgga ccttatgtgt gcatgatcaa 4740

aacccatatc gacatcattg acgacttcac ctacgccggc actgtgctcc ccctcaagga 4800

acttgctctt aagcacggtt tcttcctgtt cgaggacaga aagttcgcag atattggcaa 4860

cactgtcaag caccagtacc ggtgtcaccg aatcgccgag tggtccgata tcaccaacgc 4920

ccacggtgta cccggaaccg gaatcattgc tggcctgcga gctggtgccg aggaaactgt 4980

ctctgaacag aagaaggagg acgtctctga ctacgagaac tcccagtaca aggagttcct 5040

agtcccctct cccaacgaga agctggccag aggtctgctc atgctggccg agctgtcttg 5100

caagggctct ctggccactg gcgagtactc caagcagacc attgagcttg cccgatccga 5160

ccccgagttt gtggttggct tcattgccca gaaccgacct aagggcgact ctgaggactg 5220

gcttattctg acccccgggg tgggtcttga cgacaaggga gacgctctcg gacagcagta 5280

ccgaactgtt gaggatgtca tgtctaccgg aacggatatc ataattgtcg gccgaggtct 5340

gtacggccag aaccgagatc ctattgagga ggccaagcga taccagaagg ctggctggga 5400

ggcttaccag aagattaact gttagaggtt agactatgga tatgtaattt aactgtgtat 5460

atagagagcg tgcaagtatg gagcgcttgt tcagcttgta tgatggtcag acgacctgtc 5520

tgatcgagta tgtatgatac tgcacaacct gtgtatccgc atgatctgtc caatggggca 5580

tgttgttgtg tttctcgata cggagatgct gggtacagtg ctaatacgtt gaactactta 5640

tacttatatg aggctcgaag aaagctgact tgtgtatgac ttaattaatt tgaatcgaat 5700

cgatgagcct aaaatgaacc cgagtatatc tcataaaatt ctcggtgaga ggtctgtgac 5760

tgtcagtaca aggtgccttc attatgccct caaccttacc atacctcact gaatgtagtg 5820

tacctctaaa aatgaaatac agtgccaaaa gccaaggcac tgagctcgtc taacggactt 5880

gatatacaac caattaaaac aaatgaaaag aaatacagtt ctttgtatca tttgtaacaa 5940

ttaccctgta caaactaagg tattgaaatc ccacaatatt cccaaagtcc acccctttcc 6000

aaattgtcat gcctacaact catataccaa gcactaacct accgtttaaa cagtgtacgc 6060

agatctacta tagaggaaca tttaaattgc cccggagaag acggccaggc cgcctagatg 6120

acaaattcaa caactcacag ctgactttct gccattgcca ctaggggggg gccttttttat 6180

atggccaagc caagctctcc acgtcggttg ggctgcaccc aacaataaat gggtagggtt 6240

gcaccaacaa agggatggga tgggggggtag aagatacgag gataacgggg ctcaatggca 6300

caaataagaa cgaatactgc cattaagact cgtgatccag cgactgacac cattgcatca 6360

tctaagggcc tcaaaactac ctcggaactg ctgcgctgat ctggacacca cagaggttcc 6420

gagcacttta ggttgcacca aatgtcccac caggtgcagg cagaaaacgc tggaacagcg 6480

tgtacagttt gtcttaacaa aaagtgaggg cgctgaggtc gagcagggtg gtgtgacttg 6540

ttatagcctt tagagctgcg aaagcgcgta tggatttggc tcatcaggcc agattgaggg 6600

tctgtggaca catgtcatgt tagtgtactt caatcgcccc ctggatatag ccccgacaat 6660

aggccgtggc ctcatttttt tgccttccgc acatttccat tgctcggtac ccacaccttg 6720

cttctcctgc acttgccaac cttaatactg gtttacattg accaacatct tacaagcggg 6780

gggcttgtct agggtatata taaacagtgg ctctcccaat cggttgccag tctcttttttt 6840

cctttctttc cccacagatt cgaaatctaa actacacatc acacaatgcc tgttactgac 6900

gtccttaagc gaaagtccgg tgtcatcgtc ggcgacgatg tccgagccgt gagtatccac 6960

gacaagatca gtgtcgagac gacgcgtttt gtgtaatgac acaatccgaa agtcgctagc 7020

aacacacact ctctacacaa actaacccag ctctccatgg cacctcccaa cactatcgac 7080

gctggcttga cccagcgtca tatcaccacc acggccgccc caacctcggc caagcccgct 7140

ttcgagcgca actaccagct ccccgagttc accatcaagg agatccgaga gtgcatccct 7200

gcccactgct ttgagcgctc cggtcttcgt ggtctctgcc acgttgccat tgatctgacc 7260

tgggcctcgc tcttgttcct ggctgcaacc cagatcgaca agttcgagaa ccccttgatc 7320

cgctatctgg cctggcctgc gtactggatc atgcagggca ttgtctgcac cggcatatgg 7380

gtgctggccc acgagtgcgg tcaccagtcc ttctcgacct ccaagactct caacaacacc 7440

gtcggctgga tcctgcactc gatgctcttg gtcccctacc actcctggag aatctcgcac 7500

tcgaagcacc acaaggccac tggccacatg accaaggacc aggtctttgt tcccaagacc 7560

cgctcccagg ttggtttgcc tcccaaggag agcgctgctg ctgccgttca agaggaggac 7620

atgtccgtgc acctggatga ggaggcccct attgtgactt tgttctggat ggtgatccag 7680

ttcctgttcg gatggcctgc atacctgatc atgaacgcct ctggtcagga ctatggccgc 7740

tggacctcgc acttccacac ttactcgccc atctttgagc cccgcaactt cttcgacatt 7800

atcatctcgg atctcggtgt gttggctgcc ctcggtgccc tgatctacgc ttccatgcag 7860

ctgtcgctct tgaccgtgac caagtactac atcatcccgt acctgtttgt caacttttgg 7920

ttggtcctga ttactttctt gcagcacacc gaccccaagc tgccccatta ccgtgagggt 7980

gcctggaact tccagcgtgg agccctctgc accgttgacc gctcgtttgg caagttcttg 8040

gaccatatgt tccacggcat cgtccatacc catgtggccc atcacttgtt ctcgcagatg 8100

ccgttctacc atgctgaaga agctacctac catctcaaga aactgctggg agagtactac 8160

gtttacgacc catccccgat cgtcgttgcg gtctggaggt cgttccgcga gtgccgattc 8220

gtggaggatc agggagacgt ggtctttttc aagaagtaag cggccgcaag tgtggatggg 8280

gaagtgagtg cccggttctg tgtgcacaat tggcaatcca agatggatgg attcaacaca 8340

gggatatagc gagctacgtg gtggtgcgag gatatagcaa cggatattta tgtttgacac 8400

ttgagaatgt acgatacaag cactgtccaa gtacaatact aaacatactg tacatactca 8460

tactcgtacc cgggcaacgg tttcacttga gtgcagtggc tagtgctctt actcgtacag 8520

tgtgcaatac tgcgtatcat agtctttgat gtatatcgta ttcattcatg ttagttgcgt 8580

acgagatcgt caagggtttg tggccaactg gtatttaaat gtagctaacg gtagcaggcg 8640

aactactggt acatacctcc cccggaatat gtacaggcat aatgcgtatc tgtgggacat 8700

gtggtcgttg cgccattatg taagcagcgt gtactcctct gactgtccat atggtttgct 8760

ccatctcacc ctcatcgttt tcattgttca caggcggcca caaaaaaact gtcttctctc 8820

cttctctctt cgccttagtc tactcggacc agttttagtt tagcttggcg ccactggata 8880

aatgagacct caggccttgt gatgaggagg tcacttatga agcatgttag gaggtgcttg 8940

tatggataga gaagcaccca aaataataag aataataata aaacaggggg cgttgtcatt 9000

tcatatcgtg ttttcaccat caatacacct ccaaacaatg cccttcatgt ggccagcccc 9060

aatattgtcc tgtagttcaa ctctatgcag ctcgtatctt attgagcaag taaaactctg 9120

tcagccgata ttgcccgacc cgcgacaagg gtcaacaagg tggtgtaagg ccttcgcaga 9180

agtcaaaact gtgccaaaca aacatctaga gtctctttgg tgtttctcgc atatatttwa 9240

tcggctgtct tacgtatttg cgcctcggta ccggactaat ttcggatcat ccccaatacg 9300

ctttttcttc gcagctgtca acagtgtcca tgatctatcc acctaaatgg gtcatatgag 9360

gcgtataatt tcgtggtgct gataataatt cccatatatt tgacacaaaa cttccccccc 9420

tagacataca tctcacaatc tcacttcttg tgcttctgtc acacatctcc tccagctgac 9480

ttcaactcac acctctgccc cagttggtct acagcggtat aaggtttctc cgcatagagg 9540

tgcaccactc ctcccgatac ttgtttgtgt gacttgtggg tcacgacata tatatctaca 9600

cacattgcgc cacccttttgg ttcttccagc acaacaaaaa cacgacacgc taaccatgga 9660

gtccattgct cccttcctgc cctccaagat gcctcaggac ctgttcatgg acctcgccag 9720

cgctatcggt gtccgagctg ctccctacgt cgatcccctg gaggctgccc tggttgccca 9780

ggccgagaag tacattccca ccattgtcca tcacactcga ggcttcctgg ttgccgtgga 9840

gtctcccctg gctcgagagc tgcctctgat gaaccccttc cacgtgctcc tgatcgtgct 9900

cgcctacctg gtcaccgtgt ttgtgggtat gcagatcatg aagaactttg aacgattcga 9960

ggtcaagacc ttctccctcc tgcacaactt ctgtctggtc tccatctccg cctacatgtg 10020

cggtggcatc ctgtacgagg cttatcaggc caactatgga ctgtttgaga acgctgccga 10080

tcacaccttc aagggtctcc ctatggctaa gatgatctgg ctcttctact tctccaagat 10140

catggagttt gtcgacacca tgatcatggt cctcaagaag aacaaccgac agatttcctt 10200

tctgcacgtg taccaccact cttccatctt caccatctgg tggctggtca ccttcgttgc 10260

tcccaacggt gaagcctact tctctgctgc cctgaactcc ttcatccacg tcatcatgta 10320

cggctactac tttctgtctg ccctgggctt caagcaggtg tcgttcatca agttctacat 10380

cactcgatcc cagatgaccc agttctgcat gatgtctgtc cagtcttcct gggacatgta 10440

cgccatgaag gtccttggcc gacctggata ccccttcttc atcaccgctc tgctctggtt 10500

ctacatgtgg accatgctcg gtctcttcta caacttttac cgaaagaacg ccaagctcgc 10560

caagcaggcc aaggctgacg ctgccaagga gaaggccaga aagctccagt aagcggccgc 10620

caccgcggcc cgagattccg gcctcttcgg ccgccaagcg acccgggtgg acgtctagag 10680

gtacctagca attaacagat agtttgccgg tgataattct cttaacctcc cacactcctt 10740

tgacataacg atttatgtaa cgaaactgaa atttgaccag atattgtgtc cgcggtggag 10800

ctccaagctt ccactggtcg gcgtggtagt ggggcagagt ggggtcggtg tgctgcaggt 10860

aggtgatggc cacgagccag tggttgaccc acaggtaggg gatcaggtag tagagggtga 10920

cggaagccag gccccatcgg ttgatggagt atgcgatgac ggacatggtg ataccaatac 10980

cgacgttaga gatccagatg ttgaaccagt ccttcttctc aaacagcggg gcgttggggt 11040

tgaagtggtt gacagcccat ttgttgagct tggggtactt ctgtccggta acgtaagaca 11100

gcagatacag aggccatcca aacacctgct gggtgatgag gccgtagagg gtcatgaggg 11160

gagcgtcctc agcaagctca gaccagtcat gggcgcctcg gttctccata aactcctttc 11220

ggtccttggg cacaaacacc atatcacggg tgaggtgacc agtggacttg tggtgcatgg 11280

agtgggtcag cttccaggcg tagtaaggga ccagcatgga ggagtgcaga acccatccgg 11340

tgacgttgtt gacggtgtta gagtcggaga aagcagagtg gccacactcg tgggcaagaa 11400

cccacagacc ggtgccaaac agaccctgga caatggagta catggcccag gccacagctc 11460

ggccggaagc cgagggaata agaggcaggt acgcgtaggc catgtaggca aaaacggcga 11520

taaagaagca ggcgcgccag ctgcattaat gaatcggcca acgcgcgggg agaggcggtt 11580

tgcgtattgg gcgctcttcc gcttcctcgc tcactgactc gctgcgctcg gtcgttcggc 11640

tgcggcgagc ggtatcagct cactcaaagg cggtaatacg gttatccaca gaatcagggg 11700

ataacgcagg aaagaacatg tgagcaaaag gccagcaaaa ggccaggaac cgtaaaaagg 11760

ccgcgttgct ggcgtttttc cataggctcc gccccccctga cgagcatcac aaaaatcgac 11820

gctcaagtca gaggtggcga aacccgacag gactataaag ataccaggcg tttccccctg 11880

gaagctccct cgtgcgctct cctgttccga ccctgccgct taccggatac ctgtccgcct 11940

ttctcccttc gggaagcgtg gcgctttctc atagctcacg ctgtaggtat ctcagttcgg 12000

tgtaggtcgt tcgctccaag ctgggctgtg tgcacgaacc ccccgttcag cccgaccgct 12060

gcgccttatc cggtaactat cgtcttgagt ccaacccggt aagacacgac ttatcgccac 12120

tggcagcagc cactggtaac aggattagca gagcgaggta tgtaggcggt gctacagagt 12180

tcttgaagtg gtggcctaac tacggctaca ctagaagaac agtatttggt atctgcgctc 12240

tgctgaagcc agttaccttc ggaaaaagag ttggtagctc ttgatccggc aaacaaacca 12300

ccgctggtag cggtggtttt tttgtttgca agcagcagat tacgcgcaga aaaaaaggat 12360

ctcaagaaga tcctttgatc ttttctacgg ggtctgacgc tcagtggaac gaaaactcac 12420

gttaagggat tttggtcatg agattatcaa aaaggatctt cacctagatc cttttaaatt 12480

aaaaatgaag ttttaaatca atctaaagta tatatgagta aacttggtct gacagttacc 12540

aatgcttaat cagtgaggca cctatctcag cgatctgtct atttcgttca tccatagttg 12600

cctgactccc cgtcgtgtag ataactacga tacgggaggg cttaccatct ggccccagtg 12660

ctgcaatgat accgcgagac ccacgctcac cggctccaga tttatcagca ataaaccagc 12720

cagccggaag ggccgagcgc agaagtggtc ctgcaacttt atccgcctcc atccagtcta 12780

ttaattgttg ccgggaagct agagtaagta gttcgccagt taatagtttg cgcaacgttg 12840

ttgccattgc tacaggcatc gtggtgtcac gctcgtcgtt tggtatggct tcattcagct 12900

ccggttccca acgatcaagg cgagttacat gatcccccat gttgtgcaaa aaagcggtta 12960

gctccttcgg tcctccgatc gttgtcagaa gtaagttggc cgcagtgtta tcactcatgg 13020

ttatggcagc actgcataat tctcttactg tcatgccatc cgtaagatgc ttttctgtga 13080

ctggtgagta ctcaaccaag tcattctgag aatagtgtat gcggcgaccg agttgctctt 13140

gcccggcgtc aatacgggat aataccgcgc cacatagcag aactttaaaa gtgctcatca 13200

ttggaaaacg ttcttcgggg cgaaaactct caaggatctt accgctgttg agatccagtt 13260

cgatgtaacc cactcgtgca cccaactgat cttcagcatc ttttactttc accagcgttt 13320

c 13321

<210> 189
<211> 4313
<212> DNA
<213> Artificial Sequence

<220>
<223> Plasmid pZKUM

<400> 189
taatcgagct tggcgtaatc atggtcatag ctgtttcctg tgtgaaattg ttatccgctc 60

acaattccac acaacatacg agccggaagc ataaagtgta aagcctgggg tgcctaatga 120

gtgagctaac tcacattaat tgcgttgcgc tcactgcccg ctttccagtc gggaaacctg 180

tcgtgccagc tgcattaatg aatcggccaa cgcgcgggga gaggcggttt gcgtattggg 240

cgctcttccg cttcctcgct cactgactcg ctgcgctcgg tcgttcggct gcggcgagcg 300

gtatcagctc actcaaaggc ggtaatacgg ttatccacag aatcagggga taacgcagga 360

aagaacatgt gagcaaaagg ccagcaaaag gccaggaacc gtaaaaaggc cgcgttgctg 420

gcgtttttcc ataggctccg ccccccctgac gagcatcaca aaaatcgacg ctcaagtcag 480

aggtggcgaa acccgacagg actataaaga taccaggcgt ttcccccctgg aagctccctc 540

gtgcgctctc ctgttccgac cctgccgctt accggatacc tgtccgcctt tctcccttcg 600

ggaagcgtgg cgctttctca tagctcacgc tgtaggtatc tcagttcggt gtaggtcgtt 660

cgctccaagc tgggctgtgt gcacgaaccc cccgttcagc ccgaccgctg cgccttatcc 720

ggtaactatc gtcttgagtc caacccggta agacacgact tatcgccact ggcagcagcc 780

actggtaaca ggattagcag agcgaggtat gtaggcggtg ctacagagtt cttgaagtgg 840

tggcctaact acggctacac tagaaggaca gtatttggta tctgcgctct gctgaagcca 900

gttaccttcg gaaaaagagt tggtagctct tgatccggca aacaaaccac cgctggtagc 960

ggtggttttt ttgtttgcaa gcagcagatt acgcgcagaa aaaaaggatc tcaagaagat 1020

cctttgatct tttctacggg gtctgacgct cagtggaacg aaaactcacg ttaagggatt 1080

ttggtcatga gattatcaaa aaggatcttc acctagatcc ttttaaatta aaaatgaagt 1140

tttaaatcaa tctaaagtat atatgagtaa acttggtctg acagttacca atgcttaatc 1200

agtgaggcac ctatctcagc gatctgtcta tttcgttcat ccatagttgc ctgactcccc 1260

gtcgtgtaga taactacgat acgggagggc ttaccatctg gccccagtgc tgcaatgata 1320

ccgcgagacc cacgctcacc ggctccagat ttatcagcaa taaaccagcc agccggaagg 1380

gccgagcgca gaagtggtcc tgcaacttta tccgcctcca tccagtctat taattgttgc 1440

cgggaagcta gagtaagtag ttcgccagtt aatagtttgc gcaacgttgt tgccattgct 1500

acaggcatcg tggtgtcacg ctcgtcgttt ggtatggctt cattcagctc cggttcccaa 1560

cgatcaaggc gagttacatg atcccccatg ttgtgcaaaa aagcggttag ctccttcggt 1620

cctccgatcg ttgtcagaag taagttggcc gcagtgttat cactcatggt tatggcagca 1680

ctgcataatt ctcttactgt catgccatcc gtaagatgct tttctgtgac tggtgagtac 1740

tcaaccaagt cattctgaga atagtgtatg cggcgaccga gttgctcttg cccggcgtca 1800

atacgggata ataccgcgcc acatagcaga actttaaaag tgctcatcat tggaaaacgt 1860

tcttcggggc gaaaactctc aaggatctta ccgctgttga gatccagttc gatgtaaccc 1920

actcgtgcac ccaactgatc ttcagcatct tttactttca ccagcgtttc tgggtgagca 1980

aaaacaggaa ggcaaaatgc cgcaaaaaag ggaataaggg cgacacggaa atgttgaata 2040

ctcatactct tccttttca atattattga agcatttatc agggttattg tctcatgagc 2100

ggatacatat ttgaatgtat ttagaaaaat aaacaaatag gggttccgcg cacatttccc 2160

cgaaaagtgc cacctgacgc gccctgtagc ggcgcattaa gcgcggcggg tgtggtggtt 2220

acgcgcagcg tgaccgctac acttgccagc gccctagcgc ccgctccttt cgctttcttc 2280

ccttcctttc tcgccacgtt cgccggcttt ccccgtcaag ctctaaatcg ggggctccct 2340

ttagggttcc gatttagtgc tttacggcac ctcgacccca aaaaacttga ttagggtgat 2400

ggttcacgta gtgggccatc gccctgatag acggtttttc gccctttgac gttggagtcc 2460

acgttcttta atagtggact cttgttccaa actggaacaa cactcaaccc tatctcggtc 2520

tattcttttg atttataagg gattttgccg atttcggcct attggttaaa aaatgagctg 2580

atttaacaaa aatttaacgc gaattttaac aaaatattaa cgcttacaat ttccattcgc 2640

cattcaggct gcgcaactgt tgggaagggc gatcggtgcg ggcctcttcg ctattacgcc 2700

agctggcgaa aggggatgt gctgcaaggc gattaagttg ggtaacgcca gggttttccc 2760

agtcacgacg ttgtaaaacg acggccagtg aattgtaata cgactcacta tagggcgaat 2820

tgggtaccgg gccccccctc gaggtcgacg agtatctgtc tgactcgtca ttgccgcctt 2880

tggagtacga ctccaactat gagtgtgctt ggatcacttt gacgatacat tcttcgttgg 2940

aggctgtggg tctgacagct gcgttttcgg cgcggttggc cgacaacaat atcagctgca 3000

acgtcattgc tggctttcat catgatcaca tttttgtcgg caaaggcgac gcccagagag 3060

ccattgacgt tctttctaat ttggaccgat agccgtatag tccagtctat ctataagttc 3120

aactaactcg taactattac cataacatat acttcactgc cccagataag gttccgataa 3180

aaagttctgc agactaaatt tatttcagtc tcctcttcac 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 cgagtggtcc gatatcacca 3540

acgcccacgg tgtacccgga accggaatcg attgctggcc tgcgagctgg tgcgtacgag 3600

gaaactgtct ctgaacagaa gaaggaggac gtctctgact acgagaactc ccagtacaag 3660

gagttcctag tcccctctcc caacgagaag ctggccagag gtctgctcat gctggccgag 3720

ctgtcttgca agggctctct ggccactggc gagtactcca agcagaccat tgagcttgcc 3780

cgatccgacc ccgagtttgt ggttggcttc attgcccaga accgacctaa gggcgactct 3840

gaggactggc ttattctgac ccccggggtg ggtcttgacg acaagggaga cgctctcgga 3900

cagcagtacc gaactgttga ggatgtcatg tctaccggaa cggatatcat aattgtcggc 3960

cgaggtctgt acggccagaa ccgagatcct attgaggagg ccaagcgata ccagaaggct 4020

ggctgggagg cttaccagaa gattaactgt tagaggttag actatggata tgtaatttaa 4080

ctgtgtatat agagagcgtg caagtatgga gcgcttgttc agcttgtatg atggtcagac 4140

gacctgtctg atcgagtatg tatgatactg cacaacctgt gtatccgcat gatctgtcca 4200

atggggcatg ttgttgtgtt tctcgatacg gagatgctgg gtacagtgct aatacgttga 4260

actacttata cttatatgag gctcgaagaa agctgacttg tgtatgactt aat 4313

<210> 190
<211> 8015
<212> DNA
<213> Artificial Sequence

<220>
<223> Plasmid pMLPAT-17

<400> 190
ctccatggtc gggtccgtca cccgacccac aaaggccctg ctctatggat cagccctctt 60

cagtttctgc tcattgctca atgtggtcca ggtgttctcc atactcctgc agccgttctc 120

gaagcgtctc ttctttgaag tgaacgctcg cgtggccggc tccatgtgga aggttatgca 180

gctgattatg gagaaaaagc acaaggccgc catcaycttc tcaggagaca agatccctca 240

ccacgagagt gccatcgtct ttggcaacca ccggtccttt gtcgactttt acatgtttca 300

caccgttgct gctyggagag gcatgctcaa ctatatgaag tactttgcca aggactctct 360

gaaatacatt ccattctatg gatggggcat gtggaycatg ggaatgctat tcatcaatcg 420

caactggcag caggatcagc tcaagatcaa caagatgttt gcacggatat tggacatcca 480

agcgcccgtt tgggtcgcca gtttcttgga gggctctcgg ttgacgccca gcaaactgkc 540

tgcctctcaa aagttcatgc tgggacgcgg attgcctctg ctgtcaaacg tcatgatgcc 600

caggaccaag ggattcattg cctgtgtcaa caaattccgg ggaactcatg tgaaatgtgt 660

ttatgatttc acgttcgcct actaccacaa gaccaagggc tttggagtgy ctccagatct 720

ggtccgtgtt cacastggcc agctcagccc cgagtacaaa ttccatgttc atgtgagacg 780

ctatcagctc gacgatctgc ccacggatga ggagaagctg agcgagtggg tggtccaaaa 840

gtatgtggag aaggacgcct ttttggagca gatgaaggag aattggacag atggtattga 900

tgggggtgtg tggtcagaga actggatgtg agcggccgca agtgtggatg gggaagtgag 960

tgcccggttc tgtgtgcaca attggcaatc caagatggat ggattcaaca cagggatata 1020

gcgagctacg tggtggtgcg aggatatagc aacggatatt tatgtttgac acttgagaat 1080

gtacgataca agcactgtcc aagtacaata ctaaacatac tgtacatact catactcgta 1140

cccgggcaac ggtttcactt gagtgcagtg gctagtgctc ttactcgtac agtgtgcaat 1200

actgcgtatc atagtctttg atgtatatcg tattcattca tgttagttgc gtacgagccg 1260

gaagcataaa gtgtaaagcc tggggtgcct aatgagtgag ctaactcaca ttaattgcgt 1320

tgcgctcact gcccgctttc agtcgggaa acctgtcgtg ccagctgcat taatgaatcg 1380

gccaacgcgc ggggagaggc ggtttgcgta ttgggcgctc ttccgcttcc tcgctcactg 1440

actcgctgcg ctcggtcgtt cggctgcggc gagcggtatc agctcactca aaggcggtaa 1500

tacggttatc cacagaatca ggggataacg caggaaagaa catgtgagca aaaggccagc 1560

aaaaggccag gaaccgtaaa aaggccgcgt tgctggcgtt tttccatagg ctccgccccc 1620

ctgacgagca tcacaaaaat cgacgctcaa gtcagaggtg gcgaaacccg acaggactat 1680

aaagatacca ggcgtttccc cctggaagct ccctcgtgcg ctctcctgtt ccgaccctgc 1740

cgcttaccgg atacctgtcc gcctttctcc cttcgggaag cgtggcgctt tctcatagct 1800

cacgctgtag gtatctcagt tcggtgtagg tcgttcgctc caagctgggc tgtgtgcacg 1860

aacccccccgt tcagcccgac cgctgcgcct tatccggtaa ctatcgtctt gagtccaacc 1920

cggtaagaca cgacttatcg ccactggcag cagccactgg taacaggatt agcagagcga 1980

ggtatgtagg cggtgctaca gagttcttga agtggtggcc taactacggc tacactagaa 2040

ggacagtatt tggtatctgc gctctgctga agccagttac cttcggaaaa agagttggta 2100

gctcttgatc cggcaaacaa accaccgctg gtagcggtgg ttttttttgtt tgcaagcagc 2160

agattacgcg cagaaaaaaa ggatctcaag aagatccttt gatcttttct acggggtctg 2220

acgctcagtg gaacgaaaac tcacgttaag ggattttggt catgagatta tcaaaaagga 2280

tcttcaccta gatcctttta aattaaaaat gaagttttaa atcaatctaa agtatatatg 2340

agtaaacttg gtctgacagt taccaatgct taatcagtga ggcacctatc tcagcgatct 2400

gtctatttcg ttcatccata gttgcctgac tccccgtcgt gtagataact acgatacggg 2460

agggcttacc atctggcccc agtgctgcaa tgataccgcg agacccacgc tcaccggctc 2520

cagatttatc agcaataaac cagccagccg gaagggccga gcgcagaagt ggtcctgcaa 2580

ctttatccgc ctccatccag tctattaatt gttgccggga agctagagta agtagttcgc 2640

cagttaatag tttgcgcaac gttgttgcca ttgctacagg catcgtggtg tcacgctcgt 2700

cgtttggtat ggcttcattc agctccggtt cccaacgatc aaggcgagtt acatgatccc 2760

ccatgttgtg caaaaaagcg gttagctcct tcggtcctcc gatcgttgtc agaagtaagt 2820

tggccgcagt gttatcactc atggttatgg cagcactgca taattctctt actgtcatgc 2880

catccgtaag atgcttttct gtgactggtg agtactcaac caagtcattc tgagaatagt 2940

gtatgcggcg accgagttgc tcttgcccgg cgtcaatacg ggataatacc gcgccacata 3000

gcagaacttt aaaagtgctc atcattggaa aacgttcttc ggggcgaaaa ctctcaagga 3060

tcttaccgct gttgagatcc agttcgatgt aacccactcg tgcacccaac tgatcttcag 3120

catcttttac tttcaccagc gtttctgggt gagcaaaaac aggaaggcaa aatgccgcaa 3180

aaaagggaat aagggcgaca cggaaatgtt gaatactcat actcttcctt tttcaatatt 3240

attgaagcat ttatcagggt tattgtctca tgagcggata catatttgaa tgtatttaga 3300

aaaataaaca aatagggggtt ccgcgcacat ttccccgaaa agtgccacct gacgcgccct 3360

gtagcggcgc attaagcgcg gcgggtgtgg tggttacgcg cagcgtgacc gctacacttg 3420

ccagcgccct agcgcccgct cctttcgctt tcttcccttc ctttctcgcc acgttcgccg 3480

gctttccccg tcaagctcta aatcgggggc tccctttagg gttccgattt agtgctttac 3540

ggcacctcga ccccaaaaaa cttgattagg gtgatggttc acgtagtggg ccatcgccct 3600

gatagacggt ttttcgccct ttgacgttgg agtccacgtt ctttaatagt ggactcttgt 3660

tccaaactgg aacaacactc aaccctatct cggtctattc ttttgattta taagggattt 3720

tgccgatttc ggcctattgg ttaaaaaatg agctgattta acaaaaattt aacgcgaatt 3780

ttaacaaaat attaacgctt acaatttcca ttcgccattc aggctgcgca actgttggga 3840

agggcgatcg gtgcgggcct cttcgctatt acgccagctg gcgaaagggg gatgtgctgc 3900

aaggcgatta agttgggtaa cgccagggtt ttcccagtca cgacgttgta aaacgacggc 3960

cagtgaattg taatacgact cactataggg cgaattgggt accgggcccc ccctcgaggt 4020

cgatggtgtc gataagcttg atatcgaatt catgtcacac aaaccgatct tcgcctcaag 4080

gaaacctaat tctacatccg agagactgcc gagatccagt ctacactgat taattttcgg 4140

gccaataatt taaaaaaatc gtgttatata atattatatg tattatatat atacatcatg 4200

atgatactga cagtcatgtc ccattgctaa atagacagac tccatctgcc gcctccaact 4260

gatgttctca atatttaagg ggtcatctcg cattgtttaa taataaacag actccatcta 4320

ccgcctccaa atgatgttct caaaatatat tgtatgaact tatttttatt acttagtatt 4380

attagacaac ttacttgctt tatgaaaaac acttcctatt taggaaacaa tttataatgg 4440

cagttcgttc atttaacaat ttatgtagaa taaatgttat aaatgcgtat gggaaatctt 4500

aaatatggat agcataaatg atatctgcat tgcctaattc gaaatcaaca gcaacgaaaa 4560

aaatcccttg tacaacataa atagtcatcg agaaatatca actatcaaag aacagctatt 4620

cacacgttac tattgagatt attattggac gagaatcaca cactcaactg tctttctctc 4680

ttctagaaat acaggtacaa gtatgtacta ttctcattgt tcatacttct agtcatttca 4740

tcccacatat tccttggatt tctctccaat gaatgacatt ctatcttgca aattcaacaa 4800

ttataataag atataccaaa gtagcggtat agtggcaatc aaaaagcttc tctggtgtgc 4860

ttctcgtatt tatttttatt ctaatgatcc attaaaggta tatatttatt tcttgttata 4920

taatcctttt gtttattaca tgggctggat acataaaggt attttgattt aatttttgc 4980

ttaaattcaa tccccccctcg ttcagtgtca actgtaatgg taggaaatta ccatactttt 5040

gaagaagcaa aaaaaatgaa agaaaaaaaa aatcgtattt ccaggttaga cgttccgcag 5100

aatctagaat gcggtatgcg gtacattgtt cttcgaacgt aaaagttgcg ctccctgaga 5160

tattgtacat ttttgctttt acaagtacaa gtacatcgta caactatgta ctactgttga 5220

tgcatccaca acagtttgtt ttgttttttt ttgttttttt tttttctaat gattcattac 5280

cgctatgtat acctacttgt acttgtagta agccgggtta ttggcgttca attaatcata 5340

gacttatgaa tctgcacggt gtgcgctgcg agttactttt agcttatgca tgctacttgg 5400

gtgtaatatt gggatctgtt cggaaatcaa cggatgctca atcgatttcg acagtaatta 5460

attaagtcat acacaagtca gctttcttcg agcctcatat aagtataagt agttcaacgt 5520

attagcactg tacccagcat ctccgtatcg agaaacacaa caacatgccc cattggacag 5580

atcatgcgga tacacaggtt gtgcagtatc atacatactc gatcagacag gtcgtctgac 5640

catcatacaa gctgaacaag cgctccatac ttgcacgctc tctatataca cagttaaatt 5700

acatatccat agtctaacct ctaacagtta atcttctggt aagcctccca gccagccttc 5760

tggtatcgct tggcctcctc aataggatct cggttctggc cgtacagacc tcggccgaca 5820

attatgatat ccgttccggt agacatgaca tcctcaacag ttcggtactg ctgtccgaga 5880

gcgtctccct tgtcgtcaag acccaccccg ggggtcagaa taagccagtc ctcagagtcg 5940

cccttaggtc ggttctgggc aatgaagcca accacaaact cggggtcgga tcgggcaagc 6000

tcaatggtct gcttggagta ctcgccagtg gccagagagc ccttgcaaga cagctcggcc 6060

agcatgagca gacctctggc cagcttctcg ttgggagagg ggactaggaa ctccttgtac 6120

tgggagttct cgtagtcaga gacgtcctcc ttcttctgtt cagagacagt ttcctcggca 6180

ccagctcgca ggccagcaat gattccggtt ccgggtacac cgtgggcgtt ggtgatatcg 6240

gaccactcgg cgattcggtg acaccggtac tggtgcttga cagtgttgcc aatatctgcg 6300

aactttctgt cctcgaacag gaagaaaccg tgcttaagag caagttcctt gaggggggagc 6360

acagtgccgg cgtaggtgaa gtcgtcaatg atgtcgatat gggttttgat catgcacaca 6420

taaggtccga ccttatcggc aagctcaatg agctccttgg tggtggtaac atccagagaa 6480

gcacacaggt tggttttctt ggctgccacg agcttgagca ctcgagcggc aaaggcggac 6540

ttgtggacgt tagctcgagc ttcgtaggag ggcattttgg tggtgaagag gagactgaaa 6600

taaatttagt ctgcagaact ttttatcgga accttatctg gggcagtgaa gtatatgtta 6660

tggtaatagt tacgagttag ttgaacttat agatagactg gactatacgg ctatcggtcc 6720

aaattagaaa gaacgtcaat ggctctctgg gcgtcgcctt tgccgacaaa aatgtgatca 6780

tgatgaaagc cagcaatgac gttgcagctg atattgttgt cggccaaccg cgccgaaaac 6840

gcagctgtca gacccacagc ctccaacgaa gaatgtatcg tcaaagtgat ccaagcacac 6900

tcatagttgg agtcgtactc caaaggcggc aatgacgagt cagacagata ctcgtcgact 6960

caggcgacga cggaattcct gcagcccatc tgcagaattc aggagagacc gggttggcgg 7020

cgtatttgtg tcccaaaaaa cagccccaat tgccccggag aagacggcca ggccgcctag 7080

atgacaaatt caacaactca cagctgactt tctgccattg ccactagggg ggggccttttt 7140

tatatggcca agccaagctc tccacgtcgg ttgggctgca cccaacaata aatgggtagg 7200

gttgcaccaa caaagggatg ggatgggggg tagaagatac gaggataacg gggctcaatg 7260

gcacaaataa gaacgaatac tgccattaag actcgtgatc cagcgactga caccattgca 7320

tcatctaagg gcctcaaaac tacctcggaa ctgctgcgct gatctggaca ccacagaggt 7380

tccgagcact ttaggttgca ccaaatgtcc caccaggtgc aggcagaaaa cgctggaaca 7440

gcgtgtacag tttgtcttaa caaaaagtga gggcgctgag gtcgagcagg gtggtgtgac 7500

ttgttatagc ctttagagct gcgaaagcgc gtatggattt ggctcatcag gccagattga 7560

gggtctgtgg acacatgtca tgttagtgta cttcaatcgc cccctggata tagccccgac 7620

aataggccgt ggcctcattt ttttgccttc cgcacatttc cattgctcgg tacccacacc 7680

ttgcttctcc tgcacttgcc aaccttaata ctggtttaca ttgaccaaca tcttacaagc 7740

ggggggcttg tctagggtat atataaacag tggctctccc aatcggttgc cagtctcttt 7800

tttcctttct ttccccacag attcgaaatc taaactacac atcacacaat gcctgttact 7860

gacgtcctta agcgaaagtc cggtgtcatc gtcggcgacg atgtccgagc cgtgagtatc 7920

cacgacaaga tcagtgtcga gacgacgcgt tttgtgtaat gacacaatcc gaaagtcgct 7980

agcaacacac actctctaca caaactaacc cagct 8015

<210> 191
<211> 8411
<212> DNA
<213> Artificial Sequence

<220>
<223> Plasmid pMLPAT-Int

<400> 191
cgataccatg gtcgggcaat gagaacggca gcaactgcaa tcacagcgac atatatccaa 60

ctgttcatgt ttggttttcg gatagtcaca caccttaatt ttgatgcacg ctttatggag 120

tctctctctc tttttctctc tctcttgtcg tcgctctttc ttgtttttca accacccact 180

tgattcctgc aaacaaacta cccacactaa tttttttttcg ctgcataccc tcaaatgagc 240

ctaattggcg tgtgtctccg cacaaaaaca caccatgcac ggctgggctt gttgggaaac 300

tttgtcaggg gggtccaggg ggccattggc agacttggcc acgtgtgctc atctcggctt 360

cgtcgttatt acgtgtctgt gtaatcaaag tcgggcgttt tttgcgccat gtgtccgcat 420

gaaattggcc cctcttgaag tcccttgtgc acctacacgt gccgaaatga aggttggagt 480

cagcggggtc atgccgtggt attatgctgt ggcatgtggc attaagctgt ggcatcaagc 540

cgtggaatca agccgtggtt cacgcccttg attgcgcagg cacatggcgc cattcttgcc 600

tctgctgtaa gcccggcttt gtgtgattca gagacgctgc taccgcacaa ctgcccatac 660

tccttctcct actgtataca tccacccctc atgctgataa cattatcatc tcatctcaac 720

tcaacatttc caccaacttg ggatcaaaaa cacgttctaa tactgtacac tgtctacgat 780

ataattaccg tacagtgtgg ttataatgaa catctattag agagaattgt ttgttgctcg 840

tatcagtcat tgggaagcgg gataccatgt cattttcacc tatatcaacc atgaaactac 900

agtatgtaca gtagaagtat atactgtact gttttgtaac tatatgtaca gtagaagtat 960

atactatact gttattaact atacttgtga ctagtgttcc aaactacaag tatatactgt 1020

acttgtacac gactatccga ccagtatcca gtatacaata accaactact ctacgtacgt 1080

actaaactaa acaaatagat caatgctcaa tgtcgagctc cagcttttgt tccctttagt 1140

gagggttaat ttcgagcttg gcgtaatcat ggtcatagct gtttcctgtg tgaaattgtt 1200

atccgctcac aattccacac aacatacgag ccggaagcat aaagtgtaaa gcctggggtg 1260

cctaatgagt gagctaactc acattaattg cgttgcgctc actgcccgct ttccagtcgg 1320

gaaacctgtc gtgccagctg cattaatgaa tcggccaacg cgcggggaga ggcggtttgc 1380

gtattgggcg ctcttccgct tcctcgctca ctgactcgct gcgctcggtc gttcggctgc 1440

ggcgagcggt atcagctcac tcaaaggcgg taatacggtt atccacagaa tcaggggata 1500

acgcaggaaa gaacatgtga gcaaaaggcc agcaaaaggc caggaaccgt aaaaaggccg 1560

cgttgctggc gtttttccat aggctccgcc cccctgacga gcatcacaaa aatcgacgct 1620

caagtcagag gtggcgaaac ccgacaggac tataaagata ccaggcgttt ccccctggaa 1680

gctccctcgt gcgctctcct gttccgaccc tgccgcttac cggatacctg tccgcctttc 1740

tcccttcggg aagcgtggcg ctttctcata gctcacgctg taggtatctc agttcggtgt 1800

aggtcgttcg ctccaagctg ggctgtgtgc acgaaccccc cgttcagccc gaccgctgcg 1860

ccttatccgg taactatcgt cttgagtcca acccggtaag acacgactta tcgccactgg 1920

cagcagccac tggtaacagg attagcagag cgaggtatgt aggcggtgct acagagttct 1980

tgaagtggtg gcctaactac ggctacacta gaaggacagt atttggtatc tgcgctctgc 2040

tgaagccagt taccttcgga aaaagagttg gtagctcttg atccggcaaa caaaccaccg 2100

ctggtagcgg tggtttttt gtttgcaagc agcagattac gcgcagaaaa aaaggatctc 2160

aagaagatcc tttgatcttt tctacggggt ctgacgctca gtggaacgaa aactcacgtt 2220

aagggatttt ggtcatgaga ttatcaaaaa ggatcttcac ctagatcctt ttaaattaaa 2280

aatgaagttt taaatcaatc taaagtatat atgagtaaac ttggtctgac agttaccaat 2340

gcttaatcag tgaggcacct atctcagcga tctgtctatt tcgttcatcc atagttgcct 2400

gactccccgt cgtgtagata actacgatac gggagggctt accatctggc cccagtgctg 2460

caatgatacc gcgagaccca cgctcaccgg ctccagattt atcagcaata aaccagccag 2520

ccggaagggc cgagcgcaga agtggtcctg caactttatc cgcctccatc cagtctatta 2580

attgttgccg ggaagctaga gtaagtagtt cgccagttaa tagtttgcgc aacgttgttg 2640

ccattgctac aggcatcgtg gtgtcacgct cgtcgtttgg tatggcttca ttcagctccg 2700

gttcccaacg atcaaggcga gttacatgat cccccatgtt gtgcaaaaaa gcggttagct 2760

ccttcggtcc tccgatcgtt gtcagaagta agttggccgc agtgttatca ctcatggtta 2820

tggcagcact gcataattct cttactgtca tgccatccgt aagatgcttt tctgtgactg 2880

gtgagtactc aaccaagtca ttctgagaat agtgtatgcg gcgaccgagt tgctcttgcc 2940

cggcgtcaat acgggataat accgcgccac atagcagaac tttaaaagtg ctcatcattg 3000

gaaaacgttc ttcggggcga aaactctcaa ggatcttacc gctgttgaga tccagttcga 3060

tgtaacccac tcgtgcaccc aactgatctt cagcatcttt tactttcacc agcgtttctg 3120

ggtgagcaaa aacaggaagg caaaatgccg caaaaaaggg aataagggcg acacggaaat 3180

gttgaatact catactcttc ctttttcaat attattgaag catttatcag ggttattgtc 3240

tcatgagcgg atacatattt gaatgtattt agaaaaataa acaaataggg gttccgcgca 3300

catttccccg aaaagtgcca cctgacgcgc cctgtagcgg cgcattaagc gcggcgggtg 3360

tggtggttac gcgcagcgtg accgctacac ttgccagcgc cctagcgccc gctcctttcg 3420

ctttcttccc ttcctttctc gccacgttcg ccggctttcc ccgtcaagct ctaaatcggg 3480

ggctcccttt agggttccga tttagtgctt tacggcacct cgaccccaaa aaacttgatt 3540

agggtgatgg ttcacgtagt gggccatcgc cctgatagac ggtttttcgc cctttgacgt 3600

tggagtccac gttctttaat agtggactct tgttccaaac tggaacaaca ctcaacccta 3660

tctcggtcta ttcttttgat ttataaggga ttttgccgat ttcggcctat tggttaaaaa 3720

atgagctgat ttaacaaaaa tttaacgcga attttaacaa aatattaacg cttacaattt 3780

ccattcgcca ttcaggctgc gcaactgttg ggaagggcga tcggtgcggg cctcttcgct 3840

attacgccag ctggcgaaag ggggatgtgc tgcaaggcga ttaagttggg taacgccagg 3900

gttttcccag tcacgacgtt gtaaaacgac ggccagtgaa ttgtaatacg actcactata 3960

gggcgaattg ggtaccgggc ccccccctcga gtctagagca gggtgttgga ggggatggag 4020

aggttgagta aagcggaagt tgcggtttgc tggtcgcgtt ttaatcttct tttgtaattt 4080

aatctcggta tgccgacgtg tttcggcgag ttaattttag ctgtcaaaaa aatggatcac 4140

caggatagac aaagaagaga gtgggacaag gatctttcca gccgctaata tcggccagtt 4200

taagagcaat tcacggtccc tggagtccat ataacaacgc caatctcgca cttgtctgct 4260

ccctcgtcac ccgtccaact gtccaatcca atctatgcct ctcatgatca tcttcacact 4320

acagtagagt aggtaaaggt agacttatgg gctcctcgaa taattggact ctgtctaaac 4380

cttgtgaagt tacctcggtc ggtagagttc ggctggaagg ccatgattac taacaacatc 4440

gtacgagaga aaatcaagat gagtaataca atctcgatga gtaatacaat ctcgatgagt 4500

aatacaatct cgatgagtaa gactaggtgg tgtcacgact tttagaggaa tgagcagctt 4560

tcagggttat tatagagaca cgtccgcgga cgaagtagct ggtacatcga gcataagcat 4620

ggtacaagta ggagtagact aaaaaccaac agtttgaata gtccatgaca gtacgggcgg 4680

gtacgactga tctaagagaa ctggggtata cacgatatag cacagtacag agaaagtggg 4740

ctcgttttgc gttggtaatc gaggtagatt tcgttgctat attaatccat tcacccatag 4800

ctccacagcc aatggttcgc cgtgggtgtc gatctgaaaa atgtttcata ttacctatct 4860

ctctcctaaa gtagctacaa gcacttcttg tgctgcagtc tgcggccgct cacatccagt 4920

tctctgacca cacaccccca tcaataccat ctgtccaatt ctccttcatc tgctccaaaa 4980

aggcgtcctt ctccacatac ttttggacca cccactcgct cagcttctcc tcatccgtgg 5040

gcagatcgtc gagctgatag cgtctcacat gaacatggaa tttgtactcg gggctgagct 5100

ggccastgtg aacacggacc agatctggag rcactccaaa gcccttggtc ttgtggtagt 5160

aggcgaacgt gaaatcataa acacatttca catgagttcc ccggaatttg ttgacacagg 5220

caatgaatcc cttggtcctg ggcatcatga cgtttgacag cagaggcaat ccgcgtccca 5280

gcatgaactt ttgagaggca gmcagtttgc tgggcgtcaa ccgagagccc tccaagaaac 5340

tggcgaccca aacgggcgct tggatgtcca atatccgtgc aaacatcttg ttgatcttga 5400

gctgatcctg ctgccagttg cgattgatga atagcattcc catgrtccac atgccccatc 5460

catagaatgg aatgtatttc agagagtcct tggcaaagta cttcatatag ttgagcatgc 5520

ctctccragc agcaacggtg tgaaacatgt aaaagtcgac aaaggaccgg tggttgccaa 5580

agacgatggc actctcgtgg tgagggatct tgtctcctga gaagrtgatg gcggccttgt 5640

gctttttctc cataatcagc tgcataacct tccacatgga gccggccacg cgagcgttca 5700

cttcaaagaa gagacgcttc gagaacggct gcaggagtat ggagaacacc tggaccacat 5760

tgagcaatga gcagaaactg aagagggctg atccatagag cagggccttt gtgggtcggg 5820

tgacggaccc gaccatggag agctgggtta gtttgtgtag agagtgtgtg ttgctagcga 5880

ctttcggatt gtgtcattac acaaaacgcg tcgtctcgac actgatcttg tcgtggatac 5940

tcacggctcg gacatcgtcg ccgacgatga caccggactt tcgcttaagg acgtcagtaa 6000

caggcattgt gtgatgtgta gtttagattt cgaatctgtg gggaaagaaa ggaaaaaaga 6060

gactggcaac cgattgggag agccactgtt tatatatacc ctagacaagc cccccgcttg 6120

taagatgttg gtcaatgtaa accagtatta aggttggcaa gtgcaggaga agcaaggtgt 6180

gggtaccgag caatggaaat gtgcggaagg caaaaaaatg aggccacggc ctattgtcgg 6240

ggctatatcc agggggcgat tgaagtacac taacatgaca tgtgtccaca gaccctcaat 6300

ctggcctgat gagccaaatc catacgcgct ttcgcagctc taaaggctat aacaagtcac 6360

accaccctgc tcgacctcag cgccctcact ttttgttaag acaaactgta cacgctgttc 6420

cagcgttttc tgcctgcacc tggtgggaca tttggtgcaa cctaaagtgc tcggaacctc 6480

tgtggtgtcc agatcagcgc agcagttccg aggtagtttt gaggccctta gatgatgcaa 6540

tggtgtcagt cgctggatca cgagtcttaa tggcagtatt cgttcttatt tgtgccattg 6600

agccccgtta tcctcgtatc ttctaccccc catcccatcc ctttgttggt gcaaccctac 6660

ccatttattg ttgggtgcag cccaaccgac gtggagagct tggcttggcc atataaaaag 6720

gccccccccc agtggcaatg gcagaaagtc agctgtgagt tgttgaattt gtcatctagg 6780

cggcctggcc gtcttctccg gggcaattgg ggctgttttt tgggacacaa atacgccgcc 6840

aacccggtct ctcctgaatt ctgcagatgg gctgcaggaa ttccgtcgtc gcctgagtcg 6900

acgagtatct gtctgactcg tcattgccgc ctttggagta cgactccaac tatgagtgtg 6960

cttggatcac tttgacgata cattcttcgt tggaggctgt gggtctgaca gctgcgtttt 7020

cggcgcggtt ggccgacaac aatatcagct gcaacgtcat tgctggcttt catcatgatc 7080

acatttttgt cggcaaaggc gacgcccaga gagccattga cgttctttct aatttggacc 7140

gatagccgta tagtccagtc tatctataag ttcaactaac tcgtaactat taccataaca 7200

tatacttcac tgccccagat aaggttccga taaaaagttc tgcagactaa atttatttca 7260

gtctcctctt caccaccaaa atgccctcct acgaagctcg agctaacgtc cacaagtccg 7320

cctttgccgc tcgagtgctc aagctcgtgg cagccaagaa aaccaacctg tgtgcttctc 7380

tggatgttac caccaccaag gagctcattg agcttgccga taaggtcgga ccttatgtgt 7440

gcatgatcaa aacccatatc gacatcattg acgacttcac ctacgccggc actgtgctcc 7500

ccctcaagga acttgctctt aagcacggtt tcttcctgtt cgaggacaga aagttcgcag 7560

atattggcaa cactgtcaag caccagtacc ggtgtcaccg aatcgccgag tggtccgata 7620

tcaccaacgc ccacggtgta cccggaaccg gaatcattgc tggcctgcga gctggtgccg 7680

aggaaactgt ctctgaacag aagaaggagg acgtctctga ctacgagaac tcccagtaca 7740

aggagttcct agtcccctct cccaacgaga agctggccag aggtctgctc atgctggccg 7800

agctgtcttg caagggctct ctggccactg gcgagtactc caagcagacc attgagcttg 7860

cccgatccga ccccgagttt gtggttggct tcattgccca gaaccgacct aagggcgact 7920

ctgaggactg gcttattctg acccccgggg tgggtcttga cgacaaggga gacgctctcg 7980

gacagcagta ccgaactgtt gaggatgtca tgtctaccgg aacggatatc ataattgtcg 8040

gccgaggtct gtacggccag aaccgagatc ctattgagga ggccaagcga taccagaagg 8100

ctggctggga ggcttaccag aagattaact gttagaggtt agactatgga tatgtaattt 8160

aactgtgtat atagagagcg tgcaagtatg gagcgcttgt tcagcttgta tgatggtcag 8220

acgacctgtc tgatcgagta tgtatgatac tgcacaacct gtgtatccgc atgatctgtc 8280

caatggggca tgttgttgtg tttctcgata cggagatgct gggtacagtg ctaatacgtt 8340

gaactactta tacttatatg aggctcgaag aaagctgact tgtgtatgac ttaattaatt 8400

actgtcgaaa t 8411

<210> 192
<211> 7323
<212> DNA
<213> Artificial Sequence

<220>
<223> Plasmid pZUF-MOD-1

<400> 192
gtacgagccg gaagcataaa gtgtaaagcc tggggtgcct aatgagtgag ctaactcaca 60

ttaattgcgt tgcgctcact gcccgctttc agtcgggaa acctgtcgtg ccagctgcat 120

taatgaatcg gccaacgcgc ggggagaggc ggtttgcgta ttgggcgctc ttccgcttcc 180

tcgctcactg actcgctgcg ctcggtcgtt cggctgcggc gagcggtatc agctcactca 240

aaggcggtaa tacggttatc cacagaatca ggggataacg caggaaagaa catgtgagca 300

aaaggccagc aaaaggccag gaaccgtaaa aaggccgcgt tgctggcgtt tttccatagg 360

ctccgccccc ctgacgagca tcacaaaaat cgacgctcaa gtcagaggtg gcgaaacccg 420

acaggactat aaagatacca ggcgtttccc cctggaagct ccctcgtgcg ctctcctgtt 480

ccgaccctgc cgcttaccgg atacctgtcc gcctttctcc cttcgggaag cgtggcgctt 540

tctcatagct cacgctgtag gtatctcagt tcggtgtagg tcgttcgctc caagctgggc 600

tgtgtgcacg aaccccccgt tcagcccgac cgctgcgcct tatccggtaa ctatcgtctt 660

gagtccaacc cggtaagaca cgacttatcg ccactggcag cagccactgg taacaggatt 720

agcagagcga ggtatgtagg cggtgctaca gagttcttga agtggtggcc taactacggc 780

tacactagaa ggacagtatt tggtatctgc gctctgctga agccagttac cttcggaaaa 840

agagttggta gctcttgatc cggcaaacaa accaccgctg gtagcggtgg ttttttgtt 900

tgcaagcagc agattacgcg cagaaaaaaa ggatctcaag aagatccttt gatcttttct 960

acggggtctg acgctcagtg gaacgaaaac tcacgttaag ggattttggt catgagatta 1020

tcaaaaagga tcttcaccta gatcctttta aattaaaaat gaagttttaa atcaatctaa 1080

agtatatatg agtaaacttg gtctgacagt taccaatgct taatcagtga ggcacctatc 1140

tcagcgatct gtctatttcg ttcatccata gttgcctgac tccccgtcgt gtagataact 1200

acgatacggg agggcttacc atctggcccc agtgctgcaa tgataccgcg agacccacgc 1260

tcaccggctc cagatttatc agcaataaac cagccagccg gaagggccga gcgcagaagt 1320

ggtcctgcaa ctttatccgc ctccatccag tctattaatt gttgccggga agctagagta 1380

agtagttcgc cagttaatag tttgcgcaac gttgttgcca ttgctacagg catcgtggtg 1440

tcacgctcgt cgtttggtat ggcttcattc agctccggtt cccaacgatc aaggcgagtt 1500

acatgatccc ccatgttgtg caaaaaagcg gttagctcct tcggtcctcc gatcgttgtc 1560

agaagtaagt tggccgcagt gttatcactc atggttatgg cagcactgca taattctctt 1620

actgtcatgc catccgtaag atgcttttct gtgactggtg agtactcaac caagtcattc 1680

tgagaatagt gtatgcggcg accgagttgc tcttgcccgg cgtcaatacg ggataatacc 1740

gcgccacata gcagaacttt aaaagtgctc atcattggaa aacgttcttc ggggcgaaaa 1800

ctctcaagga tcttaccgct gttgagatcc agttcgatgt aacccactcg tgcacccaac 1860

tgatcttcag catcttttac tttcaccagc gtttctgggt gagcaaaaac aggaaggcaa 1920

aatgccgcaa aaaagggaat aagggcgaca cggaaatgtt gaatactcat actcttcctt 1980

tttcaatatt attgaagcat ttatcagggt tattgtctca tgagcggata catatttgaa 2040

tgtatttaga aaaataaaca aatagggttt ccgcgcacat ttccccgaaa agtgccacct 2100

gacgcgccct gtagcggcgc attaagcgcg gcgggtgtgg tggttacgcg cagcgtgacc 2160

gctacacttg ccagcgccct agcgcccgct cctttcgctt tcttcccttc ctttctcgcc 2220

acgttcgccg gctttccccg tcaagctcta aatcgggggc tcccctttagg gttccgattt 2280

agtgctttac ggcacctcga ccccaaaaaa cttgattagg gtgatggttc acgtagtggg 2340

ccatcgccct gatagacggt ttttcgccct ttgacgttgg agtccacgtt ctttaatagt 2400

ggactcttgt tccaaactgg aacaacactc aaccctatct cggtctattc ttttgattta 2460

taagggattt tgccgatttc ggcctattgg ttaaaaaatg agctgattta acaaaaattt 2520

aacgcgaatt ttaacaaaat attaacgctt acaatttcca ttcgccattc aggctgcgca 2580

actgttggga agggcgatcg gtgcgggcct cttcgctatt acgccagctg gcgaaagggg 2640

gatgtgctgc aaggcgatta agttgggtaa cgccagggtt ttcccagtca cgacgttgta 2700

aaacgacggc cagtgaattg taatacgact cactataggg cgaattgggt accgggcccc 2760

ccctcgaggt cgatggtgtc gataagcttg atatcgaatt catgtcacac aaaccgatct 2820

tcgcctcaag gaaacctaat tctacatccg agagactgcc gagatccagt ctacactgat 2880

taattttcgg gccaataatt taaaaaaatc gtgttatata atattatatg tattatatat 2940

atacatcatg atgatactga cagtcatgtc ccattgctaa atagacagac tccatctgcc 3000

gcctccaact gatgttctca atatttaagg ggtcatctcg cattgtttaa taataaacag 3060

actccatcta ccgcctccaa atgatgttct caaaatatat tgtatgaact tatttttatt 3120

acttagtatt attagacaac ttacttgctt tatgaaaaac acttcctatt taggaaacaa 3180

tttataatgg cagttcgttc atttaacaat ttatgtagaa taaatgttat aaatgcgtat 3240

gggaaatctt aaatatggat agcataaatg atatctgcat tgcctaattc gaaatcaaca 3300

gcaacgaaaa aaatcccttg tacaacataa atagtcatcg agaaatatca actatcaaag 3360

aacagctatt cacacgttac tattgagatt attattggac gagaatcaca cactcaactg 3420

tctttctctc ttctagaaat acaggtacaa gtatgtacta ttctcattgt tcatacttct 3480

agtcatttca tcccacatat tccttggatt tctctccaat gaatgacatt ctatcttgca 3540

aattcaacaa ttataataag atataccaaa gtagcggtat agtggcaatc aaaaagcttc 3600

tctggtgtgc ttctcgtatt tatttttatt ctaatgatcc attaaaggta tatatttatt 3660

tcttgttata taatcctttt gtttattaca tgggctggat acataaaggt attttgattt 3720

aatttttgc ttaaattcaa tccccccctcg ttcagtgtca actgtaatgg taggaaatta 3780

ccatactttt gaagaagcaa aaaaaatgaa agaaaaaaaa aatcgtattt ccaggttaga 3840

cgttccgcag aatctagaat gcggtatgcg gtacattgtt cttcgaacgt aaaagttgcg 3900

ctccctgaga tattgtacat ttttgctttt acaagtacaa gtacatcgta caactatgta 3960

ctactgttga tgcatccaca acagtttgtt ttgttttttt ttgttttttt tttttctaat 4020

gattcattac cgctatgtat acctacttgt acttgtagta agccgggtta ttggcgttca 4080

attaatcata gacttatgaa tctgcacggt gtgcgctgcg agttactttt agcttatgca 4140

tgctacttgg gtgtaatatt gggatctgtt cggaaatcaa cggatgctca atcgatttcg 4200

acagtaatta attaagtcat acacaagtca gctttcttcg agcctcatat aagtataagt 4260

agttcaacgt attagcactg tacccagcat ctccgtatcg agaaacacaa caacatgccc 4320

cattggacag atcatgcgga tacacaggtt gtgcagtatc atacatactc gatcagacag 4380

gtcgtctgac catcatacaa gctgaacaag cgctccatac ttgcacgctc tctatataca 4440

cagttaaatt acatatccat agtctaacct ctaacagtta atcttctggt aagcctccca 4500

gccagccttc tggtatcgct tggcctcctc aataggatct cggttctggc cgtacagacc 4560

tcggccgaca attatgatat ccgttccggt agacatgaca tcctcaacag ttcggtactg 4620

ctgtccgaga gcgtctccct tgtcgtcaag acccaccccg ggggtcagaa taagccagtc 4680

ctcagagtcg cccttaggtc ggttctgggc aatgaagcca accacaaact cggggtcgga 4740

tcgggcaagc tcaatggtct gcttggagta ctcgccagtg gccagagagc ccttgcaaga 4800

cagctcggcc agcatgagca gacctctggc cagcttctcg ttgggagagg ggactaggaa 4860

ctccttgtac tgggagttct cgtagtcaga gacgtcctcc ttcttctgtt cagagacagt 4920

ttcctcggca ccagctcgca ggccagcaat gattccggtt ccgggtacac cgtgggcgtt 4980

ggtgatatcg gaccactcgg cgattcggtg acaccggtac tggtgcttga cagtgttgcc 5040

aatatctgcg aactttctgt cctcgaacag gaagaaaccg tgcttaagag caagttcctt 5100

gaggggggagc acagtgccgg cgtaggtgaa gtcgtcaatg atgtcgatat gggttttgat 5160

catgcacaca taaggtccga ccttatcggc aagctcaatg agctccttgg tggtggtaac 5220

atccagagaa gcacacaggt tggttttctt ggctgccacg agcttgagca ctcgagcggc 5280

aaaggcggac ttgtggacgt tagctcgagc ttcgtaggag ggcattttgg tggtgaagag 5340

gagactgaaa taaatttagt ctgcagaact ttttatcgga accttatctg gggcagtgaa 5400

gtatatgtta tggtaatagt tacgagttag ttgaacttat agatagactg gactatacgg 5460

ctatcggtcc aaattagaaa gaacgtcaat ggctctctgg gcgtcgcctt tgccgacaaa 5520

aatgtgatca tgatgaaagc cagcaatgac gttgcagctg atattgttgt cggccaaccg 5580

cgccgaaaac gcagctgtca gacccacagc ctccaacgaa gaatgtatcg tcaaagtgat 5640

ccaagcacac tcatagttgg agtcgtactc caaaggcggc aatgacgagt cagacagata 5700

ctcgtcgact caggcgacga cggaattcct gcagcccatc tgcagaattc aggagagacc 5760

gggttggcgg cgtatttgtg tcccaaaaaa cagccccaat tgccccggag aagacggcca 5820

ggccgcctag atgacaaatt caacaactca cagctgactt tctgccattg ccactagggg 5880

ggggcctttt tatatggcca agccaagctc tccacgtcgg ttgggctgca cccaacaata 5940

aatgggtagg gttgcaccaa caaagggatg ggatgggggg tagaagatac gaggataacg 6000

gggctcaatg gcacaaataa gaacgaatac tgccattaag actcgtgatc cagcgactga 6060

caccattgca tcatctaagg gcctcaaaac tacctcggaa ctgctgcgct gatctggaca 6120

ccacagaggt tccgagcact ttaggttgca ccaaatgtcc caccaggtgc aggcagaaaa 6180

cgctggaaca gcgtgtacag tttgtcttaa caaaaagtga gggcgctgag gtcgagcagg 6240

gtggtgtgac ttgttatagc ctttagagct gcgaaagcgc gtatggattt ggctcatcag 6300

gccagattga gggtctgtgg acacatgtca tgttagtgta cttcaatcgc cccctggata 6360

tagccccgac aataggccgt ggcctcattt ttttgccttc cgcacatttc cattgctcgg 6420

tacccacacc ttgcttctcc tgcacttgcc aaccttaata ctggtttaca ttgaccaaca 6480

tcttacaagc ggggggcttg tctagggtat atataaacag tggctctccc aatcggttgc 6540

cagtctcttt tttcctttct ttccccacag attcgaaatc taaactacac atcacacaat 6600

gcctgttact gacgtcctta agcgaaagtc cggtgtcatc gtcggcgacg atgtccgagc 6660

cgtgagtatc cacgacaaga tcagtgtcga gacgacgcgt tttgtgtaat gacacaatcc 6720

gaaagtcgct agcaacacac actctctaca caaactaacc cagctctcca tggatccagg 6780

cctgttaacg gccattacgg cctgcaggat ccgaaaaaac ctcccacacc tccccctgaa 6840

cctgaaacat aaaatgaatg caattgttgt tgttaacttg tttattgcag cttataatgg 6900

ttacaaataa agcaatagca tcacaaattt cacaaataaa gcatttttt cactgcattc 6960

tagttgtggt ttgtccaaac tcatcaatgt atcttatcat gtctgcggcc gcaagtgtgg 7020

atggggaagt gagtgcccgg ttctgtgtgc acaattggca atccaagatg gatggattca 7080

acacagggat atagcgagct acgtggtggt gcgaggatat agcaacggat atttatgttt 7140

gacacttgag aatgtacgat acaagcactg tccaagtaca atactaaaca tactgtacat 7200

actcatactc gtacccgggc aacggtttca cttgagtgca gtggctagtg ctcttactcg 7260

tacagtgtgc aatactgcgt atcatagtct ttgatgtata tcgtattcat tcatgttagt 7320

tgc 7323

<210> 193
<211> 8666
<212> DNA
<213> Artificial Sequence

<220>
<223> Plasmid pMDGAT1-17

<220>
<221> misc_feature
<222> (467)..(467)
<223> n is a, c, g, or t

<400> 193
catggcagag tcgacaacaa cgacatgtgc aaaggaggag ggcattgcca acagcgctgc 60

tttgcctgac attcccccaa agatggaaga cctcaagtcc tccaggaaga ccggctcttc 120

ttacaagcac accttccccg tccatacaaa aaccatcccc agcccattgt ctaaagaggc 180

acctccagag agctatcgtg gattcgtcaa cctcggcatg ctcctacttt tcggcaacaa 240

catccgattg atcatcgaga attacctcaa atacggcttc ctgctctcaa tccctggatc 300

aagcgtctcg aagcaggact ggatcctggc tgccctcacc cacgccatcc tacccgtcaa 360

cctcatcctg gcctacaagc ttgagagctg ggccaaggag agagccgtcg gctatcgcaa 420

gcgtcgatct gacgaaccca ttgcccagga atcaaccaag gccgtgncag caggagataa 480

tgacgctatc aaaaccacaa aacccgccaa ggcccaggat ctcacacccg aggcccttgc 540

aaggaaggaa caatcgaccg tgggctggct ccatgtcttc aatctgttca ccatcgttgc 600

ctggccctcc ttcatgtcct actttatgat ctaccacccc ttcgtggcca tgtcctgcct 660

catgaacgga cttatcctct tcctcaaaat gacctccttt gcgcttgtga accaggagct 720

ccgagcagcc tacatctttg gaacacccgt ggacacgttc cagcacatgg ctaaagtgca 780

cgacatctct ggcaaggacc tgacaaagaa ggagatcttc cagtatgaca tccagtaccc 840

cgacaacatc accctcaaga acattggcta tttctggctc gcccccacgc tctgctacca 900

gccatcatac ccaaggacga ccgtcttccg caaatccttc ttcctcaagc gtgtggccga 960

gatcgtgacc tgtctgggca tgatgtactt tttagtcgag cagtacgcca cccccaccct 1020

gcagaactcg gtccgagcat tcgatgagtt ggcgttcggc accattctgg agagagtgct 1080

gaagctgagc accaccagtg tcatcatctg gctactcatg ttctacacct ttttccactc 1140

gttctttaat gctcttgcag aggcactgta ctttggagac cgtcgcttct atctcgcctg 1200

gtggaatgcc actggtgtcg gcatgtactg gaagacgtgg aactcgcccg tctacacctt 1260

cttcaaacgc cacgtatacc tgcccctgat cacctctggc acctctccca tggtcgcctc 1320

gatcgtcatc ttcctcatct cggctgtctt gcacgagatc ttgatcggct tccccactca 1380

tatgatctat ggatacgcat tcgccggcat gttcctccag atcccgctga tcattctgac 1440

ccgacccctc gaaaaatggc gaggcaccgg atcgggtctc ggcaacatga tcttctgggt 1500

ctcgttcacc atcctgggcc agccagcgtg tgcgctgctc tactactacc actggaccaa 1560

gcgccatatg gatgtttgag cggccgcaag tgtggatggg gaagtgagtg cccggttctg 1620

tgtgcacaat tggcaatcca agatggatgg attcaacaca gggatatagc gagctacgtg 1680

gtggtgcgag gatatagcaa cggatattta tgtttgacac ttgagaatgt acgatacaag 1740

cactgtccaa gtacaatact aaacatactg tacatactca tactcgtacc cgggcaacgg 1800

tttcacttga gtgcagtggc tagtgctctt actcgtacag tgtgcaatac tgcgtatcat 1860

agtctttgat gtatatcgta ttcattcatg ttagttgcgt acgagccgga agcataaagt 1920

gtaaagcctg gggtgcctaa tgagtgagct aactcacatt aattgcgttg cgctcactgc 1980

ccgctttcca gtcgggaaac ctgtcgtgcc agctgcatta atgaatcggc caacgcgcgg 2040

ggagaggcgg tttgcgtatt gggcgctctt ccgcttcctc gctcactgac tcgctgcgct 2100

cggtcgttcg gctgcggcga gcggtatcag ctcactcaaa ggcggtaata cggttatcca 2160

cagaatcagg ggataacgca ggaaagaaca tgtgagcaaa aggccagcaa aaggccagga 2220

accgtaaaaa ggccgcgttg ctggcgtttt tccataggct ccgccccccct gacgagcatc 2280

acaaaaatcg acgctcaagt cagaggtggc gaaacccgac aggactataa agataccagg 2340

cgtttccccc tggaagctcc ctcgtgcgct ctcctgttcc gaccctgccg cttaccggat 2400

acctgtccgc ctttctccct tcgggaagcg tggcgctttc tcatagctca cgctgtaggt 2460

atctcagttc ggtgtaggtc gttcgctcca agctgggctg tgtgcacgaa ccccccgttc 2520

agcccgaccg ctgcgcctta tccggtaact atcgtcttga gtccaacccg gtaagacacg 2580

acttatcgcc actggcagca gccactggta acaggattag cagagcgagg tatgtaggcg 2640

gtgctacaga gttcttgaag tggtggccta actacggcta cactagaagg acagtatttg 2700

gtatctgcgc tctgctgaag ccagttacct tcggaaaaag agttggtagc tcttgatccg 2760

gcaaacaaac caccgctggt agcggtggtt tttttgtttg caagcagcag attacgcgca 2820

gaaaaaaagg atctcaagaa gatcctttga tcttttctac ggggtctgac gctcagtgga 2880

acgaaaactc acgttaaggg attttggtca tgagattatc aaaaaggatc ttcacctaga 2940

tccttttaaa ttaaaaatga agttttaaat caatctaaag tatatatgag taaacttggt 3000

ctgacagtta ccaatgctta atcagtgagg cacctatctc agcgatctgt ctatttcgtt 3060

catccatagt tgcctgactc cccgtcgtgt agataactac gatacgggag ggcttaccat 3120

ctggccccag tgctgcaatg ataccgcgag acccacgctc accggctcca gatttatcag 3180

caataaacca gccagccgga agggccgagc gcagaagtgg tcctgcaact ttatccgcct 3240

ccatccagtc tattaattgt tgccgggaag ctagagtaag tagttcgcca gttaatagtt 3300

tgcgcaacgt tgttgccatt gctacaggca tcgtggtgtc acgctcgtcg tttggtatgg 3360

cttcattcag ctccggttcc caacgatcaa ggcgagttac atgatccccc atgttgtgca 3420

aaaaagcggt tagctccttc ggtcctccga tcgttgtcag aagtaagttg gccgcagtgt 3480

tatcactcat ggttatggca gcactgcata attctcttac tgtcatgcca tccgtaagat 3540

gcttttctgt gactggtgag tactcaacca agtcattctg agaatagtgt atgcggcgac 3600

cgagttgctc ttgcccggcg tcaatacggg ataataccgc gccacatagc agaactttaa 3660

aagtgctcat cattggaaaa cgttcttcgg ggcgaaaact ctcaaggatc ttaccgctgt 3720

tgagatccag ttcgatgtaa cccactcgtg cacccaactg atcttcagca tcttttactt 3780

tcaccagcgt ttctgggtga gcaaaaacag gaaggcaaaa tgccgcaaaa aagggaataa 3840

gggcgacacg gaaatgttga atactcatac tcttcctttt tcaatattat tgaagcattt 3900

atcagggtta ttgtctcatg agcggataca tatttgaatg tatttagaaa aataaacaaa 3960

taggggttcc gcgcacattt ccccgaaaag tgccacctga cgcgccctgt agcggcgcat 4020

taagcgcggc gggtgtggtg gttacgcgca gcgtgaccgc tacacttgcc agcgccctag 4080

cgcccgctcc tttcgctttc ttcccttcct ttctcgccac gttcgccggc tttccccgtc 4140

aagctctaaa tcgggggctc cctttagggt tccgatttag tgctttacgg cacctcgacc 4200

ccaaaaaact tgattagggt gatggttcac gtagtgggcc atcgccctga tagacggttt 4260

ttcgcccttt gacgttggag tccacgttct ttaatagtgg actcttgttc caaactggaa 4320

caacactcaa ccctatctcg gtctattctt ttgatttata agggattttg ccgatttcgg 4380

cctattggtt aaaaaatgag ctgatttaac aaaaatttaa cgcgaatttt aacaaaatat 4440

taacgcttac aatttccatt cgccattcag gctgcgcaac tgttgggaag ggcgatcggt 4500

gcgggcctct tcgctattac gccagctggc gaaaggggga tgtgctgcaa ggcgattaag 4560

ttgggtaacg ccagggtttt cccagtcacg acgttgtaaa acgacggcca gtgaattgta 4620

atacgactca ctatagggcg aattgggtac cgggcccccc ctcgaggtcg atggtgtcga 4680

taagcttgat atcgaattca tgtcacacaa accgatcttc gcctcaagga aacctaattc 4740

tacatccgag agactgccga gatccagtct acactgatta attttcgggc caataattta 4800

aaaaaatcgt gttatataat attatatgta ttatatatat acatcatgat gatactgaca 4860

gtcatgtccc attgctaaat agacagactc catctgccgc ctccaactga tgttctcaat 4920

atttaagggg tcatctcgca ttgtttaata ataaacagac tccatctacc gcctccaaat 4980

gatgttctca aaatatattg tatgaactta tttttattac ttagtattat tagacaactt 5040

acttgcttta tgaaaaacac ttcctattta ggaaacaatt tataatggca gttcgttcat 5100

ttaacaattt atgtagaata aatgttataa atgcgtatgg gaaatcttaa atatggatag 5160

cataaatgat atctgcattg cctaattcga aatcaacagc aacgaaaaaa atcccttgta 5220

caacataaat agtcatcgag aaatatcaac tatcaaagaa cagctattca cacgttacta 5280

ttgagattat tattggacga gaatcacaca ctcaactgtc tttctctctt ctagaaatac 5340

aggtacaagt atgtactatt ctcattgttc atacttctag tcatttcatc ccacatattc 5400

cttggatttc tctccaatga atgacattct atcttgcaaa ttcaacaatt ataataagat 5460

ataccaaagt agcggtatag tggcaatcaa aaagcttctc tggtgtgctt ctcgtattta 5520

tttttattct aatgatccat taaaggtata tatttatttc ttgttatata atccttttgt 5580

ttattacatg ggctggatac ataaaggtat tttgatttaa ttttttgctt aaattcaatc 5640

ccccctcgtt cagtgtcaac tgtaatggta ggaaattacc atacttttga agaagcaaaa 5700

aaaatgaaag aaaaaaaaaa tcgtatttcc aggttagacg ttccgcagaa tctagaatgc 5760

ggtatgcggt acattgttct tcgaacgtaa aagttgcgct ccctgagata ttgtacattt 5820

ttgcttttac aagtacaagt acatcgtaca actatgtact actgttgatg catccacaac 5880

agtttgtttt gttttttttt gttttttttt tttctaatga ttcattaccg ctatgtatac 5940

ctacttgtac ttgtagtaag ccgggttatt ggcgttcaat taatcataga cttatgaatc 6000

tgcacggtgt gcgctgcgag ttacttttag cttatgcatg ctacttgggt gtaatattgg 6060

gatctgttcg gaaatcaacg gatgctcaat cgatttcgac agtaattaat taagtcatac 6120

acaagtcagc tttcttcgag cctcatataa gtataagtag ttcaacgtat tagcactgta 6180

cccagcatct ccgtatcgag aaacacaaca acatgcccca ttggacagat catgcggata 6240

cacaggttgt gcagtatcat acatactcga tcagacaggt cgtctgacca tcatacaagc 6300

tgaacaagcg ctccatactt gcacgctctc tatatacaca gttaaattac atatccatag 6360

tctaacctct aacagttaat cttctggtaa gcctcccagc cagccttctg gtatcgcttg 6420

gcctcctcaa taggatctcg gttctggccg tacagacctc ggccgacaat tatgatatcc 6480

gttccggtag acatgacatc ctcaacagtt cggtactgct gtccgagagc gtctcccttg 6540

tcgtcaagac ccaccccggg ggtcagaata agccagtcct cagagtcgcc cttaggtcgg 6600

ttctgggcaa tgaagccaac cacaaactcg gggtcggatc gggcaagctc aatggtctgc 6660

ttggagtact cgccagtggc cagagagccc ttgcaagaca gctcggccag catgagcaga 6720

cctctggcca gcttctcgtt gggagagggg actaggaact ccttgtactg ggagttctcg 6780

tagtcagaga cgtcctcctt cttctgttca gagacagttt cctcggcacc agctcgcagg 6840

ccagcaatga ttccggttcc gggtacaccg tgggcgttgg tgatatcgga ccactcggcg 6900

attcggtgac accggtactg gtgcttgaca gtgttgccaa tatctgcgaa ctttctgtcc 6960

tcgaacagga agaaaccgtg cttaagagca agttccttga gggggagcac agtgccggcg 7020

taggtgaagt cgtcaatgat gtcgatatgg gttttgatca tgcacacata aggtccgacc 7080

ttatcggcaa gctcaatgag ctccttggtg gtggtaacat ccagagaagc acacaggttg 7140

gttttcttgg ctgccacgag cttgagcact cgagcggcaa aggcggactt gtggacgtta 7200

gctcgagctt cgtaggaggg cattttggtg gtgaagagga gactgaaata aatttagtct 7260

gcagaacttt ttatcggaac cttatctggg gcagtgaagt atatgttatg gtaatagtta 7320

cgagttagtt gaacttatag atagactgga ctatacggct atcggtccaa attagaaaga 7380

acgtcaatgg ctctctgggc gtcgcctttg ccgacaaaaa tgtgatcatg atgaaagcca 7440

gcaatgacgt tgcagctgat attgttgtcg gccaaccgcg ccgaaaacgc agctgtcaga 7500

cccacagcct ccaacgaaga atgtatcgtc aaagtgatcc aagcacactc atagttggag 7560

tcgtactcca aaggcggcaa tgacgagtca gacagatact cgtcgactca ggcgacgacg 7620

gaattcctgc agcccatctg cagaattcag gagagaccgg gttggcggcg tatttgtgtc 7680

ccaaaaaaca gccccaattg ccccggagaa gacggccagg ccgcctagat gacaaattca 7740

acaactcaca gctgactttc tgccattgcc actaggggggg ggccttttta tatggccaag 7800

ccaagctctc cacgtcggtt gggctgcacc caacaataaa tgggtagggt tgcaccaaca 7860

aagggatggg atgggggggta gaagatacga ggataacggg gctcaatggc acaaataaga 7920

acgaatactg ccattaagac tcgtgatcca gcgactgaca ccattgcatc atctaagggc 7980

ctcaaaacta cctcggaact gctgcgctga tctggacacc acagaggttc cgagcacttt 8040

aggttgcacc aaatgtccca ccaggtgcag gcagaaaacg ctggaacagc gtgtacagtt 8100

tgtcttaaca aaaagtgagg gcgctgaggt cgagcagggt ggtgtgactt gttatagcct 8160

ttagagctgc gaaagcgcgt atggatttgg ctcatcaggc cagattgagg gtctgtggac 8220

acatgtcatg ttagtgtact tcaatcgccc cctggatata gccccgacaa taggccgtgg 8280

cctcattttt ttgccttccg cacatttcca ttgctcggta cccacacctt gcttctcctg 8340

cacttgccaa ccttaatact ggtttacatt gaccaacatc ttacaagcgg ggggcttgtc 8400

tagggtatat ataaacagtg gctctcccaa tcggttgcca gtctcttttt tcctttcttt 8460

ccccacagat tcgaaatcta aactacacat cacacaatgc ctgttactga cgtccttaag 8520

cgaaagtccg gtgtcatcgt cggcgacgat gtccgagccg tgagtatcca cgacaagatc 8580

agtgtcgaga cgacgcgttt tgtgtaatga cacaatccga aagtcgctag caacacacac 8640

tctctacaca aactaaccca gctctc 8666

<210> 194

<211> 8084
<212> DNA
<213> Artificial Sequence

<220>
<223> Plasmid pMDGAT2-17

<400> 194
catggcgctc tttgcgcctt tacggatgcc cgtcaagcgt cgtatgcaga caggagctgt 60

cctatactgg attgcgggga tgatttactg cattggcatc tttgccttcc tctgcacgtt 120

caagatcctt cgacccttga tcatcatcta tgtcctgtgg gcctacatgc tcgaccgagc 180

accagagcgg ggtgcacgca cagtccaatg gtattgtaac tggatcggat ggaaacactt 240

tgcacagtac tttcctatga cccttgtcaa ggagggagag ctggacccat ccaagaacta 300

catctttggg tatcacccac acggaatcat ttccttgggt gccctctgcg cctttgggac 360

cgagggcctt catttctcca aacgcttccc gggtatcaag cctcatctgc tcaccattca 420

cgccaacttt cagatcccac tctatcgcga tatgatcatg gcccacggct gtgcttccgt 480

gtcgagggcc tcttgtgaac acatcctgcg gtctggcgaa ggatcctcgg tcgtgatcgt 540

tgtcgggggt gcacaagaaa gtttgtcgac tcaacctggc acgttaaatc tgacactcaa 600

gaaaagactg ggattttgca agctggcctt tgtcaatggc gcaagtctgg tacctacgtt 660

ggcctttggt gagaacgagc tctatgaggt gtaccacacc aagcccacaa gcctgatata 720

caagctccag cagttgacta aacgcacgat cggcttcaca atgcccgtgt tcaacggacg 780

aggaatcttc aattatgagt ttggactgct gccaaggagg aagcctgtct atatcgttat 840

aggaaacccc attcatgtag acaaggtcga gaacccaacg attgaacaga tgcagaaact 900

gcagtcaatt tacattgatg aggtgctaaa catttgggaa agatacaagg acaagtatgc 960

cgcaggacga actcaggaac tgtgcatcat cgaataggcg gccgcaagtg tggatgggga 1020

agtgagtgcc cggttctgtg tgcacaattg gcaatccaag atggatggat tcaacacagg 1080

gatatagcga gctacgtggt ggtgcgagga tatagcaacg gatatttatg tttgacactt 1140

gagaatgtac gatacaagca ctgtccaagt acaatactaa acatactgta catactcata 1200

ctcgtacccg ggcaacggtt tcacttgagt gcagtggcta gtgctcttac tcgtacagtg 1260

tgcaatactg cgtatcatag tctttgatgt atatcgtatt cattcatgtt agttgcgtac 1320

gagccggaag cataaagtgt aaagcctggg gtgcctaatg agtgagctaa ctcacattaa 1380

ttgcgttgcg ctcactgccc gctttccagt cgggaaacct gtcgtgccag ctgcattaat 1440

gaatcggcca acgcgcgggg agaggcggtt tgcgtattgg gcgctcttcc gcttcctcgc 1500

tcactgactc gctgcgctcg gtcgttcggc tgcggcgagc ggtatcagct cactcaaagg 1560

cggtaatacg gttatccaca gaatcagggg ataacgcagg aaagaacatg tgagcaaaag 1620

gccagcaaaa ggccaggaac cgtaaaaagg ccgcgttgct ggcgtttttc cataggctcc 1680

gccccccctga cgagcatcac aaaaatcgac gctcaagtca gaggtggcga aacccgacag 1740

gactataaag ataccaggcg tttccccctg gaagctccct cgtgcgctct cctgttccga 1800

ccctgccgct taccggatac ctgtccgcct ttctcccttc gggaagcgtg gcgctttctc 1860

atagctcacg ctgtaggtat ctcagttcgg tgtaggtcgt tcgctccaag ctgggctgtg 1920

tgcacgaacc ccccgttcag cccgaccgct gcgccttatc cggtaactat cgtcttgagt 1980

ccaacccggt aagacacgac ttatcgccac tggcagcagc cactggtaac aggattagca 2040

gagcgaggta tgtaggcggt gctacagagt tcttgaagtg gtggcctaac tacggctaca 2100

ctagaaggac agtatttggt atctgcgctc tgctgaagcc agttaccttc ggaaaaagag 2160

ttggtagctc ttgatccggc aaacaaacca ccgctggtag cggtggtttt tttgtttgca 2220

agcagcagat tacgcgcaga aaaaaaggat ctcaagaaga tcctttgatc ttttctacgg 2280

ggtctgacgc tcagtggaac gaaaactcac gttaagggat tttggtcatg agattatcaa 2340

aaaggatctt cacctagatc cttttaaatt aaaaatgaag ttttaaatca atctaaagta 2400

tatatgagta aacttggtct gacagttacc aatgcttaat cagtgaggca cctatctcag 2460

cgatctgtct atttcgttca tccatagttg cctgactccc cgtcgtgtag ataactacga 2520

tacgggaggg cttaccatct ggccccagtg ctgcaatgat accgcgagac ccacgctcac 2580

cggctccaga tttatcagca ataaaccagc cagccggaag ggccgagcgc agaagtggtc 2640

ctgcaacttt atccgcctcc atccagtcta ttaattgttg ccgggaagct agagtaagta 2700

gttcgccagt taatagtttg cgcaacgttg ttgccattgc tacaggcatc gtggtgtcac 2760

gctcgtcgtt tggtatggct tcattcagct ccggttccca acgatcaagg cgagttacat 2820

gatcccccat gttgtgcaaa aaagcggtta gctccttcgg tcctccgatc gttgtcagaa 2880

gtaagttggc cgcagtgtta tcactcatgg ttatggcagc actgcataat tctcttactg 2940

tcatgccatc cgtaagatgc ttttctgtga ctggtgagta ctcaaccaag tcattctgag 3000

aatagtgtat gcggcgaccg agttgctctt gcccggcgtc aatacgggat aataccgcgc 3060

cacatagcag aactttaaaa gtgctcatca ttggaaaacg ttcttcgggg cgaaaactct 3120

caaggatctt accgctgttg agatccagtt cgatgtaacc cactcgtgca cccaactgat 3180

cttcagcatc ttttactttc accagcgttt ctgggtgagc aaaaacagga aggcaaaatg 3240

ccgcaaaaaa gggaataagg gcgacacgga aatgttgaat actcatactc ttccttttc 3300

aatattattg aagcatttat cagggttatt gtctcatgag cggatacata tttgaatgta 3360

tttagaaaaa taaacaaata ggggttccgc gcacatttcc ccgaaaagtg ccacctgacg 3420

cgccctgtag cggcgcatta agcgcggcgg gtgtggtggt tacgcgcagc gtgaccgcta 3480

cacttgccag cgccctagcg cccgctcctt tcgctttctt cccttccttt ctcgccacgt 3540

tcgccggctt tccccgtcaa gctctaaatc gggggctccc tttagggttc cgatttagtg 3600

ctttacggca cctcgacccc aaaaaacttg attagggtga tggttcacgt agtgggccat 3660

cgccctgata gacggttttt cgcccctttga cgttggagtc cacgttcttt aatagtggac 3720

tcttgttcca aactggaaca acactcaacc ctatctcggt ctattctttt gatttataag 3780

ggattttgcc gatttcggcc tattggttaa aaaatgagct gatttaacaa aaatttaacg 3840

cgaattttaa caaaatatta acgcttacaa tttccattcg ccattcaggc tgcgcaactg 3900

ttgggaaggg cgatcggtgc gggcctcttc gctattacgc cagctggcga aaggggatg 3960

tgctgcaagg cgattaagtt gggtaacgcc agggttttcc cagtcacgac gttgtaaaac 4020

gacggccagt gaattgtaat acgactcact atagggcgaa ttgggtaccg ggcccccct 4080

cgaggtcgat ggtgtcgata agcttgatat cgaattcatg tcacacaaac cgatcttcgc 4140

ctcaaggaaa cctaattcta catccgagag actgccgaga tccagtctac actgattaat 4200

tttcgggcca ataatttaaa aaaatcgtgt tatataatat tatatgtatt atatatatac 4260

atcatgatga tactgacagt catgtcccat tgctaaatag acagactcca tctgccgcct 4320

ccaactgatg ttctcaatat ttaaggggtc atctcgcatt gtttaataat aaacagactc 4380

catctaccgc ctccaaatga tgttctcaaa atatattgta tgaacttatt tttattactt 4440

agtattatta gacaacttac ttgctttatg aaaaacactt cctatttagg aaacaattta 4500

taatggcagt tcgttcattt aacaatttat gtagaataaa tgttataaat gcgtatggga 4560

aatcttaaat atggatagca taaatgatat ctgcattgcc taattcgaaa tcaacagcaa 4620

cgaaaaaaat cccttgtaca acataaatag tcatcgagaa atatcaacta tcaaagaaca 4680

gctattcaca cgttactatt gagattatta ttggacgaga atcacacact caactgtctt 4740

tctctcttct agaaatacag gtacaagtat gtactattct cattgttcat acttctagtc 4800

atttcatccc acatattcct tggatttctc tccaatgaat gacattctat cttgcaaatt 4860

caacaattat aataagatat accaaagtag cggtatagtg gcaatcaaaa agcttctctg 4920

gtgtgcttct cgtatttatt tttattctaa tgatccatta aaggtatata tttatttctt 4980

gttatataat ccttttgttt attacatggg ctggatacat aaaggtattt tgatttaatt 5040

ttttgcttaa attcaatccc ccctcgttca gtgtcaactg taatggtagg aaattaccat 5100

acttttgaag aagcaaaaaa aatgaaagaa aaaaaaaatc gtatttccag gttagacgtt 5160

ccgcagaatc tagaatgcgg tatgcggtac attgttcttc gaacgtaaaa gttgcgctcc 5220

ctgagatatt gtacattttt gcttttacaa gtacaagtac atcgtacaac tatgtactac 5280

tgttgatgca tccacaacag tttgttttgt ttttttttgt ttttttttgt tctaatgatt 5340

cattaccgct atgtatacct acttgtactt gtagtaagcc gggttattgg cgttcaatta 5400

atcatagact tatgaatctg cacggtgtgc gctgcgagtt acttttagct tatgcatgct 5460

acttgggtgt aatattggga tctgttcgga aatcaacgga tgctcaatcg atttcgacag 5520

taattaatta agtcatacac aagtcagctt tcttcgagcc tcatataagt ataagtagtt 5580

caacgtatta gcactgtacc cagcatctcc gtatcgagaa acacaacaac atgccccatt 5640

ggacagatca tgcggataca caggttgtgc agtatcatac atactcgatc agacaggtcg 5700

tctgaccatc atacaagctg aacaagcgct ccatacttgc acgctctcta tatacacagt 5760

taaattacat atccatagtc taacctctaa cagttaatct tctggtaagc ctcccagcca 5820

gccttctggt atcgcttggc ctcctcaata ggatctcggt tctggccgta cagacctcgg 5880

ccgacaatta tgatatccgt tccggtagac atgacatcct caacagttcg gtactgctgt 5940

ccgagagcgt ctcccttgtc gtcaagaccc accccggggg tcagaataag ccagtcctca 6000

gagtcgccct taggtcggtt ctgggcaatg aagccaacca caaactcggg gtcggatcgg 6060

gcaagctcaa tggtctgctt ggagtactcg ccagtggcca gagagccctt gcaagacagc 6120

tcggccagca tgagcagacc tctggccagc ttctcgttgg gagaggggac taggaactcc 6180

ttgtactggg agttctcgta gtcagagacg tcctccttct tctgttcaga gacagtttcc 6240

tcggcaccag ctcgcaggcc agcaatgatt ccggttccgg gtacaccgtg ggcgttggtg 6300

atatcggacc actcggcgat tcggtgacac cggtactggt gcttgacagt gttgccaata 6360

tctgcgaact ttctgtcctc gaacaggaag aaaccgtgct taagagcaag ttccttgagg 6420

gggagcacag tgccggcgta ggtgaagtcg tcaatgatgt cgatatgggt tttgatcatg 6480

cacacataag gtccgacctt atcggcaagc tcaatgagct ccttggtggt ggtaacatcc 6540

agagaagcac acaggttggt tttcttggct gccacgagct tgagcactcg agcggcaaag 6600

gcggacttgt ggacgttagc tcgagcttcg taggagggca ttttggtggt gaagaggaga 6660

ctgaaataaa tttagtctgc agaacttttt atcggaacct tatctggggc agtgaagtat 6720

atgttatggt aatagttacg agttagttga acttatagat agactggact atacggctat 6780

cggtccaaat tagaaagaac gtcaatggct ctctgggcgt cgcctttgcc gacaaaaatg 6840

tgatcatgat gaaagccagc aatgacgttg cagctgatat tgttgtcggc caaccgcgcc 6900

gaaaacgcag ctgtcagacc cacagcctcc aacgaagaat gtatcgtcaa agtgatccaa 6960

gcacactcat agttggagtc gtactccaaa ggcggcaatg acgagtcaga cagatactcg 7020

tcgactcagg cgacgacgga attcctgcag cccatctgca gaattcagga gagaccgggt 7080

tggcggcgta tttgtgtccc aaaaaacagc cccaattgcc ccggagaaga cggccaggcc 7140

gcctagatga caaattcaac aactcacagc tgactttctg ccattgccac tagggggggg 7200

ccttttttata tggccaagcc aagctctcca cgtcggttgg gctgcaccca acaataaatg 7260

ggtagggttg caccaacaaa gggatgggat ggggggtaga agatacgagg ataacggggc 7320

tcaatggcac aaataagaac gaatactgcc attaagactc gtgatccagc gactgacacc 7380

attgcatcat ctaagggcct caaaactacc tcggaactgc tgcgctgatc tggacaccac 7440

agaggttccg agcactttag gttgcaccaa atgtcccacc aggtgcaggc agaaaacgct 7500

ggaacagcgt gtacagtttg tcttaacaaa aagtgagggc gctgaggtcg agcagggtgg 7560

tgtgacttgt tatagccttt agagctgcga aagcgcgtat ggatttggct catcaggcca 7620

gattgagggt ctgtggacac atgtcatgtt agtgtacttc aatcgccccc tggatatagc 7680

cccgacaata ggccgtggcc tcatttttttt gccttccgca catttccatt gctcggtacc 7740

cacaccttgc ttctcctgca cttgccaacc ttaatactgg tttacattga ccaacatctt 7800

acaagcgggg ggcttgtcta gggtatatat aaacagtggc tctcccaatc ggttgccagt 7860

ctctttttc ctttctttcc ccacagattc gaaatctaaa ctacacatca cacaatgcct 7920

gttactgacg tccttaagcg aaagtccggt gtcatcgtcg gcgacgatgt ccgagccgtg 7980

agtatccacg acaagatcag tgtcgagacg acgcgttttg tgtaatgaca caatccgaaa 8040

gtcgctagca acacacactc tctacacaaa ctaacccagc tctc 8084

<210> 195
<211> 9239
<212> DNA
<213> Artificial Sequence

<220>
<223> Plasmid pMGPAT-17

<220>
<221> misc_feature
<222> (1510)..(1510)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (1577)..(1577)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (1584)..(1584)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (1712)..(1712)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (1729)..(1729)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (1759)..(1759)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (1820)..(1820)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (1822)..(1822)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (1831)..(1831)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (1841)..(1841)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (1913)..(1913)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (1941)..(1941)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (2033)..(2033)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (2130)..(2130)
<223> n is a, c, g, or t

<400> 195
ccatggatcc tcagatctac gactttgtgt cgttcttctt cactatcctg ctcgacatct 60

tcttcaggga gattcgtccc agaggcgcac acaagattcc acaaaaaggc cccgtgatct 120

ttgtcgccgc tcctcatgcc aatcagtttg tcgatcctct cgtcttgatg cgagagtgcg 180

gccgcagagt ctcattcctt gcagccaaaa agtccatgga ccgccggtgg attggtgcaa 240

tggcacgctc gatgaatgcg attcctgttg agcgccccca ggaccttgct aaagccggct 300

cgggaatcat caaacttctg gatcgttatg gcgaccctct tcgggtcacc ggtgtcggca 360

ctaaattcac aaaggagctg cttgtgggcg accagatctc ccttccaaag gacgtcggtg 420

tcycagctgt gggcgagatc atatctgata ccgagctgat tgtcaagaag gaattcaaag 480

agctcaaggc ccttgagttg ctgaccagtg ctgaaggaac caagtacaaa tgcctacccc 540

atatggacca gacgaacgtc tacaaaactg tctttgagcg ccttaacgca ggacattgcg 600

ttggcatttt ccccgaggga ggctcccacg atcgtgctga aatgctgcca ttgaaagctg 660

gagtcaccat catggccctg ggcgcattag ccgccaaccc ttccttggat ctcaagattg 720

tcacctgcgg cctcaactac tttcacccgc atcgcttccg ctcgcgtgca gtagtcgagt 780

ttggcgagcc attgacggtt tcgcctgagc tggtcgaaat gtacaagcga ggcggggcgg 840

aaaagcgaga ggcttgcgga aaactgctcg acacgatcta tgaagctctc cgcggtgtca 900

ctctcaacgc gcctgattac gagacattga tggtcattca agcggcccgt cgcctttaca 960

agcctactca tcgcaagctg cagatctcgc aggtcgttga gttgaaccgt aggttcgtcg 1020

caggatacat gcacttcaag gacaatccaa aagtcattga agccaaggat aaagtcatgc 1080

attacaacac tcagctgcgg taccatggat tgcgggatca ccaagtgaat attcgcacca 1140

ccaggaagca cgctatcggc atgctcatct ctcggctgat ccagatgatc tttttgagtt 1200

gtctggcgct acctggaacy ctgatgaatc ttccggtcgc cattgtcgct cgtgtcatca 1260

gcaacaaaaa ggccaaagag gcgctggctg cctcgacagt caaaattgct ggaagggacg 1320

tcttggctac gtggaagttg ctggtcgctc taggattgat gcctgttctg tacttcacgt 1380

actccgtcat ggtcttcatc tattgcagcc gcttcgacct atcgttcaag tcgcgtcttt 1440

tggttgcttg ggcagcatgg gcgcttattc cttttgtaac ctacgcaagc atccgctttg 1500

gtgaagttgn tatcgatatc ttcaagtcta ttcgcccatt gttcctgtcc atcatcccgg 1560

gcgaggagag cacgatnaac gacntgcgca aggcgcgcgc agaacttcaa aagaccatca 1620

ccaatctcat caatgagcta gcgccgcaga tatatcccga ctttgattca aagcgcatcc 1680

tcgatccatc ccctgcagat cgtcccagcc gntcagcatc aggcaccanc cttgcacaga 1740

cgattttcaa cacggctgnt cagcccttga accaatggct aggcaaggat ggccgctttg 1800

aatgggagcg taccgaggan tntgatgcag ntgacgtgtt nttctttttg gacccagcga 1860

gaggaattat ggggcggtct agggcgtcgt cttggggagg tggagcattt acncctgctg 1920

ttgatgggtc gcgatcccgg natcggagca ggacaagcag cttcacgtcg ggccagatcc 1980

agctgggcga gggcttcaaa ctcgaggcac tgacggagct gccgcgggac aancctttg 2040

cagaagtgac caggcggttg agtgtgagcc gaatgcagag atatgggctg gagggtatga 2100

cgcgctcgga tacggatgaa aacgaaggcn ccgccaagtc aaaagacatt taggcggccg 2160

caagtgtgga tggggaagtg agtgcccggt tctgtgtgca caattggcaa tccaagatgg 2220

atggattcaa cacagggata tagcgagcta cgtggtggtg cgaggatata gcaacggata 2280

tttatgtttg acacttgaga atgtacgata caagcactgt ccaagtacaa tactaaacat 2340

actgtacata ctcatactcg tacccgggca acggtttcac ttgagtgcag tggctagtgc 2400

tcttactcgt acagtgtgca atactgcgta tcatagtctt tgatgtatat cgtattcatt 2460

catgttagtt gcgtacgagc cggaagcata aagtgtaaag cctggggtgc ctaatgagtg 2520

agctaactca cattaattgc gttgcgctca ctgcccgctt tccagtcggg aaacctgtcg 2580

tgccagctgc attaatgaat cggccaacgc gcggggagag gcggtttgcg tattgggcgc 2640

tcttccgctt cctcgctcac tgactcgctg cgctcggtcg ttcggctgcg gcgagcggta 2700

tcagctcact caaaggcggt aatacggtta tccacagaat caggggataa cgcaggaaag 2760

aacatgtgag caaaaggcca gcaaaaggcc aggaaccgta aaaaggccgc gttgctggcg 2820

tttttccata ggctccgccc ccctgacgag catcacaaaa atcgacgctc aagtcagagg 2880

tggcgaaacc cgacaggact ataaagatac caggcgtttc cccctggaag ctccctcgtg 2940

cgctctcctg ttccgaccct gccgcttacc ggatacctgt ccgcctttct cccttcggga 3000

agcgtggcgc tttctcatag ctcacgctgt aggtatctca gttcggtgta ggtcgttcgc 3060

tccaagctgg gctgtgtgca cgaacccccc gttcagcccg accgctgcgc cttatccggt 3120

aactatcgtc ttgagtccaa cccggtaaga cacgacttat cgccactggc agcagccact 3180

ggtaacagga ttagcagagc gaggtatgta ggcggtgcta cagagttctt gaagtggtgg 3240

cctaactacg gctacactag aaggacagta tttggtatct gcgctctgct gaagccagtt 3300

accttcggaa aaagagttgg tagctcttga tccggcaaac aaaccaccgc tggtagcggt 3360

ggtttttttg tttgcaagca gcagattacg cgcagaaaaa aaggatctca agaagatcct 3420

ttgatctttt ctacggggtc tgacgctcag tggaacgaaa actcacgtta agggattttg 3480

gtcatgagat tatcaaaaag gatcttcacc tagatccttt taaattaaaa atgaagtttt 3540

aaatcaatct aaagtatata tgagtaaact tggtctgaca gttaccaatg cttaatcagt 3600

gaggcaccta tctcagcgat ctgtctattt cgttcatcca tagttgcctg actccccgtc 3660

gtgtagataa ctacgatacg ggagggctta ccatctggcc ccagtgctgc aatgataccg 3720

cgagacccac gctcaccggc tccagattta tcagcaataa accagccagc cggaagggcc 3780

gagcgcagaa gtggtcctgc aactttatcc gcctccatcc agtctattaa ttgttgccgg 3840

gaagctagag taagtagttc gccagttaat agtttgcgca acgttgttgc cattgctaca 3900

ggcatcgtgg tgtcacgctc gtcgtttggt atggcttcat tcagctccgg ttcccaacga 3960

tcaaggcgag ttacatgatc ccccatgttg tgcaaaaaag cggttagctc cttcggtcct 4020

ccgatcgttg tcagaagtaa gttggccgca gtgttatcac tcatggttat ggcagcactg 4080

cataattctc ttactgtcat gccatccgta agatgctttt ctgtgactgg tgagtactca 4140

accaagtcat tctgagaata gtgtatgcgg cgaccgagtt gctcttgccc ggcgtcaata 4200

cgggataata ccgcgccaca tagcagaact ttaaaagtgc tcatcattgg aaaacgttct 4260

tcggggcgaa aactctcaag gatcttaccg ctgttgagat ccagttcgat gtaacccact 4320

cgtgcaccca actgatcttc agcatctttt actttcacca gcgtttctgg gtgagcaaaa 4380

acaggaaggc aaaatgccgc aaaaaaggga ataagggcga cacggaaatg ttgaatactc 4440

atactcttcc tttttcaata ttattgaagc atttatcagg gttattgtct catgagcgga 4500

tacatatttg aatgtattta gaaaaataaa caaatagggg ttccgcgcac atttccccga 4560

aaagtgccac ctgacgcgcc ctgtagcggc gcattaagcg cggcgggtgt ggtggttacg 4620

cgcagcgtga ccgctacact tgccagcgcc ctagcgcccg ctcctttcgc tttcttccct 4680

tcctttctcg ccacgttcgc cggctttccc cgtcaagctc taaatcgggg gctcccttta 4740

gggttccgat ttagtgcttt acggcacctc gaccccaaaa aacttgatta gggtgatggt 4800

tcacgtagtg ggccatcgcc ctgatagacg gtttttcgcc ctttgacgtt ggagtccacg 4860

ttctttaata gtggactctt gttccaaact ggaacaacac tcaaccctat ctcggtctat 4920

tcttttgatt tataagggat tttgccgatt tcggcctatt ggttaaaaaa tgagctgatt 4980

taacaaaaat ttaacgcgaa ttttaacaaa atattaacgc ttacaatttc cattcgccat 5040

tcaggctgcg caactgttgg gaagggcgat cggtgcgggc ctcttcgcta ttacgccagc 5100

tggcgaaagg gggatgtgct gcaaggcgat taagttgggt aacgccaggg ttttcccagt 5160

cacgacgttg taaaacgacg gccagtgaat tgtaatacga ctcactatag ggcgaattgg 5220

gtaccgggcc ccccctcgag gtcgatggtg tcgataagct tgatatcgaa ttcatgtcac 5280

acaaaccgat cttcgcctca aggaaaccta attctacatc cgagagactg ccgagatcca 5340

gtctacactg attaattttc gggccaataa tttaaaaaaa tcgtgttata taatattata 5400

tgtattatat atatacatca tgatgatact gacagtcatg tcccattgct aaatagacag 5460

actccatctg ccgcctccaa ctgatgttct caatatttaa ggggtcatct cgcattgttt 5520

aataataaac agactccatc taccgcctcc aaatgatgtt ctcaaaatat attgtatgaa 5580

cttatttta ttacttagta ttattagaca acttacttgc tttatgaaaa acacttccta 5640

tttaggaaac aatttataat ggcagttcgt tcatttaaca atttatgtag aataaatgtt 5700

ataaatgcgt atgggaaatc ttaaatatgg atagcataaa tgatatctgc attgcctaat 5760

tcgaaatcaa cagcaacgaa aaaaatccct tgtacaacat aaatagtcat cgagaaatat 5820

caactatcaa agaacagcta ttcacacgtt actattgaga ttattattgg acgagaatca 5880

cacactcaac tgtctttctc tcttctagaa atacaggtac aagtatgtac tattctcatt 5940

gttcatactt ctagtcattt catcccacat attccttgga tttctctcca atgaatgaca 6000

ttctatcttg caaattcaac aattataata agatatacca aagtagcggt atagtggcaa 6060

tcaaaaagct tctctggtgt gcttctcgta tttatttta ttctaatgat ccattaaagg 6120

tatatattta tttcttgtta tataatcctt ttgtttatta catgggctgg atacataaag 6180

gtattttgat ttaatttttt gcttaaattc aatcccccct cgttcagtgt caactgtaat 6240

ggtaggaaat taccatactt ttgaagaagc aaaaaaaatg aaagaaaaaa aaaatcgtat 6300

ttccaggtta gacgttccgc agaatctaga atgcggtatg cggtacattg ttcttcgaac 6360

gtaaaagttg cgctccctga gatattgtac atttttgctt ttacaagtac aagtacatcg 6420

tacaactatg tactactgtt gatgcatcca caacagtttg ttttgttttt ttttgttttt 6480

ttttttttcta atgattcatt accgctatgt atacctactt gtacttgtag taagccgggt 6540

tattggcgtt caattaatca tagacttatg aatctgcacg gtgtgcgctg cgagttactt 6600

ttagcttatg catgctactt gggtgtaata ttgggatctg ttcggaaatc aacggatgct 6660

caatcgattt cgacagtaat taattaagtc atacacaagt cagctttctt cgagcctcat 6720

ataagtataa gtagttcaac gtattagcac tgtacccagc atctccgtat cgagaaacac 6780

aacaacatgc cccattggac agatcatgcg gatacacagg ttgtgcagta tcatacatac 6840

tcgatcagac aggtcgtctg accatcatac aagctgaaca agcgctccat acttgcacgc 6900

tctctatata cacagttaaa ttacatatcc atagtctaac ctctaacagt taatcttctg 6960

gtaagcctcc cagccagcct tctggtatcg cttggcctcc tcaataggat ctcggttctg 7020

gccgtacaga cctcggccga caattatgat atccgttccg gtagacatga catcctcaac 7080

agttcggtac tgctgtccga gagcgtctcc cttgtcgtca agacccaccc cgggggtcag 7140

aataagccag tcctcagagt cgcccttagg tcggttctgg gcaatgaagc caaccacaaa 7200

ctcggggtcg gatcgggcaa gctcaatggt ctgcttggag tactcgccag tggccagaga 7260

gcccttgcaa gacagctcgg ccagcatgag cagacctctg gccagcttct cgttgggaga 7320

ggggactagg aactccttgt actgggagtt ctcgtagtca gagacgtcct ccttcttctg 7380

ttcagagaca gtttcctcgg caccagctcg caggccagca atgattccgg ttccgggtac 7440

accgtgggcg ttggtgatat cggaccactc ggcgattcgg tgacaccggt actggtgctt 7500

gacagtgttg ccaatatctg cgaactttct gtcctcgaac aggaagaaac cgtgcttaag 7560

agcaagttcc ttgaggggga gcacagtgcc ggcgtaggtg aagtcgtcaa tgatgtcgat 7620

atgggttttg atcatgcaca cataaggtcc gaccttatcg gcaagctcaa tgagctcctt 7680

ggtggtggta acatccagag aagcacacag gttggttttc ttggctgcca cgagcttgag 7740

cactcgagcg gcaaaggcgg acttgtggac gttagctcga gcttcgtagg agggcatttt 7800

ggtggtgaag aggagactga aataaattta gtctgcagaa ctttttatcg gaaccttatc 7860

tggggcagtg aagtatatgt tatggtaata gttacgagtt agttgaactt atagatagac 7920

tggactatac ggctatcggt ccaaattaga aagaacgtca atggctctct gggcgtcgcc 7980

tttgccgaca aaaatgtgat catgatgaaa gccagcaatg acgttgcagc tgatattgtt 8040

gtcggccaac cgcgccgaaa acgcagctgt cagacccaca gcctccaacg aagaatgtat 8100

cgtcaaagtg atccaagcac actcatagtt ggagtcgtac tccaaaggcg gcaatgacga 8160

gtcagacaga tactcgtcga ctcaggcgac gacggaattc ctgcagccca tctgcagaat 8220

tcaggagaga ccgggttggc ggcgtatttg tgtcccaaaa aacagcccca attgccccgg 8280

agaagacggc caggccgcct agatgacaaa ttcaacaact cacagctgac tttctgccat 8340

tgccactagg ggggggcctt tttatatggc caagccaagc tctccacgtc ggttgggctg 8400

cacccaacaa taaatgggta gggttgcacc aacaaaggga tgggatgggg ggtagaagat 8460

acgaggataa cggggctcaa tggcacaaat aagaacgaat actgccatta agactcgtga 8520

tccagcgact gacaccattg catcatctaa gggcctcaaa actacctcgg aactgctgcg 8580

ctgatctgga caccacagag gttccgagca ctttaggttg caccaaatgt cccaccaggt 8640

gcaggcagaa aacgctggaa cagcgtgtac agtttgtctt aacaaaaagt gagggcgctg 8700

aggtcgagca gggtggtgtg acttgttata gcctttagag ctgcgaaagc gcgtatggat 8760

ttggctcatc aggccagatt gagggtctgt ggacacatgt catgttagtg tacttcaatc 8820

gccccctgga tatagccccg acaataggcc gtggcctcat ttttttgcct tccgcacatt 8880

tccattgctc ggtacccaca ccttgcttct cctgcacttg ccaaccttaa tactggttta 8940

cattgaccaa catcttacaa gcgggggget tgtctagggt atatataaac agtggctctc 9000

ccaatcggtt gccagtctct tttttccttt ctttccccac agattcgaaa tctaaactac 9060

acatcacaca atgcctgtta ctgacgtcct taagcgaaag tccggtgtca tcgtcggcga 9120

cgatgtccga gccgtgagta tccacgacaa gatcagtgtc gagacgacgc gttttgtgta 9180

atgacacaat ccgaaagtcg ctagcaacac acactctcta cacaaactaa cccagctct 9239

<210> 196
<211> 5031
<212> DNA
<213> Artificial Sequence

<220>
<223> Plasmid pZF5T-PPC-E3

<400> 196

atggagtctg gacctatgcc tgccgggatc cccttccctg aatactatga ctttttcatg 60

gactggaaga caccctggc aattgctgcc acctacaccg ccgctgttgg gctcttcaac 120

cccaaggttg gcaaagtctc gcgcgtggta gccaagtcgg ctaacgccaa gccggcagag 180

cgcacgcagt ccggcgccgc catgaccgcc tttgtctttg tccacaacct tatcctctgc 240

gtgtactctg gaatcacctt ctactacatg ttcccagcca tggtcaagaa ctttagaaca 300

cataccctcc atgaggccta ctgcgatacg gatcagagcc tgtggaacaa cgcccttggc 360

tactggggct acctcttcta cctttcaaag ttttacgagg tcattgacac catcatcatc 420

atcttgaagg ggcgccgctc gtccctgctc cagacctacc accacgccgg cgctatgatc 480

accatgtggt ccggcatcaa ctaccaggca acgcccattt ggatttttgt cgtcttcaac 540

tcgttcatcc acaccatcat gtactgttac tatgccttca cctcaatcgg cttccacccc 600

ccargcaaga agtacctcac ctccatgcag atcacccagt ttttggtcgg catcactatc 660

gccgtctctt atctcttcgt ccctggatgt atccgcacac ccggtgctca gatggctgtc 720

tggatcaacg tcggatacct ctttcccctc acttatctct ttgtggattt tgccaagcgt 780

acttactcca agcgtagtgc catcgccgct cagaagaagg cccagtaagg ccgcattgat 840

gattggaaac acacacatgg gttatatcta ggtgagagtt agttggacag ttatatatta 900

aatcagctat gccaacggta acttcattca tgtcaacgag gaaccagtga ctgcaagtaa 960

tatagaattt gaccaccttg ccattctctt gcactccttt actatatctc atttatttct 1020

tatatacaaa tcacttcttc ttcccagcat cgagctcgga aacctcatga gcaataacat 1080

cgtggatctc gtcaatagag ggctttttgg actccttgct gttggccacc ttgtccttgc 1140

tgtctggctc attctgtttc aacgcctttt aattaatcga gcttggcgta atcatggtca 1200

tagctgtttc ctgtgtgaaa ttgttatccg ctcacaattc cacacaacat acgagccgga 1260

agcataaagt gtaaagcctg gggtgcctaa tgagtgagct aactcacatt aattgcgttg 1320

cgctcactgc ccgctttcca gtcgggaaac ctgtcgtgcc agctgcatta atgaatcggc 1380

caacgcgcgg ggagaggcgg tttgcgtatt gggcgctctt ccgcttcctc gctcactgac 1440

tcgctgcgct cggtcgttcg gctgcggcga gcggtatcag ctcactcaaa ggcggtaata 1500

cggttatcca cagaatcagg ggataacgca ggaaagaaca tgtgagcaaa aggccagcaa 1560

aaggccagga accgtaaaaa ggccgcgttg ctggcgtttt tccataggct ccgcccccct 1620

gacgagcatc acaaaaatcg acgctcaagt cagaggtggc gaaacccgac aggactataa 1680

agataccagg cgtttccccc tggaagctcc ctcgtgcgct ctcctgttcc gaccctgccg 1740

cttaccggat acctgtccgc ctttctccct tcgggaagcg tggcgctttc tcatagctca 1800

cgctgtaggt atctcagttc ggtgtaggtc gttcgctcca agctgggctg tgtgcacgaa 1860

cccccgttc agcccgaccg ctgcgcctta tccggtaact atcgtcttga gtccaacccg 1920

gtaagacacg acttatcgcc actggcagca gccactggta acaggattag cagagcgagg 1980

tatgtaggcg gtgctacaga gttcttgaag tggtggccta actacggcta cactagaagg 2040

acagtatttg gtatctgcgc tctgctgaag ccagttacct tcggaaaaag agttggtagc 2100

tcttgatccg gcaaacaaac caccgctggt agcggtggtt tttttgtttg caagcagcag 2160

attacgcgca gaaaaaaagg atctcaagaa gatcctttga tcttttctac ggggtctgac 2220

gctcagtgga acgaaaactc acgttaaggg attttggtca tgagattatc aaaaaggatc 2280

ttcacctaga tccttttaaa ttaaaaatga agttttaaat caatctaaag tatatatgag 2340

taaacttggt ctgacagtta ccaatgctta atcagtgagg cacctatctc agcgatctgt 2400

ctatttcgtt catccatagt tgcctgactc cccgtcgtgt agataactac gatacgggag 2460

ggcttaccat ctggccccag tgctgcaatg ataccgcgag acccacgctc accggctcca 2520

gatttatcag caataaacca gccagccgga agggccgagc gcagaagtgg tcctgcaact 2580

ttatccgcct ccatccagtc tattaattgt tgccgggaag ctagagtaag tagttcgcca 2640

gttaatagtt tgcgcaacgt tgttgccatt gctacaggca tcgtggtgtc acgctcgtcg 2700

tttggtatgg cttcattcag ctccggttcc caacgatcaa ggcgagttac atgatccccc 2760

atgttgtgca aaaaagcggt tagctccttc ggtcctccga tcgttgtcag aagtaagttg 2820

gccgcagtgt tatcactcat ggttatggca gcactgcata attctcttac tgtcatgcca 2880

tccgtaagat gcttttctgt gactggtgag tactcaacca agtcattctg agaatagtgt 2940

atgcggcgac cgagttgctc ttgcccggcg tcaatacggg ataataccgc gccacatagc 3000

agaactttaa aagtgctcat cattggaaaa cgttcttcgg ggcgaaaact ctcaaggatc 3060

ttaccgctgt tgagatccag ttcgatgtaa cccactcgtg cacccaactg atcttcagca 3120

tcttttactt tcaccagcgt ttctgggtga gcaaaaacag gaaggcaaaa tgccgcaaaa 3180

aagggaataa gggcgacacg gaaatgttga atactcatac tcttccttttt tcaatattat 3240

tgaagcattt atcagggtta ttgtctcatg agcggataca tatttgaatg tatttagaaa 3300

aataaacaaa taggggttcc gcgcacattt ccccgaaaag tgccacctga cgcgccctgt 3360

agcggcgcat taagcgcggc gggtgtggtg gttacgcgca gcgtgaccgc tacacttgcc 3420

agcgccctag cgcccgctcc tttcgctttc ttcccttcct ttctcgccac gttcgccggc 3480

tttccccgtc aagctctaaa tcgggggctc cctttagggt tccgatttag tgctttacgg 3540

cacctcgacc ccaaaaaact tgattagggt gatggttcac gtagtgggcc atcgccctga 3600

tagacggttt ttcgcccttt gacgttggag tccacgttct ttaatagtgg actcttgttc 3660

caaactggaa caacactcaa ccctatctcg gtctattctt ttgatttata agggattttg 3720

ccgatttcgg cctattggtt aaaaaatgag ctgatttaac aaaaatttaa cgcgaatttt 3780

aacaaaatat taacgcttac aatttccatt cgccattcag gctgcgcaac tgttgggaag 3840

ggcgatcggt gcgggcctct tcgctattac gccagctggc gaaaggggga tgtgctgcaa 3900

ggcgattaag ttgggtaacg ccagggtttt cccagtcacg acgttgtaaa acgacggcca 3960

gtgaattgta atacgactca ctatagggcg aattgggtac cgggcccccc ctcgaggtcg 4020

acgtttaaac agtgtacgca gtactataga ggaacatcga ttgccccgga gaagacggcc 4080

aggccgccta gatgacaaat tcaacaactc acagctgact ttctgccatt gccactaggg 4140

gggggccttt ttatatggcc aagccaagct ctccacgtcg gttgggctgc acccaacaat 4200

aaatgggtag ggttgcacca acaaagggat gggatggggg gtagaagata cgaggataac 4260

ggggctcaat ggcacaaata agaacgaata ctgccattaa gactcgtgat ccagcgactg 4320

acaccattgc atcatctaag ggcctcaaaa ctacctcgga actgctgcgc tgatctggac 4380

accacagagg ttccgagcac tttaggttgc accaaatgtc ccaccaggtg caggcagaaa 4440

acgctggaac agcgtgtaca gtttgtctta acaaaaagtg agggcgctga ggtcgagcag 4500

ggtggtgtga cttgttatag cctttagagc tgcgaaagcg cgtatggatt tggctcatca 4560

ggccagattg agggtctgtg gacacatgtc atgttagtgt acttcaatcg ccccctggat 4620

atagccccga caataggccg tggcctcatt tttttgcctt ccgcacattt ccattgctcg 4680

gtacccacac cttgcttctc ctgcacttgc caaccttaat actggtttac attgaccaac 4740

atcttacaag cgggggggctt gtctagggta tatataaaca gtggctctcc caatcggttg 4800

ccagtctctt ttttcctttc tttccccaca gattcgaaat ctaaactaca catcacacaa 4860

tgcctgttac tgacgtcctt aagcgaaagt ccggtgtcat cgtcggcgac gatgtccgag 4920

ccgtgagtat ccacgacaag atcagtgtcg agacgacgcg ttttgtgtaa tgacacaatc 4980

cgaaagtcgc tagcaacaca cactctctac acaaactaac ccagctctcc a 5031

<210> 197
<211> 8462
<212> DNA
<213> Artificial Sequence

<220>
<223> Plasmid pZUF6S

<400> 197
ggccgcaagt gtggatgggg aagtgagtgc ccggttctgt gtgcacaatt ggcaatccaa 60

gatggatgga ttcaacacag ggatatagcg agctacgtgg tggtgcgagg atatagcaac 120

ggatatttat gtttgacact tgagaatgta cgatacaagc actgtccaag tacaatacta 180

aacatactgt acatactcat actcgtaccc gggcaacggt ttcacttgag tgcagtggct 240

agtgctctta ctcgtacagt gtgcaatact gcgtatcata gtctttgatg tatatcgtat 300

tcattcatgt tagttgcgta cgagccggaa gcataaagtg taaagcctgg ggtgcctaat 360

gagtgagcta actcacatta attgcgttgc gctcactgcc cgctttccag tcgggaaacc 420

tgtcgtgcca gctgcattaa tgaatcggcc aacgcgcggg gagaggcggt ttgcgtattg 480

ggcgctcttc cgcttcctcg ctcactgact cgctgcgctc ggtcgttcgg ctgcggcgag 540

cggtatcagc tcactcaaag gcggtaatac ggttatccac agaatcaggg gataacgcag 600

gaaagaacat gtgagcaaaa ggccagcaaa aggccaggaa ccgtaaaaag gccgcgttgc 660

tggcgttttt ccataggctc cgccccccctg acgagcatca caaaaatcga cgctcaagtc 720

agaggtggcg aaacccgaca ggactataaa gataccaggc gtttcccccct ggaagctccc 780

tcgtgcgctc tcctgttccg accctgccgc ttaccggata cctgtccgcc tttctccctt 840

cgggaagcgt ggcgctttct catagctcac gctgtaggta tctcagttcg gtgtaggtcg 900

ttcgctccaa gctgggctgt gtgcacgaac cccccgttca gcccgaccgc tgcgccttat 960

ccggtaacta tcgtcttgag tccaacccgg taagacacga cttatcgcca ctggcagcag 1020

ccactggtaa caggattagc agagcgaggt atgtaggcgg tgctacagag ttcttgaagt 1080

ggtggcctaa ctacggctac actagaagga cagtatttgg tatctgcgct ctgctgaagc 1140

cagttacctt cggaaaaaga gttggtagct cttgatccgg caaacaaacc accgctggta 1200

gcggtggttt ttttgtttgc aagcagcaga ttacgcgcag aaaaaaagga tctcaagaag 1260

atcctttgat cttttctacg gggtctgacg ctcagtggaa 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 cccacgctca ccggctccag atttatcagc aataaaccag ccagccggaa 1620

gggccgagcg cagaagtggt cctgcaactt tatccgcctc catccagtct attaattgtt 1680

gccgggaagc tagagtaagt agttcgccag ttaatagttt gcgcaacgtt gttgccattg 1740

ctacaggcat cgtggtgtca cgctcgtcgt ttggtatggc ttcattcagc tccggttccc 1800

aacgatcaag gcgagttaca tgatccccca tgttgtgcaa aaaagcggtt agctccttcg 1860

gtcctccgat cgttgtcaga agtaagttgg ccgcagtgtt atcactcatg gttatggcag 1920

cactgcataa ttctcttact gtcatgccat ccgtaagatg cttttctgtg actggtgagt 1980

actcaaccaa gtcattctga gaatagtgta tgcggcgacc gagttgctct tgcccggcgt 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 caatattatt gaagcattta tcagggttat tgtctcatga 2340

gcggatacat atttgaatgt atttagaaaa ataaacaaat aggggttccg cgcacatttc 2400

cccgaaaagt gccacctgac gcgccctgta gcggcgcatt aagcgcggcg ggtgtggtgg 2460

ttacgcgcag cgtgaccgct acacttgcca gcgccctagc gcccgctcct ttcgctttct 2520

tcccttcctt tctcgccacg ttcgccggct ttccccgtca agctctaaat cgggggctcc 2580

ctttagggtt ccgatttagt gctttacggc acctcgaccc caaaaaactt gattagggtg 2640

atggttcacg tagtgggcca tcgccctgat agacggtttt tcgccctttg acgttggagt 2700

ccacgttctt taatagtgga ctcttgttcc 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 aaaggggatg gtgctgcaag gcgattaagt tgggtaacgc cagggttttc 3000

ccagtcacga cgttgtaaaa cgacggccag tgaattgtaa tacgactcac tatagggcga 3060

attgggtacc gggcccccccc tcgaggtcga tggtgtcgat 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 gttctcaata tttaaggggt 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 tcccttgtac aacataaata gtcatcgaga 3660

aatatcaact atcaaagaac agctattcac acgttactat tgagattatt attggacgag 3720

aatcacacac tcaactgtct ttctctcttc tagaaataca ggtacaagta tgtactattc 3780

tcattgttca tacttctagt catttcatcc cacatattcc 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 cccctcgttc agtgtcaact 4080

gtaatggtag gaaattacca tacttttgaa gaagcaaaaa aaatgaaaga aaaaaaaaat 4140

cgtatttcca ggttagacgt tccgcagaat ctagaatgcg gtatgcggta cattgttctt 4200

cgaacgtaaa agttgcgctc cctgagatat tgtacatttt tgcttttaca agtacaagta 4260

catcgtacaa ctatgtacta ctgttgatgc atccacaaca gtttgttttg ttttttttg 4320

ttttttttt ttctaatgat 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 gtctgaccat 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 caccccgggg 4980

gtcagaataa gccagtcctc agagtcgccc ttaggtcggt tctgggcaat gaagccaacc 5040

acaaactcgg ggtcggatcg ggcaagctca atggtctgct tggagtactc gccagtggcc 5100

agagagccct tgcaagacag ctcggccagc atgagcagac ctctggccag cttctcgttg 5160

ggagagggga ctaggaactc cttgtactgg gagttctcgt agtcagagac gtcctccttc 5220

ttctgttcag agacagtttc ctcggcacca gctcgcaggc cagcaatgat tccggttccg 5280

ggtacaccgt gggcgttggt gatatcggac cactcggcga ttcggtgaca ccggtactgg 5340

tgcttgacag tgttgccaat 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

attttggtgg tgaagaggag actgaaataa atttagtctg cagaactttt tatcggaacc 5700

ttatctgggg cagtgaagta tatgttatgg taatagttac gagttagttg 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 cgtactccaa aggcggcaat 6000

gacgagtcag acagatactc gtcgactcag gcgacgacgg aattcctgca gcccatctgc 6060

agaattcagg agagaccggg ttggcggcgt atttgtgtcc caaaaaacag ccccaattgc 6120

cccggagaag acggccaggc cgcctagatg acaaattcaa caactcacag ctgactttct 6180

gccattgcca ctaggggggg gccttttttat atggccaagc caagctctcc acgtcggttg 6240

ggctgcaccc aacaataaat gggtagggtt gcaccaacaa agggatggga tgggggggtag 6300

aagatacgag gataacgggg ctcaatggca caaataagaa cgaatactgc cattaagact 6360

cgtgatccag cgactgacac cattgcatca tctaagggcc tcaaaactac ctcggaactg 6420

ctgcgctgat ctggacacca cagaggttcc gagcacttta ggttgcacca aatgtcccac 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 aggccgtggc ctcatttttt tgccttccgc 6720

acatttccat tgctcggtac ccacaccttg cttctcctgc acttgccaac cttaatactg 6780

gtttacattg accaacatct tacaagcggg gggcttgtct agggtatata taaacagtgg 6840

ctctcccaat cggttgccag tctcttttttt ccttttctttc cccacagatt cgaaatctaa 6900

actacacatc acacaatgcc tgttactgac gtccttaagc gaaagtccgg tgtcatcgtc 6960

ggcgacgatg tccgagccgt gagtatccac gacaagatca gtgtcgagac gacgcgtttt 7020

gtgtaatgac acaatccgaa agtcgctagc aacacacact ctctacacaa actaacccag 7080

ctctccatgg ctgccgctcc ctctgtgcga acctttaccc gagccgaggt tctgaacgct 7140

gaggctctga acgagggcaa gaaggacgct gaggctccct tcctgatgat catcgacaac 7200

aaggtgtacg acgtccgaga gttcgtccct gaccatcctg gaggctccgt gattctcacc 7260

cacgttggca aggacggcac cgacgtcttt gacacctttc atcccgaggc tgcttgggag 7320

actctcgcca acttctacgt tggagacatt gacgagtccg accgagacat caagaacgat 7380

gactttgccg ctgaggtccg aaagctgcga accctgttcc agtctctcgg ctactacgac 7440

tcctctaagg cctactacgc cttcaaggtc tccttcaacc tctgcatctg gggactgtcc 7500

accgtcattg tggccaagtg gggtcagacc tccaccctcg ccaacgtgct ctctgctgcc 7560

ctgctcggcc tgttctggca gcagtgcgga tggctggctc acgactttct gcaccaccag 7620

gtcttccagg accgattctg gggtgatctc ttcggagcct tcctgggagg tgtctgccag 7680

ggcttctcct cttcctggtg gaaggacaag cacaacactc accatgccgc tcccaacgtg 7740

catggcgagg atcctgacat tgacacccac cctctcctga cctggtccga gcacgctctg 7800

gagatgttct ccgacgtccc cgatgaggag ctgacccgaa tgtggtctcg attcatggtc 7860

ctgaaccaga cctggttcta cttccccatt ctctccttcg ctcgactgtc ttggtgcctc 7920

cagtccattc tctttgtgct gcccaacggt caggctcaca agccctccgg agctcgagtg 7980

cccatctccc tggtcgagca gctgtccctc gccatgcact ggacctggta cctcgctacc 8040

atgttcctgt tcatcaagga tcctgtcaac atgctcgtgt acttcctggt gtctcaggct 8100

gtgtgcggaa acctgctcgc catcgtgttc tccctcaacc acaacggtat gcctgtgatc 8160

tccaaggagg aggctgtcga catggatttc tttaccaagc agatcatcac tggtcgagat 8220

gtccatcctg gactgttcgc caactggttc accggtggcc tgaactacca gatcgagcat 8280

cacctgttcc cttccatgcc tcgacacaac ttctccaaga tccagcctgc cgtcgagacc 8340

ctgtgcaaga agtacaacgt ccgataccac accactggta tgatcgaggg aactgccgag 8400

gtcttctccc gactgaacga ggtctccaag gccacctcca agatgggcaa ggctcagtaa 8460

gc 8462

<210> 198
<211> 11046
<212> DNA
<213> Artificial Sequence

<220>

<223> Plasmid pZUF6S-E3WT

<400> 198

ggccgcaagt gtggatgggg aagtgagtgc ccggttctgt gtgcacaatt ggcaatccaa 60

gatggatgga ttcaacacag ggatatagcg agctacgtgg tggtgcgagg atatagcaac 120

ggatatttat gtttgacact tgagaatgta cgatacaagc actgtccaag tacaatacta 180

aacatactgt acatactcat actcgtaccc gggcaacggt ttcacttgag tgcagtggct 240

agtgctctta ctcgtacagt gtgcaatact gcgtatcata gtctttgatg tatatcgtat 300

tcattcatgt tagttgcgta cgagccggaa gcataaagtg taaagcctgg ggtgcctaat 360

gagtgagcta actcacatta attgcgttgc gctcactgcc cgctttccag tcgggaaacc 420

tgtcgtgcca gctgcattaa tgaatcggcc aacgcgcggg gagaggcggt ttgcgtattg 480

ggcgctcttc cgcttcctcg ctcactgact cgctgcgctc ggtcgttcgg ctgcggcgag 540

cggtatcagc tcactcaaag gcggtaatac ggttatccac agaatcaggg gataacgcag 600

gaaagaacat gtgagcaaaa ggccagcaaa aggccaggaa ccgtaaaaag gccgcgttgc 660

tggcgttttt ccataggctc cgccccccctg acgagcatca caaaaatcga cgctcaagtc 720

agaggtggcg aaacccgaca ggactataaa gataccaggc gtttcccccct ggaagctccc 780

tcgtgcgctc tcctgttccg accctgccgc ttaccggata cctgtccgcc tttctccctt 840

cgggaagcgt ggcgctttct catagctcac gctgtaggta tctcagttcg gtgtaggtcg 900

ttcgctccaa gctgggctgt gtgcacgaac cccccgttca gcccgaccgc tgcgccttat 960

ccggtaacta tcgtcttgag tccaacccgg taagacacga cttatcgcca ctggcagcag 1020

ccactggtaa caggattagc agagcgaggt atgtaggcgg tgctacagag ttcttgaagt 1080

ggtggcctaa ctacggctac actagaagga cagtatttgg tatctgcgct ctgctgaagc 1140

cagttacctt cggaaaaaga gttggtagct cttgatccgg caaacaaacc accgctggta 1200

gcggtggttt ttttgtttgc aagcagcaga ttacgcgcag aaaaaaagga tctcaagaag 1260

atcctttgat cttttctacg gggtctgacg ctcagtggaa 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 cccacgctca ccggctccag atttatcagc aataaaccag ccagccggaa 1620

gggccgagcg cagaagtggt cctgcaactt tatccgcctc catccagtct attaattgtt 1680

gccgggaagc tagagtaagt agttcgccag ttaatagttt gcgcaacgtt gttgccattg 1740

ctacaggcat cgtggtgtca cgctcgtcgt ttggtatggc ttcattcagc tccggttccc 1800

aacgatcaag gcgagttaca tgatccccca tgttgtgcaa aaaagcggtt agctccttcg 1860

gtcctccgat cgttgtcaga agtaagttgg ccgcagtgtt atcactcatg gttatggcag 1920

cactgcataa ttctcttact gtcatgccat ccgtaagatg cttttctgtg actggtgagt 1980

actcaaccaa gtcattctga gaatagtgta tgcggcgacc gagttgctct tgcccggcgt 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 caatattatt gaagcattta tcagggttat tgtctcatga 2340

gcggatacat atttgaatgt atttagaaaa ataaacaaat aggggttccg cgcacatttc 2400

cccgaaaagt gccacctgac gcgccctgta gcggcgcatt aagcgcggcg ggtgtggtgg 2460

ttacgcgcag cgtgaccgct acacttgcca gcgccctagc gcccgctcct ttcgctttct 2520

tcccttcctt tctcgccacg ttcgccggct ttccccgtca agctctaaat cgggggctcc 2580

ctttagggtt ccgatttagt gctttacggc acctcgaccc caaaaaactt gattagggtg 2640

atggttcacg tagtgggcca tcgccctgat agacggtttt tcgccctttg acgttggagt 2700

ccacgttctt taatagtgga ctcttgttcc 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 aaaggggat gtgctgcaag gcgattaagt tgggtaacgc cagggttttc 3000

ccagtcacga cgttgtaaaa cgacggccag tgaattgtaa tacgactcac tatagggcga 3060

attgggtacc gggccccccc tcgaggtcga tggtgtcgat 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 gttctcaata tttaaggggt 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 tcccttgtac aacataaata gtcatcgaga 3660

aatatcaact atcaaagaac agctattcac acgttactat tgagattatt attggacgag 3720

aatcacacac tcaactgtct ttctctcttc tagaaataca ggtacaagta tgtactattc 3780

tcattgttca tacttctagt catttcatcc cacatattcc 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 cccctcgttc agtgtcaact 4080

gtaatggtag gaaattacca tacttttgaa gaagcaaaaa aaatgaaaga aaaaaaaaat 4140

cgtatttcca ggttagacgt tccgcagaat ctagaatgcg gtatgcggta cattgttctt 4200

cgaacgtaaa agttgcgctc cctgagatat tgtacatttt tgcttttaca agtacaagta 4260

catcgtacaa ctatgtacta ctgttgatgc atccacaaca gtttgttttg ttttttttg 4320

ttttttttt ttctaatgat tcattaccgc tatgtatacc tacttgtact tgtagtaagc 4380

cgggttattg gcgttcaatt aatcatagac ttatgaatct gcacggtgtg cgctgcgagt 4440

tacttttagc ttatgcatgc tacttgggtg taatattggg atctgttcgg aaatcaacgg 4500

atgctcaatc gatgcagaat tcaggagaga ccgggttggc ggcgtatttg tgtcccaaaa 4560

aacagcccca attgccccaa ttgaccccaa attgacccag tagcgggccc aaccccggcg 4620

agagccccct tcaccccaca tatcaaacct cccccggttc ccacacttgc cgttaagggc 4680

gtagggtact gcagtctgga atctacgctt gttcagactt tgtactagtt tctttgtctg 4740

gccatccggg taacccatgc cggacgcaaa atagactact gaaaattttt ttgctttgtg 4800

gttgggactt tagccaaggg tataaaagac caccgtcccc gaattacctt tcctcttctt 4860

ttctctctct ccttgtcaac tcacacccga aatcgttaag catttccttc tgagtataag 4920

aatcattcac caggagagac cgggttggcg gcgtatttgt gtcccaaaaa acagccccaa 4980

ttgccccgga gaagacggcc aggccgccta gatgacaaat tcaacaactc acagctgact 5040

ttctgccatt gccactaggg gggggccttt ttatatggcc aagccaagct ctccacgtcg 5100

gttgggctgc acccaacaat aaatgggtag ggttgcacca acaaagggat gggatggggg 5160

gtagaagata cgaggataac ggggctcaat ggcacaaata agaacgaata ctgccattaa 5220

gactcgtgat ccagcgactg acaccattgc atcatctaag ggcctcaaaa ctacctcgga 5280

actgctgcgc tgatctggac accacagagg ttccgagcac tttaggttgc accaaatgtc 5340

ccaccaggtg caggcagaaa acgctggaac agcgtgtaca gtttgtctta acaaaaagtg 5400

agggcgctga ggtcgagcag ggtggtgtga cttgttatag cctttagagc tgcgaaagcg 5460

cgtatggatt tggctcatca ggccagattg agggtctgtg gacacatgtc atgttagtgt 5520

acttcaatcg ccccctggat atagccccga caataggccg tggcctcatt tttttgcctt 5580

ccgcacattt ccattgctcg gtacccacac cttgcttctc ctgcacttgc caaccttaat 5640

actggtttac attgaccaac atcttacaag cgggggggctt gtctagggta tatataaaca 5700

gtggctctcc caatcggttg ccagtctctt ttttcctttc tttccccaca gattcgaaat 5760

ctaaactaca catcacacaa tgcctgttac tgacgtcctt aagcgaaagt ccggtgtcat 5820

cgtcggcgac gatgtccgag ccgtgagtat ccacgacaag atcagtgtcg agacgacgcg 5880

ttttgtgtaa tgacacaatc cgaaagtcgc tagcaacaca cactctctac acaaactaac 5940

ccagctatgg agtctggacc tatgcctgcc gggatcccct tccctgaata ctatgacttt 6000

ttcatggact ggaagacacc cctggcaatt gctgccacct acaccgccgc tgttgggctc 6060

ttcaacccca aggttggcaa agtctcgcgc gtggtagcca agtcggctaa cgccaagccg 6120

gcagagcgca cgcagtccgg cgccgccatg accgcctttg tctttgtcca caaccttatc 6180

ctctgcgtgt actctggaat caccttctac tacatgttcc cagccatggt caagaacttt 6240

agaacacata ccctccatga ggcctactgc gatacggatc agagcctgtg gaacaacgcc 6300

cttggctact ggggctacct cttctacctt tcaaagtttt acgaggtcat tgacaccatc 6360

atcatcatct tgaaggggcg ccgctcgtcc ctgctccaga cctaccacca cgccggcgct 6420

atgatcacca tgtggtccgg catcaactac caggcaacgc ccatttggat ttttgtcgtc 6480

ttcaactcgt tcatccacac catcatgtac tgttactatg ccttcacctc aatcggcttc 6540

cacccccar gcaagaagta cctcacctcc atgcagatca cccagttttt ggtcggcatc 6600

actatcgccg tctcttatct cttcgtccct ggatgtatcc gcacacccgg tgctcagatg 6660

gctgtctgga tcaacgtcgg atacctcttt cccctcactt atctctttgt ggattttgcc 6720

aagcgtactt actccaagcg tagtgccatc gccgctcaga agaaggccca gtaaattgat 6780

gattggaaac acacacatgg gttatatcta ggtgagagtt agttggacag ttatatatta 6840

aatcagctat gccaacggta acttcattca tgtcaacgag gaaccagtga ctgcaagtaa 6900

tatagaattt gaccaccttg ccattctctt gcactccttt actatatctc atttatttct 6960

tatatacaaa tcacttcttc ttcccagcat cgagctcgga aacctcatga gcaataacat 7020

cgtggatctc gtcaatagag ggctttttgg actccttgct gttggccacc ttgtccttgc 7080

tgtctggctc attctgtttc aacgcctttt aattaagtca tacacaagtc agctttcttc 7140

gagcctcata taagtataag tagttcaacg tattagcact gtacccagca tctccgtatc 7200

gagaaacaca acaacatgcc ccattggaca gatcatgcgg atacacaggt tgtgcagtat 7260

catacatact cgatcagaca ggtcgtctga ccatcataca agctgaacaa gcgctccata 7320

cttgcacgct ctctatatac acagttaaat tacatatcca tagtctaacc tctaacagtt 7380

aatcttctgg taagcctccc agccagcctt ctggtatcgc ttggcctcct caataggatc 7440

tcggttctgg ccgtacagac ctcggccgac aattatgata tccgttccgg tagacatgac 7500

atcctcaaca gttcggtact gctgtccgag agcgtctccc ttgtcgtcaa gacccacccc 7560

gggggtcaga ataagccagt cctcagagtc gcccttaggt cggttctggg caatgaagcc 7620

aaccacaaac tcggggtcgg atcgggcaag ctcaatggtc tgcttggagt actcgccagt 7680

ggccagagag cccttgcaag acagctcggc cagcatgagc agacctctgg ccagcttctc 7740

gttgggagag gggactagga actccttgta ctgggagttc tcgtagtcag agacgtcctc 7800

cttcttctgt tcagagacag tttcctcggc accagctcgc aggccagcaa tgattccggt 7860

tccgggtaca ccgtgggcgt tggtgatatc ggaccactcg gcgattcggt gacaccggta 7920

ctggtgcttg acagtgttgc caatatctgc gaactttctg tcctcgaaca ggaagaaacc 7980

gtgcttaaga gcaagttcct tgaggggggag cacagtgccg gcgtaggtga agtcgtcaat 8040

gatgtcgata tgggttttga tcatgcacac ataaggtccg accttatcgg caagctcaat 8100

gagctccttg gtggtggtaa catccagaga agcacacagg ttggttttct tggctgccac 8160

gagcttgagc actcgagcgg caaaggcgga cttgtggacg ttagctcgag cttcgtagga 8220

gggcattttg gtggtgaaga ggagactgaa ataaatttag tctgcagaac tttttatcgg 8280

aaccttatct ggggcagtga agtatatgtt atggtaatag ttacgagtta gttgaactta 8340

tagatagact ggactatacg gctatcggtc caaattagaa agaacgtcaa tggctctctg 8400

ggcgtcgcct ttgccgacaa aaatgtgatc atgatgaaag ccagcaatga cgttgcagct 8460

gatattgttg tcggccaacc gcgccgaaaa cgcagctgtc agacccacag cctccaacga 8520

agaatgtatc gtcaaagtga tccaagcaca ctcatagttg gagtcgtact ccaaaggcgg 8580

caatgacgag tcagacagat actcgtcgac tcaggcgacg acggaattcc tgcagcccat 8640

ctgcagaatt caggagagac cgggttggcg gcgtatttgt gtcccaaaaa acagccccaa 8700

ttgccccgga gaagacggcc aggccgccta gatgacaaat tcaacaactc acagctgact 8760

ttctgccatt gccactaggg gggggccttt ttatatggcc aagccaagct ctccacgtcg 8820

gttgggctgc acccaacaat aaatgggtag ggttgcacca acaaagggat gggatggggg 8880

gtagaagata cgaggataac ggggctcaat ggcacaaata agaacgaata ctgccattaa 8940

gactcgtgat ccagcgactg acaccattgc atcatctaag ggcctcaaaa ctacctcgga 9000

actgctgcgc tgatctggac accacagagg ttccgagcac tttaggttgc accaaatgtc 9060

ccaccaggtg caggcagaaa acgctggaac agcgtgtaca gtttgtctta acaaaaagtg 9120

agggcgctga ggtcgagcag ggtggtgtga cttgttatag cctttagagc tgcgaaagcg 9180

cgtatggatt tggctcatca ggccagattg agggtctgtg gacacatgtc atgttagtgt 9240

acttcaatcg ccccctggat atagccccga caataggccg tggcctcatt tttttgcctt 9300

ccgcacattt ccattgctcg gtacccacac cttgcttctc ctgcacttgc caaccttaat 9360

actggtttac attgaccaac atcttacaag cgggggggctt gtctagggta tatataaaca 9420

gtggctctcc caatcggttg ccagtctctt ttttcctttc tttccccaca gattcgaaat 9480

ctaaactaca catcacacaa tgcctgttac tgacgtcctt aagcgaaagt ccggtgtcat 9540

cgtcggcgac gatgtccgag ccgtgagtat ccacgacaag atcagtgtcg agacgacgcg 9600

ttttgtgtaa tgacacaatc cgaaagtcgc tagcaacaca cactctctac acaaactaac 9660

ccagctctcc atggctgccg ctccctctgt gcgaaccttt acccgagccg aggttctgaa 9720

cgctgaggct ctgaacgagg gcaagaagga cgctgaggct cccttcctga tgatcatcga 9780

caacaaggtg tacgacgtcc gagagttcgt ccctgaccat cctggaggct ccgtgattct 9840

cacccacgtt ggcaaggacg gcaccgacgt ctttgacacc tttcatcccg aggctgcttg 9900

ggagactctc gccaacttct acgttggaga cattgacgag tccgaccgag acatcaagaa 9960

cgatgacttt gccgctgagg tccgaaagct gcgaaccctg ttccagtctc tcggctacta 10020

cgactcctct aaggcctact acgccttcaa ggtctccttc aacctctgca tctggggact 10080

gtccaccgtc attgtggcca agtggggtca gacctccacc ctcgccaacg tgctctctgc 10140

tgccctgctc ggcctgttct ggcagcagtg cggatggctg gctcacgact ttctgcacca 10200

ccaggtcttc caggaccgat tctggggtga tctcttcgga gccttcctgg gaggtgtctg 10260

ccagggcttc tcctcttcct ggtggaagga caagcacaac actcaccatg ccgctcccaa 10320

cgtgcatggc gaggatcctg acattgacac ccaccctctc ctgacctggt ccgagcacgc 10380

tctggagatg ttctccgacg tccccgatga ggagctgacc cgaatgtggt ctcgattcat 10440

ggtcctgaac cagacctggt tctacttccc cattctctcc ttcgctcgac tgtcttggtg 10500

cctccagtcc attctctttg tgctgcccaa cggtcaggct cacaagccct ccggagctcg 10560

agtgcccatc tccctggtcg agcagctgtc cctcgccatg cactggacct ggtacctcgc 10620

taccatgttc ctgttcatca aggatcctgt caacatgctc gtgtacttcc tggtgtctca 10680

ggctgtgtgc ggaaacctgc tcgccatcgt gttctccctc aaccacaacg gtatgcctgt 10740

gatctccaag gaggaggctg tcgacatgga tttctttacc aagcagatca tcactggtcg 10800

agatgtccat cctggactgt tcgccaactg gttcaccggt ggcctgaact accagatcga 10860

gcatcacctg ttcccttcca tgcctcgaca caacttctcc aagatccagc ctgccgtcga 10920

gaccctgtgc aagaagtaca acgtccgata ccacaccact ggtatgatcg agggaactgc 10980

cgaggtcttc tcccgactga acgaggtctc caaggccacc tccaagatgg gcaaggctca 11040

gtaagc 11046

<210> 199
<211> 6561
<212> DNA
<213> Artificial Sequence

<220>
<223> Plasmid pZKUGPYE1-N

<400> 199
gtacgaggaa actgtctctg aacagaagaa ggaggacgtc tctgactacg agaactccca 60

gtacaaggag ttcctagtcc cctctcccaa cgagaagctg gccagaggtc tgctcatgct 120

ggccgagctg tcttgcaagg gctctctggc cactggcgag tactccaagc agaccattga 180

gcttgcccga tccgaccccg agtttgtggt tggcttcatt gcccagaacc gacctaaggg 240

cgactctgag gactggctta ttctgacccc cggggtgggt cttgacgaca agggagacgc 300

tctcggacag cagtaccgaa ctgttgagga tgtcatgtct accggaacgg atatcataat 360

tgtcggccga ggtctgtacg gccagaaccg agatcctatt gaggaggcca agcgatacca 420

gaaggctggc tgggaggctt accagaagat taactgttag aggttagact atggatatgt 480

aatttaactg tgtatataga gagcgtgcaa gtatggagcg cttgttcagc ttgtatgatg 540

gtcagacgac ctgtctgatc gagtatgtat gatactgcac aacctgtgta tccgcatgat 600

ctgtccaatg gggcatgttg ttgtgtttct cgatacggag atgctgggta cagtgctaat 660

acgttgaact acttatactt atatgaggct cgaagaaagc tgacttgtgt atgacttaat 720

taatcgagct tggcgtaatc atggtcatag ctgtttcctg tgtgaaattg ttatccgctc 780

acaattccac acaacatacg agccggaagc ataaagtgta aagcctgggg tgcctaatga 840

gtgagctaac tcacattaat tgcgttgcgc tcactgcccg ctttccagtc gggaaacctg 900

tcgtgccagc tgcattaatg aatcggccaa cgcgcgggga gaggcggttt gcgtattggg 960

cgctcttccg cttcctcgct cactgactcg ctgcgctcgg tcgttcggct gcggcgagcg 1020

gtatcagctc actcaaaggc ggtaatacgg ttatccacag aatcagggga taacgcagga 1080

aagaacatgt gagcaaaagg ccagcaaaag gccaggaacc gtaaaaaggc cgcgttgctg 1140

gcgtttttcc ataggctccg ccccccctgac gagcatcaca aaaatcgacg ctcaagtcag 1200

aggtggcgaa acccgacagg actataaaga taccaggcgt ttcccccctgg aagctccctc 1260

gtgcgctctc ctgttccgac cctgccgctt accggatacc tgtccgcctt tctcccttcg 1320

ggaagcgtgg cgctttctca tagctcacgc tgtaggtatc tcagttcggt gtaggtcgtt 1380

cgctccaagc tgggctgtgt gcacgaaccc cccgttcagc ccgaccgctg cgccttatcc 1440

ggtaactatc gtcttgagtc caacccggta agacacgact tatcgccact ggcagcagcc 1500

actggtaaca ggattagcag agcgaggtat gtaggcggtg ctacagagtt cttgaagtgg 1560

tggcctaact acggctacac tagaaggaca gtatttggta tctgcgctct gctgaagcca 1620

gttaccttcg gaaaaagagt tggtagctct tgatccggca aacaaaccac cgctggtagc 1680

ggtggttttt ttgtttgcaa gcagcagatt acgcgcagaa aaaaaggatc tcaagaagat 1740

cctttgatct tttctacggg gtctgacgct cagtggaacg aaaactcacg ttaagggatt 1800

ttggtcatga gattatcaaa aaggatcttc acctagatcc ttttaaatta aaaatgaagt 1860

tttaaatcaa tctaaagtat atatgagtaa acttggtctg acagttacca atgcttaatc 1920

agtgaggcac ctatctcagc gatctgtcta tttcgttcat ccatagttgc ctgactcccc 1980

gtcgtgtaga taactacgat acgggagggc ttaccatctg gccccagtgc tgcaatgata 2040

ccgcgagacc cacgctcacc ggctccagat ttatcagcaa taaaccagcc agccggaagg 2100

gccgagcgca gaagtggtcc tgcaacttta tccgcctcca tccagtctat taattgttgc 2160

cgggaagcta gagtaagtag ttcgccagtt aatagtttgc gcaacgttgt tgccattgct 2220

acaggcatcg tggtgtcacg ctcgtcgttt ggtatggctt cattcagctc cggttcccaa 2280

cgatcaaggc gagttacatg atcccccatg ttgtgcaaaa aagcggttag ctccttcggt 2340

cctccgatcg ttgtcagaag taagttggcc gcagtgttat cactcatggt tatggcagca 2400

ctgcataatt ctcttactgt catgccatcc gtaagatgct tttctgtgac tggtgagtac 2460

tcaaccaagt cattctgaga atagtgtatg cggcgaccga gttgctcttg cccggcgtca 2520

atacgggata ataccgcgcc acatagcaga actttaaaag tgctcatcat tggaaaacgt 2580

tcttcggggc gaaaactctc aaggatctta ccgctgttga gatccagttc gatgtaaccc 2640

actcgtgcac ccaactgatc ttcagcatct tttactttca ccagcgtttc tgggtgagca 2700

aaaacaggaa ggcaaaatgc cgcaaaaaag ggaataaggg cgacacggaa atgttgaata 2760

ctcatactct tccttttca atattattga agcatttatc agggttattg tctcatgagc 2820

ggatacatat ttgaatgtat ttagaaaaat aaacaaatag gggttccgcg cacatttccc 2880

cgaaaagtgc cacctgacgc gccctgtagc ggcgcattaa gcgcggcggg tgtggtggtt 2940

acgcgcagcg tgaccgctac acttgccagc gccctagcgc ccgctccttt cgctttcttc 3000

ccttcctttc tcgccacgtt cgccggcttt ccccgtcaag ctctaaatcg ggggctccct 3060

ttagggttcc gatttagtgc tttacggcac ctcgacccca aaaaacttga ttagggtgat 3120

ggttcacgta gtgggccatc gccctgatag acggtttttc gccctttgac gttggagtcc 3180

acgttcttta atagtggact cttgttccaa actggaacaa cactcaaccc tatctcggtc 3240

tattcttttg atttataagg gattttgccg atttcggcct attggttaaa aaatgagctg 3300

atttaacaaa aatttaacgc gaattttaac aaaatattaa cgcttacaat ttccattcgc 3360

cattcaggct gcgcaactgt tgggaagggc gatcggtgcg ggcctcttcg ctattacgcc 3420

agctggcgaa aggggatgt gctgcaaggc gattaagttg ggtaacgcca gggttttccc 3480

agtcacgacg ttgtaaaacg acggccagtg aattgtaata cgactcacta tagggcgaat 3540

tgggtaccgg gccccccctc gaggtcgacg agtatctgtc tgactcgtca ttgccgcctt 3600

tggagtacga ctccaactat gagtgtgctt ggatcacttt gacgatacat tcttcgttgg 3660

aggctgtggg tctgacagct gcgttttcgg cgcggttggc cgacaacaat atcagctgca 3720

acgtcattgc tggctttcat catgatcaca tttttgtcgg caaaggcgac gcccagagag 3780

ccattgacgt tctttctaat ttggaccgat agccgtatag tccagtctat ctataagttc 3840

aactaactcg taactattac cataacatat acttcactgc cccagataag gttccgataa 3900

aaagttctgc agactaaatt tatttcagtc tcctcttcac caccaaaatg ccctcctacg 3960

aagctcgagt gctcaagctc gtggcagcca agaaaaccaa cctgtgtgct tctctggatg 4020

ttaccaccac caaggagctc attgagcttg ccgataaggt cggaccttat gtgtgcatga 4080

tcaaaaccca tatcgacatc attgacgact tcacctacgc cggcactgtg ctccccctca 4140

aggaacttgc tcttaagcac ggtttcttcc tgttcgagga cagaaagttc gcagatattg 4200

gcaacactgt caagcaccag taccggtgtc accgaatcgc cgagtggtcc gatatcacca 4260

acgcccacgg tgtacccgga accggaatcg atgcgtatct gtgggacatg tggtcgttgc 4320

gccattatgt aagcagcgtg tactcctctg actgtccata tggtttgctc catctcaccc 4380

tcatcgtttt cattgttcac aggcggccac aaaaaaactg tcttctctcc ttctctcttc 4440

gccttagtct actcggacca gttttagttt agcttggcgc cactggataa atgagacctc 4500

aggccttgtg atgaggaggt cacttatgaa gcatgttagg aggtgcttgt atggatagag 4560

aagcacccaa aataataaga ataataataa aacaggggggc gttgtcattt catatcgtgt 4620

tttcaccatc aatacacctc caaacaatgc ccttcatgtg gccagcccca atattgtcct 4680

gtagttcaac tctatgcagc tcgtatctta ttgagcaagt aaaactctgt cagccgatat 4740

tgcccgaccc gcgacaaggg tcaacaaggt ggtgtaaggc cttcgcagaa gtcaaaactg 4800

tgccaaacaa acatctagag tctctttggt gtttctcgca tatatttwat cggctgtctt 4860

acgtatttgc gcctcggtac cggactaatt tcggatcatc cccaatacgc tttttcttcg 4920

cagctgtcaa cagtgtccat gatctatcca cctaaatggg tcatatgagg cgtataattt 4980

cgtggtgctg ataataattc ccatatattt gacacaaaac ttccccccct agacatacat 5040

ctcacaatct cacttcttgt gcttctgtca cacatctcct ccagctgact tcaactcaca 5100

cctctgcccc agttggtcta cagcggtata aggtttctcc gcatagaggt gcaccactcc 5160

tcccgatact tgtttgtgtg acttgtgggt cacgacatat atatctacac acattgcgcc 5220

acccttggt tcttccagca caacaaaaac acgacacgct aaccatggct gccgtcccta 5280

ttgaattcaa cgtcccctcc gtggaccgac cctttggtat ctacctctgg gccatctttg 5340

accaggcctg ggagaagctt ttcggctggc ccgcgtcctc tttcattttc gtgcgaaatg 5400

accccaacat ccccttttcc tctacccctc ccgtgatcat tgccatcatt gtgtactaca 5460

ttgtcatctt tggcggccga gaggtgatgc gaaacctgtc tcccatccga ctcaactggc 5520

tcttccagat ccacaacatc ttcctcaccc ttctgtccgg tatgctcctc ctcctcctcg 5580

ttgagcagct cttccccatc attgtccgac agggtatcct ctacgccatc tgcgactacg 5640

gatcttggac tcagcccatt gtcttctgct actacctcaa ctacctgacc aagtactttg 5700

agctgatcga caccgttttc cttgtgctgc gaaagaagaa gctgactttc ctccacacct 5760

accatcacgg tgccactgct cttctgtgct acacccagct cattggtaag acctcggtct 5820

cttgggtccc catcaccctt aacctgtttg tccacgttgt catgtacttc tactacttcc 5880

tggctgcgcg aggtatccga gtgtggtgga aggagtgggt cacccggctc cagatcatcc 5940

agttcgttat cgatcttgga tttgtctact ttgcctctta cacctacttc acctctacct 6000

actggccctg gatgcccaac atgggctctt gtgccggcga ggagtttgct gctatttacg 6060

gctgtggtct gctgacctct tacctcttcc tcttcatcgc cttctacatc aactcttacc 6120

gaaagccctc ttccaaggga ccttccaagc ctgttgttgc tgtcgatggc cctgttggcg 6180

gcgtcaacgc ccagactggt gcttctcgag gccagaccac tacccgatct cgacgagcat 6240

aagcggccgc aagtgtggat ggggaagtga gtgcccggtt ctgtgtgcac aattggcaat 6300

ccaagatgga tggattcaac acagggatat agcgagctac gtggtggtgc gaggatatag 6360

caacggatat ttatgtttga cacttgagaa tgtacgatac aagcactgtc caagtacaat 6420

actaaacata ctgtacatac tcatactcgt acccgggcaa cggtttcact tgagtgcagt 6480

ggctagtgct cttactcgta cagtgtgcaa tactgcgtat catagtcttt gatgtatatc 6540

gtattcattc atgttagttg c 6561

<210> 200
<211> 6498
<212> DNA
<213> Artificial Sequence

<220>
<223> Plasmid pZKUGPYE2

<400> 200
ggccgcaagt gtggatgggg aagtgagtgc ccggttctgt gtgcacaatt ggcaatccaa 60

gatggatgga ttcaacacag ggatatagcg agctacgtgg tggtgcgagg atatagcaac 120

ggatatttat gtttgacact tgagaatgta cgatacaagc actgtccaag tacaatacta 180

aacatactgt acatactcat actcgtaccc gggcaacggt ttcacttgag tgcagtggct 240

agtgctctta ctcgtacagt gtgcaatact gcgtatcata gtctttgatg tatatcgtat 300

tcattcatgt tagttgcgta cgaggaaact gtctctgaac agaagaagga ggacgtctct 360

gactacgaga actcccagta caaggagttc ctagtcccct ctcccaacga gaagctggcc 420

agaggtctgc tcatgctggc cgagctgtct tgcaagggct ctctggccac tggcgagtac 480

tccaagcaga ccattgagct tgcccgatcc gaccccgagt ttgtggttgg cttcattgcc 540

cagaaccgac ctaagggcga ctctgaggac tggcttattc tgacccccgg ggtgggtctt 600

gacgacaagg gagacgctct cggacagcag taccgaactg ttgaggatgt catgtctacc 660

ggaacggata tcataattgt cggccgaggt ctgtacggcc agaaccgaga tcctattgag 720

gaggccaagc gataccagaa ggctggctgg gaggcttacc agaagattaa ctgttagagg 780

ttagactatg gatatgtaat ttaactgtgt atatagagag cgtgcaagta tggagcgctt 840

gttcagcttg tatgatggtc agacgacctg tctgatcgag tatgtatgat actgcacaac 900

ctgtgtatcc gcatgatctg tccaatgggg catgttgttg tgtttctcga tacggagatg 960

ctgggtacag tgctaatacg ttgaactact tatacttata tgaggctcga agaaagctga 1020

cttgtgtatg acttaattaa tcgagcttgg cgtaatcatg gtcatagctg tttcctgtgt 1080

gaaattgtta tccgctcaca attccacaca acatacgagc cggaagcata aagtgtaaag 1140

cctggggtgc ctaatgagtg agctaactca cattaattgc gttgcgctca ctgcccgctt 1200

tccagtcggg aaacctgtcg tgccagctgc attaatgaat cggccaacgc gcggggagag 1260

gcggtttgcg tattgggcgc tcttccgctt cctcgctcac tgactcgctg cgctcggtcg 1320

ttcggctgcg gcgagcggta tcagctcact caaaggcggt aatacggtta tccacagaat 1380

caggggataa cgcaggaaag aacatgtgag caaaaggcca gcaaaaggcc aggaaccgta 1440

aaaaggccgc gttgctggcg tttttccata ggctccgccc ccctgacgag catcacaaaa 1500

atcgacgctc aagtcagagg tggcgaaacc cgacaggact ataaagatac caggcgtttc 1560

cccctggaag ctccctcgtg cgctctcctg ttccgaccct gccgcttacc ggatacctgt 1620

ccgcctttct cccttcggga agcgtggcgc tttctcatag ctcacgctgt aggtatctca 1680

gttcggtgta ggtcgttcgc tccaagctgg gctgtgtgca cgaaccccc gttcagcccg 1740

accgctgcgc cttatccggt aactatcgtc ttgagtccaa cccggtaaga cacgacttat 1800

cgccactggc agcagccact ggtaacagga ttagcagagc gaggtatgta ggcggtgcta 1860

cagagttctt gaagtggtgg cctaactacg gctacactag aaggacagta tttggtatct 1920

gcgctctgct gaagccagtt accttcggaa aaagagttgg tagctcttga tccggcaaac 1980

aaaccaccgc tggtagcggt ggtttttttg tttgcaagca gcagattacg cgcagaaaaa 2040

aaggatctca agaagatcct ttgatctttt ctacggggtc tgacgctcag tggaacgaaa 2100

actcacgtta agggattttg gtcatgagat tatcaaaaag gatcttcacc tagatccttt 2160

taaattaaaa atgaagtttt aaatcaatct aaagtatata tgagtaaact tggtctgaca 2220

gttaccaatg cttaatcagt gaggcaccta tctcagcgat ctgtctattt cgttcatcca 2280

tagttgcctg actccccgtc gtgtagataa ctacgatacg ggagggctta ccatctggcc 2340

ccagtgctgc aatgataccg cgagacccac gctcaccggc tccagattta tcagcaataa 2400

accagccagc cggaagggcc gagcgcagaa gtggtcctgc aactttatcc gcctccatcc 2460

agtctattaa ttgttgccgg gaagctagag taagtagttc gccagttaat agtttgcgca 2520

acgttgttgc cattgctaca ggcatcgtgg tgtcacgctc gtcgtttggt atggcttcat 2580

tcagctccgg ttcccaacga tcaaggcgag ttacatgatc ccccatgttg tgcaaaaaag 2640

cggttagctc cttcggtcct ccgatcgttg tcagaagtaa gttggccgca gtgttatcac 2700

tcatggttat ggcagcactg cataattctc ttactgtcat gccatccgta agatgctttt 2760

ctgtgactgg tgagtactca accaagtcat tctgagaata gtgtatgcgg cgaccgagtt 2820

gctcttgccc ggcgtcaata cgggataata ccgcgccaca tagcagaact ttaaaagtgc 2880

tcatcattgg aaaacgttct tcggggcgaa aactctcaag gatcttaccg ctgttgagat 2940

ccagttcgat gtaacccact cgtgcaccca actgatcttc agcatctttt actttcacca 3000

gcgtttctgg gtgagcaaaa acaggaaggc aaaatgccgc aaaaaaggga ataagggcga 3060

cacggaaatg ttgaatactc atactcttcc tttttcaata ttattgaagc atttatcagg 3120

gttattgtct catgagcgga tacatatttg aatgtattta gaaaaataaa caaatagggg 3180

ttccgcgcac atttccccga aaagtgccac ctgacgcgcc ctgtagcggc gcattaagcg 3240

cggcgggtgt ggtggttacg cgcagcgtga ccgctacact tgccagcgcc ctagcgcccg 3300

ctcctttcgc tttcttccct tcctttctcg ccacgttcgc cggctttccc cgtcaagctc 3360

taaatcgggg gctcccttta gggttccgat ttagtgcttt acggcacctc gaccccaaaa 3420

aacttgatta gggtgatggt tcacgtagtg ggccatcgcc ctgatagacg gtttttcgcc 3480

ctttgacgtt ggagtccacg ttctttaata gtggactctt gttccaaact ggaacaacac 3540

tcaaccctat ctcggtctat tcttttgatt tataagggat tttgccgatt tcggcctatt 3600

ggttaaaaaa tgagctgatt taacaaaaat ttaacgcgaa ttttaacaaa atattaacgc 3660

ttacaatttc cattcgccat tcaggctgcg caactgttgg gaagggcgat cggtgcgggc 3720

ctcttcgcta ttacgccagc tggcgaaagg gggatgtgct gcaaggcgat taagttgggt 3780

aacgccaggg ttttcccagt cacgacgttg taaaacgacg gccagtgaat tgtaatacga 3840

ctcactatag ggcgaattgg gtaccgggcc ccccctcgag gtcgacgagt atctgtctga 3900

ctcgtcattg ccgcctttgg agtacgactc caactatgag tgtgcttgga tcactttgac 3960

gatacattct tcgttggagg ctgtgggtct gacagctgcg ttttcggcgc ggttggccga 4020

caacaatatc agctgcaacg tcattgctgg ctttcatcat gatcacattt ttgtcggcaa 4080

aggcgacgcc cagagagcca ttgacgttct ttctaatttg gaccgatagc cgtatagtcc 4140

agtctatcta taagttcaac taactcgtaa ctattaccat aacatatact tcactgcccc 4200

agataaggtt ccgataaaaa gttctgcaga ctaaatttat ttcagtctcc tcttcaccac 4260

caaaatgccc tcctacgaag ctcgagtgct caagctcgtg gcagccaaga aaaccaacct 4320

gtgtgcttct ctggatgtta ccaccaccaa ggagctcatt gagcttgccg ataaggtcgg 4380

accttatgtg tgcatgatca aaacccatat cgacatcatt gacgacttca cctacgccgg 4440

cactgtgctc cccctcaagg aacttgctct taagcacggt ttcttcctgt tcgaggacag 4500

aaagttcgca gatattggca acactgtcaa gcaccagtac cggtgtcacc gaatcgccga 4560

gtggtccgat atcaccaacg cccacggtgt acccggaacc ggaatcgatg cgtatctgtg 4620

ggacatgtgg tcgttgcgcc attatgtaag cagcgtgtac tcctctgact gtccatatgg 4680

tttgctccat ctcaccctca tcgttttcat tgttcacagg cggccacaaa aaaactgtct 4740

tctctccttc tctcttcgcc ttagtctact cggaccagtt ttagtttagc ttggcgccac 4800

tggataaatg agacctcagg ccttgtgatg aggaggtcac ttatgaagca tgttaggagg 4860

tgcttgtatg gatagagaag cacccaaaat aataagaata ataataaaac agggggcgtt 4920

gtcatttcat atcgtgtttt caccatcaat acacctccaa acaatgccct tcatgtggcc 4980

agccccaata ttgtcctgta gttcaactct atgcagctcg tatcttattg agcaagtaaa 5040

actctgtcag ccgatattgc ccgacccgcg acaagggtca acaaggtggt gtaaggcctt 5100

cgcagaagtc aaaactgtgc caaacaaaca tctagagtct ctttggtgtt tctcgcatat 5160

atttwatcgg ctgtcttacg tatttgcgcc tcggtaccgg actaatttcg gatcatcccc 5220

aatacgcttt ttcttcgcag ctgtcaacag tgtccatgat ctatccacct aaatgggtca 5280

tatgaggcgt ataatttcgt ggtgctgata ataattccca tatatttgac acaaaacttc 5340

ccccccctaga catacatctc acaatctcac ttcttgtgct tctgtcacac atctcctcca 5400

gctgacttca actcacacct ctgccccagt tggtctacag cggtataagg tttctccgca 5460

tagaggtgca ccactcctcc cgatacttgt ttgtgtgact tgtgggtcac gacatatata 5520

tctacacaca ttgcgccacc ctttggttct tccagcacaa caaaaacacg acacgctaac 5580

catggtctcg tcaatctcgc ccgacctata ctcgtccttc tcgttcaaaa actcgctcgc 5640

cgaggccatg ccctccgtgc cacacgaact catcaactca aaaacactct catggatgta 5700

caatgcctct ctggacattc gggttcctct gactatcgga accatctacg ccgtctccgt 5760

gcacctgacc aactcatctg aacgaatcaa gaaacgccag cccattgcct ttgccaagac 5820

cgcactcttc aagtggctct gtgtcctcca caatgcaggt ctgtgtctct actcagcatg 5880

gacctttgtc ggtatcctca acgccgtcaa acacgcctac caaatcacag gagacagctc 5940

cgccccttc tccttcaaca ccctctgggg atcgttttgt tcacgtgact ccctctgggt 6000

caccggcctc aactactacg gatactggtt ctatctgtcc aaattctacg aagtggtgga 6060

caccatgatc atcctcgcaa agggaaaacc gtcctcaatg ctccagacat accaccacac 6120

cggcgccatg ttctccatgt gggccggcat ccgattcgcc tctcccccca tctggatctt 6180

tgtggttttc aactccctca tccacacaat catgtacttt tactacaccc tcaccaccct 6240

caagatcaag gttcccaaga tcctcaaggc atctctgacc accgcccaga tcacccagat 6300

tgtcggaggt ggcatcctgg ctgcctccca cgcctttatt tattacaagg accaccagac 6360

tgagaccgtc tgttcttgtc tcactaccca gggtcagttt ttcgctctcg ccgtcaatgt 6420

catctatctg agtcctctgg cctatctctt tattgccttc tggattcgat cttacttgaa 6480

ggccaagtcc aactaagc 6498

<210> 201

<211> 10195
<212> DNA
<213> Artificial Sequence

<220>
<223> Plasmid pZUF6TYE2

<400> 201
ggccgcatgg agcgtgtgtt ctgagtcgat gttttctatg gagttgtgag tgttagtaga 60

catgatgggt ttatatatga tgaatgaata gatgtgattt tgatttgcac gatggaattg 120

agaactttgt aaacgtacat gggaatgtat gaatgtgggg gttttgtgac tggataactg 180

acggtcagtg gacgccgttg ttcaaatatc caagagatgc gagaaacttt gggtcaagtg 240

aacatgtcct ctctgttcaa gtaaaccatc aactatgggt agtatattta gtaaggacaa 300

gagttgagat tctttggagt cctagaaacg tattttcgcg ttccaagatc aaattagtag 360

agtaatacgg gcacgggaat ccattcatag tctcaatcct gcaggtgagt taattaagtc 420

atacacaagt cagctttctt cgagcctcat ataagtataa gtagttcaac gtattagcac 480

tgtacccagc atctccgtat cgagaaacac aacaacatgc cccattggac agatcatgcg 540

gatacacagg ttgtgcagta tcatacatac tcgatcagac aggtcgtctg accatcatac 600

aagctgaaca agcgctccat acttgcacgc tctctatata cacagttaaa ttacatatcc 660

atagtctaac ctctaacagt taatcttctg gtaagcctcc cagccagcct tctggtatcg 720

cttggcctcc tcaataggat ctcggttctg gccgtacaga cctcggccga caattatgat 780

atccgttccg gtagacatga catcctcaac agttcggtac tgctgtccga gagcgtctcc 840

cttgtcgtca agacccaccc cgggggtcag aataagccag tcctcagagt cgcccttagg 900

tcggttctgg gcaatgaagc caaccacaaa ctcggggtcg gatcgggcaa gctcaatggt 960

ctgcttggag tactcgccag tggccagaga gcccttgcaa gacagctcgg ccagcatgag 1020

cagacctctg gccagcttct cgttgggaga ggggactagg aactccttgt actgggagtt 1080

ctcgtagtca gagacgtcct ccttcttctg ttcagagaca gtttcctcgg caccagctcg 1140

caggccagca atgattccgg ttccgggtac accgtgggcg ttggtgatat cggaccactc 1200

ggcgattcgg tgacaccggt actggtgctt gacagtgttg ccaatatctg cgaactttct 1260

gtcctcgaac aggaagaaac cgtgcttaag agcaagttcc ttgaggggga gcacagtgcc 1320

ggcgtaggtg aagtcgtcaa tgatgtcgat atgggttttg atcatgcaca cataaggtcc 1380

gaccttatcg gcaagctcaa tgagctcctt ggtggtggta acatccagag aagcacacag 1440

gttggttttc ttggctgcca cgagcttgag cactcgagcg gcaaaggcgg acttgtggac 1500

gttagctcga gcttcgtagg agggcatttt ggtggtgaag aggagactga aataaattta 1560

gtctgcagaa ctttttatcg gaaccttatc tggggcagtg aagtatatgt tatggtaata 1620

gttacgagtt agttgaactt atagatagac tggactatac ggctatcggt ccaaattaga 1680

aagaacgtca atggctctct gggcgtcgcc tttgccgaca aaaatgtgat catgatgaaa 1740

gccagcaatg acgttgcagc tgatattgtt gtcggccaac cgcgccgaaa acgcagctgt 1800

cagacccaca gcctccaacg aagaatgtat cgtcaaagtg atccaagcac actcatagtt 1860

ggagtcgtac tccaaaggcg gcaatgacga gtcagacaga tactcgtcga ctcaggcgac 1920

gacggaattc ctgcagccca tctgcagaat tcaggagaga ccgggttggc ggcgtatttg 1980

tgtcccaaaa aacagcccca attgccccgg agaagacggc caggccgcct agatgacaaa 2040

ttcaacaact cacagctgac tttctgccat tgccactagg ggggggcctt tttatatggc 2100

caagccaagc tctccacgtc ggttgggctg cacccaacaa taaatgggta gggttgcacc 2160

aacaaaggga tgggatgggg ggtagaagat acgaggataa cggggctcaa tggcacaaat 2220

aagaacgaat actgccatta agactcgtga tccagcgact gacaccattg catcatctaa 2280

gggcctcaaa actacctcgg aactgctgcg ctgatctgga caccacagag gttccgagca 2340

ctttaggttg caccaaatgt cccaccaggt gcaggcagaa aacgctggaa cagcgtgtac 2400

agtttgtctt aacaaaaagt gagggcgctg aggtcgagca gggtggtgtg acttgttata 2460

gcctttagag ctgcgaaagc gcgtatggat ttggctcatc aggccagatt gagggtctgt 2520

ggacacatgt catgttagtg tacttcaatc gccccctgga tatagccccg acaataggcc 2580

gtggcctcat ttttttgcct tccgcacatt tccattgctc ggtacccaca ccttgcttct 2640

cctgcacttg ccaaccttaa tactggttta cattgaccaa catcttacaa gcgggggggct 2700

tgtctagggt atatataaac agtggctctc ccaatcggtt gccagtctct tttttccttt 2760

ctttccccac agattcgaaa tctaaactac acatcacaca atgcctgtta ctgacgtcct 2820

taagcgaaag tccggtgtca tcgtcggcga cgatgtccga gccgtgagta tccacgacaa 2880

gatcagtgtc gagacgacgc gttttgtgta atgacacaat ccgaaagtcg ctagcaacac 2940

acactctcta cacaaactaa cccagctctc catggctgcc gctccctctg tgcgaacctt 3000

tacccgagcc gaggttctga acgctgaggc tctgaacgag ggcaagaagg acgctgaggc 3060

tcccttcctg atgatcatcg acaacaaggt gtacgacgtc cgagagttcg tccctgacca 3120

tcctggaggc tccgtgattc tcacccacgt tggcaaggac ggcaccgacg tctttgacac 3180

ctttcatccc gaggctgctt gggagactct cgccaacttc tacgttggag acattgacga 3240

gtccgaccga gacatcaaga acgatgactt tgccgctgag gtccgaaagc tgcgaaccct 3300

gttccagtct ctcggctact acgactcctc taaggcctac tacgccttca aggtctcctt 3360

caacctctgc atctggggac tgtccaccgt cattgtggcc aagtggggtc agacctccac 3420

cctcgccaac gtgctctctg ctgccctgct cggcctgttc tggcagcagt gcggatggct 3480

ggctcacgac tttctgcacc accaggtctt ccaggaccga ttctggggtg atctcttcgg 3540

agccttcctg ggaggtgtct gccagggctt ctcctcttcc tggtggaagg acaagcacaa 3600

cactcaccat gccgctccca acgtgcatgg cgaggatcct gacattgaca cccaccctct 3660

cctgacctgg tccgagcacg ctctggagat gttctccgac gtccccgatg aggagctgac 3720

ccgaatgtgg tctcgattca tggtcctgaa ccagacctgg ttctacttcc ccattctctc 3780

cttcgctcga ctgtcttggt gcctccagtc cattctcttt gtgctgccca acggtcaggc 3840

tcacaagccc tccggagctc gagtgcccat ctccctggtc gagcagctgt ccctcgccat 3900

gcactggacc tggtacctcg ctaccatgtt cctgttcatc aaggatcctg tcaacatgct 3960

cgtgtacttc ctggtgtctc aggctgtgtg cggaaacctg ctcgccatcg tgttctccct 4020

caaccacaac ggtatgcctg tgatctccaa ggaggaggct gtcgacatgg atttctttac 4080

caagcagatc atcactggtc gagatgtcca tcctggactg ttcgccaact ggttcaccgg 4140

tggcctgaac taccagatcg agcatcacct gttcccttcc atgcctcgac acaacttctc 4200

caagatccag cctgccgtcg agaccctgtg caagaagtac aacgtccgat accacaccac 4260

tggtatgatc gagggaactg ccgaggtctt ctcccgactg aacgaggtct ccaaggccac 4320

ctccaagatg ggcaaggctc agtaagcggc cgcaagtgtg gatggggaag tgagtgcccg 4380

gttctgtgtg cacaattggc aatccaagat ggatggattc aacacaggga tatagcgagc 4440

tacgtggtgg tgcgaggata tagcaacgga tatttatgtt tgacacttga gaatgtacga 4500

tacaagcact gtccaagtac aatactaaac atactgtaca tactcatact cgtacccggg 4560

caacggtttc acttgagtgc agtggctagt gctcttactc gtacagtgtg caatactgcg 4620

tatcatagtc tttgatgtat atcgtattca ttcatgttag ttgcgtacga gccggaagca 4680

taaagtgtaa agcctggggt gcctaatgag tgagctaact cacattaatt gcgttgcgct 4740

cactgcccgc tttccagtcg ggaaacctgt cgtgccagct gcattaatga atcggccaac 4800

gcgcggggag aggcggtttg cgtattgggc gctcttccgc ttcctcgctc actgactcgc 4860

tgcgctcggt cgttcggctg cggcgagcgg tatcagctca ctcaaaggcg gtaatacggt 4920

tatccacaga atcaggggat aacgcaggaa agaacatgtg agcaaaaggc cagcaaaagg 4980

ccaggaaccg taaaaaggcc gcgttgctgg cgtttttcca taggctccgc ccccctgacg 5040

agcatcacaa aaatcgacgc tcaagtcaga ggtggcgaaa cccgacagga ctataaagat 5100

accaggcgtt tccccctgga agctccctcg tgcgctctcc tgttccgacc ctgccgctta 5160

ccggatacct gtccgccttt ctcccttcgg gaagcgtggc gctttctcat agctcacgct 5220

gtaggtatct cagttcggtg taggtcgttc gctccaagct gggctgtgtg cacgaacccc 5280

ccgttcagcc cgaccgctgc gccttatccg gtaactatcg tcttgagtcc aacccggtaa 5340

gacacgactt atcgccactg gcagcagcca ctggtaacag gattagcaga gcgaggtatg 5400

taggcggtgc tacagagttc ttgaagtggt ggcctaacta cggctacact agaaggacag 5460

tatttggtat ctgcgctctg ctgaagccag ttaccttcgg aaaaagagtt ggtagctctt 5520

gatccggcaa acaaaccacc gctggtagcg gtggtttttt tgtttgcaag cagcagatta 5580

cgcgcagaaa aaaaggatct caagaagatc ctttgatctt ttctacgggg tctgacgctc 5640

agtggaacga aaactcacgt taagggattt tggtcatgag attatcaaaa aggatcttca 5700

cctagatcct tttaaattaa aaatgaagtt ttaaatcaat ctaaagtata tatgagtaaa 5760

cttggtctga cagttaccaa tgcttaatca gtgaggcacc tatctcagcg atctgtctat 5820

ttcgttcatc catagttgcc tgactccccg tcgtgtagat aactacgata cgggagggct 5880

taccatctgg ccccagtgct gcaatgatac cgcgagaccc acgctcaccg gctccagatt 5940

tatcagcaat aaaccagcca gccggaaggg ccgagcgcag aagtggtcct gcaactttat 6000

ccgcctccat ccagtctatt aattgttgcc gggaagctag agtaagtagt tcgccagtta 6060

atagtttgcg caacgttgtt gccattgcta caggcatcgt ggtgtcacgc tcgtcgtttg 6120

gtatggcttc attcagctcc ggttcccaac gatcaaggcg agttacatga tcccccatgt 6180

tgtgcaaaaa agcggttagc tccttcggtc ctccgatcgt tgtcagaagt aagttggccg 6240

cagtgttatc actcatggtt atggcagcac tgcataattc tcttactgtc atgccatccg 6300

taagatgctt ttctgtgact ggtgagtact caaccaagtc attctgagaa tagtgtatgc 6360

ggcgaccgag ttgctcttgc ccggcgtcaa tacgggataa taccgcgcca catagcagaa 6420

ctttaaaagt gctcatcatt ggaaaacgtt cttcggggcg aaaactctca aggatcttac 6480

cgctgttgag atccagttcg atgtaaccca ctcgtgcacc caactgatct tcagcatctt 6540

ttactttcac cagcgtttct gggtgagcaa aaacaggaag gcaaaatgcc gcaaaaaagg 6600

gaataagggc gacacggaaa tgttgaatac tcatactctt cctttttcaa tattattgaa 6660

gcatttatca gggttattgt ctcatgagcg gatacatatt tgaatgtatt tagaaaaata 6720

aacaaatagg ggttccgcgc acatttcccc gaaaagtgcc acctgacgcg ccctgtagcg 6780

gcgcattaag cgcggcgggt gtggtggtta cgcgcagcgt gaccgctaca cttgccagcg 6840

ccctagcgcc cgctcctttc gctttcttcc cttcctttct cgccacgttc gccggctttc 6900

cccgtcaagc tctaaatcgg gggctccctt tagggttccg atttagtgct ttacggcacc 6960

tcgaccccaa aaaacttgat tagggtgatg gttcacgtag tgggccatcg ccctgataga 7020

cggtttttcg ccctttgacg ttggagtcca cgttctttaa tagtggactc ttgttccaaa 7080

ctggaacaac actcaaccct atctcggtct attcttttga tttataaggg attttgccga 7140

tttcggccta ttggttaaaa aatgagctga tttaacaaaa atttaacgcg aattttaaca 7200

aaatattaac gcttacaatt tccattcgcc attcaggctg cgcaactgtt gggaagggcg 7260

atcggtgcgg gcctcttcgc tattacgcca gctggcgaaa gggggatgtg ctgcaaggcg 7320

attaagttgg gtaacgccag ggttttccca gtcacgacgt tgtaaaacga cggccagtga 7380

attgtaatac gactcactat agggcgaatt gggtaccggg ccccccctcg aggtcgatgg 7440

tgtcgataag cttgatatcg aattcatgtc acacaaaccg atcttcgcct caaggaaacc 7500

taattctaca tccgagagac tgccgagatc cagtctacac tgattaattt tcgggccaat 7560

aatttaaaaa aatcgtgtta tataatatta tatgtattat atatatacat catgatgata 7620

ctgacagtca tgtcccattg ctaaatagac agactccatc tgccgcctcc aactgatgtt 7680

ctcaatattt aaggggtcat ctcgcattgt ttaataataa acagactcca tctaccgcct 7740

ccaaatgatg ttctcaaaat atattgtatg aacttatttt tattacttag tattattaga 7800

caacttactt gctttatgaa aaacacttcc tatttaggaa acaatttata atggcagttc 7860

gttcatttaa caatttatgt agaataaatg ttataaatgc gtatgggaaa tcttaaatat 7920

ggatagcata aatgatatct gcattgccta attcgaaatc aacagcaacg aaaaaaatcc 7980

cttgtacaac ataaatagtc atcgagaaat atcaactatc aaagaacagc tattcacacg 8040

ttactattga gattattatt ggacgagaat cacacactca actgtctttc tctcttctag 8100

aaatacaggt acaagtatgt actattctca ttgttcatac ttctagtcat ttcatcccac 8160

atattccttg gatttctctc caatgaatga cattctatct tgcaaattca acaattataa 8220

taagatatac caaagtagcg gtatagtggc aatcaaaaag cttctctggt gtgcttctcg 8280

tatttatttt tattctaatg atccattaaa ggtatatatt tatttcttgt tatataatcc 8340

ttttgtttat tacatgggct ggatacataa aggtattttg atttaatttt ttgcttaaat 8400

tcaatccccc ctcgttcagt gtcaactgta atggtaggaa attaccatac ttttgaagaa 8460

gcaaaaaaaa tgaaagaaaa aaaaaatcgt atttccaggt tagacgttcc gcagaatcta 8520

gaatgcggta tgcggtacat tgttcttcga acgtaaaagt tgcgctccct gagatattgt 8580

acatttttgc ttttacaagt acaagtacat cgtacaacta tgtactactg ttgatgcatc 8640

cacaacagtt tgttttgttt ttttttgttt ttttttttttc taatgattca ttaccgctat 8700

gtatacctac ttgtacttgt agtaagccgg gttattggcg ttcaattaat catagactta 8760

tgaatctgca cggtgtgcgc tgcgagttac ttttagctta tgcatgctac ttgggtgtaa 8820

tattgggatc tgttcggaaa tcaacggatg ctcaatcgat gcagaattca ggagagaccg 8880

ggttggcggc gtatttgtgt cccaaaaaac agccccaatt gccccaattg accccaaatt 8940

gacccagtag cgggcccaac cccggcgaga gccccccttca ccccacatat caaacctccc 9000

ccggttccca cacttgccgt taagggcgta gggtactgca gtctggaatc tacgcttgtt 9060

cagactttgt actagtttct ttgtctggcc atccgggtaa cccatgccgg acgcaaaata 9120

gactactgaa aatttttttg ctttgtggtt gggactttag ccaagggtat aaaagaccac 9180

cgtccccgaa ttacctttcc tcttcttttc tctctctcct tgtcaactca cacccgaaat 9240

cgttaagcat ttccttctga gtataagaat cattcaccat ggtctcgtca atctcgcccg 9300

acctatactc gtccttctcg ttcaaaaact cgctcgccga ggccatgccc tccgtgccac 9360

acgaactcat caactcaaaa acactctcat ggatgtacaa tgcctctctg gacattcggg 9420

ttcctctgac tatcggaacc atctacgccg tctccgtgca cctgaccaac tcatctgaac 9480

gaatcaagaa acgccagccc attgcctttg ccaagaccgc actcttcaag tggctctgtg 9540

tcctccacaa tgcaggtctg tgtctctact cagcatggac ctttgtcggt atcctcaacg 9600

ccgtcaaaca cgcctaccaa atcacaggag acagctccgc ccccttctcc ttcaacaccc 9660

tctggggatc gttttgttca cgtgactccc tctgggtcac cggcctcaac tactacggat 9720

actggttcta tctgtccaaa ttctacgaag tggtggacac catgatcatc ctcgcaaagg 9780

gaaaaccgtc ctcaatgctc cagacatacc accacaccgg cgccatgttc tccatgtggg 9840

ccggcatccg attcgcctct ccccccatct ggatctttgt ggttttcaac tccctcatcc 9900

acacaatcat gtacttttac tacaccctca ccaccctcaa gatcaaggtt cccaagatcc 9960

tcaaggcatc tctgaccacc gcccagatca cccagattgt cggaggtggc atcctggctg 10020

cctcccacgc ctttatttat tacaaggacc accagactga gaccgtctgt tcttgtctca 10080

ctacccaggg tcagtttttc gctctcgccg tcaatgtcat ctatctgagt cctctggcct 10140

atctctttat tgccttctgg attcgatctt acttgaaggc caagtccaac taagc 10195

<210> 202
<211> 6561
<212> DNA
<213> Artificial Sequence

<220>
<223> Plasmid pKUGPYE1

<400> 202
gtacgaggaa actgtctctg aacagaagaa ggaggacgtc tctgactacg agaactccca 60

gtacaaggag ttcctagtcc cctctcccaa cgagaagctg gccagaggtc tgctcatgct 120

ggccgagctg tcttgcaagg gctctctggc cactggcgag tactccaagc agaccattga 180

gcttgcccga tccgaccccg agtttgtggt tggcttcatt gcccagaacc gacctaaggg 240

cgactctgag gactggctta ttctgacccc cggggtgggt cttgacgaca agggagacgc 300

tctcggacag cagtaccgaa ctgttgagga tgtcatgtct accggaacgg atatcataat 360

tgtcggccga ggtctgtacg gccagaaccg agatcctatt gaggaggcca agcgatacca 420

gaaggctggc tgggaggctt accagaagat taactgttag aggttagact atggatatgt 480

aatttaactg tgtatataga gagcgtgcaa gtatggagcg cttgttcagc ttgtatgatg 540

gtcagacgac ctgtctgatc gagtatgtat gatactgcac aacctgtgta tccgcatgat 600

ctgtccaatg gggcatgttg ttgtgtttct cgatacggag atgctgggta cagtgctaat 660

acgttgaact acttatactt atatgaggct cgaagaaagc tgacttgtgt atgacttaat 720

taatcgagct tggcgtaatc atggtcatag ctgtttcctg tgtgaaattg ttatccgctc 780

acaattccac acaacatacg agccggaagc ataaagtgta aagcctgggg tgcctaatga 840

gtgagctaac tcacattaat tgcgttgcgc tcactgcccg ctttccagtc gggaaacctg 900

tcgtgccagc tgcattaatg aatcggccaa cgcgcgggga gaggcggttt gcgtattggg 960

cgctcttccg cttcctcgct cactgactcg ctgcgctcgg tcgttcggct gcggcgagcg 1020

gtatcagctc actcaaaggc ggtaatacgg ttatccacag aatcagggga taacgcagga 1080

aagaacatgt gagcaaaagg ccagcaaaag gccaggaacc gtaaaaaggc cgcgttgctg 1140

gcgtttttcc ataggctccg ccccccctgac gagcatcaca aaaatcgacg ctcaagtcag 1200

aggtggcgaa acccgacagg actataaaga taccaggcgt ttcccccctgg aagctccctc 1260

gtgcgctctc ctgttccgac cctgccgctt accggatacc tgtccgcctt tctcccttcg 1320

ggaagcgtgg cgctttctca tagctcacgc tgtaggtatc tcagttcggt gtaggtcgtt 1380

cgctccaagc tgggctgtgt gcacgaaccc cccgttcagc ccgaccgctg cgccttatcc 1440

ggtaactatc gtcttgagtc caacccggta agacacgact tatcgccact ggcagcagcc 1500

actggtaaca ggattagcag agcgaggtat gtaggcggtg ctacagagtt cttgaagtgg 1560

tggcctaact acggctacac tagaaggaca gtatttggta tctgcgctct gctgaagcca 1620

gttaccttcg gaaaaagagt tggtagctct tgatccggca aacaaaccac cgctggtagc 1680

ggtggttttt ttgtttgcaa gcagcagatt acgcgcagaa aaaaaggatc tcaagaagat 1740

cctttgatct tttctacggg gtctgacgct cagtggaacg aaaactcacg ttaagggatt 1800

ttggtcatga gattatcaaa aaggatcttc acctagatcc ttttaaatta aaaatgaagt 1860

tttaaatcaa tctaaagtat atatgagtaa acttggtctg acagttacca atgcttaatc 1920

agtgaggcac ctatctcagc gatctgtcta tttcgttcat ccatagttgc ctgactcccc 1980

gtcgtgtaga taactacgat acgggagggc ttaccatctg gccccagtgc tgcaatgata 2040

ccgcgagacc cacgctcacc ggctccagat ttatcagcaa taaaccagcc agccggaagg 2100

gccgagcgca gaagtggtcc tgcaacttta tccgcctcca tccagtctat taattgttgc 2160

cgggaagcta gagtaagtag ttcgccagtt aatagtttgc gcaacgttgt tgccattgct 2220

acaggcatcg tggtgtcacg ctcgtcgttt ggtatggctt cattcagctc cggttcccaa 2280

cgatcaaggc gagttacatg atcccccatg ttgtgcaaaa aagcggttag ctccttcggt 2340

cctccgatcg ttgtcagaag taagttggcc gcagtgttat cactcatggt tatggcagca 2400

ctgcataatt ctcttactgt catgccatcc gtaagatgct tttctgtgac tggtgagtac 2460

tcaaccaagt cattctgaga atagtgtatg cggcgaccga gttgctcttg cccggcgtca 2520

atacgggata ataccgcgcc acatagcaga actttaaaag tgctcatcat tggaaaacgt 2580

tcttcggggc gaaaactctc aaggatctta ccgctgttga gatccagttc gatgtaaccc 2640

actcgtgcac ccaactgatc ttcagcatct tttactttca ccagcgtttc tgggtgagca 2700

aaaacaggaa ggcaaaatgc cgcaaaaaag ggaataaggg cgacacggaa atgttgaata 2760

ctcatactct tccttttttca atattattga agcatttatc agggttattg tctcatgagc 2820

ggatacatat ttgaatgtat ttagaaaaat aaacaaatag gggttccgcg cacatttccc 2880

cgaaaagtgc cacctgacgc gccctgtagc ggcgcattaa gcgcggcggg tgtggtggtt 2940

acgcgcagcg tgaccgctac acttgccagc gccctagcgc ccgctccttt cgctttcttc 3000

ccttcctttc tcgccacgtt cgccggcttt ccccgtcaag ctctaaatcg ggggctccct 3060

ttagggttcc gatttagtgc tttacggcac ctcgacccca aaaaacttga ttagggtgat 3120

ggttcacgta gtgggccatc gccctgatag acggtttttc gccctttgac gttggagtcc 3180

acgttcttta atagtggact cttgttccaa actggaacaa cactcaaccc tatctcggtc 3240

tattcttttg atttataagg gattttgccg atttcggcct attggttaaa aaatgagctg 3300

atttaacaaa aatttaacgc gaattttaac aaaatattaa cgcttacaat ttccattcgc 3360

cattcaggct gcgcaactgt tgggaagggc gatcggtgcg ggcctcttcg ctattacgcc 3420

agctggcgaa aggggatgt gctgcaaggc gattaagttg ggtaacgcca gggttttccc 3480

agtcacgacg ttgtaaaacg acggccagtg aattgtaata cgactcacta tagggcgaat 3540

tgggtaccgg gccccccctc gaggtcgacg agtatctgtc tgactcgtca ttgccgcctt 3600

tggagtacga ctccaactat gagtgtgctt ggatcacttt gacgatacat tcttcgttgg 3660

aggctgtggg tctgacagct gcgttttcgg cgcggttggc cgacaacaat atcagctgca 3720

acgtcattgc tggctttcat catgatcaca tttttgtcgg caaaggcgac gcccagagag 3780

ccattgacgt tctttctaat ttggaccgat agccgtatag tccagtctat ctataagttc 3840

aactaactcg taactattac cataacatat acttcactgc cccagataag gttccgataa 3900

aaagttctgc agactaaatt tatttcagtc tcctcttcac caccaaaatg ccctcctacg 3960

aagctcgagt gctcaagctc gtggcagcca agaaaaccaa cctgtgtgct tctctggatg 4020

ttaccaccac caaggagctc attgagcttg ccgataaggt cggaccttat gtgtgcatga 4080

tcaaaaccca tatcgacatc attgacgact tcacctacgc cggcactgtg ctccccctca 4140

aggaacttgc tcttaagcac ggtttcttcc tgttcgagga cagaaagttc gcagatattg 4200

gcaacactgt caagcaccag taccggtgtc accgaatcgc cgagtggtcc gatatcacca 4260

acgcccacgg tgtacccgga accggaatcg atgcgtatct gtgggacatg tggtcgttgc 4320

gccattatgt aagcagcgtg tactcctctg actgtccata tggtttgctc catctcaccc 4380

tcatcgtttt cattgttcac aggcggccac aaaaaaactg tcttctctcc ttctctcttc 4440

gccttagtct actcggacca gttttagttt agcttggcgc cactggataa atgagacctc 4500

aggccttgtg atgaggaggt cacttatgaa gcatgttagg aggtgcttgt atggatagag 4560

aagcacccaa aataataaga ataataataa aacaggggggc gttgtcattt catatcgtgt 4620

tttcaccatc aatacacctc caaacaatgc ccttcatgtg gccagcccca atattgtcct 4680

gtagttcaac tctatgcagc tcgtatctta ttgagcaagt aaaactctgt cagccgatat 4740

tgcccgaccc gcgacaaggg tcaacaaggt ggtgtaaggc cttcgcagaa gtcaaaactg 4800

tgccaaacaa acatctagag tctctttggt gtttctcgca tatatttwat cggctgtctt 4860

acgtatttgc gcctcggtac cggactaatt tcggatcatc cccaatacgc tttttcttcg 4920

cagctgtcaa cagtgtccat gatctatcca cctaaatggg tcatatgagg cgtataattt 4980

cgtggtgctg ataataattc ccatatattt gacacaaaac ttccccccct agacatacat 5040

ctcacaatct cacttcttgt gcttctgtca cacatctcct ccagctgact tcaactcaca 5100

cctctgcccc agttggtcta cagcggtata aggtttctcc gcatagaggt gcaccactcc 5160

tcccgatact tgtttgtgtg acttgtgggt cacgacatat atatctacac acattgcgcc 5220

acccttttggt tcttccagca caacaaaaac acgacacgct aaccatggct gccgtcccta 5280

ttgaattcaa cgtcccctcc gtggaccgac cctttggtat ctacctctgg gccatctttg 5340

accaggcctg ggagaagctt ttcggctggc ccgcgtcctc tttcattttc gtgcgaaatg 5400

accccaacat ccccttttcc tctacccctc ccgtgatcat tgccatcatt gtgtactaca 5460

ttgtcatctt tggcggccga gaggtgatgc gaaacctgtc tcccatccga ctcaactggc 5520

tcttccagat ccacaacatc ttcctcaccc ttctgtccgg tatgctcctc ctcctcctcg 5580

ttgagcagct cttccccatc attgtccgac agggtatcct ctacgccatc tgcgactacg 5640

gatcttggac tcagcccatt gtcttctgct actacctcaa ctacctgacc aagtactttg 5700

agctgatcga caccgttttc cttgtgctgc gaaagaagaa gctgactttc ctccacacct 5760

accaccatgg tgccactgct cttctgtgct acacccagct cattggtaag acctcggtct 5820

cttgggtccc catcaccctt aacctgtttg tccacgttgt catgtacttc tactacttcc 5880

tggctgcgcg aggtatccga gtgtggtgga aggagtgggt cacccggctc cagatcatcc 5940

agttcgttat cgatcttgga tttgtctact ttgcctctta cacctacttc acctctacct 6000

actggccctg gatgcccaac atgggctctt gtgccggcga ggagtttgct gctatttacg 6060

gctgtggtct gctgacctct tacctcttcc tcttcatcgc cttctacatc aactcttacc 6120

gaaagccctc ttccaaggga ccttccaagc ctgttgttgc tgtcgatggc cctgttggcg 6180

gcgtcaacgc ccagactggt gcttctcgag gccagaccac tacccgatct cgacgagcat 6240

aagcggccgc aagtgtggat ggggaagtga gtgcccggtt ctgtgtgcac aattggcaat 6300

ccaagatgga tggattcaac acagggatat agcgagctac gtggtggtgc gaggatatag 6360

caacggatat ttatgtttga cacttgagaa tgtacgatac aagcactgtc caagtacaat 6420

actaaacata ctgtacatac tcatactcgt acccgggcaa cggtttcact tgagtgcagt 6480

ggctagtgct cttactcgta cagtgtgcaa tactgcgtat catagtcttt gatgtatatc 6540

gtattcattc atgttagttg c 6561

<210> 203
<211> 10809
<212> DNA
<213> Artificial Sequence

<220>
<223> Plasmid pZUF6FYE1

<400> 203
taagtcatac acaagtcagc tttcttcgag cctcatataa gtataagtag ttcaacgtat 60

tagcactgta cccagcatct ccgtatcgag aaacacaaca acatgcccca ttggacagat 120

catgcggata cacaggttgt gcagtatcat acatactcga tcagacaggt cgtctgacca 180

tcatacaagc tgaacaagcg ctccatactt gcacgctctc tatatacaca gttaaattac 240

atatccatag tctaacctct aacagttaat cttctggtaa gcctcccagc cagccttctg 300

gtatcgcttg gcctcctcaa taggatctcg gttctggccg tacagacctc ggccgacaat 360

tatgatatcc gttccggtag acatgacatc ctcaacagtt cggtactgct gtccgagagc 420

gtctcccttg tcgtcaagac ccaccccggg ggtcagaata agccagtcct cagagtcgcc 480

cttaggtcgg ttctgggcaa tgaagccaac cacaaactcg gggtcggatc gggcaagctc 540

aatggtctgc ttggagtact cgccagtggc cagagagccc ttgcaagaca gctcggccag 600

catgagcaga cctctggcca gcttctcgtt gggagagggg actaggaact ccttgtactg 660

ggagttctcg tagtcagaga cgtcctcctt cttctgttca gagacagttt cctcggcacc 720

agctcgcagg ccagcaatga ttccggttcc gggtacaccg tgggcgttgg tgatatcgga 780

ccactcggcg attcggtgac accggtactg gtgcttgaca gtgttgccaa tatctgcgaa 840

ctttctgtcc tcgaacagga agaaaccgtg cttaagagca agttccttga gggggagcac 900

agtgccggcg taggtgaagt cgtcaatgat gtcgatatgg gttttgatca tgcacacata 960

aggtccgacc ttatcggcaa gctcaatgag ctccttggtg gtggtaacat ccagagaagc 1020

acacaggttg gttttcttgg ctgccacgag cttgagcact cgagcggcaa aggcggactt 1080

gtggacgtta gctcgagctt cgtaggaggg cattttggtg gtgaagagga gactgaaata 1140

aatttagtct gcagaacttt ttatcggaac cttatctggg gcagtgaagt atatgttatg 1200

gtaatagtta cgagttagtt gaacttatag atagactgga ctatacggct atcggtccaa 1260

attagaaaga acgtcaatgg ctctctgggc gtcgcctttg ccgacaaaaa tgtgatcatg 1320

atgaaagcca gcaatgacgt tgcagctgat attgttgtcg gccaaccgcg ccgaaaacgc 1380

agctgtcaga cccacagcct ccaacgaaga atgtatcgtc aaagtgatcc aagcacactc 1440

atagttggag tcgtactcca aaggcggcaa tgacgagtca gacagatact cgtcgactca 1500

ggcgacgacg gaattcctgc agcccatctg cagaattcag gagagaccgg gttggcggcg 1560

tatttgtgtc ccaaaaaaca gccccaattg ccccggagaa gacggccagg ccgcctagat 1620

gacaaattca acaactcaca gctgactttc tgccattgcc actaggggggg ggccttttta 1680

tatggccaag ccaagctctc cacgtcggtt gggctgcacc caacaataaa tgggtagggt 1740

tgcaccaaca aagggatggg atgggggggta gaagatacga ggataacggg gctcaatggc 1800

acaaataaga acgaatactg ccattaagac tcgtgatcca gcgactgaca ccattgcatc 1860

atctaagggc ctcaaaacta cctcggaact gctgcgctga tctggacacc acagaggttc 1920

cgagcacttt aggttgcacc aaatgtccca ccaggtgcag gcagaaaacg ctggaacagc 1980

gtgtacagtt tgtcttaaca aaaagtgagg gcgctgaggt cgagcagggt ggtgtgactt 2040

gttatagcct ttagagctgc gaaagcgcgt atggatttgg ctcatcaggc cagattgagg 2100

gtctgtggac acatgtcatg ttagtgtact tcaatcgccc cctggatata gccccgacaa 2160

taggccgtgg cctcattttt ttgccttccg cacatttcca ttgctcggta cccacacctt 2220

gcttctcctg cacttgccaa ccttaatact ggtttacatt gaccaacatc ttacaagcgg 2280

ggggcttgtc tagggtatat ataaacagtg gctctcccaa tcggttgcca gtctcttttt 2340

tcctttcttt ccccacagat tcgaaatcta aactacacat cacacaatgc ctgttactga 2400

cgtccttaag cgaaagtccg gtgtcatcgt cggcgacgat gtccgagccg tgagtatcca 2460

cgacaagatc agtgtcgaga cgacgcgttt tgtgtaatga cacaatccga aagtcgctag 2520

caacacacac tctctacaca aactaaccca gctctccatg gctgccgctc cctctgtgcg 2580

aacctttacc cgagccgagg ttctgaacgc tgaggctctg aacgagggca agaaggacgc 2640

tgaggctccc ttcctgatga tcatcgacaa caaggtgtac gacgtccgag agttcgtccc 2700

tgaccatcct ggaggctccg tgattctcac ccacgttggc aaggacggca ccgacgtctt 2760

tgacaccttt catcccgagg ctgcttggga gactctcgcc aacttctacg ttggagacat 2820

tgacgagtcc gaccgagaca tcaagaacga tgactttgcc gctgaggtcc gaaagctgcg 2880

aaccctgttc cagtctctcg gctactacga ctcctctaag gcctactacg ccttcaaggt 2940

ctccttcaac ctctgcatct ggggactgtc caccgtcatt gtggccaagt ggggtcagac 3000

ctccaccctc gccaacgtgc tctctgctgc cctgctcggc ctgttctggc agcagtgcgg 3060

atggctggct cacgactttc tgcaccacca ggtcttccag gaccgattct ggggtgatct 3120

cttcggagcc ttcctgggag gtgtctgcca gggcttctcc tcttcctggt ggaaggacaa 3180

gcacaacact caccatgccg ctcccaacgt gcatggcgag gatcctgaca ttgacaccca 3240

ccctctcctg acctggtccg agcacgctct ggagatgttc tccgacgtcc ccgatgagga 3300

gctgacccga atgtggtctc gattcatggt cctgaaccag acctggttct acttccccat 3360

tctctccttc gctcgactgt cttggtgcct ccagtccatt ctctttgtgc tgcccaacgg 3420

tcaggctcac aagccctccg gagctcgagt gcccatctcc ctggtcgagc agctgtccct 3480

cgccatgcac tggacctggt acctcgctac catgttcctg ttcatcaagg atcctgtcaa 3540

catgctcgtg tacttcctgg tgtctcaggc tgtgtgcgga aacctgctcg ccatcgtgtt 3600

ctccctcaac cacaacggta tgcctgtgat ctccaaggag gaggctgtcg acatggattt 3660

ctttaccaag cagatcatca ctggtcgaga tgtccatcct ggactgttcg ccaactggtt 3720

caccggtggc ctgaactacc agatcgagca tcacctgttc ccttccatgc ctcgacacaa 3780

cttctccaag atccagcctg ccgtcgagac cctgtgcaag aagtacaacg tccgatacca 3840

caccactggt atgatcgagg gaactgccga ggtcttctcc cgactgaacg aggtctccaa 3900

ggccacctcc aagatgggca aggctcagta agcggccgca agtgtggatg gggaagtgag 3960

tgcccggttc tgtgtgcaca attggcaatc caagatggat ggattcaaca cagggatata 4020

gcgagctacg tggtggtgcg aggatatagc aacggatatt tatgtttgac acttgagaat 4080

gtacgataca agcactgtcc aagtacaata ctaaacatac tgtacatact catactcgta 4140

cccgggcaac ggtttcactt gagtgcagtg gctagtgctc ttactcgtac agtgtgcaat 4200

actgcgtatc atagtctttg atgtatatcg tattcattca tgttagttgc gtacgagccg 4260

gaagcataaa gtgtaaagcc tggggtgcct aatgagtgag ctaactcaca ttaattgcgt 4320

tgcgctcact gcccgctttc cagtcgggaa acctgtcgtg ccagctgcat taatgaatcg 4380

gccaacgcgc ggggagaggc ggtttgcgta ttgggcgctc ttccgcttcc tcgctcactg 4440

actcgctgcg ctcggtcgtt cggctgcggc gagcggtatc agctcactca aaggcggtaa 4500

tacggttatc cacagaatca ggggataacg caggaaagaa catgtgagca aaaggccagc 4560

aaaaggccag gaaccgtaaa aaggccgcgt tgctggcgtt tttccatagg ctccgccccc 4620

ctgacgagca tcacaaaaat cgacgctcaa gtcagaggtg gcgaaacccg acaggactat 4680

aaagatacca ggcgtttccc cctggaagct ccctcgtgcg ctctcctgtt ccgaccctgc 4740

cgcttaccgg atacctgtcc gcctttctcc cttcgggaag cgtggcgctt tctcatagct 4800

cacgctgtag gtatctcagt tcggtgtagg tcgttcgctc caagctgggc tgtgtgcacg 4860

aacccccccgt tcagcccgac cgctgcgcct tatccggtaa ctatcgtctt gagtccaacc 4920

cggtaagaca cgacttatcg ccactggcag cagccactgg taacaggatt agcagagcga 4980

ggtatgtagg cggtgctaca gagttcttga agtggtggcc taactacggc tacactagaa 5040

ggacagtatt tggtatctgc gctctgctga agccagttac cttcggaaaa agagttggta 5100

gctcttgatc cggcaaacaa accaccgctg gtagcggtgg ttttttttgtt tgcaagcagc 5160

agattacgcg cagaaaaaaa ggatctcaag aagatccttt gatcttttct acggggtctg 5220

acgctcagtg gaacgaaaac tcacgttaag ggattttggt catgagatta tcaaaaagga 5280

tcttcaccta gatcctttta aattaaaaat gaagttttaa atcaatctaa agtatatatg 5340

agtaaacttg gtctgacagt taccaatgct taatcagtga ggcacctatc tcagcgatct 5400

gtctatttcg ttcatccata gttgcctgac tccccgtcgt gtagataact acgatacggg 5460

agggcttacc atctggcccc agtgctgcaa tgataccgcg agacccacgc tcaccggctc 5520

cagatttatc agcaataaac cagccagccg gaagggccga gcgcagaagt ggtcctgcaa 5580

ctttatccgc ctccatccag tctattaatt gttgccggga agctagagta agtagttcgc 5640

cagttaatag tttgcgcaac gttgttgcca ttgctacagg catcgtggtg tcacgctcgt 5700

cgtttggtat ggcttcattc agctccggtt cccaacgatc aaggcgagtt acatgatccc 5760

ccatgttgtg caaaaaagcg gttagctcct tcggtcctcc gatcgttgtc agaagtaagt 5820

tggccgcagt gttatcactc atggttatgg cagcactgca taattctctt actgtcatgc 5880

catccgtaag atgcttttct gtgactggtg agtactcaac caagtcattc tgagaatagt 5940

gtatgcggcg accgagttgc tcttgcccgg cgtcaatacg ggataatacc gcgccacata 6000

gcagaacttt aaaagtgctc atcattggaa aacgttcttc ggggcgaaaa ctctcaagga 6060

tcttaccgct gttgagatcc agttcgatgt aacccactcg tgcacccaac tgatcttcag 6120

catcttttac tttcaccagc gtttctgggt gagcaaaaac aggaaggcaa aatgccgcaa 6180

aaaagggaat aagggcgaca cggaaatgtt gaatactcat actcttcctt tttcaatatt 6240

attgaagcat ttatcagggt tattgtctca tgagcggata catatttgaa tgtatttaga 6300

aaaataaaca aatagggggtt ccgcgcacat ttccccgaaa agtgccacct gacgcgccct 6360

gtagcggcgc attaagcgcg gcgggtgtgg tggttacgcg cagcgtgacc gctacacttg 6420

ccagcgccct agcgcccgct cctttcgctt tcttcccttc ctttctcgcc acgttcgccg 6480

gctttccccg tcaagctcta aatcgggggc tccctttagg gttccgattt agtgctttac 6540

ggcacctcga ccccaaaaaa cttgattagg gtgatggttc acgtagtggg ccatcgccct 6600

gatagacggt ttttcgccct ttgacgttgg agtccacgtt ctttaatagt ggactcttgt 6660

tccaaactgg aacaacactc aaccctatct cggtctattc ttttgattta taagggattt 6720

tgccgatttc ggcctattgg ttaaaaaatg agctgattta acaaaaattt aacgcgaatt 6780

ttaacaaaat attaacgctt acaatttcca ttcgccattc aggctgcgca actgttggga 6840

agggcgatcg gtgcgggcct cttcgctatt acgccagctg gcgaaagggg gatgtgctgc 6900

aaggcgatta agttgggtaa cgccagggtt ttcccagtca cgacgttgta aaacgacggc 6960

cagtgaattg taatacgact cactataggg cgaattgggt accgggcccc ccctcgaggt 7020

cgatggtgtc gataagcttg atatcgaatt catgtcacac aaaccgatct tcgcctcaag 7080

gaaacctaat tctacatccg agagactgcc gagatccagt ctacactgat taattttcgg 7140

gccaataatt taaaaaaatc gtgttatata atattatatg tattatatat atacatcatg 7200

atgatactga cagtcatgtc ccattgctaa atagacagac tccatctgcc gcctccaact 7260

gatgttctca atatttaagg ggtcatctcg cattgtttaa taataaacag actccatcta 7320

ccgcctccaa atgatgttct caaaatatat tgtatgaact tatttttatt acttagtatt 7380

attagacaac ttacttgctt tatgaaaaac acttcctatt taggaaacaa tttataatgg 7440

cagttcgttc atttaacaat ttatgtagaa taaatgttat aaatgcgtat gggaaatctt 7500

aaatatggat agcataaatg atatctgcat tgcctaattc gaaatcaaca gcaacgaaaa 7560

aaatcccttg tacaacataa atagtcatcg agaaatatca actatcaaag aacagctatt 7620

cacacgttac tattgagatt attattggac gagaatcaca cactcaactg tctttctctc 7680

ttctagaaat acaggtacaa gtatgtacta ttctcattgt tcatacttct agtcatttca 7740

tcccacatat tccttggatt tctctccaat gaatgacatt ctatcttgca aattcaacaa 7800

ttataataag atataccaaa gtagcggtat agtggcaatc aaaaagcttc tctggtgtgc 7860

ttctcgtatt tatttttatt ctaatgatcc attaaaggta tatatttatt tcttgttata 7920

taatcctttt gtttattaca tgggctggat acataaaggt attttgattt aatttttgc 7980

ttaaattcaa tccccccctcg ttcagtgtca actgtaatgg taggaaatta ccatactttt 8040

gaagaagcaa aaaaaatgaa agaaaaaaaa aatcgtattt ccaggttaga cgttccgcag 8100

aatctagaat gcggtatgcg gtacattgtt cttcgaacgt aaaagttgcg ctccctgaga 8160

tattgtacat ttttgctttt acaagtacaa gtacatcgta caactatgta ctactgttga 8220

tgcatccaca acagtttgtt ttgttttttt ttgttttttt tttttctaat gattcattac 8280

cgctatgtat acctacttgt acttgtagta agccgggtta ttggcgttca attaatcata 8340

gacttatgaa tctgcacggt gtgcgctgcg agttactttt agcttatgca tgctacttgg 8400

gtgtaatatt gggatctgtt cggaaatcaa cggatgctca atcgattgcc ccggagaaga 8460

cggccaggcc gcctagatga caaattcaac aactcacagc tgactttctg ccattgccac 8520

taggggggggg ccttttttata tggccaagcc aagctctcca cgtcggttgg gctgcaccca 8580

acaataaatg ggtagggttg caccaacaaa gggatgggat ggggggtaga agatacgagg 8640

ataacggggc tcaatggcac aaataagaac gaatactgcc attaagactc gtgatccagc 8700

gactgacacc attgcatcat ctaagggcct caaaactacc tcggaactgc tgcgctgatc 8760

tggacaccac agaggttccg agcactttag gttgcaccaa atgtcccacc aggtgcaggc 8820

agaaaacgct ggaacagcgt gtacagtttg tcttaacaaa aagtgagggc gctgaggtcg 8880

agcagggtgg tgtgacttgt tatagccttt agagctgcga aagcgcgtat ggatttggct 8940

catcaggcca gattgagggt ctgtggacac atgtcatgtt agtgtacttc aatcgccccc 9000

tggatatagc cccgacaata ggccgtggcc tcatttttt gccttccgca catttccatt 9060

gctcggtacc cacaccttgc ttctcctgca cttgccaacc ttaatactgg tttacattga 9120

ccaacatctt acaagcgggg ggcttgtcta gggtatatat aaacagtggc tctcccaatc 9180

ggttgccagt ctctttttc ctttctttcc ccacagattc gaaatctaaa ctacacatca 9240

cacaatgcct gttactgacg tccttaagcg aaagtccggt gtcatcgtcg gcgacgatgt 9300

ccgagccgtg agtatccacg acaagatcag tgtcgagacg acgcgttttg tgtaatgaca 9360

caatccgaaa gtcgctagca acacacactc tctacacaaa ctaacccagc tctccatggc 9420

tgccgtccct attgaattca acgtcccctc cgtggaccga ccctttggta tctacctctg 9480

ggccatcttt gaccaggcct gggagaagct tttcggctgg cccgcgtcct ctttcatttt 9540

cgtgcgaaat gaccccaaca tccccttttc ctctacccct cccgtgatca ttgccatcat 9600

tgtgtactac attgtcatct ttggcggccg agaggtgatg cgaaacctgt ctcccatccg 9660

actcaactgg ctcttccaga tccacaacat cttcctcacc cttctgtccg gtatgctcct 9720

cctcctcctc gttgagcagc tcttccccat cattgtccga cagggtatcc tctacgccat 9780

ctgcgactac ggatcttgga ctcagcccat tgtcttctgc tactacctca actacctgac 9840

caagtacttt gagctgatcg acaccgtttt ccttgtgctg cgaaagaaga agctgacttt 9900

cctccacacc taccatcacg gtgccactgc tcttctgtgc tacacccagc tcattggtaa 9960

gacctcggtc tcttgggtcc ccatcaccct taacctgttt gtccacgttg tcatgtactt 10020

ctactacttc ctggctgcgc gaggtatccg agtgtggtgg aaggagtggg tcacccggct 10080

ccagatcatc cagttcgtta tcgatcttgg atttgtctac tttgcctctt acacctactt 10140

cacctctacc tactggccct ggatgcccaa catgggctct tgtgccggcg aggagtttgc 10200

tgctatttac ggctgtggtc tgctgacctc ttacctcttc ctcttcatcg ccttctacat 10260

caactcttac cgaaagccct cttccaaggg accttccaag cctgttgttg ctgtcgatgg 10320

ccctgttggc ggcgtcaacg cccagactgg tgcttctcga ggccagacca ctacccgatc 10380

tcgacgagca taagcggccg catggagcgt gtgttctgag tcgatgtttt ctatggagtt 10440

gtgagtgtta gtagacatga tgggtttata tatgatgaat gaatagatgt gattttgatt 10500

tgcacgatgg aattgagaac tttgtaaacg tacatgggaa tgtatgaatg tgggggtttt 10560

gtgactggat aactgacggt cagtggacgc cgttgttcaa atatccaaga gatgcgagaa 10620

actttgggtc aagtgaacat gtcctctctg ttcaagtaaa ccatcaacta tgggtagtat 10680

atttagtaag gacaagagtt gagattcttt ggagtcctag aaacgtattt tcgcgttcca 10740

agatcaaatt agtagagtaa tacgggcacg ggaatccatt catagtctca atcctgcagg 10800

tgagttaat 10809

<210> 204
<211> 8273
<212> DNA
<213> Artificial Sequence

<220>
<223> Plasmid pYCPT1-17

<400> 204
ggccgcaagt gtggatgggg aagtgagtgc ccggttctgt gtgcacaatt ggcaatccaa 60

gatggatgga ttcaacacag ggatatagcg agctacgtgg tggtgcgagg atatagcaac 120

ggatatttat gtttgacact tgagaatgta cgatacaagc actgtccaag tacaatacta 180

aacatactgt acatactcat actcgtaccc gggcaacggt ttcacttgag tgcagtggct 240

agtgctctta ctcgtacagt gtgcaatact gcgtatcata gtctttgatg tatatcgtat 300

tcattcatgt tagttgcgta cgagccggaa gcataaagtg taaagcctgg ggtgcctaat 360

gagtgagcta actcacatta attgcgttgc gctcactgcc cgctttccag tcgggaaacc 420

tgtcgtgcca gctgcattaa tgaatcggcc aacgcgcggg gagaggcggt ttgcgtattg 480

ggcgctcttc cgcttcctcg ctcactgact cgctgcgctc ggtcgttcgg ctgcggcgag 540

cggtatcagc tcactcaaag gcggtaatac ggttatccac agaatcaggg gataacgcag 600

gaaagaacat gtgagcaaaa ggccagcaaa aggccaggaa ccgtaaaaag gccgcgttgc 660

tggcgttttt ccataggctc cgccccccctg acgagcatca caaaaatcga cgctcaagtc 720

agaggtggcg aaacccgaca ggactataaa gataccaggc gtttcccccct ggaagctccc 780

tcgtgcgctc tcctgttccg accctgccgc ttaccggata cctgtccgcc tttctccctt 840

cgggaagcgt ggcgctttct catagctcac gctgtaggta tctcagttcg gtgtaggtcg 900

ttcgctccaa gctgggctgt gtgcacgaac cccccgttca gcccgaccgc tgcgccttat 960

ccggtaacta tcgtcttgag tccaacccgg taagacacga cttatcgcca ctggcagcag 1020

ccactggtaa caggattagc agagcgaggt atgtaggcgg tgctacagag ttcttgaagt 1080

ggtggcctaa ctacggctac actagaagga cagtatttgg tatctgcgct ctgctgaagc 1140

cagttacctt cggaaaaaga gttggtagct cttgatccgg caaacaaacc accgctggta 1200

gcggtggttt ttttgtttgc aagcagcaga ttacgcgcag aaaaaaagga tctcaagaag 1260

atcctttgat cttttctacg gggtctgacg ctcagtggaa 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 cccacgctca ccggctccag atttatcagc aataaaccag ccagccggaa 1620

gggccgagcg cagaagtggt cctgcaactt tatccgcctc catccagtct attaattgtt 1680

gccgggaagc tagagtaagt agttcgccag ttaatagttt gcgcaacgtt gttgccattg 1740

ctacaggcat cgtggtgtca cgctcgtcgt ttggtatggc ttcattcagc tccggttccc 1800

aacgatcaag gcgagttaca tgatccccca tgttgtgcaa aaaagcggtt agctccttcg 1860

gtcctccgat cgttgtcaga agtaagttgg ccgcagtgtt atcactcatg gttatggcag 1920

cactgcataa ttctcttact gtcatgccat ccgtaagatg cttttctgtg actggtgagt 1980

actcaaccaa gtcattctga gaatagtgta tgcggcgacc gagttgctct tgcccggcgt 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 caatattatt gaagcattta tcagggttat tgtctcatga 2340

gcggatacat atttgaatgt atttagaaaa ataaacaaat aggggttccg cgcacatttc 2400

cccgaaaagt gccacctgac gcgccctgta gcggcgcatt aagcgcggcg ggtgtggtgg 2460

ttacgcgcag cgtgaccgct acacttgcca gcgccctagc gcccgctcct ttcgctttct 2520

tcccttcctt tctcgccacg ttcgccggct ttccccgtca agctctaaat cgggggctcc 2580

ctttagggtt ccgatttagt gctttacggc acctcgaccc caaaaaactt gattagggtg 2640

atggttcacg tagtgggcca tcgccctgat agacggtttt tcgccctttg acgttggagt 2700

ccacgttctt taatagtgga ctcttgttcc 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 aaaggggatg gtgctgcaag gcgattaagt tgggtaacgc cagggttttc 3000

ccagtcacga cgttgtaaaa cgacggccag tgaattgtaa tacgactcac tatagggcga 3060

attgggtacc gggcccccccc tcgaggtcga tggtgtcgat 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 gttctcaata tttaaggggt 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 tcccttgtac aacataaata gtcatcgaga 3660

aatatcaact atcaaagaac agctattcac acgttactat tgagattatt attggacgag 3720

aatcacacac tcaactgtct ttctctcttc tagaaataca ggtacaagta tgtactattc 3780

tcattgttca tacttctagt catttcatcc cacatattcc 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 cccctcgttc agtgtcaact 4080

gtaatggtag gaaattacca tacttttgaa gaagcaaaaa aaatgaaaga aaaaaaaaat 4140

cgtatttcca ggttagacgt tccgcagaat ctagaatgcg gtatgcggta cattgttctt 4200

cgaacgtaaa agttgcgctc cctgagatat tgtacatttt tgcttttaca agtacaagta 4260

catcgtacaa ctatgtacta ctgttgatgc atccacaaca gtttgttttg ttttttttg 4320

ttttttttt ttctaatgat 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 gtctgaccat 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 cacccgggg 4980

gtcagaataa gccagtcctc agagtcgccc ttaggtcggt tctgggcaat gaagccaacc 5040

acaaactcgg ggtcggatcg ggcaagctca atggtctgct tggagtactc gccagtggcc 5100

agagagccct tgcaagacag ctcggccagc atgagcagac ctctggccag cttctcgttg 5160

ggagagggga ctaggaactc cttgtactgg gagttctcgt agtcagagac gtcctccttc 5220

ttctgttcag agacagtttc ctcggcacca gctcgcaggc cagcaatgat tccggttccg 5280

ggtacaccgt gggcgttggt gatatcggac cactcggcga ttcggtgaca ccggtactgg 5340

tgcttgacag tgttgccaat 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

attttggtgg tgaagaggag actgaaataa atttagtctg cagaactttt tatcggaacc 5700

ttatctgggg cagtgaagta tatgttatgg taatagttac gagttagttg 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 cgtactccaa aggcggcaat 6000

gacgagtcag acagatactc gtcgactcag gcgacgacgg aattcctgca gcccatctgc 6060

agaattcagg agagaccggg ttggcggcgt atttgtgtcc caaaaaacag ccccaattgc 6120

cccggagaag acggccaggc cgcctagatg acaaattcaa caactcacag ctgactttct 6180

gccattgcca ctaggggggg gccttttttat atggccaagc caagctctcc acgtcggttg 6240

ggctgcaccc aacaataaat gggtagggtt gcaccaacaa agggatggga tgggggggtag 6300

aagatacgag gataacgggg ctcaatggca caaataagaa cgaatactgc cattaagact 6360

cgtgatccag cgactgacac cattgcatca tctaagggcc tcaaaactac ctcggaactg 6420

ctgcgctgat ctggacacca cagaggttcc gagcacttta ggttgcacca aatgtcccac 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 aggccgtggc ctcatttttt tgccttccgc 6720

acatttccat tgctcggtac ccacaccttg cttctcctgc acttgccaac cttaatactg 6780

gtttacattg accaacatct tacaagcggg gggcttgtct agggtatata taaacagtgg 6840

ctctcccaat cggttgccag tctctttttt cctttctttc cccacagatt cgaaatctaa 6900

actacacatc acacaatgcc tgttactgac gtccttaagc gaaagtccgg tgtcatcgtc 6960

ggcgacgatg tccgagccgt gagtatccac gacaagatca gtgtcgagac gacgcgtttt 7020

gtgtaatgac acaatccgaa agtcgctagc aacacacact ctctacacaa actaacccag 7080

ctctccatgg gcgtattcat taaacaggag cagcttccgg ctctcaagaa gtacaagtac 7140

tccgccgagg atcactcgtt catctccaac aacattctgc gccccttctg gcgacagttt 7200

gtcaaaatct tccctctgtg gatggccccc aacatggtga ctctgttggg cttcttcttt 7260

gtcattgtga acttcatcac catgctcatt gttgatccca cccacgaccg cgagcctccc 7320

agatgggtct acctcaccta cgctctgggt ctgttccttt accagacatt tgatgcctgt 7380

gacggatccc atgcccgacg aactggccag agtggacccc ttggagagct gtttgaccac 7440

tgtgtcgacg ccatgaatac ctctctgatt ctcacggtgg tggtgtccac cacccatatg 7500

ggatataaca tgaagctgct gattgtgcag attgccgctc tcggaaactt ctacctgtcg 7560

acctgggaga cctaccatac cggaactctg tacctttctg gcttctctgg tcctgttgaa 7620

ggtatcttga ttctggtggc tcttttcgtc ctcaccttct tcactggtcc caacgtgtac 7680

gctctgaccg tctacgaggc tcttcccgaa tccatcactt cgctgctgcc tgccagcttc 7740

ctggacgtca ccatcaccca gatctacatt ggattcggag tgctgggcat ggtgttcaac 7800

atctacggcg cctgcggaaa cgtgatcaag tactacaaca acaagggcaa gagcgctctc 7860

cccgccattc tcggaatcgc ccccttggc atcttctacg tcggcgtctt tgcctgggcc 7920

catgttgctc ctctgcttct ctccaagtac gccatcgtct atctgtttgc cattggggct 7980

gcctttgcca tgcaagtcgg ccagatgatt cttgcccatc tcgtgcttgc tcccttcccc 8040

cactggaacg tgctgctctt cttccccttt gtgggactgg cagtgcacta cattgcaccc 8100

gtgtttggct gggacgccga tatcgtgtcg gttaacactc tcttcacctg ttttggcgcc 8160

accctctcca tttacgcctt ctttgtgctt gagatcatcg acgagatcac caactacctc 8220

gatatctggt gtctgcgaat caagtaccct caggagaaga agactgagta agc 8273

<210> 205
<211> 7822
<212> DNA
<213> Artificial Sequence

<220>
<223> Plasmid pZP2l7+Ura

<400> 205
aaacggtagg ttagtgcttg gtatatgagt tgtaggcatg acaatttgga aaggggtgga 60

ctttgggaat attgtgggat ttcaatacct tagtttgtac agggtaattg ttacaaatga 120

tacaaagaac tgtatttctt ttcatttgtt ttaattggtt gtatatcaag tccgttagac 180

gagctcagtg ccttggcttt tggcactgta tttcattttt agaggtacac tacattcagt 240

gaggtatggt aaggttgagg gcataatgaa ggcaccttgt actgacagtc acagacctct 300

caccgagaat tttatgagat atactcgggt tcattttagg ctcatcgatc aggagagacc 360

gggttggcgg cgtatttgtg tcccaaaaaa cagccccaat tgccccaatt gaccccaaat 420

tgacccagta gcgggcccaa ccccggcgag agccccccttc accccacata tcaaacctcc 480

cccggttccc acacttgccg ttaagggcgt agggtactgc agtctggaat ctacgcttgt 540

tcagactttg tactagtttc tttgtctggc catccgggta acccatgccg gacgcaaaat 600

agactactga aaatttttttt gctttgtggt tgggacttta gccaagggta taaaagacca 660

ccgtccccga attacctttc ctcttctttt ctctctctcc ttgtcaactc acacccgaaa 720

tcgttaagca tttccttctg agtataagaa tcattcacca tggctgagga taagaccaag 780

gtcgagttcc ctaccctgac tgagctgaag cactctatcc ctaacgcttg ctttgagtcc 840

aacctcggac tctcgctcta ctacactgcc cgagcgatct tcaacgcatc tgcctctgct 900

gctctgctct acgctgcccg atctactccc ttcattgccg ataacgttct gctccacgct 960

ctggtttgcg ccacctacat ctacgtgcag ggtgtcatct tctggggttt ctttaccgtc 1020

ggtcacgact gtggtcactc tgccttctcc cgataccact ccgtcaactt catcattggc 1080

tgcatcatgc actctgccat tctgactccc ttcgagtcct ggcgagtgac ccaccgacac 1140

catcacaaga acactggcaa cattgataag gacgagatct tctaccctca tcggtccgtc 1200

aaggacctcc aggacgtgcg acaatgggtc tacaccctcg gaggtgcttg gtttgtctac 1260

ctgaaggtcg gatatgctcc tcgaaccatg tcccactttg acccctggga ccctctcctg 1320

cttcgacgag cctccgctgt catcgtgtcc ctcggagtct gggctgcctt cttcgctgcc 1380

tacgcctacc tcacatactc gctcggcttt gccgtcatgg gcctctacta ctatgctcct 1440

ctctttgtct ttgcttcgtt cctcgtcatt actaccttct tgcatcacaa cgacgaagct 1500

actccctggt acggtgactc ggagtggacc tacgtcaagg gcaacctgag ctccgtcgac 1560

cgatcgtacg gagctttcgt ggacaacctg tctcaccaca ttggcaccca ccaggtccat 1620

cacttgttcc ctatcattcc ccactacaag ctcaacgaag ccaccaagca ctttgctgcc 1680

gcttaccctc acctcgtgag acgtaacgac gagcccatca ttactgcctt cttcaagacc 1740

gctcacctct ttgtcaacta cggagctgtg cccgagactg ctcagatttt caccctcaaa 1800

gagtctgccg ctgcagccaa ggccaagagc gactaagcgg ccgcaagtgt ggatggggaa 1860

gtgagtgccc ggttctgtgt gcacaattgg caatccaaga tggatggatt caacacaggg 1920

atatagcgag ctacgtggtg gtgcgaggat atagcaacgg atatttatgt ttgacacttg 1980

agaatgtacg atacaagcac tgtccaagta caatactaaa catactgtac atactcatac 2040

tcgtacccgg gcaacggttt cacttgagtg cagtggctag tgctcttact cgtacagtgt 2100

gcaatactgc gtatcatagt ctttgatgta tatcgtattc attcatgtta gttgcgtacg 2160

ggcgtcgttg cttgtgtgat ttttgaggac ccatcccttt ggtatataag tatactctgg 2220

ggttaaggtt gcccgtgtag tctaggttat agttttcatg tgaaataccg agagccgagg 2280

gagaataaac gggggtattt ggacttgttt ttttcgcgga aaagcgtcga atcaaccctg 2340

cgggccttgc accatgtcca cgacgtgttt ctcgccccaa ttcgcccctt gcacgtcaaa 2400

attaggcctc catctagacc cctccataac atgtgactgt ggggaaaagt ataagggaaa 2460

ccatgcaacc atagacgacg tgaaagacgg ggaggaacca atggaggcca aagaaatggg 2520

gtagcaacag tccaggagac agacaaggag acaaggagag ggcgcccgaa agatcggaaa 2580

aacaaacatg tccaattggg gcagtgacgg aaacgacacg gacacttcag tacaatggac 2640

cgaccatctc caagccaggg ttattccggt atcaccttgg ccgtaacctc ccgctggtac 2700

ctgatattgt acacgttcac attcaatata ctttcagcta caataagaga ggctgtttgt 2760

cgggcatgtg tgtccgtcgt atggggtgat gtccgagggc gaaattcgct acaagcttaa 2820

ctctggcgct tgtccagtat gaatagacaa gtcaagacca gtggtgccat gattgacagg 2880

gaggtacaag acttcgatac tcgagcatta ctcggacttg tggcgattga acagacgggc 2940

gatcgcttct cccccgtatt gccggcgcgc cagctgcatt aatgaatcgg ccaacgcgcg 3000

gggagaggcg gtttgcgtat tgggcgctct tccgcttcct cgctcactga ctcgctgcgc 3060

tcggtcgttc ggctgcggcg agcggtatca gctcactcaa aggcggtaat acggttatcc 3120

acagaatcag gggataacgc aggaaagaac atgtgagcaa aaggccagca aaaggccagg 3180

aaccgtaaaa aggccgcgtt gctggcgttt ttccataggc tccgcccccc tgacgagcat 3240

cacaaaaatc gacgctcaag tcagaggtgg cgaaacccga caggactata aagataccag 3300

gcgtttcccc ctggaagctc cctcgtgcgc tctcctgttc cgaccctgcc gcttaccgga 3360

tacctgtccg cctttctccc ttcgggaagc gtggcgcttt ctcatagctc acgctgtagg 3420

tatctcagtt cggtgtaggt cgttcgctcc aagctgggct gtgtgcacga acccccccgtt 3480

cagcccgacc gctgcgcctt atccggtaac tatcgtcttg agtccaaccc ggtaagacac 3540

gacttatcgc cactggcagc agccactggt aacaggatta gcagagcgag gtatgtaggc 3600

ggtgctacag agttcttgaa gtggtggcct aactacggct acactagaag aacagtattt 3660

ggtatctgcg ctctgctgaa gccagttacc ttcggaaaaa gagttggtag ctcttgatcc 3720

ggcaaacaaa ccaccgctgg tagcggtggt ttttttgttt gcaagcagca gattacgcgc 3780

agaaaaaaag gatctcaaga agatcctttg atcttttcta cggggtctga cgctcagtgg 3840

aacgaaaact cacgttaagg gattttggtc atgagattat caaaaaggat cttcacctag 3900

atccttttaa attaaaaatg aagttttaaa tcaatctaaa gtatatatga gtaaacttgg 3960

tctgacagtt accaatgctt aatcagtgag gcacctatct cagcgatctg tctatttcgt 4020

tcatccatag ttgcctgact ccccgtcgtg tagataacta cgatacggga gggcttacca 4080

tctggcccca gtgctgcaat gataccgcga gacccacgct caccggctcc agatttatca 4140

gcaataaacc agccagccgg aagggccgag cgcagaagtg gtcctgcaac tttatccgcc 4200

tccatccagt ctattaattg ttgccgggaa gctagagtaa gtagttcgcc agttaatagt 4260

ttgcgcaacg ttgttgccat tgctacaggc atcgtggtgt cacgctcgtc gtttggtatg 4320

gcttcattca gctccggttc ccaacgatca aggcgagtta catgatcccc catgttgtgc 4380

aaaaaagcgg ttagctcctt cggtcctccg atcgttgtca gaagtaagtt ggccgcagtg 4440

ttatcactca tggttatggc agcactgcat aattctctta ctgtcatgcc atccgtaaga 4500

tgcttttctg tgactggtga gtactcaacc aagtcattct gagaatagtg tatgcggcga 4560

ccgagttgct cttgcccggc gtcaatacgg gataataccg cgccacatag cagaacttta 4620

aaagtgctca tcattggaaa acgttcttcg gggcgaaaac tctcaaggat cttaccgctg 4680

ttgagatcca gttcgatgta acccactcgt gcacccaact gatcttcagc atcttttact 4740

ttcaccagcg tttctgggtg agcaaaaaca ggaaggcaaa atgccgcaaa aaagggaata 4800

agggcgacac ggaaatgttg aatactcata ctcttccttt ttcaatatta ttgaagcatt 4860

tatcagggtt attgtctcat gagcggatac atatttgaat gtatttagaa aaataaacaa 4920

atagggggttc cgcgcacatt tccccgaaaa gtgccacctg atgcggtgtg aaataccgca 4980

cagatgcgta aggagaaaat accgcatcag gaaattgtaa gcgttaatat tttgttaaaa 5040

ttcgcgttaa atttttgtta aatcagctca ttttttaacc aataggccga aatcggcaaa 5100

atcccttata aatcaaaaga atagaccgag atagggttga gtgttgttcc agtttggaac 5160

aagagtccac tattaaagaa cgtggactcc aacgtcaaag ggcgaaaaac cgtctatcag 5220

ggcgatggcc cactacgtga accatcaccc taatcaagtt ttttggggtc gaggtgccgt 5280

aaagcactaa atcggaaccc taaagggagc ccccgattta gagcttgacg gggaaagccg 5340

gcgaacgtgg cgagaaagga agggaagaaa gcgaaaggag cgggcgctag ggcgctggca 5400

agtgtagcgg tcacgctgcg cgtaaccacc acacccgccg cgcttaatgc gccgctacag 5460

ggcgcgtcca ttcgccattc aggctgcgca actgttggga agggcgatcg gtgcgggcct 5520

cttcgctatt acgccagctg gcgaaagggg gatgtgctgc aaggcgatta agttgggtaa 5580

cgccagggtt ttcccagtca cgacgttgta aaacgacggc cagtgaattg taatacgact 5640

cactataggg cgaattgggc ccgacgtcgc atgcgctgat gacactttgg tctgaaagag 5700

atgcattttg aatcccaaac ttgcagtgcc caagtgacat acatctccgc gttttggaaa 5760

atgttcagaa acagttgatt gtgttggaat ggggaatggg gaatggaaaa atgactcaag 5820

tatcaattcc aaaaacttct ctggctggca gtacctactg tccatactac tgcattttct 5880

ccagtcaggc cactctatac tcgacgacac agtagtaaaa cccagataat ttcgacataa 5940

acaagaaaac agacccaata atatttatat atagtcagcc gtttgtccag ttcagactgt 6000

aatagccgaa aaaaaatcca aagtttctat tctaggaaaa tatattccaa tatttttaat 6060

tcttaatctc atttatttta ttctagcgaa atacatttca gctacttgag acatgtgata 6120

cccacaaatc ggattcggac tcggttgttc agaagagcat atggcattcg tgctcgcttg 6180

ttcacgtatt cttcctgttc catctcttgg ccgacaatca cacaaaaatg gggtttttt 6240

tttaattcta atgattcatt acagcaaaat tgagatatag cagaccacgt attccataat 6300

caccaaggaa gttcttgggc gtcttaatta agtcatacac aagtcagctt tcttcgagcc 6360

tcatataagt ataagtagtt caacgtatta gcactgtacc cagcatctcc gtatcgagaa 6420

acacaacaac atgccccatt ggacagatca tgcggataca caggttgtgc agtatcatac 6480

atactcgatc agacaggtcg tctgaccatc atacaagctg aacaagcgct ccatacttgc 6540

acgctctcta tatacacagt taaattacat atccatagtc taacctctaa cagttaatct 6600

tctggtaagc ctcccagcca gccttctggt atcgcttggc ctcctcaata ggatctcggt 6660

tctggccgta cagacctcgg ccgacaatta tgatatccgt tccggtagac atgacatcct 6720

caacagttcg gtactgctgt ccgagagcgt ctcccttgtc gtcaagaccc accccggggg 6780

tcagaataag ccagtcctca gagtcgccct taggtcggtt ctgggcaatg aagccaacca 6840

caaactcggg gtcggatcgg gcaagctcaa tggtctgctt ggagtactcg ccagtggcca 6900

gagagccctt gcaagacagc tcggccagca tgagcagacc tctggccagc ttctcgttgg 6960

gagaggggac taggaactcc ttgtactggg agttctcgta gtcagagacg tcctccttct 7020

tctgttcaga gacagtttcc tcggcaccag ctcgcaggcc agcaatgatt ccggttccgg 7080

gtacaccgtg ggcgttggtg atatcggacc actcggcgat tcggtgacac cggtactggt 7140

gcttgacagt gttgccaata tctgcgaact ttctgtcctc gaacaggaag aaaccgtgct 7200

taagagcaag ttccttgagg gggagcacag tgccggcgta ggtgaagtcg tcaatgatgt 7260

cgatatgggt tttgatcatg cacacataag gtccgacctt atcggcaagc tcaatgagct 7320

ccttggtggt ggtaacatcc agagaagcac acaggttggt tttcttggct gccacgagct 7380

tgagcactcg agcggcaaag gcggacttgt ggacgttagc tcgagcttcg taggagggca 7440

ttttggtggt gaagaggaga ctgaaataaa tttagtctgc agaacttttt atcggaacct 7500

tatctggggc agtgaagtat atgttatggt aatagttacg agttagttga acttatagat 7560

agactggact atacggctat cggtccaaat tagaaagaac gtcaatggct ctctgggcgt 7620

cgcctttgcc gacaaaaatg tgatcatgat gaaagccagc aatgacgttg cagctgatat 7680

tgttgtcggc caaccgcgcc gaaaacgcag ctgtcagacc cacagcctcc aacgaagaat 7740

gtatcgtcaa agtgatccaa gcacactcat agttggagtc gtactccaaa ggcggcaatg 7800

acgagtcaga cagatactcg tc 7822

<210> 206
<211> 7930
<212> DNA
<213> Artificial Sequence

<220>
<223> pYCPT1-ZP2l7

<400> 206
catgggcgta ttcattaaac aggagcagct tccggctctc aagaagtaca agtactccgc 60

cgaggatcac tcgttcatct ccaacaacat tctgcgcccc ttctggcgac agtttgtcaa 120

aatcttccct ctgtggatgg cccccaacat ggtgactctg ttgggcttct tctttgtcat 180

tgtgaacttc atcaccatgc tcattgttga tcccacccac gaccgcgagc ctcccagatg 240

ggtctacctc acctacgctc tgggtctgtt cctttaccag acatttgatg cctgtgacgg 300

atcccatgcc cgacgaactg gccagagtgg acccccttgga gagctgtttg accactgtgt 360

cgacgccatg aatacctctc tgattctcac ggtggtggtg tccaccaccc atatgggata 420

taacatgaag ctgctgattg tgcagattgc cgctctcgga aacttctacc tgtcgacctg 480

ggagacctac cataccggaa ctctgtacct ttctggcttc tctggtcctg ttgaaggtat 540

cttgattctg gtggctcttt tcgtcctcac cttcttcact ggtcccaacg tgtacgctct 600

gaccgtctac gaggctcttc ccgaatccat cacttcgctg ctgcctgcca gcttcctgga 660

cgtcaccatc acccagatct acattggatt cggagtgctg ggcatggtgt tcaacatcta 720

cggcgcctgc ggaaacgtga tcaagtacta caacaacaag ggcaagagcg ctctccccgc 780

cattctcgga atcgcccccct ttggcatctt ctacgtcggc gtctttgcct gggcccatgt 840

tgctcctctg cttctctcca agtacgccat cgtctatctg tttgccattg gggctgcctt 900

tgccatgcaa gtcggccaga tgattcttgc ccatctcgtg cttgctccct tcccccactg 960

gaacgtgctg ctcttcttcc cctttgtggg actggcagtg cactacattg cacccgtgtt 1020

tggctgggac gccgatatcg tgtcggttaa cactctcttc acctgttttg gcgccaccct 1080

ctccatttac gccttctttg tgcttgagat catcgacgag atcaccaact acctcgatat 1140

ctggtgtctg cgaatcaagt accctcagga gaagaagact gagtaagcgg ccgcaagtgt 1200

ggatggggaa gtgagtgccc ggttctgtgt gcacaattgg caatccaaga tggatggatt 1260

caacacaggg atatagcgag ctacgtggtg gtgcgaggat atagcaacgg atatttatgt 1320

ttgacacttg agaatgtacg atacaagcac tgtccaagta caatactaaa catactgtac 1380

atactcatac tcgtacccgg gcaacggttt cacttgagtg cagtggctag tgctcttact 1440

cgtacagtgt gcaatactgc gtatcatagt ctttgatgta tatcgtattc attcatgtta 1500

gttgcgtacg ggcgtcgttg cttgtgtgat ttttgaggac ccatcccttt ggtatataag 1560

tatactctgg ggttaaggtt gcccgtgtag tctaggttat agttttcatg tgaaataccg 1620

agagccgagg gagaataaac gggggtattt ggacttgttt ttttcgcgga aaagcgtcga 1680

atcaaccctg cgggccttgc accatgtcca cgacgtgttt ctcgccccaa ttcgcccctt 1740

gcacgtcaaa attaggcctc catctagacc cctccataac atgtgactgt ggggaaaagt 1800

ataagggaaa ccatgcaacc atagacgacg tgaaagacgg ggaggaacca atggaggcca 1860

aagaaatggg gtagcaacag tccaggagac agacaaggag acaaggagag ggcgcccgaa 1920

agatcggaaa aacaaacatg tccaattggg gcagtgacgg aaacgacacg gacacttcag 1980

tacaatggac cgaccatctc caagccaggg ttattccggt atcaccttgg ccgtaacctc 2040

ccgctggtac ctgatattgt acacgttcac attcaatata ctttcagcta caataagaga 2100

ggctgtttgt cgggcatgtg tgtccgtcgt atggggtgat gtccgagggc gaaattcgct 2160

acaagcttaa ctctggcgct tgtccagtat gaatagacaa gtcaagacca gtggtgccat 2220

gattgacagg gaggtacaag acttcgatac tcgagcatta ctcggacttg tggcgattga 2280

acagacgggc gatcgcttct cccccgtatt gccggcgcgc cagctgcatt aatgaatcgg 2340

ccaacgcgcg gggagaggcg gtttgcgtat tgggcgctct tccgcttcct cgctcactga 2400

ctcgctgcgc tcggtcgttc ggctgcggcg agcggtatca gctcactcaa aggcggtaat 2460

acggttatcc acagaatcag gggataacgc aggaaagaac atgtgagcaa aaggccagca 2520

aaaggccagg aaccgtaaaa aggccgcgtt gctggcgttt ttccataggc tccgcccccc 2580

tgacgagcat cacaaaaatc gacgctcaag tcagaggtgg cgaaacccga caggactata 2640

aagataccag gcgtttcccc ctggaagctc cctcgtgcgc tctcctgttc cgaccctgcc 2700

gcttaccgga tacctgtccg cctttctccc ttcgggaagc gtggcgcttt ctcatagctc 2760

acgctgtagg tatctcagtt cggtgtaggt cgttcgctcc aagctgggct gtgtgcacga 2820

acccccccgtt cagcccgacc gctgcgcctt atccggtaac tatcgtcttg agtccaaccc 2880

ggtaagacac gacttatcgc cactggcagc agccactggt aacaggatta gcagagcgag 2940

gtatgtaggc ggtgctacag agttcttgaa gtggtggcct aactacggct acactagaag 3000

aacagtattt ggtatctgcg ctctgctgaa gccagttacc ttcggaaaaa gagttggtag 3060

ctcttgatcc ggcaaacaaa ccaccgctgg tagcggtggt ttttttgttt gcaagcagca 3120

gattacgcgc agaaaaaaag gatctcaaga agatcctttg atcttttcta cggggtctga 3180

cgctcagtgg aacgaaaact cacgttaagg gattttggtc atgagattat caaaaaggat 3240

cttcacctag atccttttaa attaaaaatg aagttttaaa tcaatctaaa gtatatatga 3300

gtaaacttgg tctgacagtt accaatgctt aatcagtgag gcacctatct cagcgatctg 3360

tctatttcgt tcatccatag ttgcctgact ccccgtcgtg tagataacta cgatacggga 3420

gggcttacca tctggcccca gtgctgcaat gataccgcga gacccacgct caccggctcc 3480

agatttatca gcaataaacc agccagccgg aagggccgag cgcagaagtg gtcctgcaac 3540

tttatccgcc tccatccagt ctattaattg ttgccgggaa gctagagtaa gtagttcgcc 3600

agttaatagt ttgcgcaacg ttgttgccat tgctacaggc atcgtggtgt cacgctcgtc 3660

gtttggtatg gcttcattca gctccggttc ccaacgatca aggcgagtta catgatcccc 3720

catgttgtgc aaaaaagcgg ttagctcctt cggtcctccg atcgttgtca gaagtaagtt 3780

ggccgcagtg ttatcactca tggttatggc agcactgcat aattctctta ctgtcatgcc 3840

atccgtaaga tgcttttctg tgactggtga gtactcaacc aagtcattct gagaatagtg 3900

tatgcggcga ccgagttgct cttgcccggc gtcaatacgg gataataccg cgccacatag 3960

cagaacttta aaagtgctca tcattggaaa acgttcttcg gggcgaaaac tctcaaggat 4020

cttaccgctg ttgagatcca gttcgatgta acccactcgt gcacccaact gatcttcagc 4080

atcttttact ttcaccagcg tttctgggtg agcaaaaaca ggaaggcaaa atgccgcaaa 4140

aaagggaata agggcgacac ggaaatgttg aatactcata ctcttccttt ttcaatatta 4200

ttgaagcatt tatcagggtt attgtctcat gagcggatac atatttgaat gtatttagaa 4260

aaataaacaa atagggggttc cgcgcacatt tccccgaaaa gtgccacctg atgcggtgtg 4320

aaataccgca cagatgcgta aggagaaaat accgcatcag gaaattgtaa gcgttaatat 4380

tttgttaaaa ttcgcgttaa atttttgtta aatcagctca ttttttaacc aataggccga 4440

aatcggcaaa atcccttata aatcaaaaga atagaccgag atagggttga gtgttgttcc 4500

agtttggaac aagagtccac tattaaagaa cgtggactcc aacgtcaaag ggcgaaaaac 4560

cgtctatcag ggcgatggcc cactacgtga accatcaccc taatcaagtt ttttggggtc 4620

gaggtgccgt aaagcactaa atcggaaccc taaagggagc ccccgattta gagcttgacg 4680

gggaaagccg gcgaacgtgg cgagaaagga agggaagaaa gcgaaaggag cgggcgctag 4740

ggcgctggca agtgtagcgg tcacgctgcg cgtaaccacc acacccgccg cgcttaatgc 4800

gccgctacag ggcgcgtcca ttcgccattc aggctgcgca actgttggga agggcgatcg 4860

gtgcgggcct cttcgctatt acgccagctg gcgaaagggg gatgtgctgc aaggcgatta 4920

agttgggtaa cgccagggtt ttcccagtca cgacgttgta aaacgacggc cagtgaattg 4980

taatacgact cactataggg cgaattgggc ccgacgtcgc atgcgctgat gacactttgg 5040

tctgaaagag atgcattttg aatcccaaac ttgcagtgcc caagtgacat acatctccgc 5100

gttttggaaa atgttcagaa acagttgatt gtgttggaat ggggaatggg gaatggaaaa 5160

atgactcaag tatcaattcc aaaaacttct ctggctggca gtacctactg tccatactac 5220

tgcattttct ccagtcaggc cactctatac tcgacgacac agtagtaaaa cccagataat 5280

ttcgacataa acaagaaaac agacccaata atatttatat atagtcagcc gtttgtccag 5340

ttcagactgt aatagccgaa aaaaaatcca aagtttctat tctaggaaaa tatattccaa 5400

tatttttaat tcttaatctc atttatttta ttctagcgaa atacatttca gctacttgag 5460

acatgtgata cccacaaatc ggattcggac tcggttgttc agaagagcat atggcattcg 5520

tgctcgcttg ttcacgtatt cttcctgttc catctcttgg ccgacaatca cacaaaaatg 5580

gggttttttt tttaattcta atgattcatt acagcaaaat tgagatatag cagaccacgt 5640

attccataat caccaaggaa gttcttgggc gtcttaatta agtcatacac aagtcagctt 5700

tcttcgagcc tcatataagt ataagtagtt caacgtatta gcactgtacc cagcatctcc 5760

gtatcgagaa acacaacaac atgccccatt ggacagatca tgcggataca caggttgtgc 5820

agtatcatac atactcgatc agacaggtcg tctgaccatc atacaagctg aacaagcgct 5880

ccatacttgc acgctctcta tatacacagt taaattacat atccatagtc taacctctaa 5940

cagttaatct tctggtaagc ctcccagcca gccttctggt atcgcttggc ctcctcaata 6000

ggatctcggt tctggccgta cagacctcgg ccgacaatta tgatatccgt tccggtagac 6060

atgacatcct caacagttcg gtactgctgt ccgagagcgt ctcccttgtc gtcaagaccc 6120

accccggggg tcagaataag ccagtcctca gagtcgccct taggtcggtt ctgggcaatg 6180

aagccaacca caaactcggg gtcggatcgg gcaagctcaa tggtctgctt ggagtactcg 6240

ccagtggcca gagagcccct gcaagacagc tcggccagca tgagcagacc tctggccagc 6300

ttctcgttgg gagaggggac taggaactcc ttgtactggg agttctcgta gtcagagacg 6360

tcctccttct tctgttcaga gacagtttcc tcggcaccag ctcgcaggcc agcaatgatt 6420

ccggttccgg gtacaccgtg ggcgttggtg atatcggacc actcggcgat tcggtgacac 6480

cggtactggt gcttgacagt gttgccaata tctgcgaact ttctgtcctc gaacaggaag 6540

aaaccgtgct taagagcaag ttccttgagg gggagcacag tgccggcgta ggtgaagtcg 6600

tcaatgatgt cgatatgggt tttgatcatg cacacataag gtccgacctt atcggcaagc 6660

tcaatgagct ccttggtggt ggtaacatcc agagaagcac acaggttggt tttcttggct 6720

gccacgagct tgagcactcg agcggcaaag gcggacttgt ggacgttagc tcgagcttcg 6780

taggagggca ttttggtggt gaagaggaga ctgaaataaa tttagtctgc agaacttttt 6840

atcggaacct tatctggggc agtgaagtat atgttatggt aatagttacg agttagttga 6900

acttatagat agactggact atacggctat cggtccaaat tagaaagaac gtcaatggct 6960

ctctgggcgt cgcctttgcc gacaaaaatg tgatcatgat gaaagccagc aatgacgttg 7020

cagctgatat tgttgtcggc caaccgcgcc gaaaacgcag ctgtcagacc cacagcctcc 7080

aacgaagaat gtatcgtcaa agtgatccaa gcacactcat agttggagtc gtactccaaa 7140

ggcggcaatg acgagtcaga cagatactcg tcaaacggta ggttagtgct tggtatatga 7200

gttgtaggca tgacaatttg gaaaggggtg gactttggga atattgtggg atttcaatac 7260

cttagtttgt acagggtaat tgttacaaat gatacaaaga actgtatttc ttttcatttg 7320

ttttaattgg ttgtatatca agtccgttag acgagctcag tgccttggct tttggcactg 7380

tatttcattt ttagaggtac actacattca gtgaggtatg gtaaggttga gggcataatg 7440

aaggcacctt gtactgacag tcacagacct ctcaccgaga attttatgag atatactcgg 7500

gttcatttta ggctcatcga tcaggagaga ccgggttggc ggcgtatttg tgtcccaaaa 7560

aacagcccca attgccccaa ttgaccccaa attgacccag tagcgggccc aaccccggcg 7620

agagcccccc tcaccccaca tatcaaacct cccccggttc ccacacttgc cgttaagggc 7680

gtagggtact gcagtctgga atctacgctt gttcagactt tgtactagtt tctttgtctg 7740

gccatccggg taacccatgc cggacgcaaa atagactact gaaaattttt ttgctttgtg 7800

gttgggactt tagccaaggg tataaaagac caccgtcccc gaattacctt tcctcttctt 7860

ttctctctct ccttgtcaac tcacacccga aatcgttaag catttccttc tgagtataag 7920

aatcattcac 7930

<210> 207
<211> 8179
<212> DNA
<213> Artificial Sequence

<220>
<223> Plasmid pTEF::ISC1

<400> 207
gcggccgcaa gtgtggatgg ggaagtgagt gcccggttct gtgtgcacaa ttggcaatcc 60

aagatggatg gattcaacac agggatatag cgagctacgt ggtggtgcga ggatatagca 120

acggatattt atgtttgaca cttgagaatg tacgatacaa gcactgtcca agtacaatac 180

taaacatact gtacatactc atactcgtac ccgggcaacg gtttcacttg agtgcagtgg 240

ctagtgctct tactcgtaca gtgtgcaata ctgcgtatca tagtctttga tgtatatcgt 300

attcattcat gttagttgcg tacgggcgtc gttgcttgtg tgatttttga ggacccatcc 360

ctttggtata taagtatact ctggggttaa ggttgcccgt gtagtctagg ttatagtttt 420

catgtgaaat accgagagcc gagggagaat aaacgggggt atttggactt gttttttcg 480

cggaaaagcg tcgaatcaac cctgcgggcc ttgcaccatg tccacgacgt gtttctcgcc 540

ccaattcgcc ccttgcacgt caaaattagg cctccatcta gacccctcca taacatgtga 600

ctgtggggaa aagtataagg gaaaccatgc aaccatagac gacgtgaaag acggggagga 660

accaatggag gccaaagaaa tggggtagca acagtccagg agacagacaa ggagacaagg 720

agagggcgcc cgaaagatcg gaaaaacaaa catgtccaat tggggcagtg acggaaacga 780

cacggacact tcagtacaat ggaccgacca tctccaagcc agggttattc cggtatcacc 840

ttggccgtaa cctcccgctg gtacctgata ttgtacacgt tcacattcaa tatactttca 900

gctacaataa gagaggctgt ttgtcgggca tgtgtgtccg tcgtatgggg tgatgtccga 960

gggcgaaatt cgctacaagc ttaactctgg cgcttgtcca gtatgaatag acaagtcaag 1020

accagtggtg ccatgattga cagggaggta caagacttcg atactcgagc attactcgga 1080

cttgtggcga ttgaacagac gggcgatcgc ttctcccccg tattgccggc gcgccagctg 1140

cattaatgaa tcggccaacg cgcggggaga ggcggtttgc gtattgggcg ctcttccgct 1200

tcctcgctca ctgactcgct gcgctcggtc gttcggctgc ggcgagcggt atcagctcac 1260

tcaaaggcgg taatacggtt atccacagaa tcaggggata acgcaggaaa gaacatgtga 1320

gcaaaaggcc agcaaaaggc caggaaccgt aaaaaggccg cgttgctggc gtttttccat 1380

aggctccgcc cccctgacga gcatcacaaa aatcgacgct caagtcagag gtggcgaaac 1440

ccgacaggac tataaagata ccaggcgttt ccccctggaa gctccctcgt gcgctctcct 1500

gttccgaccc tgccgcttac cggatacctg tccgcctttc tcccttcggg aagcgtggcg 1560

ctttctcata gctcacgctg taggtatctc agttcggtgt aggtcgttcg ctccaagctg 1620

ggctgtgtgc acgaaccccc cgttcagccc gaccgctgcg ccttatccgg taactatcgt 1680

cttgagtcca acccggtaag acacgactta tcgccactgg cagcagccac tggtaacagg 1740

attagcagag cgaggtatgt aggcggtgct acagagttct tgaagtggtg gcctaactac 1800

ggctacacta gaagaacagt atttggtatc tgcgctctgc tgaagccagt taccttcgga 1860

aaaagagttg gtagctcttg atccggcaaa caaaccaccg ctggtagcgg tggttttttt 1920

gtttgcaagc agcagattac gcgcagaaaa aaaggatctc aagaagatcc tttgatcttt 1980

tctacggggt ctgacgctca gtggaacgaa aactcacgtt aagggatttt ggtcatgaga 2040

ttatcaaaaa ggatcttcac ctagatcctt ttaaattaaa aatgaagttt taaatcaatc 2100

taaagtatat atgagtaaac ttggtctgac agttaccaat gcttaatcag tgaggcacct 2160

atctcagcga tctgtctatt tcgttcatcc atagttgcct gactccccgt cgtgtagata 2220

actacgatac gggagggctt accatctggc cccagtgctg caatgatacc gcgagaccca 2280

cgctcaccgg ctccagattt atcagcaata aaccagccag ccggaagggc cgagcgcaga 2340

agtggtcctg caactttatc cgcctccatc cagtctatta attgttgccg ggaagctaga 2400

gtaagtagtt cgccagttaa tagtttgcgc aacgttgttg ccattgctac aggcatcgtg 2460

gtgtcacgct cgtcgtttgg tatggcttca ttcagctccg gttcccaacg atcaaggcga 2520

gttacatgat cccccatgtt gtgcaaaaaa gcggttagct ccttcggtcc tccgatcgtt 2580

gtcagaagta agttggccgc agtgttatca ctcatggtta tggcagcact gcataattct 2640

cttactgtca tgccatccgt aagatgcttt tctgtgactg gtgagtactc aaccaagtca 2700

ttctgagaat agtgtatgcg gcgaccgagt tgctcttgcc cggcgtcaat acgggataat 2760

accgcgccac atagcagaac tttaaaagtg ctcatcattg gaaaacgttc ttcggggcga 2820

aaactctcaa ggatcttacc gctgttgaga tccagttcga tgtaacccac tcgtgcaccc 2880

aactgatctt cagcatcttt tactttcacc agcgtttctg ggtgagcaaa aacaggaagg 2940

caaaatgccg caaaaaaggg aataagggcg acacggaaat gttgaatact catactcttc 3000

ctttttcaat attattgaag catttatcag ggttattgtc tcatgagcgg atacatattt 3060

gaatgtattt agaaaaataa acaaataggg gttccgcgca catttccccg aaaagtgcca 3120

cctgatgcgg tgtgaaatac cgcacagatg cgtaaggaga aaataccgca tcaggaaatt 3180

gtaagcgtta atattttgtt aaaattcgcg ttaaattttt gttaaatcag ctcatttttt 3240

aaccaatagg ccgaaatcgg caaaatccct tataaatcaa aagaatagac cgagataggg 3300

ttgagtgttg ttccagtttg gaacaagagt ccactattaa agaacgtgga ctccaacgtc 3360

aaagggcgaa aaaccgtcta tcagggcgat ggcccactac gtgaaccatc accctaatca 3420

agtttttttgg ggtcgaggtg ccgtaaagca ctaaatcgga accctaaagg gagcccccga 3480

tttagagctt gacggggaaa gccggcgaac gtggcgagaa aggaagggaa gaaagcgaaa 3540

ggagcgggcg ctagggcgct ggcaagtgta gcggtcacgc tgcgcgtaac caccacaccc 3600

gccgcgctta atgcgccgct acagggcgcg tccattcgcc attcaggctg cgcaactgtt 3660

gggaagggcg atcggtgcgg gcctcttcgc tattacgcca gctggcgaaa gggggatgtg 3720

ctgcaaggcg attaagttgg gtaacgccag ggttttccca gtcacgacgt tgtaaaacga 3780

cggccagtga attgtaatac gactcactat agggcgaatt gggcccgacg tcgcatgcgc 3840

tgatgacact ttggtctgaa agagatgcat tttgaatccc aaacttgcag tgcccaagtg 3900

acatacatct ccgcgttttg gaaaatgttc agaaacagtt gattgtgttg gaatggggaa 3960

tggggaatgg aaaaatgact caagtatcaa ttccaaaaac ttctctggct ggcagtacct 4020

actgtccata ctactgcatt ttctccagtc aggccactct atactcgacg acacagtagt 4080

aaaacccaga taatttcgac ataaacaaga aaacagaccc aataatattt atatatagtc 4140

agccgtttgt ccagttcaga ctgtaatagc cgaaaaaaaa tccaaagttt ctattctagg 4200

aaaatatatt ccaatatttt taattcttaa tctcatttat tttattctag cgaaatacat 4260

ttcagctact tgagacatgt gatacccaca aatcggattc ggactcggtt gttcagaaga 4320

gcatatggca ttcgtgctcg cttgttcacg tattcttcct gttccatctc ttggccgaca 4380

atcacacaaa aatggggttt ttttttaat tctaatgatt cattacagca aaattgagat 4440

atagcagacc acgtattcca taatcaccaa ggaagttctt gggcgtctta attaagtcat 4500

acacaagtca gctttcttcg agcctcatat aagtataagt agttcaacgt attagcactg 4560

tacccagcat ctccgtatcg agaaacacaa caacatgccc cattggacag atcatgcgga 4620

tacacaggtt gtgcagtatc atacatactc gatcagacag gtcgtctgac catcatacaa 4680

gctgaacaag cgctccatac ttgcacgctc tctatataca cagttaaatt acatatccat 4740

agtctaacct ctaacagtta atcttctggt aagcctccca gccagccttc tggtatcgct 4800

tggcctcctc aataggatct cggttctggc cgtacagacc tcggccgaca attatgatat 4860

ccgttccggt agacatgaca tcctcaacag ttcggtactg ctgtccgaga gcgtctccct 4920

tgtcgtcaag acccaccccg ggggtcagaa taagccagtc ctcagagtcg cccttaggtc 4980

ggttctgggc aatgaagcca accacaaact cggggtcgga tcgggcaagc tcaatggtct 5040

gcttggagta ctcgccagtg gccagagagc ccttgcaaga cagctcggcc agcatgagca 5100

gacctctggc cagcttctcg ttgggagagg ggactaggaa ctccttgtac tgggagttct 5160

cgtagtcaga gacgtcctcc ttcttctgtt cagagacagt ttcctcggca ccagctcgca 5220

ggccagcaat gattccggtt ccgggtacac cgtgggcgtt ggtgatatcg gaccactcgg 5280

cgattcggtg acaccggtac tggtgcttga cagtgttgcc aatatctgcg aactttctgt 5340

cctcgaacag gaagaaaccg tgcttaagag caagttcctt gaggggggagc acagtgccgg 5400

cgtaggtgaa gtcgtcaatg atgtcgatat gggttttgat catgcacaca taaggtccga 5460

ccttatcggc aagctcaatg agctccttgg tggtggtaac atccagagaa gcacacaggt 5520

tggttttctt ggctgccacg agcttgagca ctcgagcggc aaaggcggac ttgtggacgt 5580

tagctcgagc ttcgtaggag ggcattttgg tggtgaagag gagactgaaa taaatttagt 5640

ctgcagaact ttttatcgga accttatctg gggcagtgaa gtatatgtta tggtaatagt 5700

tacgagttag ttgaacttat agatagactg gactatacgg ctatcggtcc aaattagaaa 5760

gaacgtcaat ggctctctgg gcgtcgcctt tgccgacaaa aatgtgatca tgatgaaagc 5820

cagcaatgac gttgcagctg atattgttgt cggccaaccg cgccgaaaac gcagctgtca 5880

gacccacagc ctccaacgaa gaatgtatcg tcaaagtgat ccaagcacac tcatagttgg 5940

agtcgtactc caaaggcggc aatgacgagt cagacagata ctcgtcaaac ggtaggttag 6000

tgcttggtat atgagttgta ggcatgacaa tttggaaagg ggtggacttt gggaatattg 6060

tgggatttca ataccttagt ttgtacaggg taattgttac aaatgataca aagaactgta 6120

tttcttttca tttgttttaa ttggttgtat atcaagtccg ttagacgagc tcagtgcctt 6180

ggcttttggc actgtatttc atttttagag gtacactaca ttcagtgagg tatggtaagg 6240

ttgagggcat aatgaaggca ccttgtactg acagtcacag acctctcacc gagaatttta 6300

tgagatatac tcgggttcat tttaggctca tcgatcagga gagaccgggt tggcggcgta 6360

tttgtgtccc aaaaaacagc cccaattgcc ccaattgacc ccaaattgac ccagtagcgg 6420

gcccaacccc ggcgagagcc cccttcaccc cacatatcaa acctccccg gttcccacac 6480

ttgccgttaa gggcgtaggg tactgcagtc tggaatctac gcttgttcag actttgtact 6540

agtttctttg tctggccatc cgggtaaccc atgccggacg caaaatagac tactgaaaat 6600

ttttttgctt tgtggttggg actttagcca agggtataaa agaccaccgt ccccgaatta 6660

cctttcctct tcttttctct ctctccttgt caactcacac ccgaaatcgt taagcatttc 6720

cttctgagta taagaatcat tcaccatgga caacagaaaa gacagagatg ttcacgagag 6780

gaaggaagat ggtcaatctg agtttgaagc actgaacggg accaacgcaa ttatgtccga 6840

taatagtaaa gcgtattcca taaagtttct gaccttcaat acatgggggt taaaatacgt 6900

ctccaaacac cgtaaagaaa gactcagagc aattgctgat aaattggcgg gccactcaat 6960

gcttacgcca atatctgacg agttgttgcc caatggtgga gatagtaatg aaaacgaaga 7020

ttacgacgtg attgccttac aagaaatctg gtgtgtggaa gactggaagt atctagcttc 7080

tgcgtgtgcc tccaagtatc cgtatcagcg tttgttccat tctggtattc tgacggggcc 7140

tgggttggcc atactgtcca aggtcccgat agagtcgacc tttctttacc ggttcccgat 7200

aaacggtaga ccgagtgcgg tgttccgtgg cgactggtac gtagggaaat ctatagcaat 7260

caccgtattg aacacaggaa cccgccccat tgcaataatg aacagtcaca tgcacgcccc 7320

atacgccaag cagggtgatg ccgcctactt gtgccacaga tcttgtcagg cctgggattt 7380

cagcaggctc attaagcttt acaggcaggc cggttatgcg gtgattgtgg tgggtgactt 7440

aaactccaga ccgggctcac tgccccacaa atttctcacg caggaggccg gcctggtcga 7500

ctcctgggag caattgcatg ggaagcaaga cttggcggtg atcgctcgtc tgtctccatt 7560

gcaacaattg cttaagggct gtaccacgtg cgattcgctg ctcaacacat ggagggccca 7620

aagacaaccc gatgaggcat gcaggttgga ttatgctctt atcgaccctg atttcttgca 7680

aacagtagac gcaggtgtca ggttcactga acggatccct cacctggact gcagtgtctc 7740

tgaccatttt gcatactcat gcacccttaa catcgtccca cagggcacag agtcccgtcc 7800

atccacctcc gttaagcgtg cgaagactca tgatagagag ctgatcttgc agagatactc 7860

caactacgaa accatgatag aatgcatcca cacgtacttg aagacagccc aaagacagaa 7920

atttttccgt ggcctacatt tctgggcctc aatacttctc ctaatagcgt cgttggtcgt 7980

gacaacgttt actgcaaaca aggcaggctg gtcctccatc ttctgggtcc ttttcgctat 8040

tgctgtctcc atctcgggca ccatcgacgg tgccatctcc ttcttgtttg gcaggtctga 8100

aatcagagcc ctcatcgaag tcgaacaaga ggttctggac gcggagcacc acctgcaaac 8160

tttcttgagc gagaaatga 8179

<210> 208
<211> 8196
<212> DNA
<213> Artificial Sequence

<220>
<223> Plasmid pY20

<400> 208
tggagctcca gcttttgttc cctttagtga gggttaattt cgagcttggc gtaatcatgg 60

tcatagctgt ttcctgtgtg aaattgttat ccgctcacaa ttccacacaa catacgagcc 120

ggaagcataa agtgtaaagc ctggggtgcc taatgagtga gctaactcac attaattgcg 180

ttgcgctcac tgcccgcttt ccagtcggga aacctgtcgt gccagctgca ttaatgaatc 240

ggccaacgcg cggggagagg cggtttgcgt attgggcgct cttccgcttc ctcgctcact 300

gactcgctgc gctcggtcgt tcggctgcgg cgagcggtat cagctcactc aaaggcggta 360

atacggttat ccacagaatc aggggataac gcaggaaaga acatgtgagc aaaaggccag 420

caaaaggcca ggaaccgtaa aaaggccgcg ttgctggcgt ttttccatag gctccgcccc 480

cctgacgagc atcacaaaaa tcgacgctca agtcagaggt ggcgaaaccc gacaggacta 540

taaagatacc aggcgtttcc ccctggaagc tccctcgtgc gctctcctgt tccgaccctg 600

ccgcttaccg gatacctgtc cgcctttctc ccttcgggaa gcgtggcgct ttctcatagc 660

tcacgctgta ggtatctcag ttcggtgtag gtcgttcgct ccaagctggg ctgtgtgcac 720

gaaccccccg ttcagcccga ccgctgcgcc ttatccggta actatcgtct tgagtccaac 780

ccggtaagac acgacttatc gccactggca gcagccactg gtaacaggat tagcagagcg 840

aggtatgtag gcggtgctac agagttcttg aagtggtggc ctaactacgg ctacactaga 900

aggacagtat ttggtatctg cgctctgctg aagccagtta ccttcggaaa aagagttggt 960

agctcttgat ccggcaaaca aaccaccgct ggtagcggtg gtttttttgt ttgcaagcag 1020

cagattacgc gcagaaaaaa aggatctcaa gaagatcctt tgatcttttc tacggggtct 1080

gacgctcagt ggaacgaaaa ctcacgttaa gggattttgg tcatgagatt atcaaaaagg 1140

atcttcacct agatcctttt aaattaaaaa tgaagtttta aatcaatcta aagtatatat 1200

gagtaaactt ggtctgacag ttaccaatgc ttaatcagtg aggcacctat ctcagcgatc 1260

tgtctatttc gttcatccat agttgcctga ctccccgtcg tgtagataac tacgatacgg 1320

gagggcttac catctggccc cagtgctgca atgataccgc gagacccacg ctcaccggct 1380

ccagatttat cagcaataaa ccagccagcc ggaagggccg agcgcagaag tggtcctgca 1440

actttatccg cctccatcca gtctattaat tgttgccggg aagctagagt aagtagttcg 1500

ccagttaata gtttgcgcaa cgttgttgcc attgctacag gcatcgtggt gtcacgctcg 1560

tcgtttggta tggcttcatt cagctccggt tcccaacgat caaggcgagt tacatgatcc 1620

cccatgttgt gcaaaaaagc ggttagctcc ttcggtcctc cgatcgttgt cagaagtaag 1680

ttggccgcag tgttatcact catggttatg gcagcactgc ataattctct tactgtcatg 1740

ccatccgtaa gatgcttttc tgtgactggt gagtactcaa ccaagtcatt ctgagaatag 1800

tgtatgcggc gaccgagttg ctcttgcccg gcgtcaatac gggataatac cgcgccacat 1860

agcagaactt taaaagtgct catcattgga aaacgttctt cggggcgaaa actctcaagg 1920

atcttaccgc tgttgagatc cagttcgatg taacccactc gtgcacccaa ctgatcttca 1980

gcatctttta ctttcaccag cgtttctggg tgagcaaaaa caggaaggca aaatgccgca 2040

aaaaagggaa taagggcgac acggaaatgt tgaatactca tactcttcct ttttcaatat 2100

tattgaagca tttatcaggg ttattgtctc atgagcggat acatatttga atgtatttag 2160

aaaaataaac aaataggggt tccgcgcaca tttccccgaa aagtgccacc tgacgcgccc 2220

tgtagcggcg cattaagcgc ggcgggtgtg gtggttacgc gcagcgtgac cgctacactt 2280

gccagcgccc tagcgcccgc tcctttcgct ttcttccctt cctttctcgc cacgttcgcc 2340

ggctttcccc gtcaagctct aaatcggggg ctccctttag ggttccgatt tagtgcttta 2400

cggcacctcg accccaaaaa acttgattag ggtgatggtt cacgtagtgg gccatcgccc 2460

tgatagacgg tttttcgccc tttgacgttg gagtccacgt tctttaatag tggactcttg 2520

ttccaaactg gaacaacact caaccctatc tcggtctatt cttttgattt ataagggatt 2580

ttgccgattt cggcctattg gttaaaaaat gagctgattt aacaaaaatt taacgcgaat 2640

tttaacaaaa tattaacgct tacaatttcc attcgccatt caggctgcgc aactgttggg 2700

aagggcgatc ggtgcgggcc tcttcgctat tacgccagct ggcgaaaggg ggatgtgctg 2760

caaggcgatt aagttgggta acgccagggt tttcccagtc acgacgttgt aaaacgacgg 2820

ccagtgaatt gtaatacgac tcactatagg gcgaattggg taccgggccc cccctcgagg 2880

tcgacggtat cgataagctt gatatcgaat tcatgtcaca caaaccgatc ttcgcctcaa 2940

ggaaacctaa ttctacatcc gagagactgc cgagatccag tctacactga ttaattttcg 3000

ggccaataat ttaaaaaaat cgtgttatat aatattatat gtattatata tatacatcat 3060

gatgatactg acagtcatgt cccattgcta aatagacaga ctccatctgc cgcctccaac 3120

tgatgttctc aatatttaag gggtcatctc gcattgttta ataataaaca gactccatct 3180

accgcctcca aatgatgttc tcaaaatata ttgtatgaac ttatttttat tacttagtat 3240

tattagacaa cttacttgct ttatgaaaaa cacttcctat ttaggaaaca atttataatg 3300

gcagttcgtt catttaacaa tttatgtaga ataaatgtta taaatgcgta tgggaaatct 3360

taaatatgga tagcataaat gatatctgca ttgcctaatt cgaaatcaac agcaacgaaa 3420

aaaatccctt gtacaacata aatagtcatc gagaaatatc aactatcaaa gaacagctat 3480

tcacacgtta ctattgagat tattattgga cgagaatcac acactcaact gtctttctct 3540

cttctagaaa tacaggtaca agtatgtact attctcattg ttcatacttc tagtcatttc 3600

atcccacata ttccttggat ttctctccaa tgaatgacat tctatcttgc aaattcaaca 3660

attataataa gatataccaa agtagcggta tagtggcaat caaaaagctt ctctggtgtg 3720

cttctcgtat ttatttttat tctaatgatc cattaaaggt atatatttat ttcttgttat 3780

ataatccttt tgtttattac atgggctgga tacataaagg tattttgatt taatttttg 3840

cttaaattca atccccccctc gttcagtgtc aactgtaatg gtaggaaatt accatacttt 3900

tgaagaagca aaaaaaatga aagaaaaaaa aaatcgtatt tccaggttag acgttccgca 3960

gaatctagaa tgcggtatgc ggtacattgt tcttcgaacg taaaagttgc gctccctgag 4020

atattgtaca tttttgcttt tacaagtaca agtacatcgt acaactatgt actactgttg 4080

atgcatccac aacagtttgt tttgtttttt tttgtttttt ttttttctaa tgattcatta 4140

ccgctatgta tacctacttg tacttgtagt aagccgggtt attggcgttc aattaatcat 4200

agacttatga atctgcacgg tgtgcgctgc gagttacttt tagcttatgc atgctacttg 4260

ggtgtaatat tgggatctgt tcggaaatca acggatgctc aaccgatttc gacagtaata 4320

atttgaatcg aatcggagcc taaaatgaac ccgagtatat ctcataaaat tctcggtgag 4380

aggtctgtga ctgtcagtac aaggtgcctt cattatgccc tcaaccttac catacctcac 4440

tgaatgtagt gtacctctaa aaatgaaata cagtgccaaa agccatggca ctgagctcgt 4500

ctaacggact tgatatacaa ccaattaaaa caaatgaaaa gaaatacagt tctttgtatc 4560

atttgtaaca attaccctgt acaaactaag gtattgaaat cccacaatat tcccaaagtc 4620

cacccctttc caaattgtca tgcctacaac tcatatacca agcactaacc taccaaacac 4680

cactaaaacc ccacaaaata tatcttaccg aatatacagt aacaagctac caccacactc 4740

gttgggtgca gtcgccagct taaagatatc tatccacatc agccacaact cccttccttt 4800

aataaaccga ctacaccctt ggctattgag gttatgagtg aatatactgt agacaagaca 4860

ctttcaagaa gactgtttcc aaaacgtacc actgtcctcc actacaaaca cacccaatct 4920

gcttcttcta gtcaaggttg ctacaccggt aaattataaa tcatcatttc attagcaggg 4980

cagggcccttt tttatagagt cttatacact agcggaccct gccggtagac caacccgcag 5040

gcgcgtcagt ttgctccttc catcaatgcg tcgtagaaac gacttactcc ttcttgagca 5100

gctccttgac cttgttggca acaagtctcc gacctcggag gtggaggaag agcctccgat 5160

atcggcggta gtgataccag cctcgacgga ctccttgacg gcagcctcaa cagcgtcacc 5220

ggcgggcttc atgttaagag agaacttgag catcatggcg gcagacagaa tggtggcaat 5280

ggggttgacc ttctgcttgc cgagatcggg ggcagatccg tgacagggct cgtacagacc 5340

gaacgcctcg ttggtgtcgg gcagagaagc cagagaggcg gagggcagca gacccagaga 5400

accggggatg acggaggcct cgtcggagat gatatcgcca aacatgttgg tggtgatgat 5460

gataccattc atcttggagg gctgcttgat gaggatcatg gcggccgagt cgatcagctg 5520

gtggttgagc tcgagctggg ggaattcgtc cttgaggact cgagtgacag tctttcgcca 5580

aagtcgagag gaggccagca cgttggcctt gtcaagagac cacacgggaa gaggggggtt 5640

gtgctgaagg gccaggaagg cggccattcg ggcaattcgc tcaacctcag gaacggagta 5700

ggtctcggtg tcggaagcga cgccagatcc gtcatcctcc tttcgctctc caaagtagat 5760

acctccgacg agctctcgga caatgatgaa gtcggtgccc tcaacgtttc ggatggggga 5820

gagatcggcg agcttgggcg acagcagctg gcagggtcgc aggttggcgt acaggttcag 5880

gtcctttcgc agcttgagga gaccctgctc gggtcgcacg tcggttcgtc cgtcgggagt 5940

ggtccatacg gtgttggcag cgcctccgac agcaccgagc ataatagagt cagcctttcg 6000

gcagatgtcg agagtagcgt cggtgatggg ctcgccctcc ttctcaatgg cagctcctcc 6060

aatgagtcgg tcctcaaaca caaactcggt gccggaggcc tcagcaacag acttgagcac 6120

cttgacggcc tcggcaatca cctcggggcc acagaagtcg ccgccgagaa gaacaatctt 6180

cttggagtca gtcttggtct tcttagtttc gggttccatt gtggatgtgt gtggttgtat 6240

gtgtgatgtg gtgtgtggag tgaaaatctg tggctggcaa acgctcttgt atatatacgc 6300

actttttgccc gtgctatgtg gaagactaaa cctccgaaga ttgtgactca ggtagtgcgg 6360

tatcggctag ggacccaaac cttgtcgatg ccgatagcgc tatcgaacgt accccagccg 6420

gccgggagta tgtcggaggg gacatacgag atcgtcaagg gtttgtggcc aactggtaaa 6480

taaatgatga ctcaggcgac gacggaattc ctgcagccca tctgcagaat tcaggagaga 6540

ccgggttggc ggcgtatttg tgtcccaaaa aacagcccca attgccccaa ttgaccccaa 6600

attgacccag tagcgggccc aaccccggcg agagcccccc tcaccccaca tatcaaacct 6660

cccccggttc ccacacttgc cgttaagggc gtagggtact gcagtctgga atctacgctt 6720

gttcagactt tgtactagtt tctttgtctg gccatccggg taacccatgc cggacgcaaa 6780

atagactact gaaaattttt ttgctttgtg gttgggactt tagccaaggg tataaaagac 6840

caccgtcccc gaattacctt tcctcttctt ttctctctct ccttgtcaac tcacacccga 6900

aatcgttaag catttccttc tgagtataag aatcattcaa aggatccact agttctagag 6960

cggccgctta aaccatgaaa aagcctgaac tcaccgcgac gtctgtcgag aagtttctga 7020

tcgaaaagtt cgacagcgtc tccgacctga tgcagctctc ggagggcgaa gaatctcgtg 7080

ctttcagctt cgatgtagga gggcgtggat atgtcctgcg ggtaaatagc tgcgccgatg 7140

gtttctacaa agatcgttat gtttatcggc actttgcatc ggccgcgctc ccgattccgg 7200

aagtgcttga cattggggaa ttcagcgaga gcctgaccta ttgcatctcc cgccgtgcac 7260

agggtgtcac gttgcaagac ctgcctgaaa ccgaactgcc cgctgttctg cagccggtcg 7320

cggaggccat ggatgcgatc gctgcggccg atcttagcca gacgagcggg ttcggcccat 7380

tcggaccgca aggaatcggt caatacacta catggcgtga tttcatatgc gcgattgctg 7440

atccccatgt gtatcactgg caaactgtga tggacgacac cgtcagtgcg tccgtcgcgc 7500

aggctctcga tgagctgatg ctttgggccg aggactgccc cgaagtccgg cacctcgtgc 7560

acgcggattt cggctccaac aatgtcctga cggacaatgg ccgcataaca gcggtcattg 7620

actggagcga ggcgatgttc ggggattccc aatacgaggt cgccaacatc ttcttctgga 7680

ggccgtggtt ggcttgtatg gagcagcaga cgcgctactt cgagcggagg catccggagc 7740

ttgcaggatc gccgcggctc cgggcgtata tgctccgcat tggtcttgac caactctatc 7800

agagcttggt tgacggcaat ttcgatgatg cagcttgggc gcagggtcga tgcgacgcaa 7860

tcgtccgatc cggagccggg actgtcgggc gtacacaaat cgcccgcaga agcgcggccg 7920

tctggaccga tggctgtgta gaagtactcg ccgatagtgg aaaccgacgc cccagcactc 7980

gtccgagggc aaaggaatag tgaggtacct aaagcggccg ccaccgcggc ccgagattcc 8040

ggcctcttcg gccgccaagc gacccgggtg gacgtctaga ggtacctagc aattaacaga 8100

tagtttgccg gtgataattc tcttaacctc ccacactcct ttgacataac gatttatgta 8160

acgaaactga aatttgacca gatattgtgt ccgcgg 8196

<210> 209
<211> 5105
<212> DNA
<213> Artificial Sequence

<220>
<223> Plasmid pLV13

<220>
<221> misc_feature
<222> (4446)..(4446)
<223> n is a, c, g, or t

<400> 209
gtacgagccg gaagcataaa gtgtaaagcc tggggtgcct aatgagtgag ctaactcaca 60

ttaattgcgt tgcgctcact gcccgctttc cagtcgggaa acctgtcgtg ccagctgcat 120

taatgaatcg gccaacgcgc ggggagaggc ggtttgcgta ttgggcgctc ttccgcttcc 180

tcgctcactg actcgctgcg ctcggtcgtt cggctgcggc gagcggtatc agctcactca 240

aaggcggtaa tacggttatc cacagaatca ggggataacg caggaaagaa catgtgagca 300

aaaggccagc aaaaggccag gaaccgtaaa aaggccgcgt tgctggcgtt tttccatagg 360

ctccgccccc ctgacgagca tcacaaaaat cgacgctcaa gtcagaggtg gcgaaacccg 420

acaggactat aaagatacca ggcgtttccc cctggaagct ccctcgtgcg ctctcctgtt 480

ccgaccctgc cgcttaccgg atacctgtcc gcctttctcc cttcgggaag cgtggcgctt 540

tctcatagct cacgctgtag gtatctcagt tcggtgtagg tcgttcgctc caagctgggc 600

tgtgtgcacg aaccccccgt tcagcccgac cgctgcgcct tatccggtaa ctatcgtctt 660

gagtccaacc cggtaagaca cgacttatcg ccactggcag cagccactgg taacaggatt 720

agcagagcga ggtatgtagg cggtgctaca gagttcttga agtggtggcc taactacggc 780

tacactagaa ggacagtatt tggtatctgc gctctgctga agccagttac cttcggaaaa 840

agagttggta gctcttgatc cggcaaacaa accaccgctg gtagcggtgg ttttttttgtt 900

tgcaagcagc agattacgcg cagaaaaaaa ggatctcaag aagatccttt gatcttttct 960

acggggtctg acgctcagtg gaacgaaaac tcacgttaag ggattttggt catgagatta 1020

tcaaaaagga tcttcaccta gatcctttta aattaaaaat gaagttttaa atcaatctaa 1080

agtatatatg agtaaacttg gtctgacagt taccaatgct taatcagtga ggcacctatc 1140

tcagcgatct gtctatttcg ttcatccata gttgcctgac tccccgtcgt gtagataact 1200

acgatacggg agggcttacc atctggcccc agtgctgcaa tgataccgcg agacccacgc 1260

tcaccggctc cagatttatc agcaataaac cagccagccg gaagggccga gcgcagaagt 1320

ggtcctgcaa ctttatccgc ctccatccag tctattaatt gttgccggga agctagagta 1380

agtagttcgc cagttaatag tttgcgcaac gttgttgcca ttgctacagg catcgtggtg 1440

tcacgctcgt cgtttggtat ggcttcattc agctccggtt cccaacgatc aaggcgagtt 1500

acatgatccc ccatgttgtg caaaaaagcg gttagctcct tcggtcctcc gatcgttgtc 1560

agaagtaagt tggccgcagt gttatcactc atggttatgg cagcactgca taattctctt 1620

actgtcatgc catccgtaag atgcttttct gtgactggtg agtactcaac caagtcattc 1680

tgagaatagt gtatgcggcg accgagttgc tcttgcccgg cgtcaatacg ggataatacc 1740

gcgccacata gcagaacttt aaaagtgctc atcattggaa aacgttcttc ggggcgaaaa 1800

ctctcaagga tcttaccgct gttgagatcc agttcgatgt aacccactcg tgcacccaac 1860

tgatcttcag catcttttac tttcaccagc gtttctgggt gagcaaaaac aggaaggcaa 1920

aatgccgcaa aaaagggaat aagggcgaca cggaaatgtt gaatactcat actcttcctt 1980

tttcaatatt attgaagcat ttatcagggt tattgtctca tgagcggata catatttgaa 2040

tgtatttaga aaaataaaca aatagggggtt ccgcgcacat ttccccgaaa agtgccacct 2100

gacgcgccct gtagcggcgc attaagcgcg gcgggtgtgg tggttacgcg cagcgtgacc 2160

gctacacttg ccagcgccct agcgcccgct cctttcgctt tcttcccttc ctttctcgcc 2220

acgttcgccg gctttccccg tcaagctcta aatcgggggc tcccttttagg gttccgattt 2280

agtgctttac ggcacctcga ccccaaaaaa cttgattagg gtgatggttc acgtagtggg 2340

ccatcgccct gatagacggt ttttcgccct ttgacgttgg agtccacgtt ctttaatagt 2400

ggactcttgt tccaaactgg aacaacactc aaccctatct cggtctattc ttttgattta 2460

taagggattt tgccgatttc ggcctattgg ttaaaaaatg agctgattta acaaaaattt 2520

aacgcgaatt ttaacaaaat attaacgctt acaatttcca ttcgccattc aggctgcgca 2580

actgttggga agggcgatta agtcatacac aagtcagctt tcttcgagcc tcatataagt 2640

ataagtagtt caacgtatta gcactgtacc cagcatctcc gtatcgagaa acacaacaac 2700

atgccccatt ggacagatca tgcggataca caggttgtgc agtatcatac atactcgatc 2760

agacaggtcg tctgaccatc atacaagctg aacaagcgct ccatacttgc acgctctcta 2820

tatacacagt taaattacat atccatagtc taacctctaa cagttaatct tctggtaagc 2880

ctcccagcca gccttctggt atcgcttggc ctcctcaata ggatctcggt tctggccgta 2940

cagacctcgg ccgacaatta tgatatccgt tccggtagac atgacatcct caacagttcg 3000

gtactgctgt ccgagagcgt ctcccttgtc gtcaagaccc accccggggg tcagaataag 3060

ccagtcctca gagtcgccct taggtcggtt ctgggcaatg aagccaacca caaactcggg 3120

gtcggatcgg gcaagctcaa tggtctgctt ggagtactcg ccagtggcca gagagccctt 3180

gcaagacagc tcggccagca tgagcagacc tctggccagc ttctcgttgg gagaggggac 3240

taggaactcc ttgtactggg agttctcgta gtcagagacg tcctccttct tctgttcaga 3300

gacagtttcc tcggcaccag ctcgcaggcc agcaatgatt ccggttccgg gtacaccgtg 3360

ggcgttggtg atatcggacc actcggcgat tcggtgacac cggtactggt gcttgacagt 3420

gttgccaata tctgcgaact ttctgtcctc gaacaggaag aaaccgtgct taagagcaag 3480

ttccttgagg gggagcacag tgccggcgta ggtgaagtcg tcaatgatgt cgatatgggt 3540

tttgatcatg cacacataag gtccgacctt atcggcaagc tcaatgagct ccttggtggt 3600

ggtaacatcc agagaagcac acaggttggt tttcttggct gccacgagct tgagcactcg 3660

agcggcaaag gcggacttgt ggacgttagc tcgagcttcg taggagggca ttttggtggt 3720

gaagaggaga ctgaaataaa tttagtctgc agaacttttt atcggaacct tatctggggc 3780

agtgaagtat atgttatggt aatagttacg agttagttga acttatagat agactggact 3840

atacggctat cggtccaaat tagaaagaac gtcaatggct ctctgggcgt cgcctttgcc 3900

gacaaaaatg tgatcatgat gaaagccagc aatgacgttg cagctgatat tgttgtcggc 3960

caaccgcgcc gaaaacgcag ctgtcagacc cacagcctcc aacgaagaat gtatcgtcaa 4020

agtgatccaa gcacactcat agttggagtc gtactccaaa ggcggcaatg acgagtcaga 4080

cagatactcg tcgactcagg cgacgacgga attcagcttt gagcttggag aagtatccgt 4140

ctcggtgctc caaatcaggg taggacagtc tccagtcgta cgccgcagca gacattgtat 4200

ccgtatcgta tccaataaca gccaggttct cgagcagctt gttccacagc cagtagcctg 4260

ccataaagaa gtcggcggag gcaaatccct gggcggctcg cagcttgaaa tggggagggt 4320

ctagaccggt ctctgtgtcc agcatcaggt tctgcagcca gcagtacttg tccagcagca 4380

tgactcggat catgtaccag gagccccaca ttcgctttct gaagtgcgac tcggtgggac 4440

actccnyggt tccctccagg gaccagctct ccagtcccgt ggagatgacg ccgggcacca 4500

gcaccaccgg gtacttggcg ttgagtccct cgctcttcat agccttgccc acagcaaaag 4560

gggcagcctt ctgttcgctt ttctggtgct tctcgatctt gagatctaga atacagaagt 4620

caacggttca tccatggtga ggtcttagca gagtactcgg tcaagactcc ttgaacttga 4680

tgaagttgcc aaaggtgaca ttctggccgg gctcgtcatc aggggctcca gaatgatcac 4740

cccagatagc cttaccaccc ttgggaatca tggaagcaat tcctccccat gttcgcagca 4800

gatcggctcg ctctcgtcga gagaagaact gctccagtcc atacacagcc atcgcgttca 4860

gctgcacggt atccttcatt tctccagaca gaagagcaac cagggtcttg ggagtaccca 4920

gcatggagcc ggaaatgtcg acaaaggatt caatatggtc attgacccag ttgggacctc 4980

ctcctccata tccctcggcc tcagcccact tcatgaagta gaagatgacc tgggagccca 5040

tggaatggcc cgtcagaact gtcttctcac ctgtcatacg cttagtctct tcgattgagg 5100

atccg 5105

<210> 210
<211> 971
<212> DNA
<213> Yarrowia lipolytica

<220>
<221> misc_feature

<223> Promoter GPD

<400> 210
gacgcagtag gatgtcctgc acgggtcttt ttgtggggtg tggagaaagg ggtgcttgga 60

gatggaagcc ggtagaaccg ggctgcttgt gcttggagat ggaagccggt agaaccgggc 120

tgcttggggg gatttggggc cgctgggctc caaagagggg taggcatttc gttggggtta 180

cgtaattgcg gcatttgggt cctgcgcgca tgtcccattg gtcagaatta gtccggatag 240

gagacttatc agccaatcac agcgccggat ccacctgtag gttgggttgg gtgggagcac 300

ccctccacag agtagagtca aacagcagca gcaacatgat agttgggggt gtgcgtgtta 360

aaggaaaaaa aagaagcttg ggttatattc ccgctctatt tagaggttgc gggatagacg 420

ccgacggagg gcaatggcgc catggaacct tgcggatatc gatacgccgc ggcggactgc 480

gtccgaacca gctccagcag cgtttttttcc gggccattga gccgactgcg accccgccaa 540

cgtgtcttgg cccacgcact catgtcatgt tggtgttggg aggccacttt ttaagtagca 600

caaggcacct agctcgcagc aaggtgtccg aaccaaagaa gcggctgcag tggtgcaaac 660

ggggcggaaa cggcgggaaa aagccacggg ggcacgaatt gaggcacgcc ctcgaatttg 720

agacgagtca cggccccatt cgcccgcgca atggctcgcc aacgcccggt cttttgcacc 780

acatcaggtt accccaagcc aaacctttgt gttaaaaagc ttaacatatt ataccgaacg 840

taggtttggg cgggcttgct ccgtctgtcc aaggcaacat ttatataagg gtctgcatcg 900

ccggctcaat tgaatctttt ttcttcttct cttctctata ttcattcttg aattaaacac 960

acatcaacat g 971

<210> 211
<211> 1174
<212> DNA
<213> Yarrowia lipolytica

<220>
<221> misc_feature
<223> Promoter GPDIN

<400> 211
aaacacgcag taggatgtcc tgcacgggtc tttttgtggg gtgtggagaa aggggtgctt 60

ggagatggaa gccggtagaa ccgggctgct tgtgcttgga gatggaagcc ggtagaaccg 120

ggctgcttgg ggggatttgg ggccgctggg ctccaaagag gggtaggcat ttcgttgggg 180

ttacgtaatt gcggcatttg ggtcctgcgc gcatgtccca ttggtcagaa ttagtccgga 240

taggagactt atcagccaat cacagcgccg gatccacctg taggttgggt tgggtgggag 300

caccccctcca cagagtagag tcaaacagca gcagcaacat gatagttggg ggtgtgcgtg 360

ttaaaggaaa aaaaagaagc ttgggttata ttcccgctct atttagaggt tgcgggatag 420

acgccgacgg agggcaatgg cgctatggaa ccttgcggat atccatacgc cgcggcggac 480

tgcgtccgaa ccagctccag cagcgttttt tccgggccat tgagccgact gcgaccccgc 540

caacgtgtct tggcccacgc actcatgtca tgttggtgtt gggaggccac tttttaagta 600

gcacaaggca cctagctcgc agcaaggtgt ccgaaccaaa gaagcggctg cagtggtgca 660

aacggggcgg aaacggcggg aaaaagccac gggggcacga attgaggcac gccctcgaat 720

ttgagacgag tcacggcccc attcgcccgc gcaatggctc gccaacgccc ggtcttttgc 780

accacatcag gttaccccaa gccaaacctt tgtgttaaaa agcttaacat attataccga 840

acgtaggttt gggcgggctt gctccgtctg tccaaggcaa catttatata agggtctgca 900

tcgccggctc aattgaatct tttttcttct tctcttctct atattcattc ttgaattaaa 960

cacacatcaa catggccatc aaagtcggta ttaacggatt cgggcgaatc ggacgaattg 1020

tgagtaccat agaaggtgat ggaaacatga cccaacagaa acagatgaca agtgtcatcg 1080

acccaccaga gcccaattga gctcatacta acagtcgaca acctgtcgaa ccaattgatg 1140

actccccgac aatgtactaa cacaggtcct gccc 1174

<210> 212
<211> 878
<212> DNA
<213> Yarrowia lipolytica

<220>
<221> misc_feature
<223> Promoter GPM

<400> 212
gcctctgaat actttcaaca agttacaccc ttcattaatt ctcacgtgac acagattatt 60

aacgtctcgt accaaccaca gattacgacc cattcgcagt cacagttcac tagggtttgg 120

gttgcatccg ttgagagcgg tttgttttta accttctcca tgtgctcact caggttttgg 180

gttcagatca aatcaaggcg tgaaccactt tgtttgagga caaatgtgac acaaccaacc 240

agtgtcaggg gcaagtccgt gacaaagggg aagatacaat gcaattactg acagttacag 300

actgcctcga tgccctaacc ttgccccaaa ataagacaac tgtcctcgtt taagcgcaac 360

cctattcagc gtcacgtcat aatagcgttt ggatagcact agtctatgag gagcgtttta 420

tgttgcggtg agggcgattg gtgctcatat gggttcaatt gaggtggcgg aacgagctta 480

gtcttcaatt gaggtgcgag cgacacaatt gggtgtcacg tggcctaatt gacctcgggt 540

cgtggagtcc ccagttatac agcaaccacg aggtgcatgg gtaggagacg tcaccagaca 600

atagggtttt ttttggactg gagagggttg ggcaaaagcg ctcaacgggc tgtttgggga 660

gctgtggggg aggaattggc gatatttgtg aggttaacgg ctccgatttg cgtgttttgt 720

cgctcctgca tctccccata cccatatctt ccctccccac ctctttccac gataatttta 780

cggatcagca ataaggttcc ttctcctagt ttccacgtcc atatatatct atgctgcgtc 840

gtccttttcg tgacatcacc aaaacacata caaaaatg 878

<210> 213
<211> 1001
<212> DNA
<213> Yarrowia lipolytica

<220>
<221> misc_feature
<223> Promoter FBA

<400> 213
taaacagtgt acgcagtact atagaggaac aattgccccg gagaagacgg ccaggccgcc 60

tagatgacaa attcaacaac tcacagctga ctttctgcca ttgccactag gggggggcctt 120

tttatatggc caagccaagc tctccacgtc ggttgggctg cacccaacaa taaatgggta 180

gggttgcacc aacaaaggga tgggatgggg ggtagaagat acgaggataa cggggctcaa 240

tggcacaaat aagaacgaat actgccatta agactcgtga tccagcgact gacaccattg 300

catcatctaa gggcctcaaa actacctcgg aactgctgcg ctgatctgga caccacagag 360

gttccgagca ctttaggttg caccaaatgt cccaccaggt gcaggcagaa aacgctggaa 420

cagcgtgtac agtttgtctt aacaaaaagt gagggcgctg aggtcgagca gggtggtgtg 480

acttgttata gcctttagag ctgcgaaagc gcgtatggat ttggctcatc aggccagatt 540

gagggtctgt ggacacatgt catgttagtg tacttcaatc gccccctgga tatagccccg 600

acaataggcc gtggcctcat ttttttgcct tccgcacatt tccattgctc ggtacccaca 660

ccttgcttct cctgcacttg ccaaccttaa tactggttta cattgaccaa catcttacaa 720

gcgggggggct tgtctagggt atatataaac agtggctctc ccaatcggtt gccagtctct 780

tttttccttt ctttccccac agattcgaaa tctaaactac acatcacaca atgcctgtta 840

ctgacgtcct taagcgaaag tccggtgtca tcgtcggcga cgatgtccga gccgtgagta 900

tccacgacaa gatcagtgtc gagacgacgc gttttgtgta atgacacaat ccgaaagtcg 960

ctagcaacac acactctcta cacaaactaa cccagctctc c 1001

<210> 214
<211> 973
<212> DNA
<213> Yarrowia lipolytica

<220>
<221> misc_feature
<223> Promoter FBAIN

<400> 214
aaattgcccc ggagaagacg gccaggccgc ctagatgaca aattcaacaa ctcacagctg 60

actttctgcc attgccacta gggggggggcc tttttatatg gccaagccaa gctctccacg 120

tcggttgggc tgcacccaac aataaatggg tagggttgca ccaacaaagg gatgggatgg 180

ggggtagaag atacgaggat aacggggctc aatggcacaa ataagaacga atactgccat 240

taagactcgt gatccagcga ctgacaccat tgcatcatct aagggcctca aaactacctc 300

ggaactgctg cgctgatctg gacaccacag aggttccgag cactttaggt tgcaccaaat 360

gtcccaccag gtgcaggcag aaaacgctgg aacagcgtgt acagtttgtc ttaacaaaaa 420

gtgagggcgc tgaggtcgag cagggtggtg tgacttgtta tagcctttag agctgcgaaa 480

gcgcgtatgg atttggctca tcaggccaga ttgagggtct gtggacacat gtcatgttag 540

tgtacttcaa tcgccccctg gatatagccc cgacaatagg ccgtggcctc atttttttgc 600

cttccgcaca tttccattgc tcggtaccca caccttgctt ctcctgcact tgccaacctt 660

aatactggtt tacattgacc aacatcttac aagcgggggg cttgtctagg gtatatataa 720

acagtggctc tcccaatcgg ttgccagtct ctttttttcct ttctttcccc acagattcga 780

aatctaaact acacatcaca caatgcctgt tactgacgtc cttaagcgaa agtccggtgt 840

catcgtcggc gacgatgtcc gagccgtgag tatccacgac aagatcagtg tcgagacgac 900

gcgttttgtg taatgacaca atccgaaagt cgctagcaac acacactctc tacacaaact 960

aacccagctc tcc 973

<210> 215
<211> 924
<212> DNA
<213> Yarrowia lipolytica

<220>
<221> misc_feature
<223> Promoter FBAINm

<400> 215
aaattgcccc ggagaagacg gccaggccgc ctagatgaca aattcaacaa ctcacagctg 60

actttctgcc attgccacta gggggggcc tttttatatg gccaagccaa gctctccacg 120

tcggttgggc tgcacccaac aataaatggg tagggttgca ccaacaaagg gatgggatgg 180

ggggtagaag atacgaggat aacggggctc aatggcacaa ataagaacga atactgccat 240

taagactcgt gatccagcga ctgacaccat tgcatcatct aagggcctca aaactacctc 300

ggaactgctg cgctgatctg gacaccacag aggttccgag cactttaggt tgcaccaaat 360

gtcccaccag gtgcaggcag aaaacgctgg aacagcgtgt acagtttgtc ttaacaaaaa 420

gtgagggcgc tgaggtcgag cagggtggtg tgacttgtta tagcctttag agctgcgaaa 480

gcgcgtatgg atttggctca tcaggccaga ttgagggtct gtggacacat gtcatgttag 540

tgtacttcaa tcgccccctg gatatagccc cgacaatagg ccgtggcctc atttttttgc 600

cttccgcaca tttccattgc tcggtaccca caccttgctt ctcctgcact tgccaacctt 660

aatactggtt tacattgacc aacatcttac aagcgggggg cttgtctagg gtatatataa 720

acagtggctc tcccaatcgg ttgccagtct ctttttttcct ttctttcccc acagattcga 780

aatctaaact acacatcaca gaattccgag ccgtgagtat ccacgacaag atcagtgtcg 840

agacgacgcg ttttgtgtaa tgacacaatc cgaaagtcgc tagcaacaca cactctctac 900

acaaactaac ccagctctgg tacc 924

<210> 216
<211> 1130
<212> DNA
<213> Yarrowia lipolytica

<400> 216
caacttttct tgtcgacctg agataccgag gttgcgcagg ggatcaactt ttgtgtctca 60

gagggaccca agtgcgtacg gagagtacag tacatactgt agctaacggt agcaggcgaa 120

ctactggtac atacctcccc cggaatatgt acaggcataa tgcgtatctg tgggacatgt 180

ggtcgttgcg ccattatgta agcagcgtgt actcctctga ctgtccatat ggtttgctcc 240

atctcaccct catcgttttc attgttcaca ggcggccaca aaaaaactgt cttctctcct 300

tctctcttcg ccttagtcta ctcggaccag ttttagttta gcttggcgcc actggataaa 360

tgagacctca ggccttgtga tgaggaggtc acttatgaag catgttagga ggtgcttgta 420

tggatagaga agcacccaaa ataataagaa taataataaa acaggggggcg ttgtcatttc 480

atatcgtgtt ttcaccatca atacacctcc aaacaatgcc cttcatgtgg ccagccccaa 540

tattgtcctg tagttcaact ctatgcagct cgtatcttat tgagcaagta aaactctgtc 600

agccgatatt gcccgacccg cgacaagggt caacaaggtg gtgtaaggcc ttcgcagaag 660

tcaaaactgt gccaaacaaa catctagagt ctctttggtg tttctcgcat atatttaatc 720

ggctgtctta cgtatttggc ctcggtaccg gactaatttc ggatcatccc caatacgctt 780

tttcttcgca gctgtcaaca gtgtccatga tctatccacc taaatgggtc atatgaggcg 840

tataatttcg tggtgctgat aataattccc atatatttga cacaaaactt ccccccctag 900

acatacatct cacaatctca cttcttgtgc ttctgtcaca catctcctcc agctgacttc 960

aactcacacc tctgccccag ttggtctaca gcggtataag gtttctccgc atagaggtgc 1020

accactcctc ccgatacttg tttgtgtgac ttgtgggtca cgacatatat atctacacac 1080

attgcgccac cctttggttc ttccagcaca acaaaaacac gacacgctaa 1130

<210> 217
<211> 778
<212> DNA
<213> Yarrowia lipolytica

<220>
<221> misc_feature
<223> Promoter YAT1

<400> 217
ataagtttgc aaaaagatcg tattatagtt ggagcaaggg agaaatgtag agtgtgaaag 60

actcactatg gtccgggctt atctcgacca atagccaaag tctggagttt ctgagagaaa 120

aaggcaagat acgtatgtaa caaagcgacg catggtacaa taataccgga ggcatgtatc 180

atagagagtt agtggttcga tgatggcact ggtgcctggt atgactttat acggctgact 240

acatatttgt cctcagacat acaattacag tcaagcactt acccttggac atctgtaggt 300

acccccccggc caagacgatc tcagcgtgtc gtatgtcgga ttggcgtagc tccctcgctc 360

gtcaattggc tcccatctac tttcttctgc ttggctacac ccagcatgtc tgccatggct 420

cgttttcgtg ccttatctat cctcccagta ttaccaactc taaatgacat gatgtgattg 480

ggtctacact ttcatatcag agataaggag tagcacagtt gcataaaaag cccaactcta 540

atcagcttct tcctttcttg taattagtac aaaggtgatt agcgaaatct ggaagcttag 600

ttggccctaa aaaaatcaaa aaaagcaaaa aacgaaaaac gaaaaaccac agttttgaga 660

acagggaggt aacgaaggat cgtatatata tatatatata tatataccca cggatcccga 720

gaccggcctt tgattcttcc ctacaaccaa ccattctcac caccctaatt cacaaatg 778

<210> 218
<211> 436
<212> DNA
<213> Yarrowia lipolytica

<220>
<221> misc_feature
<223> Promoter TEF

<400> 218
gaccgggttg gcggcgtatt tgtgtcccaa aaaacagccc caattgcccc aattgacccc 60

aaattgaccc agtagcgggc ccaaccccgg cgagagcccc cttcacccca catatcaaac 120

ctcccccggt tcccacactt gccgttaagg gcgtagggta ctgcagtctg gaatctacgc 180

ttgttcagac tttgtactag tttctttgtc tggccatccg ggtaacccat gccggacgca 240

aaatagacta ctgaaaattt ttttgctttg tggttgggac tttagccaag ggtataaaag 300

accaccgtcc ccgaattacc tttcctcttc ttttctctct ctccttgtca actcacaccc 360

gaaatcgtta agcatttcct tctgagtata agaatcattc aaaggatcca ctagttctag 420

agcggccgct taaacc 436

<210> 219
<211> 1020
<212> DNA
<213> Artificial Sequence

<220>
<223> GPM::FBAIN chimeric Yarrowia lipolytica promoter

<400> 219
gtcgaccatt aattctcacg tgacacagat tattaacgtc tcgtaccaac cacagattac 60

gacccattcg cagtcacagt tcactagggt ttgggttgca tccgttgaga gcggtttgtt 120

tttaaccttc tccatgtgct cactcaggtt ttgggttcag atcaaatcaa ggcgtgaacc 180

actttgtttg aggacaaatg tgacacaacc aaccagtgtc aggggcaagt ccgtgacaaa 240

ggggaagata caatgcaatt actgacagtt acagactgcc tcgatgccct aaccttgccc 300

caaaataaga caactgtcct cgtttaagcg caaccctatt cagcgtcacg tcataatagc 360

gtttggatag cactagtcta tgaggagcgt tttatgttgc ggtgagggcg attggtgctc 420

atatgggttc aattgaggtg gcggaacgag cttagtcttc aattgaggtg cgagcgacac 480

aattgggtgt cacgtggcct aattgacctc gggtcgtgga gtccccagtt atacagcaac 540

cacgaggtgc atgggtagga gacgtcacca gacaataggg ttttttttgg actggagagg 600

gttgggcaaa agcgctcaac gggctgtttg gggagctgtg ggggaggaat tggcgatatt 660

tgtgaggtta acggctccga tttgcgtgtt ttgtcgctcc tgcatctccc catacccata 720

tcttccctcc ccacctcttt ccacgataat tttacggatc agcaataagg ttccttctcc 780

tagtttccac gtccatatat atctatgctg cgtcgtcctt ttcgtgacat caccaaaaca 840

catacaacca tggctgttac tgacgtcctt aagcgaaagt ccggtgtcat cgtcggcgac 900

gatgtccgag ccgtgagtat ccacgacaag atcagtgtcg agacgacgcg ttttgtgtaa 960

tgacacaatc cgaaagtcgc tagcaacaca cactctctac acaaactaac ccagctctcc 1020

<210> 220
<211> 1052
<212> DNA
<213> Artificial Sequence

<220>
<223> GPM::GPDIN chimeric Yarrowia lipolytica promoter

<400> 220
gtcgaccatt aattctcacg tgacacagat tattaacgtc tcgtaccaac cacagattac 60

gacccattcg cagtcacagt tcactagggt ttgggttgca tccgttgaga gcggtttgtt 120

tttaaccttc tccatgtgct cactcaggtt ttgggttcag atcaaatcaa ggcgtgaacc 180

actttgtttg aggacaaatg tgacacaacc aaccagtgtc aggggcaagt ccgtgacaaa 240

ggggaagata caatgcaatt actgacagtt acagactgcc tcgatgccct aaccttgccc 300

caaaataaga caactgtcct cgtttaagcg caaccctatt cagcgtcacg tcataatagc 360

gtttggatag cactagtcta tgaggagcgt tttatgttgc ggtgagggcg attggtgctc 420

atatgggttc aattgaggtg gcggaacgag cttagtcttc aattgaggtg cgagcgacac 480

aattgggtgt cacgtggcct aattgacctc gggtcgtgga gtccccagtt atacagcaac 540

cacgaggtgc atgggtagga gacgtcacca gacaataggg tttttttgg actggagagg 600

gttgggcaaa agcgctcaac gggctgtttg gggagctgtg ggggaggaat tggcgatatt 660

tgtgaggtta acggctccga tttgcgtgtt ttgtcgctcc tgcatctccc catacccata 720

tcttccctcc ccacctcttt ccacgataat tttacggatc agcaataagg ttccttctcc 780

tagtttccac gtccatatat atctatgctg cgtcgtcctt ttcgtgacat caccaaaaca 840

catacaacca tggccatcaa agtcggtatt aacggattcg ggcgaatcgg acgaattgtg 900

agtaccatag aaggtgatgg aaacatgacc caacagaaac agatgacaag tgtcatcgac 960

ccaccagagc ccaattgagc tcatactaac agtcgacaac ctgtcgaacc aattgatgac 1020

tccccgacaa tgtactaaca caggtcctgc cc 1052

<210> 221

<211> 1000
<212> DNA
<213> Yarrowia lipolytica

<220>
<221> misc_feature
<223> Promoter EXP1

<400> 221
ggagtttggc gcccgttttt tcgagcccca cacgtttcgg tgagtatgag cggcggcaga 60

ttcgagcgtt tccggtttcc gcggctggac gagagcccat gatgggggct cccaccacca 120

gcaatcaggg ccctgattac acacccacct gtaatgtcat gctgttcatc gtggttaatg 180

ctgctgtgtg ctgtgtgtgt gtgttgtttg gcgctcattg ttgcgttatg cagcgtacac 240

cacaatattg gaagcttatt agcctttcta ttttttcgtt tgcaaggctt aacaacattg 300

ctgtggagag ggatggggat atggaggccg ctggagggag tcggagaggc gttttggagc 360

ggcttggcct ggcgcccagc tcgcgaaacg cacctaggac cctttggcac gccgaaatgt 420

gccacttttc agtctagtaa cgccttacct acgtcattcc atgcatgcat gtttgcgcct 480

tttttccctt gcccttgatc gccacacagt acagtgcact gtacagtgga ggttttgggg 540

gggtcttaga tgggagctaa aagcggccta gcggtacact agtgggattg tatggagtgg 600

catggagcct aggtggagcc tgacaggacg cacgaccggc tagcccgtga cagacgatgg 660

gtggctcctg ttgtccaccg cgtacaaatg tttgggccaa agtcttgtca gccttgcttg 720

cgaacctaat tcccaatttt gtcacttcgc acccccattg atcgagccct aacccctgcc 780

catcaggcaa tccaattaag ctcgcattgt ctgccttgtt tagtttggct cctgcccgtt 840

tcggcgtcca cttgcacaaa cacaaacaag cattatatat aaggctcgtc tctccctccc 900

aaccacactc actttttgc ccgtcttccc ttgctaacac aaaagtcaag aacacaaaca 960

accaccccaa ccccccttaca cacaagacat atctacagca 1000

<210> 222
<211> 5
<212> PRT
<213> Artificial Sequence

<220>
<223> His Box 1

<400> 222

His Glu Cys Gly His
1 5

<210> 223
<211> 8
<212> PRT
<213> Artificial Sequence

<220>
<223> Delta-15 desaturase motif

<220>
<221> misc_feature
<222> (2)..(3)
<223> Xaa can be any naturally occurring amino acid

<400> 223

Ile Xaa Xaa His Glu Cys Gly His
1 5

<210> 224
<211> 8
<212> PRT
<213> Artificial Sequence

<220>
<223> Delta-12 desaturase motif

<220>
<221> misc_feature
<222> (2)..(3)
<223> Xaa can be any naturally occurring amino acid

<400> 224

Val Xaa Xaa His Glu Cys Gly His
1 5

<210> 225
<211> 30
<212> DNA
<213> Artificial Sequence

<220>

<223> Primer YL211

<400> 225
tttgtcgacg cagtaggatg tcctgcacgg 30

<210> 226
<211> 34
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer YL212

<400> 226
tttccatggt tgatgtgtgt ttaattcaag aatg 34

<210> 227
<211> 27
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer YL376

<400> 227
aatctgcagt ggtgcaaacg gggcgga 27

<210> 228
<211> 32
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer YL377

<400> 228
tttccatggg caggacctgt gttagtacat tg 32

<210> 229
<211> 33
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer YL203

<400> 229
tttccatggt tgtatgtgtt ttggtgatgt cac 33

<210> 230
<211> 33
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer YL204

<400> 230
tttgtcgacc gtttaagcgc aaccctattc agc 33

<210> 231
<211> 22
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer GPAT-5-1

<400> 231
caacttttct tgtcgacctg ag 22

<210> 232
<211> 34
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer GPAT-5-2

<400> 232
gatcccatgg ttagcgtgtc gtgtttttgt tgtg 34

<210> 233
<211> 33
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer ODMW314

<400> 233
cgtgccatgg tgtgatgtgt agtttagatt tcg 33

<210> 234
<211> 39
<212> DNA

<213> Artificial Sequence

<220>
<223> Primer YL341

<400> 234
tttgtcgacg tttaaacagt gtacgcagta ctatagagg 39

<210> 235
<211> 34
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer ODMW320

<400> 235
cgtgccatgg agagctgggt tagtttgtgt agag 34

<210> 236
<211> 39
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer ODMW341

<400> 236
tttgtcgacg tttaaacagt gtacgcagta ctatagagg 39

<210> 237
<211> 32
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer 27203-F

<400> 237
gatcgtcgac ataagtttgc aaaaagatcg ta 32

<210> 238
<211> 34
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer 27203-R

<400> 238
gatcccatgg ttgtgaatta gggtggtgag aatg 34

<210> 239
<211> 21
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer YL-URA-16F

<400> 239
gctcgagcta acgtccacaa g 21

<210> 240
<211> 18
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer YL-URA-78R

<400> 240
cttggctgcc acgagctt 18

<210> 241
<211> 22
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer GUS-767F

<400> 241
ccaaaagcca gacagagtgt ga 22

<210> 242
<211> 22
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer GUS-891R

<400> 242
ttcatgacga ccaaagccag ta 22

<210> 243
<211> 2987
<212> DNA
<213> Yarrowia lipolytica

<220>
<221> misc_feature
<223> mutant acetohydroxyacid synthase (AHAS) with W497L mutation

<400> 243
ttccctagtc ccagtgtaca cccgccgata tcgcttaccc tgcagccgga ttaaggttgg 60

caatttttca cgtccttgtc tccgcaatta ctcaccgggt ggtttataag attgcaagcg 120

tcttgatttg tctctgtata ctaacatgca atcgcgactc gcccgacggg ccactaacct 180

ggccagaatc tccagatcca agtattctct tggtctgcga tatgtttcca acacaaaagc 240

ccctgctgcc cagccggcaa ctgctgagtg agtattcctt gccataaacg acccagaacc 300

actgtatagt gtttggaagc actagtcaga agaccagcga aaacaggtgg aaaaaactga 360

gacgaaaagc aacgaccaga aatgtaatgt gtggaaaagc gacacacaca gagcagataa 420

agaggtgaca aataacgaca aatgaaatat cagtatcttc ccacaatcac tacctctcag 480

ctgtctgaag gtgcggctga tatatccatc ccacgtctaa cgtatggagt gtgatagaat 540

atgacgacac aagcatgaga actcgctctc tatccaacca ccgaaacact gtcactacag 600

ccgttcttgt tgctccattc gcttttgtga ttccatgcct tctctggtga ctgacaacat 660

tccttccttt tctccagccc tgttgttatc tgctcatgac ctacggccac tctctatcgc 720

atactaacat agacgatccc agcccgctcc ccacttccag ggcaccgttg gcaagcctcc 780

tatcctcaag aaggctgagg ctgccaacgc tgacatggac gagtccttca tcggaatgtc 840

tggaggagag atcttccacg agatgatgct gcgacacaac gtcgacactg tcttcggtta 900

ccccggtgga gccattctcc ccgtctttga cgccattcac aactctgagt acttcaactt 960

tgtgctccct cgacacgagc agggtgccgg ccacatggcc gagggctacg ctcgagcctc 1020

tggtaagccc ggtgtcgttc tcgtcacctc tggccccggt gccaccaacg tcatcacccc 1080

catgcaggac gctctttccg atggtacccc catggttgtc ttcaccggtc aggtcctgac 1140

ctccgttatc ggcactgacg ccttccagga ggccgatgtt gtcggcatct cccgatcttg 1200

caccaagtgg aacgtcatgg tcaagaacgt tgctgagctc ccccgacgaa tcaacgaggc 1260

ctttgagatt gctacttccg gccgacccgg tcccgttctc gtcgatctgc ccaaggatgt 1320

tactgctgcc atcctgcgag agcccatccc caccaagtcc accattccct cgcattctct 1380

gaccaacctc acctctgccg ccgccaccga gttccagaag caggctatcc agcgagccgc 1440

caacctcatc aaccagtcca agaagcccgt cctttacgtc ggacagggta tccttggctc 1500

cgaggagggt cctaagctgc ttaaggagct ggctgagaag gccgagattc ccgtcaccac 1560

tactctgcag ggtcttggtg cctttgacga gcgagacccc aagtctctgc acatgctcgg 1620

tatgcacggt tccggctacg ccaacatggc catgcagaac gctgactgta tcattgctct 1680

cggcgcccga tttgatgacc gagttaccgg ctccatcccc aagtttgccc ccgaggctcg 1740

agccgctgcc cttgagggtc gaggtggtat tgttcacttt gagatccagg ccaagaacat 1800

caacaaggtt gttcaggcca ccgaagccgt tgagggagac gttaccgagt ctgtccgaca 1860

gctcatcccc ctcatcaaca aggtctctgc cgctgagcga gctccctgga ctgagactat 1920

ccagtcctgg aagcagcagt tccccttcct cttcgaggct gaaggtgagg atggtgttat 1980

caagccccag tccgtcattg ctctgctctc tgacctgaca gagaacaaca aggacaagac 2040

catcatcacc accggtgttg gtcagcatca gatgtggact gcccagcatt tccgatggcg 2100

acaccctcga accatgatca cttctggtgg tcttggaact atgggttacg gcctgcccgc 2160

cgctatcggc gccaaggttg cccgacctga ctgcgacgtc attgacatcg atggtgacgc 2220

ttctttcaac atgactctga ccgagctgtc caccgccgtt cagttcaaca ttggcgtcaa 2280

ggctattgtc ctcaacaacg aggaacaggg tatggtcacc cagctgcagt ctctcttcta 2340

cgagaaccga tactgccaca ctcatcagaa gaaccccgac ttcatgaagc tggccgagtc 2400

catgggcatg aagggtatcc gaatcactca cattgaccag ctggaggccg gtctcaagga 2460

gatgctcgca tacaagggcc ctgtgctcgt tgaggttgtt gtcgacaaga agatccccgt 2520

tcttcccatg gttcccgctg gtaaggcttt gcatgagttc cttgtctacg acgctgacgc 2580

cgaggctgct tctcgacccg atcgactgaa gaatgccccc gcccctcacg tccaccagac 2640

caccttttgag aactaagtgg aaaggaacac aagcaatccg aaccaaaaat aattggggtc 2700

ccgtgcccac agagtctagt gcagacctaa aatgaccaca gtaaattata gctgttatta 2760

aacatgagat tttgaccaac aagagcgtag gaatgttatt agctactact tgtacataca 2820

cagcatttgt tttaaataat gttgcctcca ggggcagtga gatcaggacc cagatccgtg 2880

gccagctctc tgacttcaga ccgcttgtac ttaagcagct cgcaacactg ttgtcgagga 2940

ttgaacttgc catattcgat tttgtggtca tgaatccagc acacctc 2987

<210> 244
<211> 25
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer 410

<400> 244
ttccctagtc ccagtgtaca cccgc 25

<210> 245
<211> 25
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer 411

<400> 245
agagagactg cagctgggtg accat 25

<210> 246
<211> 27
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer 412

<400> 246
gtatggtcac ccagctgcag tctctct 27

<210> 247
<211> 25
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer 413

<400> 247
gaggtgtgct ggattcatga ccaca 25

<210> 248
<211> 44
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer 414

<400> 248
ctggtaaata aatgatgttc cctagtccca gtgtacaccc gccg 44

<210> 249
<211> 42
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer 415

<400> 249
tatgcttccg gctcgtacga ggtgtgctgg attcatgacc ac 42

<210> 250
<211> 105
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer D4-1A

<400> 250
ggtcggcttc gacgagaccg tcactatgga caccgtccga aaccacaaca tgcccgacga 60

tgcctggtgc gctatccacg gcaccgtgta cgacatcacc aagtt 105

<210> 251
<211> 107
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer D4-1B

<400> 251

ggagaacttg gtgatgtcgt acacggtgcc gtggatagcg caccaggcat cgtcgggcat 60

gttgtggttt cggacggtgt ccatagtgac ggtctcgtcg aagccga 107

<210> 252
<211> 107
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer D4-2A

<400> 252
ctccaaggtg catcctggtg gagacattat catgctggcc gctggcaagg aggccaccat 60

cctgttcgag acctaccaca tcaagggtgt tcccgacgct gtcctgc 107

<210> 253
<211> 107
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer D4-2B

<400> 253
tttcgcagga cagcgtcggg aacacccttg atgtggtagg tctcgaacag gatggtggcc 60

tccttgccag cggccagcat gataatgtct ccaccaggat gcacctt 107

<210> 254
<211> 106
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer D4-3A

<400> 254
gaaagtacaa ggtcggcaag ctcccctcagg gcaagaaggg tgagacctcc cacatgccca 60

ccggactcga ctctgcctcc tactactcct gggactctga gttcta 106

<210> 255
<211> 106
<212> DNA
<213> Artificial Sequence

<220>

<223> Primer D4-3B

<400> 255
tcggtagaac tcagagtccc aggagtagta ggaggcagag tcgagtccgg tgggcatgtg 60

ggaggtctca cccttcttgc cctgagggag cttgccgacc ttgtac 106

<210> 256
<211> 106
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer D4-4A

<400> 256
ccgagtcctc cgagagcgag tcgccaagaa gctggccgag cccggtctca tgcagcgagc 60

tcgaatggag ctctgggcca aggctatctt tctcctggct ggtttc 106

<210> 257
<211> 104
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer D4-4B

<400> 257
cagaaaccag ccaggagaaa gatagccttg gcccagagct ccattcgagc tcgctgcatg 60

agaccgggct cggccagctt cttggcgact cgctctcgga ggac 104

<210> 258
<211> 89
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer D4-5A

<400> 258
ctgtacgcca tgtgcgtgct cgaccctcac ggaggtgcta tggtcgctgc cgtcaccctc 60

ggtgtgtttg ctgccttcgt cggaacctg 89

<210> 259
<211> 91
<212> DNA

<213> Artificial Sequence

<220>
<223> Primer D4-5B

<400> 259
gatgcaggtt ccgacgaagg cagcaaacac accgagggtg acggcagcga ccatagcacc 60

tccgtgaggg tcgagcacgc acatggcgta c 91

<210> 260
<211> 99
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer D4-6A

<400> 260
catccagcac gacggctccc acggagcctt ctccaagtct cgattcatga acaaggctgc 60

cggttggacc ctcgacatga ttggtgcctc cgctatgac 99

<210> 261
<211> 99
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer D4-6B

<400> 261
ccaggtcata gcggaggcac caatcatgtc gagggtccaa ccggcagcct tgttcatgaa 60

tcgagacttg gagaaggctc cgtgggagcc gtcgtgctg 99

<210> 262
<211> 95
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer D4-7A

<400> 262
ctgggagatg cagcacgtcc tcggtcacca tccctacacc aacctgatcg agatggagaa 60

cggtctggcc aaggtcaagg gtgccgacgt ggacc 95

<210> 263
<211> 95
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer D4-7B

<400> 263
ttggggtcca cgtcggcacc cttgaccttg gccagaccgt tctccatctc gatcaggttg 60

gtgtagggat ggtgaccgag gacgtgctgc atctc 95

<210> 264
<211> 88
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer D4-8A

<400> 264
ccaagaaggt cgaccaggag tccgaccctg acgtcttctc tacctaccct atgctgcgac 60

tgcatccctg gcaccgacag cgattcta 88

<210> 265
<211> 88
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer D4-8B

<400> 265
gtggtagaat cgctgtcggt gccagggatg cagtcgcagc atagggtagg tagagaagac 60

gtcagggtcg gactcctggt cgaccttc 88

<210> 266
<211> 91
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer D4-9A

<400> 266

cagcacctgt acgctcccct gatcttcggc ttcatgacca tcaacaaggt gatctcccag 60

gacgtcggtg tcgtgctgcg aaagcgactg t 91

<210> 267
<211> 91
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer D4-9B

<400> 267
tgaaacagtc gctttcgcag cacgacaccg acgtcctggg agatcacctt gttgatggtc 60

atgaagccga agatcagggg agcgtacagg t 91

<210> 268
<211> 89
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer D4-10A

<400> 268
ttcagatcga cgccaactgc cgatatggct ctccctggaa cgtggctcga ttctggatca 60

tgaagctcct gaccactctc tacatggtg 89

<210> 269
<211> 89
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer D4-10B

<400> 269
gagccaccat gtagagagtg gtcaggagct tcatgatcca gaatcgagcc acgttccagg 60

gagagccata tcggcagttg gcgtcgatc 89

<210> 270
<211> 84
<212> DNA
<213> Artificial Sequence

<220>

<223> Primer D4-11A

<400> 270
gctctgccca tgtacatgca gggacccgct cagggtctga agctgttctt tatggctcac 60

ttcacctgcg gagaggtcct cgcc 84

<210> 271
<211> 84
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer D4-11B

<400> 271
tggtggcgag gacctctccg caggtgaagt gagccataaa gaacagcttc agaccctgag 60

cgggtccctg catgtacatg ggca 84

<210> 272
<211> 85
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer D4-12A

<400> 272
accatgttca ttgttaacca cattatcgag ggtgtctcct acgcctccaa ggacgctgtc 60

aagggcgtca tggctcctcc ccgaa 85

<210> 273
<211> 83
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer D4-12B

<400> 273
ggttcgggga ggagccatga cgcccttgac agcgtccttg gaggcgtagg agacaccctc 60

gataatgtgg ttaacaatga aca 83

<210> 274
<211> 87
<212> DNA

<213> Artificial Sequence

<220>
<223> Primer D4-13A

<400> 274
ggtgtcactc ccatgcaggt gacccagaag gctctctctg ctgccgagtc caccaagtct 60

gacgctgaca agaccactat gatccct 87

<210> 275
<211> 89
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer D4-13B

<400> 275
tgagagggat catagtggtc ttgtcagcgt cagacttggt ggactcggca gcagagagag 60

ccttctgggt cacctgcatg ggagtgaca 89

<210> 276
<211> 88
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer D4-14A

<400> 276
ctcaacgact gggctgccgt ccagtgccag acctccgtca actgggctgt cggatcctgg 60

ttctggaacc acttctctgg aggtctca 88

<210> 277
<211> 88
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer D4-14B

<400> 277
tggttgagac ctccagagaa gtggttccag aaccaggatc cgacagccca gttgacggag 60

gtctggcact ggacggcagc ccagtcgt 88

<210> 278
<211> 99
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer D4-15A

<400> 278
accaccagat tgagcaccat tgctttcctc agaaccccca caccgtcaac gtctacatct 60

ccggcatcgt caaggagacc tgcgaggagt acggcgtgc 99

<210> 279
<211> 99
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer D4-15B

<400> 279
tagggcacgc cgtactcctc gcaggtctcc ttgacgatgc cggagatgta gacgttgacg 60

gtgtgggggt tctgaggaaa gcaatggtgc tcaatctgg 99

<210> 280
<211> 95
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer D4-16A

<400> 280
cctaccaggc tgagatctcc ctcttctctg cctatttcaa gatgctgtcc cacctccgaa 60

ccctcggcaa cgaggacctc accgcctggt ccacc 95

<210> 281
<211> 94
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer D4-16B

<400> 281

ttaggtggac caggcggtga ggtcctcgtt gccgagggtt cggaggtggg acagcatctt 60

gaaataggca gagaagaggg agatctcagc ctgg 94

<210> 282
<211> 28
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer D4-1F

<400> 282
ttccatggtc ggcttcgacg agaccgtc 28

<210> 283
<211> 34
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer D4-4R

<400> 283
ttggatcccc agaaaccagc caggagaaag atag 34

<210> 284
<211> 29
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer D4-5F

<400> 284
ttggatccct gtacgccatg tgcgtgctc 29

<210> 285
<211> 31
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer D4-8R

<400> 285
ccaagcttgt ggtagaatcg ctgtcggtgc c 31

<210> 286
<211> 32
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer D4-9F

<400> 286
ccaagcttca gcacctgtac gctcccctga tc 32

<210> 287
<211> 34
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer D4-12R

<400> 287
ttgtgcacgg ttcggggagg agccatgacg ccct 34

<210> 288
<211> 29
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer D4-13F

<400> 288
ttgtgcacgg tgtcactccc atgcaggtg 29

<210> 289
<211> 34
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer D4-16R

<400> 289
ttgcggccgc ttaggtggac caggcggtga ggtc 34

<210> 290
<211> 33
<212> DNA

<213> Artificial Sequence

<220>
<223> Primer YL251

<400> 290
cgattctacc acaagttcca gcacctgtac gct 33

<210> 291
<211> 33
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer YL252

<400> 291
agcgtacagg tgctggaact tgtggtagaa tcg 33

<210> 292
<211> 101
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer IL3-1A

<400> 292
gccaacgacg ctggcgagcg aatctgggct gccgtcaccg atcccgaaat cctcattggc 60

accttctcct acctgctcct gaagcctctc ctgcgaaact c 101

<210> 293
<211> 101
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer IL3-1B

<400> 293
accagagttt cgcaggagag gcttcaggag caggtaggag aaggtgccaa tgaggatttc 60

gggatcggtg acggcagccc agattcgctc gccagcgtcg t 101

<210> 294
<211> 100
<212> DNA

<213> Artificial Sequence

<220>
<223> Primer IL3-2A

<400> 294
tggtctcgtg gacgagaaga aaggagccta ccgaacctcc atgatctggt acaacgtcct 60

cctggctctc ttctctgccc tgtccttcta cgtgactgcc 100

<210> 295
<211> 100
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer IL3-2B

<400> 295
cggtggcagt cacgtagaag gacagggcag agaagagagc caggaggacg ttgtaccaga 60

tcatggaggt tcggtaggct cctttcttct cgtccacgag 100

<210> 296
<211> 100
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer IL3-3A

<400> 296
accgctctcg gctgggacta cggtactgga gcctggctgc gaagacagac cggtgatact 60

ccccagcctc tctttcagtg tccctctcct gtctgggact 100

<210> 297
<211> 100
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer IL3-3B

<400> 297
ttggagtccc agacaggaga gggacactga aagagaggct ggggagtatc accggtctgt 60

cttcgcagcc aggctccagt accgtagtcc cagccgagag 100

<210> 298
<211> 100
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer IL3-4A

<400> 298
ccaagctgtt cacctggact gccaaggcct tctactattc taagtacgtg gagtacctcg 60

acaccgcttg gctggtcctc aagggcaagc gagtgtcctt 100

<210> 299
<211> 100
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer IL3-4B

<400> 299
cagaaaggac actcgcttgc ccttgaggac cagccaagcg gtgtcgaggt actccacgta 60

cttagaatag tagaaggcct tggcagtcca ggtgaacagc 100

<210> 300
<211> 89
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer IL3-5A

<400> 300
ttccatcact ttggagctcc ctgggacgtc tacctcggca ttcgactgca caacgagggt 60

gtgtggatct tcatgttctt taactcgtt 89

<210> 301
<211> 89
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer IL3-5B

<400> 301

aatgaacgag ttaaagaaca tgaagatcca cacaccctcg ttgtgcagtc gaatgccgag 60

gtagacgtcc cagggagctc caaagtgat 89

<210> 302
<211> 91
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer IL3-6A

<400> 302
cattcacacc atcatgtaca cctactatgg actgactgcc gctggctaca agttcaaggc 60

caagcctctg atcactgcca tgcagatttg c 91

<210> 303
<211> 91
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer IL3-6B

<400> 303
actggcaaat ctgcatggca gtgatcagag gcttggcctt gaacttgtag ccagcggcag 60

tcagtccata gtaggtgtac atgatggtgt g 91

<210> 304
<211> 94
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer IL3-7A

<400> 304
cagttcgtcg gtggctttct cctggtctgg gactacatca acgttccctg cttcaactct 60

gacaagggca agctgttctc ctgggctttc aact 94

<210> 305
<211> 94
<212> DNA
<213> Artificial Sequence

<220>

<223> Primer IL3-7B

<400> 305
gcgtagttga aagcccagga gaacagcttg cccttgtcag agttgaagca gggaacgttg 60

atgtagtccc agaccaggag aaagccaccg acga 94

<210> 306
<211> 91
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer IL3-8A

<400> 306
acgcctacgt cggatctgtc tttctcctgt tctgtcactt cttttaccag gacaacctgg 60

ccaccaagaa atccgctaag gctggtaagc a 91

<210> 307
<211> 91
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer IL3-8B

<400> 307
aagctgctta ccagccttag cggatttctt ggtggccagg ttgtcctggt aaaagaagtg 60

acagaacagg agaaagacag atccgacgta g 91

<210> 308
<211> 41
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer IL3-1F

<400> 308
tttccatggc tctggccaac gacgctggcg agcgaatctg g 41

<210> 309
<211> 36
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer IL3-4R

<400> 309
tttctgcaga aaggacactc gcttgccctt gaggac 36

<210> 310
<211> 41
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer IL3-5F

<400> 310
tttctgcagg ccttccatca ctttggagct ccctgggacg t 41

<210> 311
<211> 42
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer IL3-8R

<400> 311
tttgcggccg ctaaagctgc ttaccagcct tagcggattt ct 42

<210> 312
<211> 417
<212> DNA
<213> Artificial Sequence

<220>
<223> 417 bp NcoI/PstI fragment pT9(1-4)

<400> 312
catggctctg gccaacgacg ctggcgagcg aatctgggct gccgtcaccg atcccgaaat 60

cctcattggc accttctcct acctgctcct gaagcctctc ctgcgaaact ctggtctcgt 120

ggacgagaag aaaggagcct accgaacctc catgatctgg tacaacgtcc tcctggctct 180

cttctctgcc ctgtccttct acgtgactgc caccgctctc ggctgggact acggtactgg 240

agcctggctg cgaagacaga ccggtgatac tccccagcct ctctttcagt gtccctctcc 300

tgtctgggac tccaagctgt tcacctggac tgccaaggcc ttctactatt ctaagtacgt 360

ggagtacctc gacaccgctt ggctggtcct caagggcaag cgagtgtcct ttctgca 417

<210> 313
<211> 377
<212> DNA
<213> Artificial Sequence

<220>
<223> 377 bp PstI/Not1 fragment pT9(5-8)

<400> 313
ggccttccat cactttggag ctccctggga cgtctacctc ggcattcgac tgcacaacga 60

gggtgtgtgg atcttcatgt tctttaactc gttcattcac accatcatgt acacctacta 120

tggactgact gccgctggct acaagttcaa ggccaagcct ctgatcactg ccatgcagat 180

ttgccagttc gtcggtggct ttctcctggt ctgggactac atcaacgttc cctgcttcaa 240

ctctgacaag ggcaagctgt tctcctgggc tttcaactac gcctacgtcg gatctgtctt 300

tctcctgttc tgtcacttct tttaccagga caacctggcc accaagaaat ccgctaaggc 360

tggtaagcag ctttagc 377

<210> 314
<211> 104
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer D8-1A

<400> 314
atggagtcca agcgacaggc tctgtctccc ctccagctga tggaacagac ctacgacgtc 60

gtgaacttcc accctggtgg agctgaaatc attgagaact acca 104

<210> 315
<211> 104
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer D8-1B

<400> 315
tccctggtag ttctcaatga tttcagctcc accagggtgg aagttcacga cgtcgtaggt 60

ctgttccatc agctggaggg gagacagagc ctgtcgcttg gact 104

<210> 316
<211> 102
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer D8-2A

<400> 316
gggacgagat gctactgacg ccttcatggt tatgcacttt caggaagcct tcgacaagct 60

caagcgaatg cccaagatca acccctcctt tgagctgcct cc 102

<210> 317
<211> 102
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer D8-2B

<400> 317
ctggggaggc agctcaaagg aggggttgat cttgggcatt cgcttgagct tgtcgaaggc 60

ttcctgaaag tgcataacca tgaaggcgtc agtagcatct cg 102

<210> 318
<211> 101
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer D8-3A

<400> 318
ccaggctgcc gtcaacgaag ctcaggagga tttccgaaag ctccgagaag agctgatcgc 60

cactggcatg tttgacgcct ctcccctctg gtactcgtac a 101

<210> 319
<211> 101
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer D8-3B

<400> 319

atcttgtacg agtaccagag gggagaggcg tcaaacatgc cagtggcgat cagctcttct 60

cggagctttc ggaaatcctc ctgagcttcg ttgacggcag c 101

<210> 320
<211> 101
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer D8-4A

<400> 320
ccaccaccct gggtcttggc gtgcttggat acttcctgat ggtccagtac cagatgtact 60

tcattggtgc tgtgctgctc ggtatgcact accagcaaat g 101

<210> 321
<211> 101
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer D8-4B

<400> 321
atcccatttg ctggtagtgc ataccgagca gcacagcacc aatgaagtac atctggtact 60

ggaccatcag gaagtatcca agcacgccaa gacccagggt g 101

<210> 322
<211> 104
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer D8-5A

<400> 322
ggatggctgt ctcatgacat ctgccaccac cagaccttca agaaccgaaa ctggaataac 60

ctcgtgggtc tggtctttgg caacggactc cagggcttct ccgt 104

<210> 323
<211> 104
<212> DNA
<213> Artificial Sequence

<220>

<223> Primer D8-5B

<400> 323
ggtcacggag aagccctgga gtccgttgcc aaagaccaga cccacgaggt tattccagtt 60

tcggttcttg aaggtctggt ggtggcagat gtcatgagac agcc 104

<210> 324
<211> 101
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer D8-6A

<400> 324
gacctgttgg aaggacagac acaacgccca tcattctgct accaacgttc agggtcacga 60

tcccgacatt gataacctgc ctccctcgc ctggtccgag g 101

<210> 325
<211> 101
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer D8-6B

<400> 325
tcgtcctcgg accaggcgag gggaggcagg ttatcaatgt cgggatcgtg accctgaacg 60

ttggtagcag aatgatgggc gttgtgtctg tccttccaac a 101

<210> 326
<211> 95
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer D8-7A

<400> 326
tcactcgagc ttctcccatc tcccgaaagc tcattcagtt ccaacagtac tatttcctgg 60

tcatctgtat tctcctgcga ttcatctggt gtttc 95

<210> 327
<211> 95
<212> DNA

<213> Artificial Sequence

<220>
<223> Primer D8-7B

<400> 327
actggaaaca ccagatgaat cgcaggagaa tacagatgac caggaaatag tactgttgga 60

actgaatgag ctttcgggag atgggagaag ctcga 95

<210> 328
<211> 89
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer D8-8A

<400> 328
cagtgcgtgc tgaccgttcg atccctcaag gaccgagaca accagttcta ccgatctcag 60

tacaagaaag aggccattgg actcgctct 89

<210> 329
<211> 89
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer D8-8B

<400> 329
gtgcagagcg agtccaatgg cctctttctt gtactgagat cggtagaact ggttgtctcg 60

gtccttgagg gatcgaacgg tcagcacgc 89

<210> 330
<211> 85
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer D8-9A

<400> 330
gcactggact ctcaaggctc tgttccacct cttctttatg ccctccatcc tgacctcgct 60

cctggtgttc tttgtttccg agctc 85

<210> 331
<211> 85
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer D8-9B

<400> 331
cgacgagctc ggaaacaaag aacaccagga gcgaggtcag gatggagggc ataaagaaga 60

ggtggaacag agccttgaga gtcca 85

<210> 332
<211> 91
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer D8-10A

<400> 332
gtcggtggct tcggaattgc catcgtggtc ttcatgaacc actaccctct ggagaagatc 60

ggtgatcccg tctgggacgg acatggcttc t 91

<210> 333
<211> 91
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer D8-10B

<400> 333
acagagaagc catgtccgtc ccagacggga tcaccgatct tctccagagg gtagtggttc 60

atgaagacca cgatggcaat tccgaagcca c 91

<210> 334
<211> 92
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer D8-11A

<400> 334

ctgtgggtca gatccatgag accatgaaca ttcgacgagg catcattact gactggttct 60

ttggaggcct gaactaccag atcgagcacc at 92

<210> 335
<211> 92
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer D8-11B

<400> 335
agagatggtg ctcgatctgg tagttcaggc ctccaaagaa ccagtcagta atgatgcctc 60

gtcgaatgtt catggtctca tggatctgac cc 92

<210> 336
<211> 93
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer D8-12A

<400> 336
ctctggccca ctctgcctcg acacaacctc actgccgttt cctaccaggt ggaacagctg 60

tgccagaagc acaacctccc ctaccgaaac cct 93

<210> 337
<211> 93
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer D8-12B

<400> 337
gcagagggtt tcggtagggg aggttgtgct tctggcacag ctgttccacc tggtaggaaa 60

cggcagtgag gttgtgtcga ggcagagtgg gcc 93

<210> 338
<211> 90
<212> DNA
<213> Artificial Sequence

<220>

<223> Primer D8-13A

<400> 338
ctgccccatg aaggtctcgt catcctgctc cgatacctgg ccgtgttcgc tcgaatggcc 60

gagaagcagc ccgctggcaa ggctctctaa 90

<210> 339
<211> 90
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer D8-13B

<400> 339
ccgcttagag agccttgcca gcgggctgct tctcggccat tcgagcgaac acggccaggt 60

atcggagcag gatgacgaga ccttcatggg 90

<210> 340
<211> 38
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer D8-1F

<400> 340
tttccatgga gtccaagcga caggctctgt ctcccctc 38

<210> 341
<211> 37
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer D8-3R

<400> 341
tttagatctt gtacgagtac cagaggggag aggcgtc 37

<210> 342
<211> 41
<212> DNA
<213> Artificial Sequence

<220>

<223> Primer D8-4F

<400> 342
acaagatctc caccaccctg ggtcttggcg tgcttggata c 41

<210> 343
<211> 43
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer D8-6R

<400> 343
tttctcgagt gacatcgtcc tcggaccagg cgaggggagg cag 43

<210> 344
<211> 32
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer D8-7F

<400> 344
tcactcgagc ttctcccatc tcccgaaagc tc 32

<210> 345
<211> 29
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer D8-9R

<400> 345
cgacgagctc ggaaacaaag aacaccagg 29

<210> 346
<211> 39
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer D8-10F

<400> 346
tttgagctcg tcggtggctt cggaattgcc atcgtggtc 39

<210> 347
<211> 36
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer D8-13R

<400> 347
tttgcggccg cttagagagc cttgccagcg ggctgc 36

<210> 348
<211> 309
<212> DNA
<213> Artificial Sequence

<220>
<223> 309 bp NcoI/BglII fragment of pT8(1-3)

<400> 348
catggagtcc aagcgacagg ctctgtctcc cctccagctg atggaacaga cctacgacgt 60

cgtgaacttc caccctggtg gagctgaaat cattgagaac taccagggac gagatgctac 120

tgacgccttc atggttatgc actttcagga agccttcgac aagctcaagc gaatgcccaa 180

gatcaacccc tcctttgagc tgcctcccca ggctgccgtc aacgaagctc aggaggattt 240

ccgaaagctc cgagaagagc tgatcgccac tggcatgttt gacgcctctc ccctctggta 300

ctcgtacaa 309

<210> 349
<211> 321
<212> DNA
<213> Artificial Sequence

<220>
<223> 321 bp BglII/XhoI fragment of pT8(4-6)

<400> 349
gatctccacc accctgggtc ttggcgtgct tggatacttc ctgatggtcc agtaccagat 60

gtacttcatt ggtgctgtgc tgctcggtat gcactaccag caaatgggat ggctgtctca 120

tgacatctgc caccaccaga ccttcaagaa ccgaaactgg aataacctcg tgggtctggt 180

ctttggcaac ggactccagg gcttctccgt gacctgttgg aaggacagac acaacgccca 240

tcattctgct accaacgttc agggtcacga tcccgacatt gataacctgc ctcccctcgc 300

ctggtccgag gacgatgtca c 321

<210> 350
<211> 264
<212> DNA
<213> Artificial Sequence

<220>
<223> 264 bp XhoI/SacI fragment of pT8(7-9)

<400> 350
tcgagcttct cccatctccc gaaagctcat tcagttccaa cagtactatt tcctggtcat 60

ctgtattctc ctgcgattca tctggtgttt ccagtgcgtg ctgaccgttc gatccctcaa 120

ggaccgagac aaccagttct accgatctca gtacaagaaa gaggccattg gactcgctct 180

gcactggact ctcaaggctc tgttccacct cttctttatg ccctccatcc tgacctcgct 240

cctggtgttc tttgtttccg agct 264

<210> 351
<211> 369
<212> DNA
<213> Artificial Sequence

<220>
<223> 369 bp Sac1/Not1 fragment of pT8(10-13)

<400> 351
cgtcggtggc ttcggaattg ccatcgtggt cttcatgaac cactaccctc tggagaagat 60

cggtgatccc gtctgggacg gacatggctt ctctgtgggt cagatccatg agaccatgaa 120

cattcgacga ggcatcatta ctgactggtt ctttggaggc ctgaactacc agatcgagca 180

ccatctctgg cccaccctgc ctcgacacaa cctcactgcc gtttcctacc aggtggaaca 240

gctgtgccag aagcacaacc tcccctaccg aaaccctctg ccccatgaag gtctcgtcat 300

cctgctccga tacctggccg tgttcgctcg aatggccgag aagcagcccg ctggcaaggc 360

tctctaagc 369

<210> 352
<211> 34
<212> DNA

<213> Artificial Sequence

<220>
<223> Primer ODMW390

<400> 352
aagaatcatt caccatgaag tccaagcgac aggc 34

<210> 353
<211> 34
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer ODMW391

<400> 353
gcctgtcgct tggacttcat ggtgaatgat tctt 34

<210> 354
<211> 1852
<212> DNA
<213> Artificial Sequence

<220>
<223> chimeric D8S-1::XPR gene

<400> 354
cgatcaggag agaccgggtt ggcggcgtat ttgtgtccca aaaaacagcc ccaattgccc 60

caattgaccc caaattgacc cagtagcggg cccaaccccg gcgagagccc ccttcacccc 120

acatatcaaa cctccccccgg ttcccacact tgccgttaag ggcgtagggt actgcagtct 180

ggaatctacg cttgttcaga ctttgtacta gtttctttgt ctggccatcc gggtaaccca 240

tgccggacgc aaaatagact actgaaaatt tttttgcttt gtggttggga ctttagccaa 300

gggtataaaa gaccaccgtc cccgaattac ctttcctctt cttttctctc tctccttgtc 360

aactcacacc cgaaatcgtt aagcatttcc ttctgagtat aagaatcatt caccatggag 420

tccaagcgac aggctctgtc tcccctccag ctgatggaac agacctacga cgtcgtgaac 480

ttccaccctg gtggagctga aatcattgag aactaccagg gacgagatgc tactgacgcc 540

ttcatggtta tgcactttca ggaagccttc gacaagctca agcgaatgcc caagatcaac 600

ccctcctttg agctgcctcc ccaggctgcc gtcaacgaag ctcaggagga tttccgaaag 660

ctccgagaag agctgatcgc cactggcatg tttgacgcct ctcccctctg gtactcgtac 720

aagatctcca ccaccctggg tcttggcgtg cttggatact tcctgatggt ccagtaccag 780

atgtacttca ttggtgctgt gctgctcggt atgcactacc agcaaatggg atggctgtct 840

catgacatct gccaccacca gaccttcaag aaccgaaact ggaataacct cgtgggtctg 900

gtctttggca acggactcca gggcttctcc gtgacctgtt ggaaggacag acacaacgcc 960

catcattctg ctaccaacgt tcagggtcac gatcccgaca ttgataacct gcctcccctc 1020

gcctggtccg aggacgatgt cactcgagct tctcccatct cccgaaagct cattcagttc 1080

caacagtact atttcctggt catctgtatt ctcctgcgat tcatctggtg tttccagtgc 1140

gtgctgaccg ttcgatccct caaggaccga gacaaccagt tctaccgatc tcagtacaag 1200

aaagaggcca ttggactcgc tctgcactgg actctcaagg ctctgttcca cctcttcttt 1260

atgccctcca tcctgacctc gctcctggtg ttctttgttt ccgagctcgt cggtggcttc 1320

ggaattgcca tcgtggtctt catgaaccac taccctctgg agaagatcgg tgatcccgtc 1380

tgggacggac atggcttctc tgtgggtcag atccatgaga ccatgaacat tcgacgaggc 1440

atcattactg actggttctt tggaggcctg aactaccaga tcgagcacca tctctggccc 1500

accctgcctc gacacaacct cactgccgtt tcctaccagg tggaacagct gtgccagaag 1560

cacaacctcc cctaccgaaa ccctctgccc catgaaggtc tcgtcatcct gctccgatac 1620

ctggccgtgt tcgctcgaat ggccgagaag cagcccgctg gcaaggctct ctaagcggcc 1680

gccaccgccg agattccggc ctcttcggcc gccaagcgac ccgggtggac gtctagaggt 1740

acctagcaat taacagatag tttgccggtg ataattctct taacctccca cactcctttg 1800

acataacgat ttatgtaacg aaactgaaat ttgaccagat attgtgtccg cg 1852

<210> 355
<211> 1898
<212> DNA
<213> Artificial Sequence

<220>
<223> chimeric D8S-2::XPR gene

<400> 355
cgatcaggag agaccgggtt ggcggcgtat ttgtgtccca aaaaacagcc ccaattgccc 60

caattgaccc caaattgacc cagtagcggg cccaaccccg gcgagagccc ccttcacccc 120

acatatcaaa cctccccccgg ttcccacact tgccgttaag ggcgtagggt actgcagtct 180

ggaatctacg cttgttcaga ctttgtacta gtttctttgt ctggccatcc gggtaaccca 240

tgccggacgc aaaatagact actgaaaatt tttttgcttt gtggttggga ctttagccaa 300

gggtataaaa gaccaccgtc cccgaattac ctttcctctt cttttctctc tctccttgtc 360

aactcacacc cgaaatcgtt aagcatttcc ttctgagtat aagaatcatt caccatgaag 420

tccaagcgac aggctctgtc tcccctccag ctgatggaac agacctacga cgtcgtgaac 480

ttccaccctg gtggagctga aatcattgag aactaccagg gacgagatgc tactgacgcc 540

ttcatggtta tgcactttca ggaagccttc gacaagctca agcgaatgcc caagatcaac 600

ccctcctttg agctgcctcc ccaggctgcc gtcaacgaag ctcaggagga tttccgaaag 660

ctccgagaag agctgatcgc cactggcatg tttgacgcct ctcccctctg gtactcgtac 720

aagatctcca ccaccctggg tcttggcgtg cttggatact tcctgatggt ccagtaccag 780

atgtacttca ttggtgctgt gctgctcggt atgcactacc agcaaatggg atggctgtct 840

catgacatct gccaccacca gaccttcaag aaccgaaact ggaataacct cgtgggtctg 900

gtctttggca acggactcca gggcttctcc gtgacctgtt ggaaggacag acacaacgcc 960

catcattctg ctaccaacgt tcagggtcac gatcccgaca ttgataacct gcctcccctc 1020

gcctggtccg aggacgatgt cactcgagct tctcccatct cccgaaagct cattcagttc 1080

caacagtact atttcctggt catctgtatt ctcctgcgat tcatctggtg tttccagtgc 1140

gtgctgaccg ttcgatccct caaggaccga gacaaccagt tctaccgatc tcagtacaag 1200

aaagaggcca ttggactcgc tctgcactgg actctcaagg ctctgttcca cctcttcttt 1260

atgccctcca tcctgacctc gctcctggtg ttctttgttt ccgagctcgt cggtggcttc 1320

ggaattgcca tcgtggtctt catgaaccac taccctctgg agaagatcgg tgatcccgtc 1380

tgggacggac atggcttctc tgtgggtcag atccatgaga ccatgaacat tcgacgaggc 1440

atcattactg actggttctt tggaggcctg aactaccaga tcgagcacca tctctggccc 1500

accctgcctc gacacaacct cactgccgtt tcctaccagg tggaacagct gtgccagaag 1560

cacaacctcc cctaccgaaa ccctctgccc catgaaggtc tcgtcatcct gctccgatac 1620

ctggccgtgt tcgctcgaat ggccgagaag cagcccgctg gcaaggctct ctaagcggcc 1680

gccaccgcgg cccgagattc cggcctcttc ggccgccaag cgacccgggt ggacgtctag 1740

aggtacctag caattaacag atagtttgcc ggtgataatt ctcttaacct cccacactcc 1800

tttgacataa cgatttatgt aacgaaactg aaatttgacc agatattgtg tccgcggtgg 1860

agctccagct tttgttccct ttagtgaggg ttaattaa 1898

<210> 356
<211> 45
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer ODMW392

<400> 356
gaacagacct acgacgtctc cgcttgggtg aacttccacc ctggt 45

<210> 357
<211> 45
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer ODMW393

<400> 357
accagggtgg aagttcaccc aagcggagac gtcgtaggtc tgttc 45

<210> 358
<211> 19
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer Eg5-1

<400> 358
gaaatgaagt caaagcgcc 19

<210> 359
<211> 19
<212> DNA
<213> Artificial Sequence

<220>

<223> Primer Eg3-3

<400> 359
ccttatagag ccttccccg 19

<210> 360
<211> 22
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer T7

<400> 360
gtaatacgac tcactatagg gc 22

<210> 361
<211> 19
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer M13-28Rev

<400> 361
ggaaacagct atgaccatg 19

<210> 362
<211> 19
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer Eg3-2

<400> 362
ttggcaatgg tctgcaagg 19

<210> 363
<211> 19
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer Eg5-2

<400> 363
aatgttcatg gtctcatgg 19

<210> 364
<211> 29
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer ODMW404

<400> 364
cctggtacca tgaagtccaa gcgacaggc 29

<210> 365
<211> 1272
<212> DNA
<213> Artificial Sequence

<220>
<223> chimeric gene

<400> 365
catgaagtcc aagcgacagg ctctgtctcc cctccagctg atggaacaga cctacgacgt 60

ctccgcttgg gtgaacttcc accctggtgg agctgaaatc attgagaact accagggacg 120

agatgctact gacgccttca tggttatgca ctttcaggaa gccttcgaca agctcaagcg 180

aatgcccaag atcaacccct cctttgagct gcctccccag gctgccgtca acgaagctca 240

ggaggatttc cgaaagctcc gagaagagct gatcgccact ggcatgtttg acgcctctcc 300

cctctggtac tcgtacaaga tctccaccac cctgggtctt ggcgtgcttg gatacttcct 360

gatggtccag taccagatgt acttcattgg tgctgtgctg ctcggtatgc actaccagca 420

aatgggatgg ctgtctcatg acatctgcca ccaccagacc ttcaagaacc gaaactggaa 480

taacctcgtg ggtctggtct ttggcaacgg actccagggc ttctccgtga cctgttggaa 540

ggacagacac aacgcccatc attctgctac caacgttcag ggtcacgatc ccgacattga 600

taacctgcct cccctcgcct ggtccgagga cgatgtcact cgagcttctc ccatctcccg 660

aaagctcatt cagttccaac agtactattt cctggtcatc tgtattctcc tgcgattcat 720

ctggtgtttc cagtgcgtgc tgaccgttcg atccctcaag gaccgagaca accagttcta 780

ccgatctcag tacaagaaag aggccattgg actcgctctg cactggactc tcaaggctct 840

gttccacctc ttctttatgc cctccatcct gacctcgctc ctggtgttct ttgtttccga 900

gctcgtcggt ggcttcggaa ttgccatcgt ggtcttcatg aaccactacc ctctggagaa 960

gatcggtgat cccgtctggg acggacatgg cttctctgtg ggtcagatcc atgagaccat 1020

gaacattcga cgaggcatca ttactgactg gttctttgga ggcctgaact accagatcga 1080

gcaccatctc tggcccaccc tgcctcgaca caacctcact gccgtttcct accaggtgga 1140

acagctgtgc cagaagcaca acctccccta ccgaaaccct ctgccccatg aaggtctcgt 1200

catcctgctc cgatacctgg ccgtgttcgc tcgaatggcc gagaagcagc ccgctggcaa 1260

ggctctctaa gc 1272

<210> 366
<211> 80
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer YL521

<400> 366
tttccatggt gaagtccaag cgacaggctc tgcccctcac catcgacgga actacctacg 60

acgtctccgc ttgggtgaac 80

<210> 367
<211> 30
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer YL522

<400> 367
tggagatctt gtacgagtac cagaggggag 30

<210> 368
<211> 37
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer YL525

<400> 368
ccttcatggt tatgcactct caggaagcct tcgacaa 37

<210> 369
<211> 37
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer YL526

<400> 369
ttgtcgaagg cttcctgaga gtgcataacc atgaagg 37

<210> 370
<211> 38
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer YL527

<400> 370
ccaagatcaa cccctcctcc gagctgcctc cccaggct 38

<210> 371
<211> 38
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer YL528

<400> 371
agcctgggga ggcagctcgg aggaggggtt gatcttgg 38

<210> 372
<211> 37
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer YL529

<400> 372
gggcttctcc gtgacctggt ggaaggacag acacaac 37

<210> 373
<211> 37
<212> DNA

<213> Artificial Sequence

<220>
<223> Primer YL530

<400> 373
gttgtgtctg tccttccacc aggtcacgga gaagccc 37

<210> 374
<211> 38
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer YL531

<400> 374
acattgataa cctgcctctg ctcgcctggt ccgaggac 38

<210> 375
<211> 38
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer YL532

<400> 375
gtcctcggac caggcgagca gaggcaggtt atcaatgt 38

<210> 376
<211> 38
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer YL533

<400> 376
tcatctggtg tttccagtct gtgctgaccg ttcgatcc 38

<210> 377
<211> 38
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer YL534

<400> 377
ggatcgaacg gtcagcacag actggaaaca ccagatga 38

<210> 378
<211> 39
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer YL535

<400> 378
ctgcactgga ctctcaagac cctgttccac ctcttcttt 39

<210> 379
<211> 39
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer YL536

<400> 379
aaagaagagg tggaacaggg tcttgagagt ccagtgcag 39

<210> 380
<211> 37
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer YL537

<400> 380
ctggagaaga tcggtgattc cgtctgggac ggacatg 37

<210> 381
<211> 37
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer YL538

<400> 381
catgtccgtc ccagacggaa tcaccgatct tctccag 37

<210> 382
<211> 34
<212> DNA
<213> cow cat

<400> 382
ataacttcgt ataatgtatg ctatacgaag ttat 34

<210> 383
<211> 45
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer 436

<400> 383
tatctgctgt cgtggttaat taacgcggac acaatatctg gtcaa 45

<210> 384
<211> 40
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer 437

<400> 384
ctgcaggtaa aggccggcct cgacgcagta ggatgtcctg 40

<210> 385
<211> 38
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer 475

<400> 385
ccatccacac ttgcggccgc ttatttgttg agatacgc 38

<210> 386
<211> 38
<212> DNA
<213> Artificial Sequence

<220>

<223> Primer 477

<400> 386
acaagcttgc ggccgcttat ttgttgagat acgcatcc 38

<210> 387
<211> 35
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer 478

<400> 387
agtggctctc cataaacctg acgaaattaa agagg 35

<210> 388
<211> 40
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer 476

<400> 388
ctagtggatc cgaattccat gggtggcgga ggacagcaga 40

<210> 389
<211> 36
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer 505

<400> 389
catcatctct ttcccacagt gcctcggcac aactac 36

<210> 390
<211> 36
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer 506

<400> 390
gtagttgtgc cgaggcactg tgggaaagag atgatg 36

<210> 391
<211> 39
<212> DNA
<213> Artificial Sequence

<220>
<223> SMART IV oligonucleotide

<400> 391
aagcagtggt atcaacgcag agtggccatt acggccggg 39

<210> 392
<211> 59
<212> DNA
<213> Artificial Sequence

<220>
<223> CDSIII/3'PCR primer

<220>
<221> misc_feature
<222> (28)..(57)
<223> thymidine (dT); see BD Biosciences Clontech's SMART cDNA technology

<220>
<221> misc_feature
<222> (59)..(59)
<223> n is a, c, g, or t

<400> 392
attctagagg ccgaggcggc cgacatgttt ttttttttt ttttttttt ttttttvn 59

<210> 393
<211> 23
<212> DNA
<213> Artificial Sequence

<220>
<223> 5'-PCR primer

<400> 393
aagcagtggt atcaacgcag agt 23

<210> 394
<211> 18
<212> DNA

<213> Artificial Sequence

<220>
<223> Primer M13

<400> 394
tgtaaaacga cggccagt 18

<210> 395
<211> 33
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer MLPAT-F

<400> 395
gatcccatgg tcgggtccgt cacccgaccc aca 33

<210> 396
<211> 34
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer MLPAT-R

<400> 396
gatcgcggcc gctcacatcc agttctctga ccac 34

<210> 397
<211> 32
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer LPAT-Re-5-1

<400> 397
gatcgagctc gacattgagc attgatctat tt 32

<210> 398
<211> 41
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer LPAT-Re-5-2

<400> 398
gactatcgat accatggtcg ggcaatgaga acggcagcaa g 41

<210> 399
<211> 1129
<212> DNA
<213> Yarrowia lipolytica

<400> 399
gatcgagctc gacattgagc attgatctat ttgtttagtt tagtacgtac gtagagtagt 60

tggttattgt atactggata ctggtcggat agtcgtgtac aagtacagta tatacttgta 120

gtttggaaca ctagtcacaa gtatagttaa taacagtata gtatatactt ctactgtaca 180

tatagttaca aaacagtaca gtatatactt ctactgtaca tactgtagtt tcatggttga 240

tataggtgaa aatgacatgg tatcccgctt cccaatgact gatacgagca acaaacaatt 300

ctctctaata gatgttcatt ataaccacac tgtacggtaa ttatatcgta gacagtgtac 360

agtattagaa cgtgtttttg atcccaagtt ggtggaaatg ttgagttgag atgagatgat 420

aatgttatca gcatgagggg tggatgtata cagtaggaga aggagtatgg gcagttgtgc 480

ggtagcagcg tctctgaatc acacaaagcc gggcttacag cagaggcaag aatggcgcca 540

tgtgcctgcg caatcaaggg cgtgaaccac ggcttgattc cacggcttga tgccacagct 600

taatgccaca tgccacagca taataccacg gcatgacccc gctgactcca accttcattt 660

cggcacgtgt aggtgcacaa gggacttcaa gaggggccaa tttcatgcgg acacatggcg 720

caaaaaacgc ccgactttga ttacacagac acgtaataac gacgaagccg agatgagcac 780

acgtggccaa gtctgccaat ggccccctgg acccccctga caaagtttcc caacaagccc 840

agccgtgcat ggtgtgtttt tgtgcggaga cacacgccaa ttaggctcat ttgagggtat 900

gcagcgaaaa aaaattagtg tgggtagttt gtttgcagga atcaagtggg tggttgaaaa 960

acaagaaaga gcgacgacaa gagagagaga aaaagagaga gagactccat aaagcgtgca 1020

tcaaaattaa ggtgtgtgac tatccgaaaa ccaaacatga acagttggat atatgtcgct 1080

gtgattgcag ttgctgccgt tctcattgcc cgaccatgga tcgatgatc 1129

<210> 400
<211> 40
<212> DNA

<213> Artificial Sequence

<220>
<223> Primer LPAT-Re-3-1

<400> 400
gactatcgat gcggccgcag actgcagcac aagaagtgct 40

<210> 401
<211> 40
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer LPAT-Re-3-2

<400> 401
gatcctcgag tctagagcag ggtgttggag gggatggaga 40

<210> 402
<211> 938
<212> DNA
<213> Yarrowia lipolytica

<400> 402
gactatcgat gcggccgcag actgcagcac aagaagtgct tgtagctact ttaggagaga 60

gataggtaat atgaaacatt tttcagatcg acacccacgg cgaaccattg gctgtggagc 120

tatgggtgaa tggattaata tagcaacgaa atctacctcg attaccaacg caaaacgagc 180

ccactttctc tgtactgtgc tatatcgtgt ataccccagt tctcttagat cagtcgtacc 240

cgcccgtact gtcatggact attcaaactg ttggttttta gtctactcct acttgtacca 300

tgcttatgct cgatgtacca gctacttcgt ccgcggacgt gtctctataa taaccctgaa 360

agctgctcat tcctctaaaa gtcgtgacac cacctagtct tactcatcga gattgtatta 420

ctcatcgaga ttgtattact catcgagatt gtattactca tcttgatttt ctctcgtacg 480

atgttgttag taatcatggc cttccagccg aactctaccg accgaggtaa cttcacaagg 540

tttagacaga gtccaattat tcgaggagcc cataagtcta cctttaccta ctctactgta 600

gtgtgaagat gatcatgaga ggcatagatt ggattggaca gttggacggg tgacgaggga 660

gcagacaagt gcgagattgg cgttgttata tggactccag ggaccgtgaa ttgctcttaa 720

actggccgat attagcggct ggaaagatcc ttgtcccact ctcttctttg tctatcctgg 780

tgatccattt ttttgacagc taaaattaac tcgccgaaac acgtcggcat accgagatta 840

aattacaaaa gaagattaaa acgcgaccag caaaccgcaa cttccgcttt actcaacctc 900

tccatcccct ccaacaccct gctctagact cgaggatc 938

<210> 403
<211> 28
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer PZUF-mod1

<400> 403
gatcccatgg atccaggcct gttaacgg 28

<210> 404
<211> 32
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer PZUF-mod2

<400> 404
gatcgcggcc gcagacatga taagatacat tg 32

<210> 405
<211> 32
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer MACAT-F1

<400> 405
gatcccatgg cagagtcgac aacaacgaca tg 32

<210> 406
<211> 34
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer MACAT-R

<400> 406
gatcgcggcc gctcaaacat ccatatggcg cttg 34

<210> 407
<211> 28
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer MDGAT-F

<400> 407
gatcccatgg cgctctttgc gcctttac 28

<210> 408
<211> 35
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer MDGAT-R1

<400> 408
gatcgcggcc gcctattcga tgatgcacag ttcct 35

<210> 409
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer MGPAT-N1

<220>
<221> misc_feature
<222> (3)..(3)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (9)..(9)
<223> n is a, c, g, or t

<400> 409
ccncaygcna aycarttygt 20

<210> 410
<211> 20
<212> DNA

<213> Artificial Sequence

<220>
<223> Primer MGPAT-NR5

<220>
<221> misc_feature
<222> (6)..(6)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (9)..(9)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (15)..(15)
<223> n is a, c, g, or t

<400> 410
ttccangtng ccatntcrtc 20

<210> 411
<211> 22
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer MGPAT-5N1

<400> 411
catctctcgg ctgatccaga tg 22

<210> 412
<211> 21
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer MGPAT-5N2

<400> 412
ttgtctggcg ctacctggaa c 21

<210> 413
<211> 22
<212> DNA

<213> Artificial Sequence

<220>
<223> Primer MGPAT-5N3

<400> 413
cattgtcgct cgtgtcatca gc 22

<210> 414
<211> 44
<212> DNA
<213> Artificial Sequence

<220>
<223> Genome Walker adaptor-1

<400> 414
gtaatacgac tatagggcac gcgtggtcga cggcccgggc tggt 44

<210> 415
<211> 8
<212> DNA
<213> Artificial Sequence

<220>
<223> Genome Walker adaptor-2

<220>
<221> misc_feature
<222> (1)..(1)
<223> 5' end is associated with a -PO4 group

<220>
<221> misc_feature
<222> (8)..(8)
<223> 3' end is associated with a -H2N group

<400> 415
accagccc 8

<210> 416
<211> 22
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer MGPAT-5-1A

<400> 416
cagctccttt gtgaatttag tg 22

<210> 417
<211> 22
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer AP1

<400> 417
gtaatacgac tcactatagg gc 22

<210> 418
<211> 23
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer MGPAT-3N1

<400> 418
cataacgatc cagaagtttg atg 23

<210> 419
<211> 19
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer AP2

<400> 419
actatagggc acgcgtggt 19

<210> 420
<211> 32
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer mgpat-cdna-5

<400> 420
gatcggatcc tcagatctac gactttgtgt cg 32

<210> 421
<211> 33
<212> DNA

<213> Artificial Sequence

<220>
<223> Primer MGPAT-cDNA-R

<400> 421
gatcgcggcc gcctaaatgt cttttgactt ggc 33

<210> 422
<211> 25
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer MA Elong 3'1

<400> 422
gatacggatc agagcctgtg gaaca 25

<210> 423
<211> 25
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer MA elong 3'2

<400> 423
cttctacctt tcaaagtttt acgag 25

<210> 424
<211> 25
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer MA Elong 5'1

<400> 424
gcataggtcc agactccatt atgcg 25

<210> 425
<211> 25
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer MA Elong 5'2

<400> 425
cagaggaata taggtcggca ccaaa 25

<210> 426
<211> 24
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer MA ELONG 5' NcoI 3

<400> 426
gatcccatgg agtctggacc tatg 24

<210> 427
<211> 30
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer MA ELONG 3' NotI 1

<400> 427
gatcgcggcc gcttactggg ccttcttctg 30

<210> 428
<211> 28
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer YL597

<400> 428
tttccatggt ctcgtcaatc tcgccgac 28

<210> 429
<211> 33
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer YL598

<400> 429
tttgcggccg cttagatgga cttggccttc aag 33

<210> 430
<211> 32
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer YL325

<400> 430
tttgcggccg catggagcgt gtgttctgag tc 32

<210> 431
<211> 46
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer YL326

<400> 431
cccttaatta actcacctgc aggattgaga ctatgaatgg attccc 46

<210> 432
<211> 36
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer YL567

<400> 432
tttccatggc tgccgtccct attgaattca acgtcc 36

<210> 433
<211> 24
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer YL568

<400> 433
gatcagctca aagtacttgg tcag 24

<210> 434
<211> 24
<212> DNA

<213> Artificial Sequence

<220>
<223> Primer YL569

<400> 434
ctgaccaagt actttgagct gatc 24

<210> 435
<211> 36
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer YL570

<400> 435
tttgcggccg cttatgctcg tcgagatcgg gtagtg 36

<210> 436
<211> 36
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer YL571

<400> 436
tcctccacac ctaccatcac ggtgccactg ctcttc 36

<210> 437
<211> 36
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer YL572

<400> 437
gaagagcagt ggcaccgtga tggtaggtgt ggagga 36

<210> 438
<211> 28
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer CPT1-5'-NcoI

<400> 438
acgtccatgg gcgtattcat taaacagg 28

<210> 439
<211> 33
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer CPT1-3'-NotI

<400> 439
acgtgcggcc gcttattcgg tcttcttctc ctg 33

<210> 440
<211> 30
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer Isc1F

<400> 440
ataccatgga caacagaaaa gacagagatg 30

<210> 441
<211> 31
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer Isc1R

<400> 441
tgcggccgct catttctcgc tcaagaaagt t 31

<210> 442
<211> 30
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer Pcl1F

<400> 442
ataccatggg tgaatacagc aaaggctctt 30

<210> 443
<211> 33
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer Pcl1R

<400> 443
tgcggccgct taaaacccat gttgactcat gat 33

<210> 444
<211> 29
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer P95

<400> 444
ggcaagctta ttgtcgttgg tggagcaca 29

<210> 445
<211> 35
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer P96

<400> 445
aattccacca gatctgtcgt ggtattcgga cactt 35

<210> 446
<211> 39
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer P97

<400> 446
ataccacgac agatctggtg gaattgccac cgagggagc 39

<210> 447
<211> 30
<212> DNA

<213> Artificial Sequence

<220>
<223> Primer P98

<400> 447
gcggaattcg cagatagact ggtttcgctt 30

<210> 448
<211> 22
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer P115

<400> 448
aactacatct tcggctatca cc 22

<210> 449
<211> 22
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer P116

<400> 449
tgaacaagcg tagattccag ac 22

<210> 450
<211> 22
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer P112

<400> 450
caccettgct cggcgatgta tc 22

<210> 451
<211> 34
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer P39

<400> 451
ggcggtaccg gatcctcaat cgaagagact aagc 34

<210> 452
<211> 41
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer P41

<400> 452
cttctgtatt ctagatctca agatcgagaa gcaccagaaa a 41

<210> 453
<211> 48
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer P40

<400> 453
gcttctcgat cttgagatct agaatacaga agtcaacggt tcatccat 48

<210> 454
<211> 30
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer P42

<400> 454
ccggaattca gctttgagct tggagaagta 30

<210> 455
<211> 22
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer P51

<400> 455
tagatagact ggactatacg gc 22

<210> 456
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer P52

<400> 456
gactgtccta ccctgatttg 20

<210> 457
<211> 33
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer P37

<400> 457
ccaggtacca agatcgagaa gcaccagaaa agc 33

<210> 458
<211> 36
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer P38

<400> 458
ctcgaattca gaatacagaa gtcaacggtt catcca 36

<210> 459
<211> 41
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer P201

<400> 459
ctcgcggccg ccatggaggt ccgacgacga aarathgayg t 41

<210> 460
<211> 46
<212> DNA

<213> Artificial Sequence

<220>
<223> Primer P203

<220>
<221> misc_feature
<222> (38)..(38)
<223> n is a, c, g, or t

<400> 460
gaggcggccg ctactggttc tgcttgtagt tgtaggcnar rtarta 46

<210> 461
<211> 44
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer P214

<400> 461
atctcgacaa tcgtcgcagc cctcctcaag agcatcgtcc agaa 44

<210> 462
<211> 42
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer P215

<400> 462
cgaacagatc ccaatattac atgaggcgag tttgagagac ag 42

<210> 463
<211> 44
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer P216

<400> 463
aactggtatt taaatgatgt ccccaagacg gagcgtattc gacc 44

<210> 464

<211> 40
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer P217

<400> 464
gtagttgtag gctaggtagt aaaggaatgc acaagtgggt 40

<210> 465
<211> 40
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer P219

<400> 465
cgaatacgct ccgtcttggg gacatcattt aaataccagt 40

<210> 466
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer P226

<400> 466
cgacgaaaaa ttgacgtgct 20

<210> 467
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer P227

<400> 467
gatttccgaa cagatcccaa 20